

Ministério da Saúde

**FIOCRUZ**  
**Fundação Oswaldo Cruz**

## **INSTITUTO OSWALDO CRUZ**

### **Pós-Graduação em Biologia Celular e Molecular**

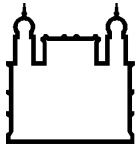
**AUTOR: HARRISON MAGDINIER GOMES**

Identificação Molecular de Cepas do Complexo *Mycobacterium tuberculosis* e Detecção de Resistência a Antibióticos

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Biologia Celular e Molecular

**Orientador:** Dr. Philip Noel Suffys

**Rio de Janeiro**  
2006



Ministério da Saúde

**FIOCRUZ**

**Fundação Oswaldo Cruz**

**INSTITUTO OSWALDO CRUZ**  
**Pós-Graduação em Biologia Celular e Molecular**

***AUTOR: HARRISON MAGDINIER GOMES***

**IDENTIFICAÇÃO MOLECULAR DE CEPAS DO COMPLEXO *Mycobacterium tuberculosis* E DETECÇÃO DE RESISTÊNCIA A ANTIBIÓTICOS**

**ORIENTADOR: Dr. Philip Noel Suffys**

**Aprovada em: 10/ 08/ 2006**

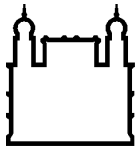
**EXAMINADORES:**

**Dr. Adalberto Rezende Santos - Presidente**

**Dr. Douglas MacIntosh**

**Dr. Luis Cláudio de Oliveira Lazzarini**

Rio de Janeiro, 10 de Agosto de 2006



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## INSTITUTO OSWALDO CRUZ

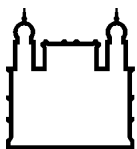
### IDENTIFICAÇÃO MOLECULAR DE CEPAS DO COMPLEXO *Mycobacterium tuberculosis* E DETECÇÃO DE RESISTÊNCIA A ANTIBIÓTICOS

#### RESUMO

#### TESE DE DOUTORADO

Harrison Magdinier Gomes

A tuberculose mata aproximadamente 3 milhões de pessoas no mundo a cada ano. A pobreza, a falta de informação, a demora no diagnóstico de cepas resistentes e multi-resistentes e fatores de risco associados à infecção por *Mycobacterium tuberculosis* são os principais responsáveis pela manutenção de sua transmissão. Em nosso estudo, utilizamos duas abordagens moleculares, baseadas na reação de PCR, “Single Strand Conformational Polymorphism-SSCP” e “Reverse Line Blot Hybridization”-RLBH, para verificarmos genótipos mutantes e a frequência de mutações no genoma de *M. tuberculosis* associadas à resistência a drogas, assim como avaliarmos a RLBH para a detecção rápida de resistência. Utilizando o SSCP, verificamos que a versão radioativa apresentou uma melhor resolução que a não radioativa. Na versão não-radioativa uma discordância foi observada (9,1%) ao comparar os resultados com a análise por sequenciamento. Na RLBH, uma membrana contendo sondas para detecção de resistência a: rifampicina, isoniazida, estreptomicina e etambutol foi avaliada em um estudo multi-cêntrico e 50% das cepas resistentes a rifampicina apresentavam mutações em *rpoB* nos códons 531-TTG (43%) e 526-GTC (7%) enquanto que 49% das cepas resistentes a etambutol apresentavam mutações no códon 306-GTG (34%) e ATA (14%) do gene *embB*. Uma segunda versão da RLBH, para detecção de resistência a rifampicina com sondas específicas para o gene *rpoB* foi avaliada. Analisamos 50 cepas de *Mycobacterium Bovis*, sensíveis e 157 cepas de *M. tuberculosis* sensíveis e resistentes a rifampicina. Todas as cepas de *M. bovis* apresentaram o genótipo de sensível e entre as cepas resistentes de *M. tuberculosis*, 93% apresentaram mutações na região de 157pb do gene *rpoB*. O Inno-LiPA (teste comercial), baseado em RLBH, para detecção de cepas de *M. tuberculosis* resistente a rifampicina, também foi analisado, apresentando uma sensibilidade de 97,6%. Concluímos que a hibridação reversa tem a sensibilidade adequada para detectar cepas de *M. tuberculosis* resistentes a rifampicina e que os mecanismos relacionados à resistência a estreptomicina e isoniazida necessitam de uma melhor compreensão. Na parte deste estudo referente à variabilidade genética de cepas brasileiras de *M. tuberculosis*, avaliamos a composição de uma região específica do Complexo *M. tuberculosis*, a “Direct repeat” (DR), pela técnica de “spoligotyping”. Através desta técnica determinamos os genótipos de cepas de *M. tuberculosis*, de pacientes, provenientes de 11 estados brasileiros e os comparamos com “spoligotypes” de 122 países depositados no Banco de Dados Internacional de “spoligotypes”. 80% das cepas brasileiras pertencem às famílias dos “spoligotypes” “Latin-American and Mediterrânea” (LAM; 46,5%), T (19,7%) and Haarlem (13,3 %). 302 “spoligotypes” novos estão distribuídos entre todas as famílias. Dentro da família LAM, uma deleção denominada de RD<sup>Rio</sup> de 26 kB foi detectada e esta deleção está relacionada a 30% dos casos de TB na cidade do Rio de Janeiro. A distribuição de cepas de *M. tuberculosis* no Brasil parece estar relacionada à colonização brasileira e ao moderno fluxo migratório.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## INSTITUTO OSWALDO CRUZ

### IDENTIFICAÇÃO MOLECULAR DE CEPAS DO COMPLEXO *Mycobacterium tuberculosis* E DETECÇÃO DE RESISTÊNCIA A ANTIBIÓTICOS

#### ABSTRACT

#### TESE DE DOUTORADO

Harrison Magdinier Gomes

Each year, tuberculosis kills, approximately, 3 million people worldwide. Poverty, lack of knowledge (information), late diagnosis of resistant and multi-resistant strains and risk factors associated to *Mycobacterium tuberculosis* infection are the main reasons (principally responsible) for the maintenance of the (ongoing) transmission. In our study, two molecular approaches, based on PCR amplification, have been performed, “Single Strand Conformational Polymorphism - SSCP” and “Reverse Line Blot Hybridization - RLBH”, for the identification of mutant genotypes and the frequency of mutations present in *M. tuberculosis* genome associated to drug resistance, as well as the evaluation of RLBH for the rapid detection of resistance. The radioactive version of SSCP presented a better resolution than the non radioactive version. In the non-radioactive version, discordance of 9.1% was observed when the results were compared to those obtained by sequencing. Concerning the RLBH, a membrane containing probes designed to detect resistance to: rifampin, isoniazid, streptomycin and ethambutol was evaluated in a multi-center study and 50% of the rifampin resistant strains showed mutations in *rpoB* in codons 531-TTG (43%) and 526-GTC (7%), while 49% of the resistant strains to ethambutol presented mutations in codons 306-GTG (34%) and ATA (14%) from *embB* gene. A second version of RLBH, designed for the detection of rifampin resistance by specific probes for the *rpoB* gene, was evaluated. Fifty sensitive strains of *Mycobacterium bovis* and 157 sensitive and rifampin-resistant strains of *M. tuberculosis* were analyzed. All *M. bovis* strains presented a sensitive genotype, while 93% of the *M. tuberculosis* resistant strains presented mutations in the 157 bp region of the *rpoB* gene. The Inno-LiPA (commercial test), based in the RLBH for the detection of strains resistant to rifampin, was also analyzed presenting a sensitivity of 97.6%. We have concluded that the RLBH has an adequate sensitivity to detect rifampin-resistant strains and that the mechanisms involved in resistance to streptomycin and isoniazid need a better understanding. Concerning the part of this study related to the genotypic variability of the Brazilian strains of *M. tuberculosis*, we have evaluated the composition of the *M. tuberculosis* complex specific region, the “Direct repeat” (DR) locus, by the spoligotyping technique. Through this technique, we have determined the genotypes of *M. tuberculosis* strains from patients attended in 11 Brazilian states, which were compared to spoligotypes from 122 countries deposited in the International Database of Spoligotypes. Eighty percent of the Brazilian strains belong to the spoligotyping families “Latin-American and Mediterranean” (LAM; 46.5%), T (19.7%) and Haarlem (13.3%). Three-hundred-and-two new spoligotypes are distributed in all families. Within the LAM family, a deletion called RD<sup>Rio</sup> was detected and it is related to 30% of TB cases in the city of Rio de Janeiro. The distribution of strains of *M. tuberculosis* in Brazil seems to be related to the Brazilian colonization and the modern migratory flux.

*“O meu coração contemplou abundantemente a sabedoria e a ciência. E apliquei o meu coração a conhecer a sabedoria e a conhecer os desvarios e as loucuras, e vim, a saber, que também isto era aflição de espírito. Porque na muita sabedoria há muito enfado, e o que aumenta em ciência, aumenta em trabalho”.* *Eclesiastes*

**Sonho dos poetas:** *porque gado a gente tange, ferra, engorda e mata, mas com gente é diferente.*

*Disparada, GV & TB*

**A realidade de um povo:** *a diferença está no direito do gado à engorda.*

*À Deus que cuidadosamente moldou a molécula da vida*

*Aos que sofrem... Que sofrem desta e de outras mazelas*

*Aos meus queridos pais Noêmia e João*

*As Mulheres de minha vida: Isis, Thaís e Marianne*

## **Agradecimentos**

Os meus agradecimentos são para os queridos (as) companheiros (as) de toda a minha vida, os de “longa” e os de “curta” data. Para aqueles que como eu, durante 25 anos, acreditaram na coerência de um partido e para aqueles que também com coerência se mantiveram a direita ou no centro ou que preferiram não estar em lugar nenhum ou não acreditam em nenhum destes rumos.



## Lista de Ilustrações

	Página
<b>Figura 1.1</b> - Estimativa das taxas de incidência de TB, OMS, 2003. Reproduzido de Global Tuberculosis Control (WHO Report 2005).	3
<b>Figura 1.2</b> - Mutações localizadas entre os códons 507 e 533 do gene <i>rpoB</i> de 478 cepas de <i>M. tuberculosis</i> resistentes a rifampicina. Aqui estão incluídas mutações missense, inserções e deleções. O número dos códons estão com base no alinhamento da seqüência de <i>rpoB</i> de <i>E. coli</i> com a seqüência homóloga do <i>rpoB</i> de <i>M. tuberculosis</i> .	6
<b>Figura 1.3</b> - Cromossomo de uma cepa <i>x</i> hipotética de <i>M. tuberculosis</i> e genotipagem (IS6110-RFLP, MIRU-VNTR e spoligotyping) do <i>M. bovis</i> BCG, de <i>M. tuberculosis</i> H37Rv e da cepa <i>x</i> (adaptada de Barnes e Cave 2003).	12
<b>Figura 1.4</b> - Posição dos 41 MIRU-VNTR no cromossomo do <i>M. tuberculosis</i> H37Rv. O primeiro número indica a posição de cada locus ocupado por um MIRU-VNTR. A letra c indica orientação contrária a estabelecida por Cole et al, em 1998. Algarismos romanos indicam o tipo de MIRU (I, II ou III). Os números seguintes indicam a localização de cada loci. A esfera em preto indica os 12 loci com número variável de MIRU-VNTR.	19
<b>Figura 1.5</b> - Cenário evolutivo dos organismos do complexo <i>M. tuberculosis</i> . (Sreevatsan et al, 1997).	22
<b>Figura 1.6</b> - Esquema proposto para evolução do bacilo da TB ilustrando as deleções (retângulos cinzas). O esquema é baseado na presença ou ausência de regiões deletadas conservadas e no polimorfismo em 5 genes selecionados. Em azul isolados caracterizados como pertencentes ao grupo 1, em verde isolados pertencentes ao grupo 2 e em vermelho isolados pertencentes ao grupo 3. (Brosch et al. 2002).	23
<b>Tabela 1.1</b> - Análise do banco de dados SpolDB3 mostrando os “spoligotypes” protótipos, regras para o reconhecimento visual e descrição binária, adaptado de Filliol et al, (2002).	17

## Abreviaturas

CAS	<i>Central Asia</i> Asia Central
DR	<i>Direct Repeat</i>
EAI	<i>East African and Indian</i> Leste Africano e Índia
EUA	Estados Unidos da América
INNO-LiPa	Innogenetics Line Probe Assay
IUATLD	<i>International Union Against Tuberculosis and Lung Disease</i> União Internacional Contra Tuberculose e Doenças Pulmonares
Kb	Kilobase
LAM	<i>Latin American and Mediterranean</i> América Latina e Mediterrâneo
<i>M. africanum</i>	<i>Mycobacterium africanum</i>
<i>M. bovis</i>	<i>Mycobacterium bovis</i>
<i>M. canettii</i>	<i>Mycobacterium canettii</i>
<i>M. microti</i>	<i>Mycobacterium microti</i>
MDR	<i>Multi Drug Resistance</i> Resistência a Múltiplas Drogas
MGIT	<i>Mycobacterial Growth Indicator Tube</i>
MIC	<i>Minimal Inhibitory Concentration</i> Concentração Inibitória Mínima
MIRU	<i>Micobacterial Interspersed Repetitive Unit</i>
MS	Ministério da Saúde
pb	par de base
PCR	<i>Polymerase Chain Reaction</i> - Reação em Cadeia da Polimerase
PE/PPE	<i>Proline-Glutamic Acid / Proline-Proline-Glutamic Acid</i>
PGG ou MGG	<i>Principal Genetic Group</i> ou <i>Major Genetic Group</i> Grupo Genético Principal
RD	Região de Diferença
RFLP	<i>Restriction Fragment Length Polymorphism</i> Polimorfismo do Comprimento do Fragmento de Restrição

RLBH	<i>Reverse Line Blot Hybridization</i> Hibridação Reversa em Linha
SSCP	<i>Single Strand Conformation Polymorphism</i> Polimorfismo Conformacional de Fita Simples
TB	Tuberculose
VNTR	<i>Variable Number and Tandem Repeat</i>
WHO	<i>World Healthy Organization</i> Organização Mundial da Saúde

## Sumário

		Página
	<b>Resumo</b>	iii
	<b>Abstract</b>	iv
	<b>Lista de Ilustrações</b>	ix
	<b>Lista de Abreviaturas</b>	xii - xi
	<b>Sumário</b>	xii-xiv
Capítulo I	<b>Introdução</b>	1-28
I.1.	Generalidades	2-4
I.2.	Resistência em <i>M. tuberculosis</i> às drogas utilizadas no tratamento da tuberculose	4-7
I.3.	Métodos de detecção de resistência em cepas de <i>M. tuberculosis</i> às drogas utilizadas no tratamneto da tuberculose	7
I.3.1.	Métodos fenotípicos de detecção de resistência às drogas utilizadas no tratamento da tuberculose	8-9
I.3.2.	Métodos genotípicos de detecção de resistência às drogas utilizadas no tratamento da tuberculose	9-10
I.4.	Aspectos Gerais da genotipagem de cepas do complexo <i>M. tuberculosis</i>	10-21
I.5.	Filogenia e evolução de cepas do complexo <i>M. tuberculosis</i>	21-28
Capítulo II	<b>Objetivos</b>	29-30
II.1.	Detecção de resistência a antibióticos em cepas do complexo <i>M. tuberculosis</i>	30
II.2.	Caracterização molecular de cepas do complexo <i>M. tuberculosis</i>	30
Capítulo III	<b>Materiais, métodos e resultados</b>	31-143
Parte 1	<b>Detecção de resistência em cepas de <i>M. tuberculosis</i></b>	31-73

III.1.	<b>Artigo 1</b> - Evaluation of polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP) analysis for the detection of the <i>rpoB</i> mutations associated with resistance to rifampicin in <i>M. tuberculosis</i>	32-39
III.2.	<b>Artigo 2</b> - Multicenter evaluation of reverse line blot assay for detection of drug resistance in <i>M. tuberculosis</i> clinical isolates	40-53
III.3.	<b>Artigo 3</b> - Rapid detection of resistance against rifampicin in isolates of <i>M. tuberculosis</i> from Brazilian patients using a reverse phase hybridization assay	54-62
III.4.	<b>Artigo 4</b> - In house reverse line blot assay for rapid detection of susceptibility to rifampicin in isolates of <i>M. tuberculosis</i>	63-68
III.5.	<b>Artigo 5</b> - Drug susceptibility of Brazilian strains of <i>M. bovis</i> using traditional and molecular techniques	69-73
Parte 2.	<b>Tipagem molecular de cepas de <i>M. tuberculosis</i></b>	74-142
III.6.	<b>Artigo 6</b> - <i>M. tuberculosis</i> complex genetic diversity: mining the fourth international spoligotyping database for classification, population genetics and epidemiology	75-92
III.7.	<b>Manuscrito 1</b> - Discovery of a new <i>M. tuberculosis</i> genotype that is the most predominant cause of TB in Rio de Janeiro, Brazil	93-127
III.8	<b>Manuscrito 2</b> - Distribution of <i>M. tuberculosis</i> strains in Brazil as reflected by spoligotyping	128-143
Capítulo IV	<b>Discussão e Conclusão</b>	144-162
IV.1	Discussão	145-161
IV.2	Conclusão	162-163
Capítulo V	Bibliografia	164-172

Anexo I	Página 1 do Banco de Dados dos “spoligotypes” brasileiros	173-174
Anexo II	Página 1 do Banco de Dados Internacional de “spoligotypes” SpolDB4	175-176

*Capitulo I - Introdução*

## 1.1. Generalidades

Em 24 de Março de 1882, em uma reunião na “*Berlin Physiological Society*”, Robert Koch anunciou o isolamento do micróbio causador da tuberculose (TB): o *Mycobacterium tuberculosis*. Desde então o estudo da TB e do *M. tuberculosis* desenvolveu-se de forma a resultar na prevenção, no tratamento e em sua cura.

A TB, como muitas outras doenças que afligem a humanidade, está diretamente relacionada aos aspectos sociais, econômicos e políticos das diversas populações. Após várias epidemias na Europa e nos Estados Unidos da América houve um declínio de TB, até o início dos anos 80, na maioria dos países industrializados, no entanto a baixa qualidade de vida nos chamados países do terceiro mundo e a co-infecção com HIV tornaram-se significativos fatores relacionados à presença contínua da TB em vários continentes (Roberts *et al.*, 1998).

*M. tuberculosis* é a espécie micobacteriana responsável pela maioria dos casos de TB e pertence ao Complexo *M. tuberculosis*, que reúne seis espécies com identidade genômica de 99,9%: *M. tuberculosis*, *M. bovis*, *M. africanum* subtipo I (grupo A, concentrado na África Ocidental e *M. bovis-like*) e subtipo II (grupo B, África Oriental e *M. tuberculosis-like*) *M. pinnipedii*, *M. microtti* e *M. canettii* (Kremer *et al.*, 1999). Entre estas, *M. bovis* causa TB em humanos e em animais enquanto *M. tuberculosis* e *M. africanum* infectam principalmente o homem, com raros casos em outras espécies animais (Gutiérrez *et al.*, 1999).

O sucesso de algumas destas espécies em causar TB é tão significativo que uma estimativa global prediz que uma em cada três pessoas está infectada por *M. tuberculosis*, e este fato continuará representando um sério problema à vida se as abordagens sociais, econômicas e políticas dos governos, em seus diferentes extratos, não mudarem seus rumos. Atualmente, 8 milhões de pessoas são infectadas por *M. tuberculosis*, com 2 a 3 milhões de mortes por ano. Caso as condições continuem as mesmas, estima-se que mais de um bilhão de pessoas serão infectadas por *M. tuberculosis*, 200 milhões irão adoecer e 35 milhões morrerão. Para enfrentar esta tendência, em meados dos anos 90, a Organização Mundial da Saúde (OMS) passou considerar a TB, como emergência mundial.

O cenário acima descrito se reflete de forma dramática, principalmente, nos países pobres e em desenvolvimento, e no contexto deste trabalho é imperativo descrever a situação da TB no Brasil. Nosso país hoje ocupa a 14ª posição entre os 23 países que concentram 80% dos casos desta doença, e de acordo com o Ministério da Saúde, 50 milhões de brasileiros (as) estão infectados (as) por *M. tuberculosis*, com 111.000 novos casos, e 6.000 mortes por ano, com uma prevalência de 58 casos/100.000 habitantes (II Consenso Brasileiro de Tuberculose, 2004).

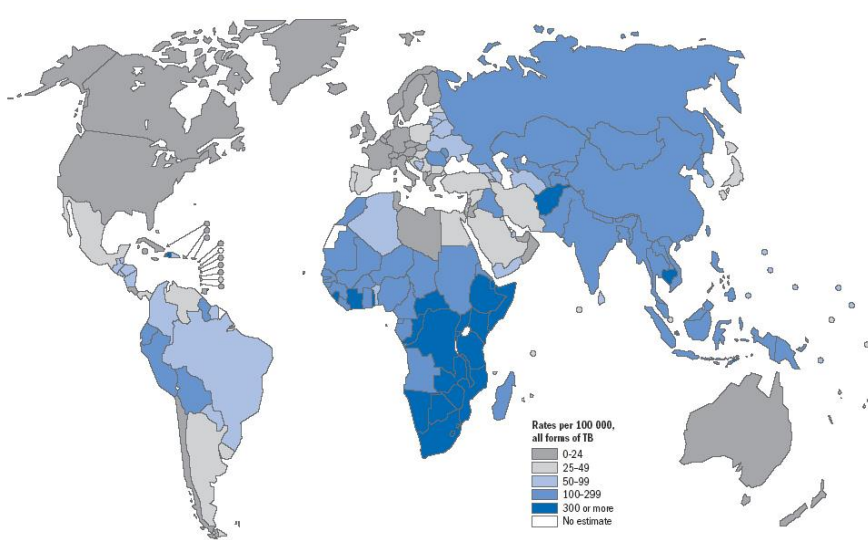
No final do século 20, três importantes Estados brasileiros: Rio de Janeiro, São Paulo e Paraná com taxas de incidência de TB de 160, 55 e 28 casos por 100.000 habitantes respectivamente, destacam as disparidades observadas nas diferentes regiões do Brasil, estas disparidades são o reflexo da condição social, política e econômica do Brasil, além de ser o resultado direto do desmonte do Programa Nacional de Controle da



TB no início dos anos 90 (Kritski *et al.*, 2000). Nestes grandes centros urbanos a epidemia causada pelo Vírus da Imunodeficiência Humana (HIV) foi também a responsável pelo aumento dos casos de TB, que está presente em 25% dos indivíduos infectados pelo HIV (Kritski *et al.*, 2000; Burger *et al.*, 1998). Entre março de 2000 e dezembro de 2003 foram notificados, no Brasil, 1469 casos de MDR-TB (*Multidrug-resistant tuberculosis*), com 975 (66,4%) casos concentrados na Região Sudeste, e destes, 617 (42%) casos somente no Estado do Rio de Janeiro e a frequência de associação com a infecção pelo HIV foi de 6,8% (II Consenso Brasileiro de Tuberculose, 2004).

O mapa abaixo (Figura I.1) mostra claramente a situação do Brasil no mundo, que, através de suas instâncias federais decidiram, no início da década de 1990, que a TB estava sobre controle, representando pouco impacto na comunidade (Kritski *et al.*, 2000).

Figura I.1. Estimativa das taxas de incidência de TB, OMS, 2003.



Reproduzido de: Global Tuberculosis Control. WHO Report 2005.

Em 1998, o Ministério da Saúde brasileiro declarou o controle da TB como prioridade máxima, desde então o governo federal e os governos estaduais e municipais, além de organizações civis, vem trabalhando para o efetivo controle da TB.

## I.2. Resistência de *M. tuberculosis* as Drogas Utilizadas no Tratamento da Tuberculose

Antes da Segunda Guerra Mundial e do desenvolvimento da moderna utilização dos antibióticos a TB matava de 50% a 60% dos casos que desenvolviam a doença. Entretanto, a descoberta das propriedades antimicrobianas da estreptomicina em 1944 e da isoniazida e pirazinamida em 1952, conduziu para a redução da

mortalidade em todo o mundo. Esta tendência progrediu de 1953, quando foram instituídos os programas de controle da TB nos Estados Unidos até meados dos anos 80, quando o número de novos casos voltou a aumentar nos Estados Unidos. Qual foi a causa do retorno desta doença? São várias as possibilidades: o surgimento do vírus da AIDS no início dos anos 80 foi um importante fator, assim como a deteriorização da qualidade de vida dentro das áreas urbanas. Aumento da população de rua e a má qualidade dos serviços de saúde e dos programas de controle da TB. Em muitos casos a negligência com a quimioterapia anti-TB pode resultar no aparecimento da resistência a drogas.

O tratamento anti-TB, consiste da administração das chamadas drogas de primeira linha: isoniazida, rifampicina e pirazinamida por dois meses, seguidas pela administração, por um período de quatro meses, de isoniazida e rifampicina. Em caso de falha deste esquema, o tratamento terá continuidade com a administração das chamadas drogas de segunda linha, são elas: estreptomicina, etambutol, etionamida e pirazinamida por três meses, seguidas por etambutol e etionamida por mais nove meses.

A falha de um esquema terapêutico relaciona-se à resistência de *M. tuberculosis* às drogas utilizadas no tratamento da TB ou a particularidades da interação parasito-hospedeiro, e esta resistência surgirá quando um paciente é submetido a um esquema terapêutico inadequado, ou quando não observa as recomendações para a ingestão do medicamento, ou abandona o tratamento ou por fatores ainda desconhecidos.

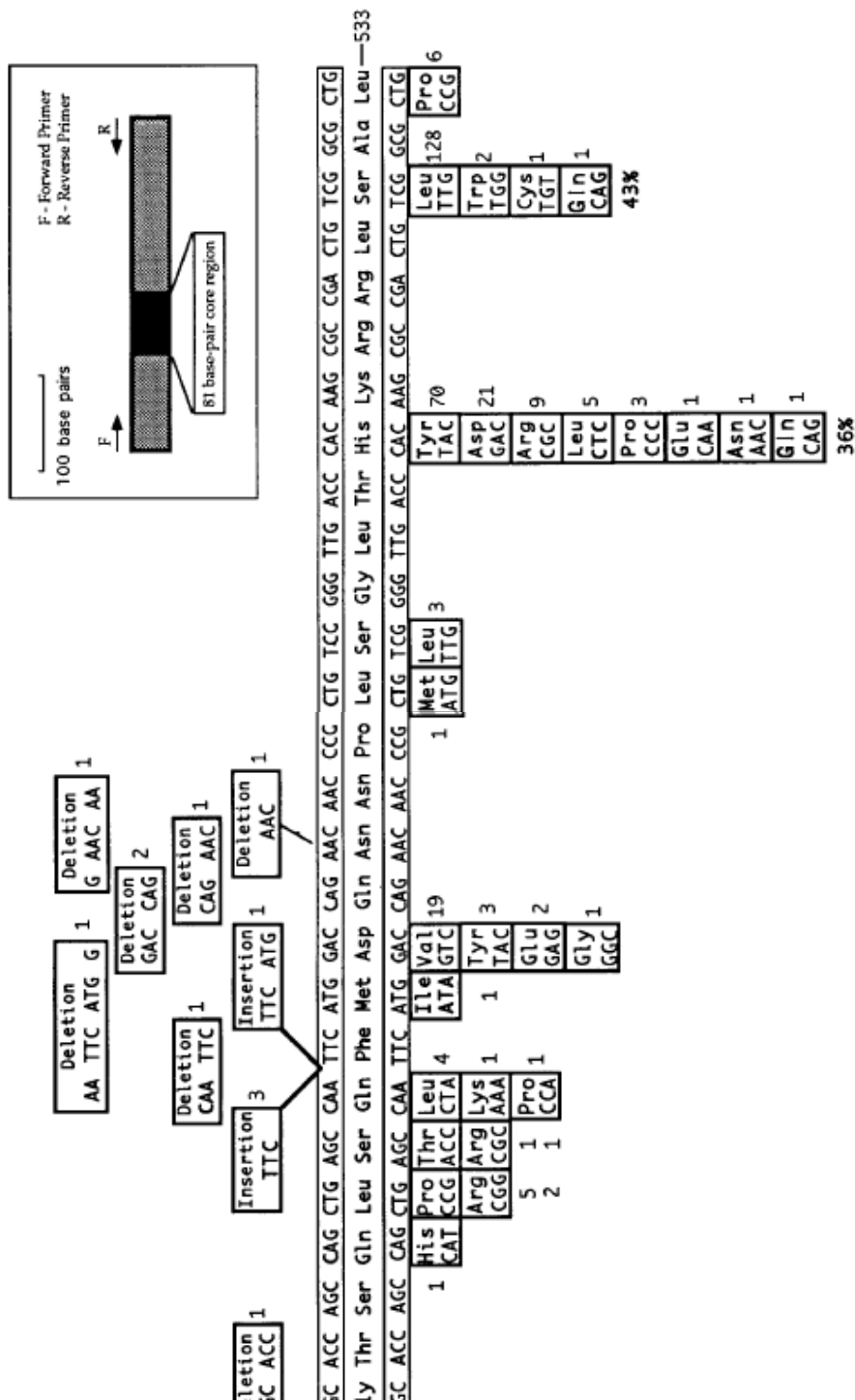
A resistência de *M. tuberculosis* às drogas é caracterizada como resistência primária, quando cepas resistentes são isoladas de pacientes novos que não foram tratados com drogas anti-TB ou que se infectaram por uma cepa resistente proveniente de outro paciente tratado com droga. A resistência primária também é conhecida como resistência natural (Sharma e Moham, 2004 e Cohn *et al.*, 1997).

Parte dos mecanismos moleculares de resistência de *M. tuberculosis* contra as principais drogas utilizadas no tratamento da TB foi elucidada na década passada. Mutações que levam à alteração da subunidade  $\beta$  da RNA polimerase conferem resistência à rifampicina, enquanto que mutações no gene *katG*, que codifica uma catalase-peroxidase, conferem resistência à isoniazida (a catalase-peroxidase converte a isoniazida à sua forma biologicamente ativa); mutações nos genes *inhA* e *kasA*, também estão relacionadas à resistência à isoniazida (Ramaswamy e Musser, 1998). Os genes que codificam para a subunidade 16S do RNA ribossomal e para a proteína ribossomal S12 (*rpsL*), quando sofrem mutações, conferem resistência em *M. tuberculosis* à estreptomicina, enquanto que mutações no gene *pncA* e *embB* conferem resistência à pirazinamida e ao etambutol respectivamente.

Outro importante grupo de drogas utilizadas no tratamento da TB são as fluoroquinolonas. A resistência de *M. tuberculosis* às fluoroquinolonas se dá quando ocorrem mutações nos genes *gyrA* e *gyrB*. Estes genes codificam para as subunidades da DNA girase (Ramaswamy e Musser, 1998).

Dos genes envolvidos na resistência a drogas, o gene *rpoB* é o que está melhor caracterizado. Este gene codifica para a subunidade  $\beta$  da RNA polimerase e as principais mutações relacionadas à resistência neste gene, estão localizadas entre os códons 507 e 533, que totaliza 81pb. A ligação da rifampicina a subunidade  $\beta$  da RNA

polimerase inibe a iniciação da transcrição, portanto mutações “nonsense” ou pequenas inserções e deleções na região central do gene resultam na modificação da interação entre a droga e a enzima e em consequência a atividade transcricional da RNA polimerase é abolida (Ramaswamy e Musser, 1998). Na figura I.2 podemos observar as mutações “nonsense” e deleções/inserções descritas na região de 81pb do gene *rpoB*, relacionadas a resistência de *M. tuberculosis* a rifampicina. É também importante ressaltar que um estudo com 329 cepas resistentes a rifampicina, mostrou que 4% das cepas resistentes a esta droga não apresentavam mutações neste gene, isto significa dizer que diferentes mecanismos de resistência estão disponíveis em cepas de *M. tuberculosis* resistentes a drogas (Ramaswamy e Musser, 1998).



**Figura 1.2** - Mutações localizadas entre os códons 507 e 533 do gene *rpoB* de 478 cepas de *M. tuberculosis* resistentes a rifampicina. Aqui estão incluídas mutações missense, inserções e deleções. O número dos códons está de acordo com o alinhamento da sequência de *rpoB* de *E. coli* com a sequência homóloga do *rpoB* de *M. tuberculosis* (Musser, 1995).

A resistência simultânea de cepas de *M. tuberculosis* a rifampicina e a isoniazida é conhecida como TB resistente a múltiplas drogas (MDR-TB) (Bifani et al., 1996). As cepas MDR aumentam significativamente o risco de morte do paciente e contribuem para a dispersão da TB na comunidade (Sharma e Mohan, 2004).

A gravidade do problema da MDR-TB foi documentada pela WHO-IUATLD (*World Health Organization – International Union Against Tuberculosis and Lung Disease*) que avaliou a resistência a drogas de 1ª linha entre 1994 e 1997 em 35 países. Esta análise mostrou que a resistência estava presente nos 35 países, o que destacou a relevância global do problema (WHO, 1997). Um recente levantamento mostrou que 3,2% dos casos de MDR-TB são casos novos (resistência primária) de TB. São considerados com resistência primária aqueles pacientes sem contato prévio com medicamentos anti-TB ou que nunca foram tratados.

### **I.3. Métodos de Detecção de Resistência às Drogas em Cepas de *M. tuberculosis* Utilizadas no Tratamento da Tuberculose**

Métodos para detecção de resistência estão sendo desenvolvidos e novas versões de métodos são comparadas às anteriores. Os testes de sensibilidade são recomendados, naqueles casos em que houve falência do tratamento, em casos de re-tratamento, em casos de suspeita de resistência primária ou em caso de contatos de pacientes portadores de TB resistente.

As metodologias de detecção de resistência são classificadas como: Fenotípicas e Genotípicas. Uma importante diferença entre estes dois grupos é o tempo necessário para obtenção dos resultados, que é crucial no tratamento da TB. O Sistema Bactec, os métodos que utilizam bacteriófagos, o E-test, o MIC e outros são fenotípicos.

E entre os métodos genotípicos, podemos destacar: o método de DNA heteroduplex, INNO-Lipa, SSCP e o seqüenciamento de DNA. Estes métodos serão descritos abaixo.

### I.3.1. Métodos Fenotípicos de Detecção de Resistência às Drogas Utilizadas no Tratamento da Tuberculose

Está entre as metodologias fenotípicas, o Método das Proporções descrito por Canetti e colaboradores em 1963, foi aprovado pela OMS e é atualmente o mais utilizado no Brasil, além de recomendado no Manual de Bacteriologia da Tuberculose, de publicação do Ministério da Saúde brasileiro (Ministério da Saúde, 1994 e II Consenso Brasileiro de Tuberculose, 2004). Este método determina a relação proporcional entre cepas resistentes e sensíveis presentes na amostra que contém *M. tuberculosis*. Esta amostra é semeada em meio Löwenstein-Jensen contendo diferentes concentrações das drogas utilizadas no tratamento anti-TB e com base no número de unidades formadoras de colônias no meio sem a droga e no meio contendo a droga, determina-se a proporção entre cepas sensíveis e resistentes ao antibiótico. Para cada droga está definida a proporção de mutantes resistentes em uma população bacteriana, em valores iguais ou acima desta proporção, a amostra é considerada resistente (Ministério da Saúde, 1994). O método das proporções é o “padrão ouro” para detecção de resistência, sensível e de baixo custo com resultados disponibilizados em 60 dias.

Existem outros métodos de detecção de resistência aprovados pela OMS e utilizados no Brasil, como o da concentração absoluta e o da razão de resistência. Estes métodos são baseados no crescimento de *M. tuberculosis* em meio de cultura, o que resulta em um tempo demasiado longo (aproximadamente dois meses) para se definir o padrão de suscetibilidade e conseqüentemente, para o estabelecimento do esquema terapêutico adequado e/ou para a interrupção da cadeia de transmissão de cepas resistentes.

O MGIT (*Mycobacterial Growth Indicator Tube*) consiste de um tubo (fundo em “U”) contendo meio 7H9 com 0,25% de glicerol. No fundo deste tubo repousa uma camada de silicone contendo um indicador fluorescente que é inibido na presença de oxigênio. Uma cepa de *M. tuberculosis* em crescimento ativo no meio 7H9 consome o oxigênio do tubo e então o indicador fluoresce (Bergmann *et al.*, 1997).

O sistema Bactec 460<sup>®</sup> TB (*Becton e Dickinson Instruments System*, Sparks, MD, USA), detecta o crescimento micobacteriano e baseia-se na captação do CO<sub>2</sub> radioativo liberado pela Micobactéria após o consumo do ácido palmítico (C<sup>14</sup>) presente no meio 7H12A e 7H12B (Ministério da Saúde, 1994). O Bactec-MGIT 960<sup>®</sup> utiliza a mesma tecnologia de fluorescência do MGIT e do Sistema Bactec 9000<sup>®</sup> MB. O Bactec 460<sup>®</sup> e o Bactec 960<sup>®</sup> estão validados pela OMS e ANVISA (Agência Nacional de Vigilância Sanitária).

O LRP (*Luciferase Reporter Phage*), o FASTPlaqueTB<sup>®</sup> (Biotech Laboratories LTD, Ipswich, Reino Unido) e o Phage Tek<sup>®</sup>MB (BioMérieux, Marcy l’Etoile, França) são métodos que utilizam bacteriófagos na identificação e detecção de resistência, através da determinação da viabilidade do *M. tuberculosis*. No LRP a viabilidade das micobactérias é detectada através da emissão de luz, e está validado para detecção de resistência a rifampicina e isoniazida. O FASTPlaqueTB<sup>®</sup> e Phage Tek<sup>®</sup>MB estão disponíveis em versão comercializada enquanto o LRP está em fase de análise. Estes métodos foram descritos recentemente e a maioria não está sendo aplicada de rotina.

No E-test, uma fita de nitrocelulose contendo um gradiente de concentração de uma determinada droga é

colocada sobre uma placa de Middlebrook 7H11 suplementado com OADC e previamente semeada com a cepa de *M. tuberculosis* a ser testada. Após cinco a sete dias de incubação, através da observação do halo de inibição próxima a fita, é possível estabelecer o MIC para cada droga utilizada (Wanger e Mills, 1996).

### **I.3.2. Métodos Genotípicos de Detecção de Resistência às Drogas Utilizadas no Tratamento da Tuberculose**

O método de heteroduplex de DNA consiste de uma mistura de proporções iguais do produto de PCR, de uma cepa sensível e de uma cepa resistente, que é desnaturado e em seguida hibridado. O híbrido é fracionado em gel de agarose e visualizado com brometo de etídio. O heteroduplex, resultante da hibridação da fita de DNA da cepa sensível com o da cepa resistente, exibe uma mobilidade eletroforética diferente quando comparada a um híbrido homoduplex.

Alguns “kits” comerciais também estão disponíveis no mercado para detecção de resistência, dentre estes, destacamos um método baseado em Lipa (*Line Probe Assay*), capaz de detectar mutações que causam resistência a rifampicina (INNO-Lipa Rif TB, Innogenetics N.V., Gent, Bélgica). Neste método, oligonucleotídeos, que correspondem à região de 81pb do gene *rpoB* de *M. tuberculosis*, estão covalentemente ligados a uma membrana de nitrocelulose e são hibridados com os produtos amplificados pela PCR das amostras a serem testadas (De Beenhouwer *et al.*, 1995).

O método SSCP (*Single Strand Conformation Polymorphism*) baseia-se na diferença de migração, em um gel de poliacrilamida não-desnaturante, de uma molécula de DNA fita simples, que assume diferentes conformações de acordo com sua composição de bases. Uma mutação pontual gera uma diferença na estrutura do fragmento de DNA, que influencia sua migração no gel (Telenti *et al.*, 1993).

O seqüenciamento de DNA tem sido utilizado com sucesso na identificação de espécies, na tipagem molecular de diferentes espécies, na identificação de marcadores moleculares, na identificação de “*single nucleotide polymorphisms*” e na detecção de mutações relacionadas à resistência em micobactérias. A desvantagem deste método é o fato de requerer grande investimento em equipamentos e insumos, contudo é o padrão para validar outros métodos indiretos de detecção de mutações (Telenti *et al.*, 1993).

Os métodos de detecção de resistência representam um avanço no combate a TB, auxiliando na cura e interrompendo a sua cadeia de transmissão, porém os métodos fenotípicos exigem um longo tempo até a liberação do resultado, enquanto que os métodos genotípicos necessitam de um preciso estabelecimento da relação custo-efetividade que permita sua utilização no dia-a-dia do diagnóstico da TB.

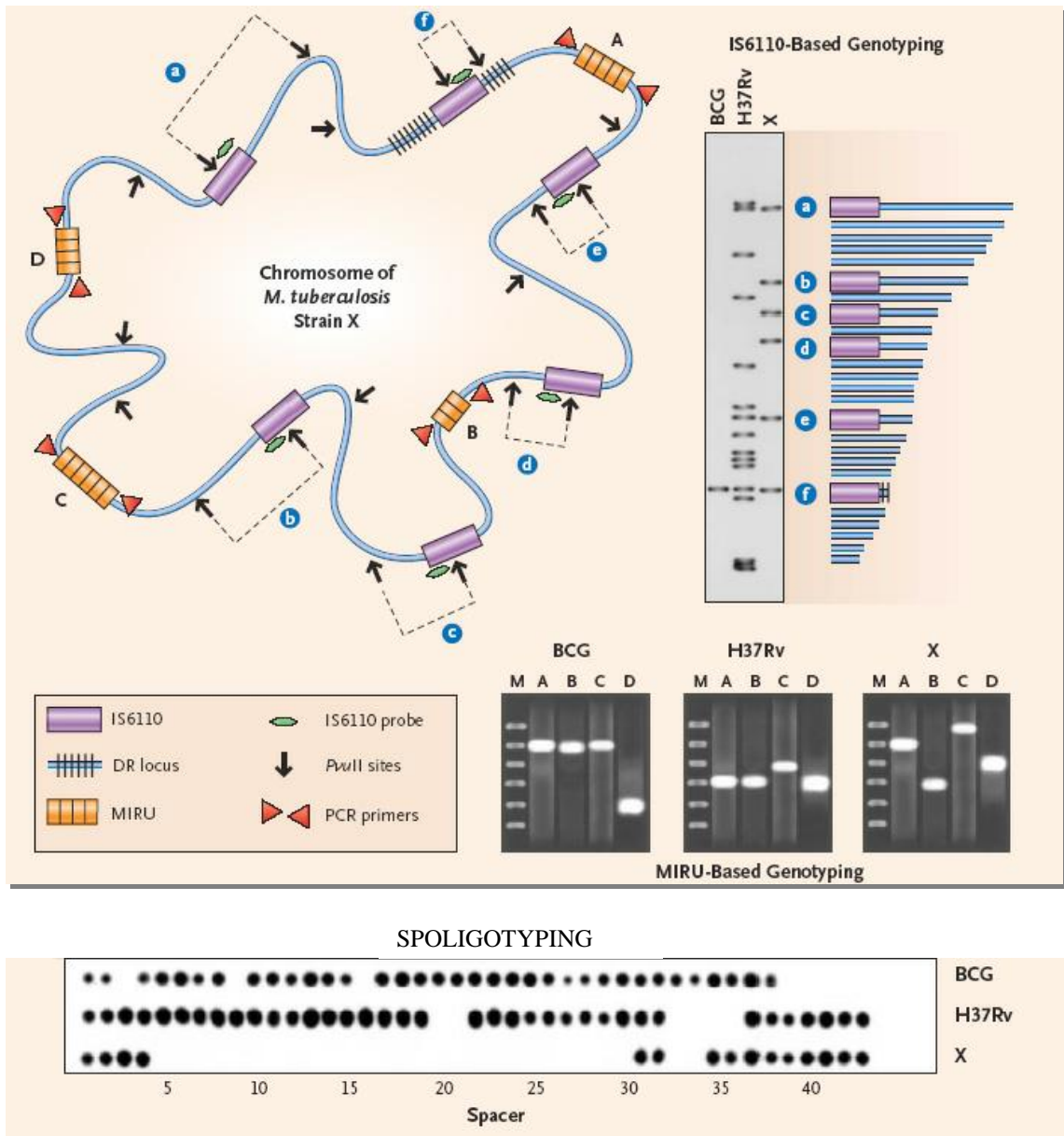
Outras abordagens, além das relacionadas à resistência aos fármacos, vêm sendo desenvolvidas para uma melhor compreensão da TB e de *M. tuberculosis*. A redução no tempo de diagnóstico de resistência e a interrupção da cadeia de transmissão da doença são importantes aspectos estudados pela Epidemiologia Molecular, que aborda o problema através de várias metodologias de genotipagem, portanto a Biologia Molecular tem sido utilizada com êxito no diagnóstico de resistência a drogas de cepas de *M. tuberculosis*.

#### **I.4. Aspectos Gerais da Genotipagem de Cepas do Complexo *M. tuberculosis***

Na genotipagem de *M. tuberculosis*, O IS6110-RFLP (*Restriction-Fragment-Length Polimorphism*), o “*spoligotyping*” (*Spacer Oligonucleotide Typing*) e “MIRU-VNTR typing” (*Mycobacterial Interspersed Repeat Units – Variable Number Tandem Repeat Typing*) são as metodologias que melhor se adequaram ao seu estudo e as mais utilizadas. O IS6110-RFLP, padrão ouro para genotipagem de *M. tuberculosis*, analisa a distribuição da seqüência de inserção 6110 (IS6110) no genoma bacteriano (van Embden *et al.*, 1993). O “*spoligotyping*” baseia-se na visualização do polimorfismo da região DR (*Direct Repeat*), presente exclusivamente no genoma de cepas do complexo *M. tuberculosis* (Kameebek *et al.*, 1997). Os MIRU-VNTR são VNTRs (minissatélites) que estão presentes em muitas cópias, no genoma de *M. tuberculosis*, variando em sua seqüência e comprimento (Supply *et al.*, 2001) (Figura I.3).

A capacidade de distinguir cepas pertencentes ao complexo *M. tuberculosis*, tornou-se mais efetiva quando estudos de epidemiologia molecular foram conduzidos associando o IS6110-RFLP e o “*spoligotyping*” (Horgen *et al.*, 1998). Adicionalmente, observou-se que a associação de duas metodologias de tipagem, “*spoligotyping*” e DRE-PCR (*Double-Repetitive-Element-PCR*), para a epidemiologia molecular da TB, apresenta uma habilidade discriminatória similar ao método de referência IS6110-RFLP (Sola *et al.*, 1998).





**Figura 1.3** - Cromossomo de uma cepa *x* hipotética de *M. tuberculosis* e genotipagem (IS6110-RFLP, MIRU-VNTR e spoligotyping) de *M. bovis* BCG, de *M. tuberculosis* H37Rv e da cepa *x* (adaptada de Barnes e Cave, 2003).

O transposon IS6110 é um dos responsáveis pelo polimorfismo no genoma do Complexo *M. tuberculosis*, esta seqüência de 1361pb é exclusiva deste complexo, e embora seu número de cópias possa variar entre 0 e 20, a maioria das cepas carrega entre oito e 15 cópias em seu genoma (Thierry *et al.*, 1990, Supply *et al.*, 1997). Devido a estas características a IS6110 tem sido amplamente utilizada como marcador molecular para diferenciar

isolado de *M. tuberculosis*, de diversas origens (Kremer *et al.*, 1999).

O IS6110-RFLP exibe com clareza as relações entre cepas de *M. tuberculosis* em investigações epidemiológicas realizadas em indivíduos HIV-positivos, em prisões, em hospitais e albergues e na população em geral. Apesar deste tipo de investigação ser possível, graças ao alto poder discriminatório deste método, cepas que apresentam menos que seis cópias da IS6110 são inadequadamente caracterizadas, nestes casos as relações epidemiológicas podem ser estabelecidas com uma segunda metodologia de genotipagem.

Até o início da década de 1990 considerava-se que a maioria dos casos de TB nos países em desenvolvimento resultava da reativação de uma infecção anteriormente adquirida e nos Estados Unidos da América, 90% dos casos de TB eram relacionados à reativação. Esta estimativa foi questionada em um dos primeiros estudos utilizando o RFLP na genotipagem de um grupo de pacientes tuberculosos, da cidade de Nova Iorque; neste estudo 62,5% dos casos estavam relacionados a perfis órfãos (reativação da infecção) enquanto que os 37,5% dos casos restantes foram agrupados em 12 “clusters” (transmissão recente) (Alland *et al.*, 1994).

Um estudo com os isolados de todos os casos de TB diagnosticados entre 1991 e 1992, em São Francisco, resultou em uma taxa de transmissão recente, semelhante a do estudo anterior, embora o número de amostras analisadas nesta cidade tenha sido quase cinco vezes superior ao de Nova Iorque (Small *et al.*, 1994).

Algumas condições podem ser estabelecidas para a definição de transmissão recente, tais como: cepas com ligação epidemiológica que apresentam padrões genotípicos idênticos e ocorrem ao longo de um tempo limitado, cepas com perfil genotípico idêntico e mesmo padrão de resistência e o aparecimento de crianças com TB na comunidade. Observando estes critérios, podemos considerar que a transmissão recente estimada nestes estudos pode não representar a realidade, já que um “cluster” pode incluir um caso índice fruto de uma reativação ou mesmo um caso índice qualquer (Small *et al.*, 1994).

A partir daí, estudos que identificam a transmissão recente da TB e seus fatores de risco, em diferentes países, utilizando o RFLP, contribuíram para melhorar a compreensão de como a TB é transmitida. Devido ao grande número de trabalhos apresentando dados obtidos com métodos de tipagem molecular, podemos afirmar que tanto em áreas de baixa incidência, como em áreas de alta incidência de TB estas metodologias confirmam ou estabelecem: a possibilidade de transmissão recente, a cadeia de transmissão pessoa a pessoa, fatores de risco envolvidos, os locais de dispersão, assim como se as ligações epidemiológicas são estáveis ou temporárias. (Narayanan *et al.*, 2004; Small *et al.*, 1994; Pfyfer *et al.*, 1998; van Deutekon *et al.*, 1997).

No Brasil, analisando 11 diferentes postos de saúde da cidade do Rio de Janeiro (140 amostras), não foi possível estabelecer um fator de risco para os “clusters” obtidos (Oelemann, submetido), do mesmo modo que em um estudo incluindo pacientes com TB do HUCFF (Hospital Universitário Clementino Fraga Filho da UFRJ) a taxa de “cluster” não se relacionou com: gênero, idade, infecção pelo HIV, tipo de residência, albergados, indigentes ou história prévia de TB, no entanto esta taxa esteve relacionada a cepas MDR (Fandinho *et al.*, 2000).

A identificação de fatores de risco para transmissão da TB, na cidade de Bauru (São Paulo), mostrou que, de 57 pacientes submetidos ao IS6110-RFLP, 26,3% estavam em “clusters”, e o número destes “clusters” não se mostrou diferente entre pacientes provenientes de hospitais e de postos de saúde. E não se estabeleceu relação entre esta taxa de “cluster” e fatores de risco para transmissão da TB (Baptista *et al.*, 2002), diferente de uma área endêmica no Brasil, onde as maiores taxas de “clusters” estavam relacionados a pacientes HIV+ (38%) e a TB-MDR (63%) (Ferrazoli *et al.*, 2000).

A TB como um importante problema de saúde pública nas cidades acima citadas, exige que medidas de controle e interrupção da cadeia de transmissão sejam implantadas. Em Nova Iorque, a implantação destas medidas foi responsável pela significativa redução do número de casos, inclusive os casos de resistência a drogas, além disso, a utilização da estratégia de terapia diretamente observada (DOTS), na qual o profissional de saúde vê o paciente tomando o medicamento, foi primordial nesta interrupção (Hamburg *et al.*, 1995). De semelhante forma após a identificação de uma região com alta taxa de “cluster” na cidade do Rio de Janeiro (Área Programática 1), estabeleceu-se como estratégia de controle da transmissão, a implantação da estratégia DOT (Oelemann, submetido).

A genotipagem de cepas de *M. tuberculosis* por RFLP pode consumir mais de cinco semanas e necessita de grande quantidade de DNA, além disto, o IS6110-RFLP é incapaz de detectar polimorfismos em cepas que não carregam ou carregam entre um e cinco cópias do transposon IS6110 (Sola *et al.*, 1998), portanto métodos baseados na reação de PCR, que não necessitam de grandes quantidades de DNA para a genotipagem, e podem ser aplicados a amostras clínicas, como biopsias e esfregaços em lâminas coradas pelo método de Ziehl-Neelsen, são alternativas viáveis para a genotipagem de cepas de *M. tuberculosis*.

O “*spoligotyping*” como um destes métodos detecta e tipa cepas do Complexo *M. tuberculosis*, baseado no polimorfismo do locus DR, que é formado por uma seqüência conservada de 36 pb, denominada DR, que está presente em um número de 49 cópias em BCG e 39 em *M. tuberculosis* H37rv, entre as seqüências DRs estão localizadas as seqüências espaçadoras (“spacers”) não repetitivas cujo tamanho varia de 34 a 41pb (Hermans *et al.*, 1991).

As cepas de *M. tuberculosis* variam de acordo com o número de DRs e com a presença ou ausência das seqüências espaçadoras. As bases moleculares deste polimorfismo provavelmente estão relacionadas à recombinação homóloga entre as DRs próximas ou distantes e ao rearranjo ocorrido devido a presença do transposon IS6110 na região DR da maioria das cepas de *M. tuberculosis* (Kamerbeek *et al.*, 1997).

As seqüências DRs são usadas como alvo de “primers” para amplificação das seqüências espaçadoras, o polimorfismo desta região é observado através da hibridação com 43 seqüências espaçadoras sintéticas, derivadas do locus DR de *M. tuberculosis* e *M. bovis* BCG, covalentemente ligadas a uma membrana de náilon. Para uma melhor compreensão deste princípio veja a figura I.3.

O “*spoligotyping*” e o IS6110-RFLP, apresentam diferenças em suas capacidades de distinguir cepas, e embora o IS6110-RFLP seja o padrão ouro, cepas com uma ou poucas cópias do transposon IS6110 são

melhores discriminadas pelo “*spoligotyping*” (Kamerbeek *et al.*, 1997 e de la Salmonière *et al.*, 1997).

O “*spoligotyping*” testa simultaneamente até 43 amostras, dispensando o cultivo do *M. tuberculosis*, com base na reação de PCR. Devido a esta versatilidade é a técnica adequada para genotipagem em casos de surtos ou em situações em que não se pode cultivar o bacilo. Além do reconhecimento das relações epidemiológicas entre as cepas de *M. tuberculosis*. O “*spoligotyping*” foi reconhecido como capaz de distinguir especificidades geográficas entre os perfis, de modo que algumas cepas agrupadas em famílias se relacionam a uma determinada população, região geográfica e até a própria história da dispersão da TB no mundo (Sola *et al.*, 1999). Atualmente, há um grande investimento para definição e reconhecimento de famílias genotípicas de *M. tuberculosis*, um exemplo é a família Beijing, responsável por 85% dos casos de TB em Beijing, capital da China, e que está também distribuída em outros países da Ásia e no resto do mundo. Na cidade de Nova Iorque esta cepa ficou conhecida como cepa W e foi responsável por um surto de TB-MDR (Bifani *et al.*, 2002).

A grande capacidade de dispersão da família Beijing está associada à facilidade de adquirir resistência a drogas, a capacidade de escapar da resposta imune (induzida ou não pela vacina BCG), a alta virulência (inclusive velocidade de crescimento e grande capacidade de crescer em macrófagos), a capacidade de induzir resposta imune diferencial e sua alta transmissibilidade, (Kremer *et al.*, 2005). No “*spoligotyping*” esta família é definida pela ausência dos espaçadores de 1 a 35, para um reconhecimento visual desta família consulte a tabela da página seguinte (Tabela 1).

Utilizando os “*Spoligotypes*” para compreender a distribuição de *M. tuberculosis* foi possível estabelecer um caráter biogeográfico para as cepas que foram agrupadas em famílias, então uma importante função do SpolDB (Banco de Dados Internacional de “*spoligotypes*”) é o reconhecimento e enquadramento destas cepas dentro de grupos específicos, apresentados na tabela 1. As regras estabelecidas enquadram a família EAI (*East-African Indian*) como a mais próxima do grupo ancestral do Complexo *M. tuberculosis*, esta família é dividida em pelo menos cinco classes, também está definida a família Haarlem com três principais padrões de “*spoligotypes*”, duas famílias denominadas CAS1 e CAS2 (*Central and Middle Eastern Asia*), a Família X dividida em três classes bem definidas, a família T (T1-T4) ainda por ser melhor caracterizada e por último a família LAM (*Latin-American and Mediterranean*) dividida em 11 classes (LAM1 a LAM11).





Dentre as famílias de *M. tuberculosis* definidas por “*spoligotyping*” e disponíveis no SpolDB3, 35% dos clusters estão compostos por cepas pertencentes às famílias: Beijing (11%), LAM (9,3%), Haarlem (7,5%) e X (7%). Portanto, a análise do SpolDB3 mostra importantes diferenças na estrutura populacional de *M. tuberculosis* dentro dos oito subcontinentes estudados (África, Europa, América do Norte, América Central, América do Sul, Ásia Central e Oriente Médio, Extremo Oriente e Oceania) (Filliol *et al.*, 2002). A sobreposição de perfis foi observada para os vários subcontinentes, o que é explicado através da compreensão do fluxo migratório das populações humanas assim como do longo processo de globalização (Filliol *et al.*, 2003).

O “*spoligotyping*” além de ser uma metodologia versátil na tipagem de micobactérias que causam tuberculose, é também capaz de diferenciar organismos do complexo *M. tuberculosis*; um exemplo na tabela 1 é *M. bovis* com ausência dos espaçadores 39 ao 43. E na genotipagem de *M. bovis* por IS6110-RFLP muitos grupos relatam a presença de uma única cópia de IS6110 em seu genoma (Cousins *et al.*, 1998), enquanto van Soolingen *et al.*, (1994), destacam a presença de *M. bovis*, com múltiplas cópias de IS6110, provenientes de animais selvagens ou confinados em zoológicos (van Soolingen *et al.*, 1994). Portanto, a abordagem mais adequada para a epidemiologia molecular de cepas de *M. bovis*, à semelhança de *M. tuberculosis*, também inclui uma associação de metodologias de tipagem molecular.

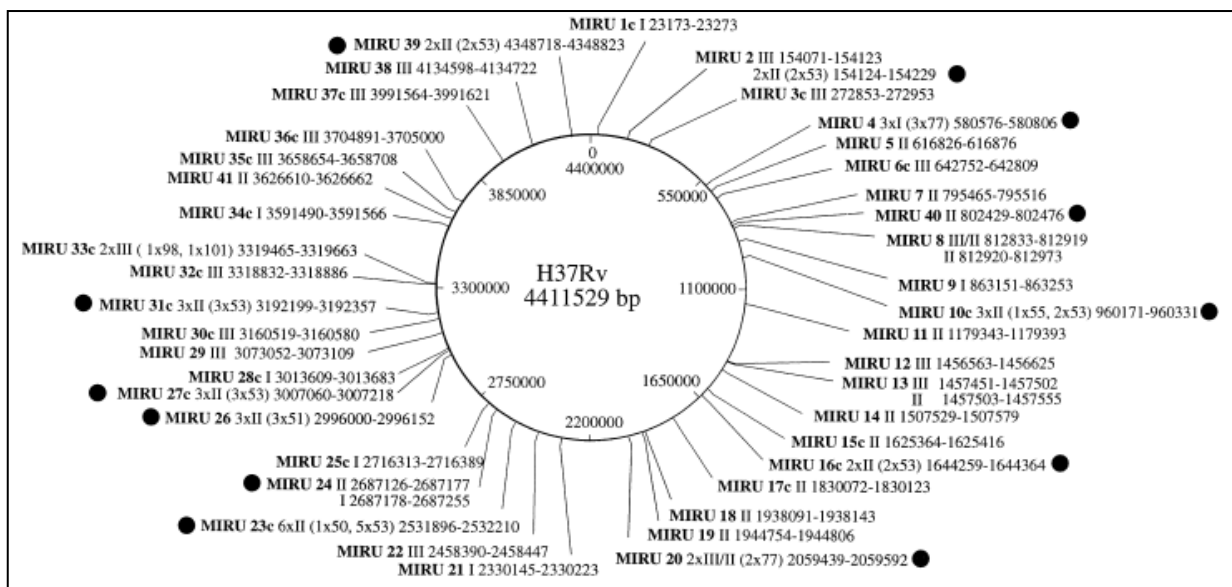
A associação de métodos e a busca de novos marcadores moleculares para a epidemiologia molecular de *M. tuberculosis*, que resultem no rápido reconhecimento de cepas e que apresentem uma resolução semelhante ou superior ao IS6110-RFLP fez com que no final dos anos 90 Supply e colaboradores (2001), identificassem no genoma de *M. tuberculosis*, uma região semelhante aos minissatélites, variáveis, presentes no genoma humano, que foi denominada *Mycobacterial Interspersed Repetitive Units* (MIRU-VNTR).

DNA minissatélites estão aparentemente dispersos em milhares de cópias no genoma de todos os eucariotos superiores, com tamanhos que variam entre 10 e 100pb, muitos destes loci são hipervariáveis em humanos e em outros animais, sendo então denominados de *Variable Number Tandem Repeats* (VNTR). Estes VNTR são valiosos marcadores para estudos evolutivos e de relações filogenéticas. Apesar de regiões microsatélites (1 a 13pb) terem sido descritas em muitos procariotos, e utilizadas como marcadores populacionais, não são estruturalmente similares àquelas dos VNTR humanos.

Neste contexto, os MIRU-VNTR são seqüências de DNA homólogas dispersas dentro das regiões intergênicas do genoma de *M. tuberculosis*, de *M. leprae* e de outras espécies micobacterianas, compostas por 40 a 100pb e localizados em 41 posições dispersas ao longo do genoma de *M. tuberculosis* H37Rv. Destas 41 posições, 12 mostraram um polimorfismo no número de cópias de MIRU entre isolados de *M. tuberculosis* não relacionados (Supply *et al.*, 2001).

Ao comparar os 41 Loci MIRU de *M. tuberculosis* H37Rv com *M. tuberculosis* CDC 1551 e com *M. bovis* AF2122/97, observou-se a presença dos 41 MIRU-VNTR nestas cepas, com exceção do locus 21, que junto com sua região flanqueadora, não foi encontrado em *M. bovis*. Esta região corresponde a *Region of Difference 9* (RD9), ausente em *M. bovis* e *M. bovis* “*Bacille de Calmette-Guérin*” (BCG). Os genes que

flanqueiam estes MIRU-VNTR estão relacionados às vias metabólicas ou regulatórias. No entanto, nenhuma relação foi estabelecida entre estes genes e os MIRU-VNTR.



**Figura 1.4** - Posição dos 41 MIRU-VNTR no cromossomo do *M. tuberculosis* H37Rv. O primeiro número indica a posição de cada locus ocupado por um MIRU-VNTR. A letra c indica orientação contrária a estabelecida por Cole *et al.*, em 1998. Algarismos romanos indicam o tipo de MIRU (I, II ou III). Os números seguintes indicam a localização de cada loci. A esfera em preto indica os 12 loci com número variável de MIRU-VNTR (Supply *et al.*, 2000).

Para compreender a dinâmica do MIRU-VNTR, os MIRUs, das cepas de *M. bovis* BCG Pasteur, Glaxo e Japão foram analisados. Esta análise mostrou que o MIRU 21 estava ausente nas três linhagens e que, somente o locus 4 foi polimórfico. Estes dados mostram que MIRU-VNTR são extremamente estáveis em culturas axênicas, já que estas cepas são resultantes de várias passagens da cultura original de Calmette e Guérin, desde 1921.

Os métodos de genotipagem, que tem como base a reação de PCR, tornaram a epidemiologia molecular da TB rápida, confiável e versátil para a realização de um estudo global. A detecção de DNA de *M. tuberculosis* proveniente de material de lâminas coradas por Ziehl-Neelsen e de biopsias de pacientes com TB, associados aos métodos acima citados possibilitou análises retrospectivas do perfil genotípico daquelas amostras que durante anos estiveram “esquecidas” nos “velhos” arquivos da saúde.

A tipagem por “*spoligotyping*” de DNA extraído de biopsia de pulmão parafinada, coletadas entre 1956 e 1990 na cidade de Beijing, mostrou que 92% das amostras positivas pertenciam ao clássico perfil da família Beijing, indicando que este perfil tem sido predominante naquela região (Qian *et al.*, 1999). Além da alta prevalência no sudeste asiático esta cepa tem mostrado uma preferência geográfica pelo hemisfério norte, e o fato de ter causado surtos de TB-MDR em Nova Iorque e outras cidades dos Estados Unidos da América (EUA)



tornou-a alvo de importantes estudos. (Kremer *et al.*, 1999).

A região DR, MIRU e o IS6110 são marcadores que melhor definem isolados do complexo *M. tuberculosis* e as suas adequações nestas definições estão diretamente relacionadas às suas capacidades de classificar os isolados apropriadamente. Um bom exemplo é a caracterização molecular das cepas pertencentes à família Beijing, pois através da análise destes três marcadores, foi confirmada a homogeneidade genética desta família e a sua diferenciação de outras linhagens (Kremer *et al.*, 1999).

A análise da família Beijing foi utilizada para melhor estabelecer o poder discriminatório do “*spoligotyping*”, VNTR (14 loci) e IS6110-RFLP. Esta análise mostra uma discreta vantagem do IS6110-RFLP sobre o VNTR, quanto ao seu poder discriminatório (Kremer *et al.*, 2005).

Como mencionado anteriormente, no caso da família Beijing, os marcadores utilizados para traçar a epidemiologia molecular da TB, e que permitem um reconhecimento de modo a intervir com eficiência na cadeia de transmissão da doença, são também indicados para a identificação das características genéticas de uma família específica de *M. tuberculosis*. Um estudo com isolados da República dos Camarões, que apresenta alta incidência de TB relacionada ao *M. tuberculosis* (no passado o *M. africanum* era a cepa prevalente), mostrou que 40% dos casos TB pulmonar, escarro-positivos, são causados pela família “Cameroon”. Uma análise molecular poderá definir as características específicas desta família de cepas, e este conhecimento será de grande importância no combate a TB, já que poderá esclarecer os motivos pelo qual esta classe se distribui com tanto sucesso naquela região (Niobe-Eyangoh *et al.*, 2004).

### **I.5. Filogenia e Evolução de Cepas do Complexo *M. tuberculosis***

O conhecimento de alguns aspectos do cenário evolutivo de *M. tuberculosis* e de seu relacionamento com o complexo e com outras espécies de micobactérias, fez com que a análise genética permitisse uma melhor definição das relações evolucionárias entre as cepas desta espécie. Os marcadores moleculares de *M. tuberculosis*, mencionados anteriormente, têm sido utilizados, juntamente com outros, para revelar as relações filogenéticas deste microorganismo, assim como para compreender a sua origem e sua evolução.

As subespécies que pertencem ao complexo *M. tuberculosis* apresentam uma similaridade de 99,9% em suas seqüências nucleotídicas e a maioria dos genes, utilizados para diferenciar espécies de micobactérias, são idênticos entre os organismos do complexo. No entanto, estas subespécies apresentam grandes diferenças em suas preferências pelos hospedeiros, nas características fenotípicas e nos diferentes graus de patogenicidade. Muitas perguntas relacionadas ao complexo *M. tuberculosis* poderão nos ajudar a compreender melhor a epidemiologia da TB e o seu comportamento no presente e no futuro. Uma destas perguntas relaciona-se a origem e evolução dos organismos do complexo *M. tuberculosis*.

Na história evolutiva do *M. tuberculosis*, a baixa frequência de mutações neutras em comparação com outros microorganismos deu origem à hipótese, de que *M. tuberculosis* divergiu de seu ancestral, em um tempo considerado recente, caracterizando-o, como uma espécie jovem, do mesmo modo que a sua distribuição ao redor

do mundo. Em trabalho histórico, com base em dois polimorfismos, um que ocorre no gene que codifica para uma catalase-peroxidase (*katG*) e outro para a subunidade A da girase (*gyrA*), três grupos genéticos principais (PGGs ou MGGs), foram descritos (figura I.5).

Grupos Genéticos do complexo *M. tuberculosis*

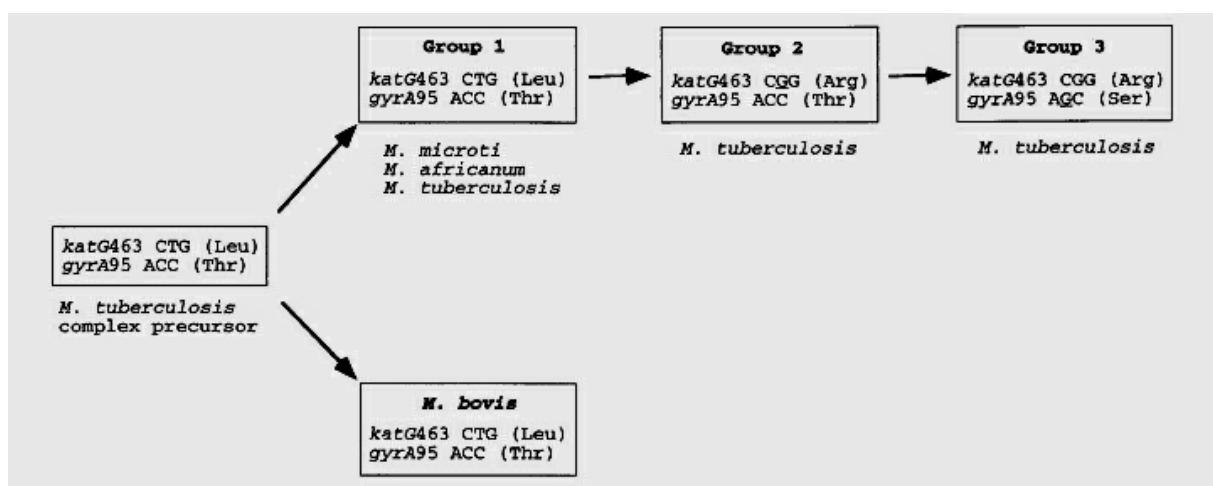
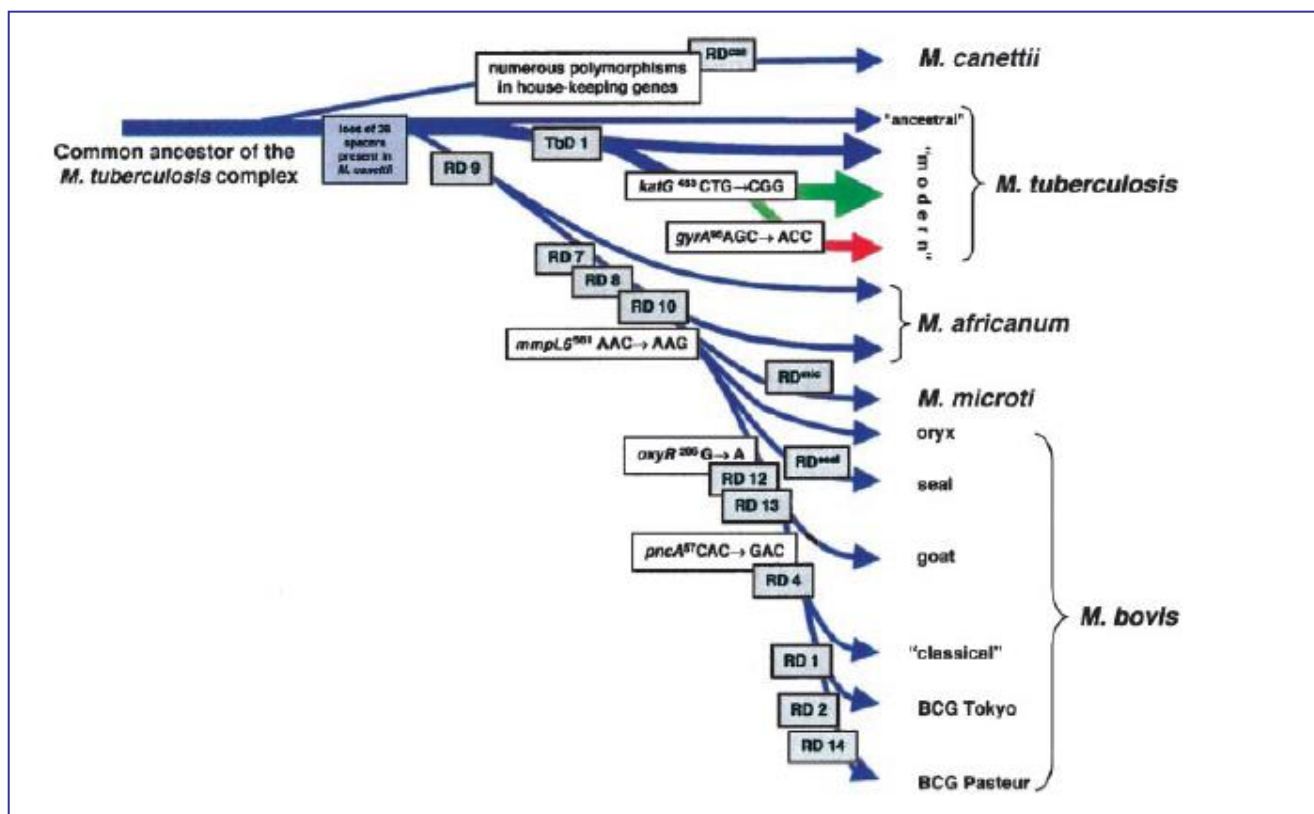


Figura I.5 - Cenário evolutivo dos organismos do complexo *M. tuberculosis* (Sreevatsan *et al.*, 1997).

Os organismos do complexo *M. tuberculosis* apresentam uma variação alélica menor que aquela observada em outras micobactérias e demais bactérias patogênicas, no entanto as variações observadas em *katG*<sup>463</sup> e *gyrA*<sup>95</sup> estão presentes em alta frequência, registrando a história da divergência entre os grupos 1, 2 e 3. Uma análise, por IS6110-RFLP, de isolados pertencentes a estes grupos, mostrou uma variabilidade entre estes perfis, sugerindo que mutações silenciosas seriam as responsáveis por esta diferenciação cromossomal (Sreevatsan *et al.*, 1997). O polimorfismo em *katG*, *gyrA* e a conservação de sítio de inserção da IS6110 em cepas do grupo 1 sugerem que estes estão mais próximos do progenitor do complexo *M. tuberculosis*. Esta hipótese é corroborada pela maior variabilidade genética observada no grupo 1 em relação aos grupos 2 e 3. O PGG1 apresenta um maior número de cópias de IS6110 e reúne 73% das mutações sinônimas identificadas, além disto, todas as variações presentes no pseudogene *oxyR* são encontradas somente no grupo 1. A possibilidade de um grupo genético ser caracterizado por uma maior virulência foi sugerida, e embora não tenha

sido confirmada, pode ser fruto de interações regiões-específicas entre parasito-hospedeiro.

### Cenário Evolutivo do *M. tuberculosis*



**Figura 1.6** - Esquema proposto para evolução do bacilo da TB ilustrando as deleções (retângulos cinzas). O esquema é baseado na presença ou ausência de regiões deletadas conservadas e no polimorfismo em cinco genes seleccionados. Em azul isolados caracterizados como pertencentes ao grupo 1, em verde isolados pertencentes ao grupo 2 e em vermelho isolados pertencentes ao grupo 3 (Brosch *et al.*, 2002).

O locus DR pode conter uma ou duas cópias da IS6110, esta inserção pode ser simétrica com uma DR dividida em dois fragmentos iguais ou assimétrica gerando um segmento com 6 pb e outro com 30 pb. O significado disto é que a caracterização de isolados clínicos de *M. tuberculosis* como tipo 50 (ausência do

espaçador 31) pode estar relacionada à inserção de uma cópia de IS6110 entre os espaçadores 31 e 32. Esta cópia divide a DVR 31, gerando uma DR de 6 pb, onde não é possível a hibridação do oligo DRa, tornando impossível a detecção do espaçador adjacente (Legrand *et al.*, 2001). São três os mecanismos responsáveis pela evolução do locus DR: recombinação homóloga, falha no deslize da polimerase (“slippage”) e transposição mediada por seqüência de inserção.

Através do “*spoligotyping*” foi possível identificar perfis encontrados em diferentes regiões do mundo, assim como, identificar aqueles que, quase exclusivamente, estão em uma única região ou em um grupo específico de regiões. Deste modo, o “*spoligotyping*” é capaz de estabelecer o caráter biogeográfico de cada cepa. Como exemplo, podemos citar o tipo 17 (LAM2), que é encontrado apenas em seis países do continente americano e o tipo 33 (LAM3), que apresenta uma freqüência de 21,4% em algumas comunidades da África do Sul, podendo também ser encontrado em outros três países do continente Africano (dois na África Sub-Saariana e um na África do Norte). O tipo 33 também está amplamente distribuído na Europa Ocidental, com exceção de Portugal, Suécia e Noruega e no continente Americano está presente em 11 países, inclusive no Brasil (Victor *et al.*, 2004).

O banco de dados internacional de “*spoligotypes*” (spolDB) está atualmente em sua 3ª versão (spolDB3.0), e reúne 13.008 padrões, provenientes de mais de 90 países, com 11.708 padrões em “clusters” e 1.300 padrões órfãos (isolados clínicos detectados uma única vez). Na análise global, do spolDB3.0, foram definidas 36 superfamílias e nove superfamílias de cepas (*M. africanum*, Beijing, *M. bovis*, EAI, CAS, T, Haarlem, X, LAM) (Tabela 1) (Sola *et al.*, 2001, Filliol *et al.*, 2003). A minuciosa análise do spolDB3.0 nos conduz a uma clara compreensão da distribuição global de *M. tuberculosis*. Os 20 isolados clínicos mais freqüentes totalizaram 5865 isolados, estando incluídos nesta freqüência os tipos 683 e 481 para *M. bovis* e o tipo 482 para *M. bovis* BCG, além dos tipos: 2 (Haarlem2), 33 (LAM3), 197, 11 (EAI3), 48 (EAI1), 19 (EAI2), 20 (LAM1), 52 (T2), 61 (LAM10), 139 (EAI4), 47 (Haarlem1), 119 (X1), 42 (LAM9) e 137 (X2), 50 (Haarlem3), 53 (T1) e 1 (Beijing), sendo estes três últimos, os mais freqüentes. O genótipo mais freqüente é o Beijing (cepa tipo 1) com 1282 isolados, seguidos pela cepa tipo 47 ou 50 que, devido à ausência do espaçador 31, estão incluídas na família Haarlem.

O continente Europeu, os Estados Unidos da América e a África apresentaram as maiores diversidades de “clusters”, enquanto que a menor foi observada no Extremo Oriente. Neste mesmo contexto, a maior taxa de “*spoligotypes*” compartilhados está presente entre a Europa e a América do Norte, sugerindo um nível mais alto de migração entre os dois continentes (Sola *et al.*, 2001).

A análise filogenética de *M. tuberculosis*, a partir de uma árvore construída com base no algoritmo “Neighbor-Joining”, mostrou que um ramo da árvore altamente ramificado, mas homogêneo, caracterizou-se pela concentração de cepas que apresentavam simultânea ausência dos espaçadores 21 a 24 e 33 a 36 e por pertencerem ao Grupo Genético Principal 2. As cepas desta família são frequentemente encontradas na região Sul e Central da Europa e na América do Sul, tendo como tipo ancestral o Tipo 42, de onde se originam os demais membros da família LAM. (Sola *et al.*, 2001, Sreevatsan *et al.*, 1997).

A distribuição Global de “*spoligotypes*” apresenta a frequência da família LAM da seguinte forma: América do Sul com 22% (tipo 42 - 9%, 17 - 6%, 64 - 3%, 93 - 2%, 33 - 2%), África 13% (tipo 20 - 8%, 42 - 5%), América Central 8% (tipo 42 - 5%, 17 - 3%), Europa 3% (tipo 42), América do Norte 2% (tipo42), Oceania <2% (tipo42), Extremo Oriente 0% e Oriente Médio e Ásia Central 0% (Filliol *et al.*, 2003).

Particularidades da distribuição global da família LAM estão em diversas publicações. Exemplos importantes têm na região da Sicília, Itália, onde 26% do total de isolados estavam incluídos nesta família (Sola *et al.*, 2001), enquanto que na Guiana Francesa 15 (14%) cepas de um total de 94 isolados analisados pertenciam à família LAM (tipo 42 - cinco isolados; tipos 93, 20 e 17 – um isolado cada; tipo 64 – três isolados; tipo 95 – um isolado; tipo órfão – um isolado; tipo 33 – dois isolados) (Filliol *et al.*, 2002). Em um estudo de reconstrução filogenética englobando quatro regiões do caribe: Guadalupe e Martinica (França), Cuba e Haiti utilizando como marcador o locus DR, foi possível observar que dos 321 isolados, 97 foram classificados como pertencentes à família LAM, e destes, 48 foram encontrados em Cuba. Somente a família LAM, tipo 17 foi encontrado em maior frequência em Guadeloupe e Martinica que em Cuba (Duchêne *et al.*, 2004). Um marcante caráter filogeográfico para a família LAM foi recentemente desenhado, a partir, da genotipagem por “*spoligotyping*” de 245 isolados clínicos de *M. tuberculosis*, provenientes de duas províncias turcas (Malatya n=147 e Ancara n=98). O tipo 41 (LAM7), como classificado pelo SpolDB3, representou a quarta parte de todos os isolados analisados (Zozio *et al.*, 2005).

Os genótipos com maior frequência na Espanha são os pertencentes à família LAM, T1 e Haarlem3, neste país duas novas famílias relacionadas à população espanhola: Família Madrid1 e Madrid2. A família Madrid1, por sua estreita relação com a família LAM foi classificada como LAM11 (Viedma *et al.*, 2005).

A preferência biogeográfica documentada para os isolados de *M. tuberculosis*, pertencentes às famílias Beijing e LAM, é também observada para o *M. africanum*. Esta espécie está relacionada à TB na África Sub-Saariana ou em pacientes provenientes destas regiões. Este conhecimento fragmentou-se quando foram identificados pacientes com TB, causada por *M. africanum* na Europa, e que não possuíam relação epidemiológica com o continente Africano (Viana-Niero *et al.*, 2001).

O *M. africanum* é responsável por mais de 60% dos casos de TB pulmonar na África, no entanto, a frequência de distribuição destes casos sofre extrema variação entre países e esta mesma variação pode ser observada dentro de um mesmo país: 5% em Costa do Marfim, ~10% em Camarões com 60% dos casos em sua capital Yaoundé, ~60% em Guiné-Bissau, ~20% no Senegal com 9% dos casos na região de Casamance e 47% nas proximidades do Rio Senegal (Viana-Niero *et al.*, 2001, Niemann *et al.*, 2002, Niemann *et al.*, 2004).

A ausência dos espaçadores 8, 9 e 39 e a presença dos espaçadores 33 a 36 e 40 a 43 faz parte de uma tentativa de se estabelecer uma assinatura molecular específica do *M. africanum*, além disto, como demonstrado, o *M. africanum*, possui no códon 463 do gene *katG* uma Leucina (CTG), o que coloca esta espécie no grupo ancestral do Complexo *M. tuberculosis* (Figura I.5).

*M. africanum* através de suas características bioquímicas está dividido em dois principais grupos: o

subtipo I e o subtipo II, freqüentes respectivamente no oeste e leste da África. A análise numérica das características bioquímicas desta espécie mostra que o subtipo I compartilha características com *M. bovis* e o subtipo II com o *M. tuberculosis* (Niemann *et al.*, 2002) e de um modo geral o “spoligotype” do *M. africanum* apresenta um perfil intermediário entre *M. tuberculosis* e *M. bovis* (Viana-Niero *et al.*, 2001).

Os marcadores moleculares caracterizam o *M. africanum*, assim como as demais espécies do Complexo e permitem o desenvolvimento de novas metodologias de genotipagem e contribuem diretamente para a compreensão filogenética destas espécies.

Análises Filogenéticas baseadas na região DR são comparadas com as obtidas avaliando outros marcadores. As famílias “East African-Indian” (EA-I) e “Latin American and Mediterranean” (LAM) são exemplos de recente caracterização baseada no polimorfismo do locus DR e que posteriormente foram confirmados pela análise de VNTR (Variable Number of Tandem Repeats) e *katG-gyrA*. Nestes dois casos, as árvores construídas baseadas no algoritmo “Neighbor-Joining” usando o locus DR ou VNTR, agruparam os isolados de acordo com suas origens geográficas (Sola *et al.*, 2001). Nesta árvore, a família Beijing posicionou-se de modo homogêneo, em um ramo distante dos demais isolados de *M. tuberculosis*, enquanto que a família Haarlem está aparentemente agrupada com isolados de origem africana e próxima, ao ramo que agrega isolado de *M. bovis* e ao ramo de LAM. A homogeneidade da família Beijing é demonstrada pela presença de um único padrão de VNTR (42432), tendo como exceção, um isolado da Rússia com VNTR (42452) (Sola *et al.*, 2001).

O RFLP, MIRU-VNTR e “spoligotyping” são metodologias para epidemiologia molecular da TB, que por sua ampla utilização, tornaram-se clássicas, no entanto, análises baseadas no Polimorfismo de Base Única (SNP) e na Diversidade Sinônima de Base, têm estabelecido um profundo conhecimento da diversidade genética, da filogenia e da evolução do complexo *M. tuberculosis*.

A partir da identificação de regiões de diferença (RDs) do bacilo da TB, foi possível desenhar uma bela árvore evolutiva (Figura I.6), sugerindo uma filogenia para o complexo *M. tuberculosis*, que tem como raiz *M. tuberculosis* ou *M. canettii*. São 14 as RD, que variam de 2 a 12,7 Kb e seis as regiões de deleção relacionadas ao H37Rv: RvD1, 2, 3, 4, 5 e TbD1 (“*M. tuberculosis* specific deletion 1”), uma deleção específica de *M. tuberculosis* de 2153pb. As RDs 1, 2, 4, 7, 8, 9, 10, 12, 13 e 14 estão bem conservadas em isolados de *M. tuberculosis* e podem ser utilizadas na diferenciação desta espécie das demais do Complexo *M. tuberculosis* (Mostowy *et al.*, 2005).

Se a TB surgiu no mesmo continente onde a espécie humana “nasceu”, nós podemos acreditar que a sua dispersão pelo mundo esteve associada às sucessivas migrações humanas e embora este microorganismo se relacione com a espécie humana desde seus primórdios, foi somente a partir do reconhecimento das Regiões de Diferença (RD) que um cenário evolutivo para o complexo *M. tuberculosis* foi estabelecido, conduzindo-nos há 20.000-35.000 anos atrás, quando o *M. tuberculosis* e o *M. bovis* divergiram de um progenitor comum, contrariando a afirmação de que o *M. bovis* seria o precursor evolucionário do *M. tuberculosis* (Mostowy *et al.*, 2005). A hipótese da existência de um precursor comum foi confirmada recentemente, onde um isolado geneticamente relacionado ao *M. tuberculosis/M. africanum* foi identificado em um bisão de 17.000 anos.

Também foi observado, através da avaliação de deleções em isolados que mostram que a TB humana antecede aquela observada em focas, cabras e no gado bovino, tipicamente infectado por *M. pinipedi* e *M. bovis* (Mostowy *et al.*, 2002).

A introdução do MIRU aos já conhecidos “*spoligotyping*” e VNTR resultou em um suporte para a compreensão do cenário evolutivo do complexo *M. tuberculosis*. A análise de 115 isolados mostrou um cenário evolutivo do Complexo *M. tuberculosis* semelhante ao descrito por Brosch *et al.*, (2002). Neste novo cenário o *M. canettii* assemelha-se aquela espécie que divergiu do ancestral comum seguida pela presença da família EAI grupo homogêneo que pode ser dividido em várias subfamílias, no entanto os grupos T1 e T2 necessitam de marcadores adicionais para uma melhor definição (Sola *et al.*, 2003).

Em nível de seqüência, e não considerando variações induzidas por elementos móveis ou regiões repetitivas, o genoma de *M. tuberculosis* é altamente conservado. A comparação da seqüência completa de H37Rv e da cepa CDC1551 mostrou a presença de apenas 1075 SNPs entre H37rv e *M. Bovis* cepa AF2122, observou-se apenas a existência de 2437 SNPs (Filliol *et al.*, 2006)

A construção filogenética de isolados de *M. tuberculosis* com base em SNP, apresentou-se exata e com pouca tendência a sofrer distorções por pressão seletiva, como é comum com outros marcadores. Análise filogenética de isolados clínicos com base em SNPs agrupou, por “Neighbor-Joining”, sete clusters divididos claramente em seis grupos, contendo isolados de *M. tuberculosis* e um grupo formado por *M. bovis*. Estes grupos tornaram-se mais aparentes quando se ampliou os tipos de SNPs analisadas (sinônima - sSNP, não sinônima - nsSNP e intergênica - igSNP) (Gutacker *et al.*, 2006). A análise filogenética que reuniu os grupos genéticos do complexo *M. tuberculosis* 1, 2 e 3 com aquela estabelecida por SNPs, resultou em estreita correlação. A composição filogenética estabelecida por SNP e sua relação com os “clades” definidos pelos “*spoligotyping*” foi clara para a família “Beijing” que se concentrou exclusivamente no grupo 2. A família EAI concentrou seus isolados no grupo 1 e o grupo 3a reuniu isolados da Família CAS. Outras famílias se dispersaram em maior ou menor grau entre os demais grupos (Filliol *et al.*, 2006).

A genotipagem, a filogenia e a evolução do complexo *M. tuberculosis* mostra como a relação entre o *Homo sapiens sapiens* e *Mycobacterium tuberculosis* está desenhando o perfil de distribuição global desta micobactéria. A filogeografia, identificada principalmente pelo “*spoligotyping*” e mais recentemente também através da análise de SNPs, permiti-nos inferir as rotas ancestrais de distribuição de *M. tuberculosis*, que estão diretamente relacionadas aos fluxos migratórios das populações humanas e de outros animais ao longo dos anos. Este fluxo foi o responsável pelo sucesso da dispersão do *M. tuberculosis* no passado. Nas duas últimas décadas do século 20 e no início do século 21 o sucesso do *M. tuberculosis*, anteriormente relacionado à dispersão da espécie humana através dos continentes, é atribuído à globalização, que cada vez mais “reduz as distâncias” e aproxima povos e culturas.

## Capítulo II – Objetivos



### **II.1. Detecção de resistência a antibióticos em cepas de *Mycobacterium tuberculosis***

- 1 - Avaliação do SSCP como método genotípico na detecção de mutações no gene *rpoB*, de *Mycobacterium tuberculosis*, relacionadas à resistência a rifampicina.
- 2 - Avaliar a hibridação reversa (RLB- “reverse line blot assay”) para a detecção de mutações nos genes *rpoB*, *inhA*, *ahpC*, *rpsL* e *embB*, relacionadas a resistência de cepas de *Mycobacterium tuberculosis* à rifampicina, isoniazida, estreptomicina e etambutol.
- 3 - Avaliar a capacidade da RLB para a detecção de resistência à rifampicina em cepas de *M. tuberculosis* isoladas no Brasil.
- 4 - Avaliar a RLB para a detecção de resistência a rifampicina em isolados de *M. bovis*.

### **II.2. Caracterização molecular de cepas do complexo *Mycobacterium tuberculosis***

- 1 – Identificar, através da técnica de “*spoligo*typing”, o perfil genotípico das cepas de *Mycobacterium tuberculosis* isoladas de pacientes com tuberculose do Brasil.
- 2 - Avaliar a importância das outras espécies do complexo *Mycobacterium tuberculosis* na transmissão da tuberculose no Brasil.
- 3 - Estabelecer o perfil filogeográfico dos “*spoligo*types” para o Brasil.
- 4 - Identificar a existência de classes ou tipos de cepas de *Mycobacterium tuberculosis* com maior impacto na transmissão da tuberculose no Brasil, assim como a sua distribuição nos diferentes estados ou regiões brasileiras.

**Capítulo III – Materiais, Métodos e Resultados**

**Parte 1 - Detecção de Resistência em Cepas de *Mycobacterium tuberculosis***

### III.1. Artigo 1

Evaluation of Polymerase Chain Reaction-Single Strand Conformation Polymorphism (PCR-SSCP) Analysis for the Detection of the *rpoB* Mutations Associated with Resistance to Rifampicin in *Mycobacterium tuberculosis*.

Lee H *et al.*, 2003. World Journal of Nuclear Medicine, 2: 45 - 51.



Os resultados deste artigo iniciam as nossas investigações sobre mutações, presentes no genoma de cepas de *M. tuberculosis* resistentes às drogas e empregadas no tratamento da tuberculose, utilizando para isto uma abordagem que, naquele momento, se mostrava promissora (SSCP).

### **III.7. Manuscrito 1**

Discovery of a New *Mycobacterium tuberculosis* Genotype that is the Most Predominant Cause of TB in Rio de Janeiro, Brazil

Lazzarini LCO. et al.

### III.2. Artigo 2

Multicenter Evaluation of Reverse Line Blot Assay for Detection of Drug Resistance in *Mycobacterium tuberculosis* Clinical Isolates

Igor Mokrousov et al. 2004. Journal of Microbiological Methods. 57:323-335



Este artigo é a base para a compreensão da detecção de resistência, em *M. tuberculosis* às drogas utilizadas no tratamento da tuberculose, através da metodologia da hibridação reversa, e contribuiu para esta tese, pelo fato de que a partir dos conhecimentos nele estabelecidos foi possível determinar as mutações relacionadas à resistência a drogas, no genoma de *M. tuberculosis*, em cepas brasileiras.

### III.3. Artigo 3

Rapid Detection of Resistance Against Rifampicin in Isolates of *Mycobacterium tuberculosis* from Brazilian Patients Using a Reverses Phase Hybridization Assay.

De Oliveira et al. 2003. Journal of Microbiological Methods. 53: 323-335



Através da comparação dos resultados deste artigo, com aqueles apresentados no artigo 2 (Mokrousov et al. 2004), foi possível construir uma relação precisa das mutações em *M. tuberculosis* que estão relacionadas com resistência à rifampicina, cuja a investigação esta incluída nos objetivos desta tese.

#### III.4. Artigo 4

In House Reverse Line Blot Assay for Rapid Detection of Susceptibility to Rifampicin in Isolates of *Mycobacterium tuberculosis*.

Senna SG et al. 2006. Journal of Microbiological Methods. *In press*



Este artigo encerra a abordagem desenvolvida para a detecção de mutações relacionadas à resistência de *M. tuberculosis* as drogas utilizadas no tratamento da tuberculose. Os artigos anteriores, principalmente o artigo 2 (Mokrousov et al. 2004) e o artigo 3 (de Oliveira et al. 2003) nos conduzem a uma melhor investigação da resistência a rifampicina em cepas de *M. tuberculosis* brasileiras que está descrita neste artigo (artigo 4).

### III. 5. Artigo 5

Drug Susceptibility of Brazilian Strains of *Mycobacterium bovis* Using Traditional and Molecular Techniques  
Parreiras PM et al. 2004. Memórias do Instituto Oswaldo Cruz. 99: 749 - 752



Este artigo tem como objetivo mostrar que a hibridação reversa em linha pode também ser utilizada na investigação de resistência de *M. bovis* a rifampicina, embora amostras resistentes a rifampicina precisem ser incluídas nos ensaios posteriores.



## **Parte 2 - Tipagem Molecular de Cepas *M. tuberculosis***

Os trabalhos a seguir compõem a análise genotípica de cepas do complexo *M. tuberculosis* através do método de genotipagem, denominado de “spoligotyping”. Com estes trabalhos foi possível construir o perfil genotípico (“spoligotypes”) das cepas de *M. tuberculosis* prevalentes no Brasil (manuscrito 2, em preparação) e estabelecer a relação destes “spoligotypes” com aqueles distribuídos em mais de 122 países, através e nossa participação no Banco de Dados Internacional de “Spoligotypes” (artigo 6, Brudey et al. 2006). Uma importante colaboração desta seção foi estabelecer a relação das cepas da família LAM com uma deleção de 26Kb do genoma de *M. tuberculosis* em alta frequência em cepas provenientes de pacientes da cidade do Rio de Janeiro (manuscrito 1, em preparação). Estes trabalhos respondem as questões estabelecidas nos objetivos desta tese.

### **III. 6. Artigo 6**

*Mycobacterium tuberculosis* Complex genetic Diversity: Mining the Fourth International Spoligotyping Database (SpolDB4) for Classification, Population Genetics and Epidemiology

Brudey K et al. 2006. Biomedical Central. 6: 23.

**Discovery of a New *Mycobacterium tuberculosis* Genotype That is The Most Predominant Cause of Tuberculosis in Rio de Janeiro, Brazil.**

Luiz Claudio Oliveira Lazzarini<sup>\*†¶</sup>, Richard C. Huard<sup>\*¶</sup>, Neio L. Boechat<sup>†</sup>, Harrison M. Gomes<sup>‡</sup>, Maranibia C. Oelemann<sup>‡</sup>, Fernanda Q. Mello<sup>†</sup>, Milena J. Virginio<sup>‡</sup>, Ana Grazia Marsico<sup>†</sup>, Ray Butler<sup>§</sup>, Philip N. Suffys<sup>‡</sup>, Jose Roberto Lapa e Silva<sup>†</sup>, John L. Ho<sup>\*#</sup>.

\* Department of Medicine, Division of International Medicine and Infectious Diseases, Weill Medical College of Cornell University, New York, USA.

† Institute of Thoracic Diseases. Federal University of Rio de Janeiro, Brazil.

‡ Laboratory of Molecular Biology Applied to Mycobacteria, Oswaldo Cruz Institute, Fiocruz, Rio de Janeiro, Brazil.

§ Centers for Disease Control and Prevention, Atlanta, GA, USA.

¶ Authors contributed equally to the work

# Corresponding author. Mailing address: Cornell University, Joan and Sanford I. Weill Medical College, Department of Medicine, Division of International Medicine and Infectious Diseases, Room A-421, 525 East 68th St., New York, NY 10021.

Phone: (212) 746-6316. Fax: (212) 746-8675. E-mail: jlho@med.cornell.edu.

Abstract

Gene deletion in *M. tuberculosis* (Mtb) and in the Mtb complex (MTC) is a determinant for biological diversity. To evaluate whether other members of the MTC cause tuberculosis (TB) in Rio de Janeiro, we applied the MTC-PCR typing panel to speciate 407 clinical isolates from 314 TB patients collected during 2002 and 2003. Only *Mtb stricto sensu* was found but interestingly, 30% of the Mtb isolates

( $n=122$ ) unexpectedly failed to amplify the IS1561' locus. PCR bridging the segment showed a deletion of unprecedented length of >26 Kb for all the IS1561' absent Mtb strains in which we named RD<sup>Rio</sup>. Sequence analysis of RD<sup>Rio</sup> strains suggested that homologous recombination was the mechanism responsible for the deletion of eight genes, including 2 PPE genes encoding proteins believed to confer antigenic variability while remaining infectious and transmissible as shown by epidemiologically linked transmission. TB caused by RD<sup>Rio</sup> strains manifested with a more severe illness by having higher frequency of hemoptysis and weight loss, as well as, higher bacillary load. RD<sup>Rio</sup> Mtb genotype is not limited to Brazil but also found in USA, Europe and Africa. All RD<sup>Rio</sup> strains 1) are typed as principal genetic group 2, 2) are members of the Latin American and Mediterranean (LAM) spoligotyping family and 3) share a specific MIRU-VNTR profile. Phylogenetic analyses suggested evolutionary divergence of RD<sup>Rio</sup> from other Mtb likely from a single progenitor. These data raises the need to evaluate the worldwide prevalence of TB caused by RD<sup>Rio</sup> and the disease spectra caused by this genotype.

## INTRODUCTION

*Mycobacterium tuberculosis* (Mtb) is one of the most important human pathogens, believed to infect a third of the humankind and present virtually in any region of the world (1). There is an enormous biodiversity of Mtb strains, showing different geographic preferences, virulence, transmissibility, antibiotic susceptibility and ability to induce or evade the immune system (2-6), and probably most of them are related to differences in their genomes. Mtb manifest very little genome sequence diversity in the form of single nucleotide polymorphisms (SNPs), whereas large sequence polymorphisms (LSPs) appear to be the major source of variability (7). The presence of insertions and deletions within the *Mycobacterium tuberculosis* complex (MTC) genome serves as species-specific markers as well as a way to study phylogenetic evolution. One of these region is the non-functional insertion sequence IS1561' (Rv3349c), described to be present in all Mtb strains but *M. microti* and *M. pinnipedii* (8,9). Although deletion is usually considered detrimental to the bacteria (10,11), it sometime turns them more virulent (12,13), so the net effect depends mostly on which genes are deleted or disrupted. The knowledge that isolates from patients infected with epidemiologically related strains (clusters) have

identical gene sequence whereas unrelated strains differs is the basis for the use of genotyping methods. Although IS6110-based restriction fragment length polymorphisms (RFLP) typing is considered the gold standard technique for genotyping (14,15), it has several drawbacks and combined use of two PCR-based techniques, spoligotyping and mycobacterial interspaced repetitive units and variable numbers of tandem repeats (MIRU-VNTR) have been proposed as an adequate alternative (16). They are faster, cheaper, less labor-consuming and easier to be adapted to automated high-throughput, having MIRU-VNTR an discriminatory power almost similar to IS6110-RFLP (14,16-19). An additional advantage of spoligotyping is that it can also classify Mtb into families, allowing better understanding of phylogeny and population structure on a region-specific or global level (20). In this context, detection of Mtb strains with specific genetic features and/or clinical outcome could advance our understanding of the role of specific genes as well as change epidemiological strategies to fight this pathogen. In this paper we describe a new family of Mtb designated as Mtb RD<sup>Rio</sup> which have a never described deletion (RD<sup>Rio</sup> deletion) and particular spoligotype and MIRU-VNTR profiles. It is the predominant Mtb family causing TB in Rio de Janeiro but is also observed in other continents and infection with RD<sup>Rio</sup> strains seems to cause a more severe form of the disease. The discovery of this family will hopefully lead to better understanding of Mtb pathogenesis, epidemiology, phylogeny and gene function.

## MATERIALS AND METHODS

**Data collection.** Four hundred thirty mycobacterial clinical isolates were randomly selected from Clementino Fraga Filho University Hospital of the Federal University of Rio de Janeiro, Brazil, between 2002 and 2003. They were cultured from patients with pulmonary and extra-pulmonary disease and the same patient could have more than one sample analyzed when they were collected from diverse sites or from the same site at different periods of time, but always during the same TB episode. The isolates were identified as Mtb, based on the morphology of the colonies in Löwenstein-Jensen media, AFB staining and by standard biochemical tests or as mycobacteria other than tuberculosis (MOTT), which was blindly included to evaluate the ability of the PCR system to differentiate Mtb from MOTT.

**Cell Thermolysis.** All cultures were grown in Löwenstein-Jensen slants and when the growth was considered optimal, 1 ml of TE buffer (10 mM Tris, 1 mM EDTA, pH 7,5) was added and the colonies

were detached from the medium using a loop. The liquid was transferred into a 1.5 ml eppendorf tube containing 6 glass beads of 3 mm, vortexed for 2 minutes inside the hood, then 10 µl of proteinase K (Fermentas Inc, Hanover, MD) at 20 mg/ml was added, vortexed again and incubate at 60°C for 4 h until all the clumps were completely dissolved, with vortexing every half hour during this process. Finally, the tubes were incubated at 80°C for additional 30 minutes.

**PCR amplification, primers and conditions.** The Mtb PCR system was used as previously described (9,21). Each PCR reaction of 50 µl was prepared using 25 µl of High Fidelity PCR Master (Roche Diagnostics, Indianapolis, IN), 2.5 µl of dimethyl sulfoxide (DMSO), 1 µl of each primer at 20 µM and 10 µl of the lysate. Due to its high GC content, the amount of DMSO was increased to 5 µl when primers within the PE\_PGRS50 gene were used. The PCR amplifications were performed in a Gene Amp PCR System 9700 (PE Applied Biosystems, Foster City, CA) using an initial denaturation step of 5 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 65°C and 4 min at 72°C and final extension step for 10 min at 72°C. In ten isolates, the annealing temperature was decreased to 60°C to increase the PCR yield. PCR products were visualized by electrophoresis in 1.5% agarose gel and ethidium bromide staining and their size estimated by comparing to a 100 base-pair ladder (Invitrogen, Carlsbad, CA). Images were captured using the Nighthawk Imaging System (PDI Inc., Huntington Station, NY) and Quality One software package (PDI Inc., NY). All negative results were repeated at least once more. The identification of MOTT was confirmed by biochemical tests and by DNA sequence analysis of the *16S rRNA* and *rpoB* genes.

**PCR product restriction enzyme digest analysis (PCR-RFLP) and DNA sequencing.** Presence of polymorphisms in *katG* and *gyrA* genes permit members of the MTC to be assigned into one of the three principal genetic groups (PGG) (7). The restriction enzymes *Bst*NI (New England Biolab, Beverly, Mass.) and *A*/el (New England Biolab, Beverly, Mass.) were used, respectively, to digest the *katG* and *gyrA* amplicons as previously described (9,21). Digested products were visualized by electrophoresis in 3% agarose gel and ethidium bromide staining. DNA sequence of the PCR fragments was performed at Cornell University BioResource Center (Ithaca, NY) using a BigDye Terminator Kit (PE Applied Biosystems) and an ABI 3700 DNA Sequencer.

**Identification of the deletion length in RD<sup>Rio</sup> strains.** To identify the boundaries of the RD<sup>Rio</sup> deletion, primer pairs were designed flanking, and increasingly outward from either side of the absent IS1561' locus using the described sequence of Mtb H37Rv (available at <http://www.ncbi.nlm.nih.gov> and <http://genolist.pasteur.fr/Tuberculist>) and primers were generated using the program DNASTAR (DNASTAR, Inc., Madison, Wis). The genes upstream and downstream to IS 1561' were tested in PCR experiments until a clear amplification was obtained in both wildtype (WT) and RD<sup>Rio</sup> strains and they were used to bridge the deletion. All amplicons were sequenced and compared with the available database.

**Clinical, radiological and laboratorial data.** When available, the clinical records of the patients were reviewed to complete a standardized questionnaire including gender, age, HIV status, PPD skin reactivity, clinical symptoms and site of TB. Radiology data evaluated the presence of pulmonary cavitation, extension of lung infiltrates, mediastinal adenomegaly and pleural effusion. Bacteriology data evaluated presence of acid-fast stain in smears, quantity of colonies on cultures and sensitivity to anti-TB drugs. The research team conducting the medical record had no knowledge of the Mtb genotype results.

**Spoligotyping.** All Mtb isolates were submitted to spoligotyping as described by Kamerbeek et al (22). Again, the person who performed the spoligotyping was not informed on the genotype results. For amplification, 10 to 20 ng of Mtb lysates was added to a PCR mix containing 20 pmol of the primers Dra and Drb, 0.2 mM of each dNTP, 1 U of Taq polymerase (Invitrogen), and PCR buffer to obtain a final concentration of 1.5 mM MgCl<sub>2</sub> in a total volume of 50 µl. For each experiment, two negative controls and a sample of *M. bovis* BCG Moureau and Mtb H37Rv were included. Samples were heated to 96°C for 3 min, followed by 35 cycles of 1 min at 96°C, 1 min at 55°C and 30 sec at 72°C, and a final extension step of 5 min at 72°C. For hybridization to the membrane (Isogen, The Netherlands), 30 µl of each PCR

product was diluted in 150 µl of 2X SSPE (0.2 M Na<sub>2</sub>HPO<sub>4</sub>, 3.6 M NaCl, 20 mM EDTA, pH 7.4)/0,1% SDS, denatured during 10 min at 99°C and put on ice. The membrane was mounted in the MN45 miniblottedter (Isogen) and after washing with 200 ml 2X SSPE/0.5% SDS, PCR products were added and incubated during 1 hour at 60°C. The membrane was washed twice with 200 ml of 2X SSPE/0.5% SDS at 60°C for 5 min, incubated with 10 ml 2X SSPE/0,5% SDS containing 10 µL streptavidin-peroxidase conjugate during 45 to 60 min at 42°C and washed twice with 2X SSPE/0,5% SDS at 42°C for 10 min and rinsed twice with 2X SSPE at RT. Hybridization was visualized using the ECL luminescence kit (Amersham Pharmacia Biotech, New Jersey, USA) for two hours. For analysis, spoligotypes were introduced into Excel and Bionumerics (Version 4.0, Applied Maths, Belgium).

**MIRU-VNTR.** MIRU-VNTR was performed as described by Supply et al (19) in at least one isolate from every patient in the RD<sup>Rio</sup> group, from patients in the WT group whose spoligotypes were shared with RD<sup>Rio</sup> and on over 60 randomly selected Mtb WT strains. Briefly, PCR reactions contained 10 to 20 ng of target DNA, 100 pmol of each primer pair, 0.2 mM of each dNTP, 1 U of Taq polymerase and PCR buffer to obtain a final concentration of MgCl<sub>2</sub> of 1 mM for MIRUs 20 e 26, 1.5 mM for MIRUs 16, 24 e 27, 2.5 mM for MIRU 39 and 2 mM for the rest, in a final volume of 25 µl. For amplification, samples were denatured at 95°C during 5 min, followed by 40 cycles of 1 min at 95°C, 1 min at 60°C (55°C for MIRU 24 and 26) and 1.5 min at 72°C, followed by a final extension step of 10 min at 72°C. Amplified products were visualised by electrophoresis in 2% agarose and ethidium bromide staining. Amplicon size was estimated by comparison against a 100-bp ladder. After determination of the number of repeats for each MIRU by comparison of the product size, the 12 number code of each isolate were introduced into Excel for grouping of identical patterns.

**Phylogenetic analysis.** The Bionumerics software (Version 4.0, Applied Maths, St. Martens Laetem, Belgium) was used for cluster analysis and evolutionary tree building. Spoligotyping and MIRU-VNTR patterns were introduced into Bionumerics using an Open Database Connectivity (ODBC) protocol and



a similarity matrix constructed using respectively the dice and the categorical index. UPGMA and neighbor-joining algorithms were used on both fingerprint patterns individually and the “average from experiments” option for analysis of the composite spoligo- and MIRU pattern dataset. Minimum Spanning Trees (MST) from spoligotype and MIRU patterns separately were also built for better understanding of the generation of the deletion within the analyzed strain set. For this purpose, either a binary (spoligopatterns) or a categorical (MIRU patterns) similarity coefficient was used and in both cases, priority for linking was given for spoligo- or MIRU patterns that had the highest number of single locus variants (SLVs). The creation of hypothetical types was allowed during construction of both MSTs.

### Statistical analysis

Statistical analysis were performed using the program STATA 7.0. Categorical variables were compared by the chi-square test, two-tail Fisher’s exact test or Yates. Student’s *t* test was used for continuous variables. To non-Gaussian variables, Mann-Whitney or Kruskal-Wallis were used and *p* value considered significant when  $< 0.05$ .

## RESULTS

**PCR panel results.** To determine the sensitivity of the Mtb PCR typing panel, we first performed PCR reactions using different dilutions of the DNA lysate and obtained amplification signals with volumes corresponding to 1 to 44 cfu, depending on the member of the MTC (see supplementary table 1 as supporting document on the PNAS website). The PCR panel was then applied to 430 mycobacterial clinical isolates and a clear amplification was achieved in 426 isolates (99%). The PCR typing panel clearly differentiate Mtb from MOTT, in accordance to the results later obtained from the mycobacterial laboratory and in the case of MOTT also by *16S rRNA* and *rpoB* sequence analysis: 407 isolates were Mtb (corresponding to 314 patients), 17 isolates were MOTT and in 2 isolates (from the same patient) a mix Mtb and MOTT was suggested by an unusual intensity pattern and later confirmed to be a mix of Mtb and *M. fortuitum* (see all MOTT in the supplementary table 2 on the PNAS website). Among the 407 Mtb isolates, 69% ( $n=281$ ) showed the typical Mtb profile (amplification of all primers) and were named wildtype (WT) genotype while in 30% ( $n=122$ ), the *IS1561*’ locus was missing and were named RD<sup>Rio</sup> genotype (Figures 1A and 1B). Approximately 1% ( $n=4$ ) of the Mtb isolates

was mixed WT/RD<sup>Rio</sup> (see below). Genotype frequencies among patients had similar proportion: WT in 69% ( $n=218$ ), RD<sup>Rio</sup> in 30% ( $n=93$ ) and mixed WT/RD<sup>Rio</sup> in 1% ( $n=3$ ) of the cases. The month proportion of each genotype was similar during the study period and no relationship between genotype and site of infection (i.e, respiratory versus non-respiratory) was observed (data not shown). Significantly, no other member of the MTC but *Mtb sensu stricto* was found.

**Characterization of the RD<sup>Rio</sup> deletion.** Primer sets were designed with increasing distance away from the IS1561' locus and a successful amplification in the RD<sup>Rio</sup> strains was obtained with primers targeting amplicon PE\_PGRS50 (Rv3345c) and fold (Rv3356c). Indeed, using PE\_PGRS50 as the forward primer and fold as the reverse, PCR successfully bridged the deleted segment for all ( $n=122$ ) IS1561'-deleted *Mtb* strains but none ( $n=281$ ) of the IS1561'-positive *Mtb* strains (Figure 1C and 1D, see all primers in supplementary table 3 on the PNAS website). The bridge amplicons of all RD<sup>Rio</sup> strains were sequenced and revealed an identical nucleotide sequence. Because the fact that Rv3346c and Rv3355c are 85% analogous in their gene sequences (accession number NP\_217863 at <http://genolist.pasteur.fr/tTubercuList>), homologous recombination was the mechanism for the deletion generation, creating a chimeric gene Rv3346/55c (Figure 2A). Neither tandem repeats typical for an insertion sequence mediated recombination event nor a clear breakpoint suggestive for DNA polymerase errors were present. The chimeric gene is composed of three parts: the first is identical only to Rv3346c, the middle is identical to both genes (125 nucleotides long) and the last is identical only to Rv3355c (Figure 2B). The homologous recombination did not create a new protein as the amino acid sequence of the chimeric protein is identical to the sequence of Rv3355c (data not shown). All genes located between Rv3346c and Rv3355c were removed in the RD<sup>Rio</sup> strains, generating a deletion of 26,314 kb (0.6% of the genome). Among the deleted segments were two PPE genes (PPE55 and PPE56), two non-functional transposable elements (Rv3348 and Rv3349c), three putative intermediate metabolism genes (Rv3351c, Rv3352c, and Rv3353c) hypothesized to have been an oxidoreductase that underwent frameshift mutations, an unknown function gene (Rv3354), recently described as one of the most conserved genes in the *Mycobacterium* genus (23) and portions of the Rv3346c and Rv3355c excised during the homologous recombination (the location and descriptions of the deleted genes are in the supplementary table 4 on the PNAS website).

**Characterization of mixed RD<sup>Rio</sup>/WT infection.** A mixed RD<sup>Rio</sup>/WT isolate was defined when all PCR systems of the panel were positive (WT profile) but also amplified in the PCR system for bridging the deletion (RD<sup>Rio</sup> profile). This was observed in 4 isolates obtained from three patients (0.6% of patients), with the two isolates from the same patient been collected in different days. We recultured these two isolates from their original tubes and submitted ten colonies from each subculture to the forementioned PCR systems, together with spoligotyping and MIRU-VNTR. All the mixed isolates were through this procedure segregated as either RD<sup>Rio</sup> or WT colonies, confirming the presence of different strains in the same patient (see supplementary table 5 on the PNAS website). They were excluded from additional analysis.

**Clinical and epidemiological results.** Patients infected by the Mtb RD<sup>Rio</sup> demonstrated hemoptysis ( $p=0,03$ ) and weight loss ( $p=0,06$ ) more frequently and were more likely to have high mycobacteria load as shown by the number of colonies on the cultures ( $p=0,05$ ) (Tables 1 and 2). Unfortunately, because only 146 patients (47%) had available medical records for review [46% WT (100/218) and 49% RD<sup>Rio</sup> (46/93)], these results should be interpreted with caution and a prospective study involving more patients is underway to confirm the findings. The geographic distribution of WT and RD<sup>Rio</sup> carrying patients was evaluated by plotting their home address onto a map of the City of Rio de Janeiro and no clustering of either RD<sup>Rio</sup> or WT strains could be observed as both were evenly distributed throughout the city (see supplementary figure on the PNAS website). To analyze whether RD<sup>Rio</sup> strains were present in older samples from Rio de Janeiro, 46 Mtb clinical isolates collected in 1998 from the same hospital were evaluated. One third of cases ( $n=15$ , 33%) presented the RD<sup>Rio</sup> genotype, a rate nearly similar to the current strain collection. The RD<sup>Rio</sup> isolates recovered from 1998 had the same chimeric gene in the seven ones sequenced; these isolates were not included in any further analysis.

**Identification of the RD<sup>Rio</sup> Mtb in international clinical isolates.** We have identified seven RD<sup>Rio</sup> strains outside Brazil and again, all have a chimeric gene identical to the Brazilian RD<sup>Rio</sup> strains. The RD<sup>Rio</sup> genotype was detected in strains from the USA ( $n=3$ ) (Ray Butler, CDC), the Netherlands ( $n=3$ ) (kindly provided by Dick van Soolingen) and from Djibouti, Africa ( $n=1$ ) (kindly provided by Michael Fabre). Of the three USA isolates, two were from the mother and her son who immigrated to Arizona from Mexico and the third isolate was a non-related homeless from Indiana. The Netherlands isolates were from a zookeeper and the related chimpanzee, and a third, non-related Dutch patient. The African

isolate belonged to a 2 year-old girl, member of a native tribe.

### **Classification of WT and RD<sup>Rio</sup> Mtb strains by molecular typing methods**

**Principal genetic groups.** From the four-hundred and seven isolates, 364 were classified as group 2 (89%), 40 as group 3 (10%) and 3 as group 1b (1%). Notably, the WT Mtb strains were represented in all three groups whereas all RD<sup>Rio</sup> Mtb strains were found only in group 2. Isolates from the same patient always belonged to the same genetic group.

**Spoligotyping.** The most predominant spoligotype, present in almost 60% of all TB cases in the strain population from Rio de Janeiro was the Latin American and Mediterranean (LAM) type (see supplementary table 5 on the PNAS website). All RD<sup>Rio</sup> strains belonged exclusively to this family, characterized by the simultaneous absence of at least spacers 21 to 24 and 33 to 36 (2). The RD<sup>Rio</sup> strains were subtyped mainly as LAM 9 (22%), LAM 1 (16%) and LAM 2 (18%), and in a lesser amount, LAM 4 (2.5%), LAM 5 (2.5%), LAM 6 (1.6%), five LAM types (5%) and twenty-three “new LAM types” (32.5%). The seven international RD<sup>Rio</sup> isolates belonged to LAM 9 (3 cases), LAM 4 (1 case), LAM type-125 (1 case) and one LAM “new type” (2 cases). Together, LAM 1, 2 and 9 accounts for almost 60% of all RD<sup>Rio</sup> strains. The WT genotype was more heterogeneous, the most prevalent families being LAM (40%), Haarlem (20%) and T (9%). Other families included S (2.7%), Beijing (0.7%) and East African Indian type (EAI) (0.3%), nineteen “type” (23%) and nine “new type” (5%). Among the LAM-WT, the predominant type was LAM 3 (7.5%), and, in smaller frequency, LAM 6 (5.3%), LAM 9 (2.1%), LAM 5 (2.1%), LAM 4 (0.7%), ten LAM types (8.5%) and twenty-three “new LAM types” (13%). Interestingly, LAM 1 and 2 were found exclusively represented in the RD<sup>Rio</sup> genotype while LAM 3 was solely represented in WT genotype. In contrast, RD<sup>Rio</sup> and WT genotypes were both represented in LAM 4, 5, 6 and 9 while none of the other LAM types or any “new LAM type” were shared by both groups. Overall, 118 different spoligotypes were observed, with 77 unique spoligotypes and 41 shared spoligotypes (cluster size ranging from 2 to 22 strains); 47 new spoligotypes were observed when compared to the spolDB3 database (available at <http://www.pasteur-guadeloupe.fr/tb/spolddb3>). When analyzing both genotypes separately, RD<sup>Rio</sup> accounted for 37 different spoligotypes (13 were shared and 24 were unique) whereas WT accounted for 85 spoligotypes (32 shared and 53 unique). To see the distribution of each spoligotype in both groups, go to supplementary tables 5 and 6 on the PNAS website.

**MIRU-VNTR.** MIRU-VNTR analysis was used to further discriminate both WT and RD<sup>Rio</sup> genotypes. A consensus pattern was observed in 5 different loci for the RD<sup>Rio</sup> strains: 86 of 93 patients carrying RD<sup>Rio</sup> isolates, had a 2-2-2-1-1 pattern respectively for MIRUs 2, 4, 20, 24 and 40. The most conserved repeat was MIRU 40, demonstrating a single copy in all RD<sup>Rio</sup> isolates while all LAM-WT strains had more than one copy in this allele. Another allele strongly associated with RD<sup>Rio</sup> was MIRU 2, presenting 2 copies in 98% (91 out of 93) of the RD<sup>Rio</sup> isolates, while in LAM-WT strains it was found in 33% (24 out of 71) of the isolates. Importantly, among the 19 patients in the LAM-WT whose spoligotypes were shared with RD<sup>Rio</sup>, 18 had only 1 copy in this MIRU. So, the RD<sup>Rio</sup> MIRU-VNTR pattern (2-2-X-X-2-X-1-X-X-X-X-1) in a LAM isolate was unique to RD<sup>Rio</sup> isolates as in none of LAM WT strains this was found, although it can be found in some non-LAM WT strains. In the case of RD<sup>Rio</sup> strains, both exceptions had just one copy in MIRU 2 but still presenting one copy in MIRU 40. Overall, 158 different MIRU-VNTR patterns were observed, with 126 unique MIRU-VNTR and 32 shared MIRU-VNTR, confirming the higher discriminatory power of MIRU-VNTR when compared to spoligotyping. When analyzing both genotypes separately, 58 different MIRU-VNTR patterns were found in the RD<sup>Rio</sup> group (62% of patients) and 100 MIRU-VNTR patterns in the WT group (75% of patients). However, when comparing MIRU pattern diversity only in LAM types in both groups, 30% of the 57 MIRU-VNTR analyzed LAM patterns in the RD<sup>Rio</sup> are shared when compared to 16% of the 43 LAM patterns in WT, suggesting that RD<sup>Rio</sup> is more closely related and are probably part of a large cluster. The predominate MIRU-VNTR pattern (224226153321) in LAM strains, shared by strains from 10 different patients, was unique to RD<sup>Rio</sup> isolates as in none of LAM-WT strains this profile was observed. Remarkably, MIRUs 2 and 40 can easily differentiate the four LAM subtypes shared by both RD<sup>Rio</sup> and WT. This reinforces that although they share some spoligotypes, RD<sup>Rio</sup> and WT LAMs have a diverse phylogenetic origin. An additional finding is that in the eight patients where multiple samples were analyzed, a concordant MIRU-VNTR between isolates was found in seven. The only patient whose samples gave different MIRU-VNTR results also had different spoligotypes, raising a question of additional different strains in the same patient. To see the MIRU-VNTR patterns and their frequency in patients infected by RD<sup>Rio</sup> and WT (LAM and Non-LAM) isolates, go to supplementary tables 5 and 7 on the PNAS website.

**Phylogenetic analysis using spoligotyping and MIRU-VNTR.** To visualize the generation of the deletion within the bacterial population, initially a MST was constructed considering all LAM-type spoligopatterns (68 different patterns from 239 strains). This tree however presented a large number of patterns outside of the main complex, due to the existence of nearest differences of patterns up to 9

spacers (data not shown), suggesting the existence of a very incomplete dataset, causing the creation of a lot of inconsistent spoligotype evolution proposals. Therefore, a MST using only the main spoligotypes and those being part of the main complex (37 spoligotypes from 186 Mtb strains) was created (Figure 3). The tree strongly suggests that the deletion was generated in a LAM 9 strain, forming the central WT and RD<sup>Rio</sup> containing precursor of all other LAM patterns observed in the present study. The presence of three other WT and RD<sup>Rio</sup> strain containing groups, LAM 4, LAM 5 and LAM 6, all derived from LAM 9 by a single spacer deletion, suggests that the spacer deletion events occurred several times in different strains. The LAM 3 group was apparently derived from one or more WT LAM 9 strains through the T81 intermediate while LAM 1 strains in Rio de Janeiro seems to have been generated from a deleted LAM 9 strain, generating LAM 2 by a second single spacer deletion event. Evolution of LAM 2 from LAM 1 and not from LAM 5 is, besides from the lack of WT strains in the former groups, also suggested by the distribution of number of strains among the different genotypes. Finally, all “next generation” LAM patterns, derived from the formerly mentioned patterns are constituted of either WT or RD<sup>Rio</sup> strains; no mixture of both genotypes is observed further than 1 spacer deletion events from LAM 9.

As a second approach, a MST was built from 107 different MIRU-VNTR patterns from 165 Mtb strains. Initially, a MST was built using MIRU-VNTR types from all strains (n=222) that had MIRU-VNTR typing performed, but this tree did not add any extra information that is already evidenced by the tree presented here (data not shown). Therefore, a MST using MIRU-VNTR only from all LAM strains was created (Figure 4). Interestingly, this MST shows that MIRU-VNTR patterns from WT strains are in completely different branches than those formed by RD<sup>Rio</sup> isolates, except for 3 outliers, one WT strain that was located in the RD<sup>Rio</sup> group and 2 RD<sup>Rio</sup> strains with a 1 locus difference of which one was linked to the WT branch. The reason for the misclassified WT strain is probably related to convergence of the MIRU-VNTR pattern as this strain had more than one copy in MIRU 40, and if using this MIRU allele for data mining, would have been considered WT. The two RD<sup>Rio</sup> isolates were misclassified because they had just one copy in MIRU 2, although both have one copy in MIRU 40. If only MIRU 40 was considered, all strains would have been correctly classified but the algorithm does not allow segregation based on the number of copies by a specific MIRU. Furthermore, because the program does not differentiate “gains” from “losses” and, in the case of MIRU, no importance is given to the copy number change at a certain locus, this representation should not be considered as an evolutionary analysis.

## DISCUSSION

Since *M. tuberculosis* strain H37Rv was sequenced in 1998, a lot of progress has been made in understanding the role of many genes, with putative functions being attributed to 52% of them and the remaining 48% being conserved hypothetical or unknown (24). Although the majority of deletions are considered deleterious to the bacteria (10,11,25), some deletions may bring advantages by escaping the immune system or increasing Mtb virulence (12,13). The transposition of mobile elements and homologous recombination between IS6110 elements are the major events contributing to Mtb diversity (8,26) and more recently, recombination between PPE genes was also described (23). Here we describe a large chromosome deletion resulting from homologous recombination generated by an undescribed mechanism in Mtb. The RD<sup>Rio</sup> deletion removed >26 kb and is the largest deletion described so far in any Mtb to our knowledge. Among the genes deleted were six genes which function still remains undiscovered and two PPE genes (PPE55 and PPE56). The PE/PPE family are mycobacteria-restricted genes comprising about 8% of their genome, postulated to be expressed on their surface and to be putative antigens in mycobacteria virulence or host immune response, promoting antigenic diversity, interference with antigen processing, immune recognition and survival inside the granuloma (27-30). The importance of variability within these elements is further supported by a 60% polymorphism rate between PE/PPE genes from *M. tuberculosis* and the closely related *M. bovis* (31) and these genes are also the main source of variation between Mtb strains H37Rv and CDC1551 (32). The two PPE genes removed in RD<sup>Rio</sup> were recently shown to be functional as they are induced *in vitro* when Mtb is exposed to IFN- $\gamma$  activated-macrophages and by the oxidative agent diamide (30). Interestingly, these two PPE genes are also deleted in CDC1551, a Mtb strain described as having caused outbreaks (33) and to elicit an early and vigorous cytokine response (3). We speculate that the loss of these PPE genes (and maybe others not yet defined) altered the antigenic presentation of the RD<sup>Rio</sup> strains to the immune system, making it harder for the defense mechanisms to recognize and eliminate the bacilli. Indeed, antigenic variation as a mechanism to evade the immune system has been described for a number of microbial pathogens (34,35). This hypothesis becomes even more likely after a recent observation that PPE55 is a strong immunogenic protein that elicits antibody production (36) and, therefore, its loss probably interferes with the immune response against Mtb RD<sup>Rio</sup>. The finding that all RD<sup>Rio</sup> Mtb strains are PGG2, have an identical recombinant gene sequence, belong exclusively to the LAM spoligofamily and have a particular MIRU-VNTR profile, not

shared with any other non-RD<sup>Rio</sup> strains, strengthen the concept that the deletion took place as a single event in an ancestral strain and that this genetic related advantage allowed these strains to thrive and spread.

The LAM family is a modern (PGG 2) Mtb, part of group VI SNP cluster (37) and is widespread in Europe and in Central and South America (2,20). In the largest international analysis of spoligotype distribution (20), the LAM family accounts for 22% of all Mtb in South America and was also observed in Central America (8% of Mtb isolates), Africa (5%), Europe (3%), North America (2%), and Oceania (1%). Moreover, the subfamilies LAM 9 (ST-42), LAM 1 (ST-20) and LAM 2 (ST-17) are considered ubiquitous, so the deletion observed in RD<sup>Rio</sup> strains is not restricted to Brazil, confirmed by the discovery of RD<sup>Rio</sup> strains in the United States, the Netherlands and in Africa. Furthermore, our finding that all LAM 1 (octal 677777607760771) and LAM 2 (octal 677737607760771) in our study are exclusively RD<sup>Rio</sup> suggests that these subtypes represent only this population. Therefore, the detection of LAM 1 or LAM 2 in countries like, Australia, Austria, Cameroon, Cuba, Ecuador, France, French Guyana, Germany, Great Britain and Ireland, Guadeloupe, Guinea-Bissau, Haiti, Indonesia, Martinique, Namibia, the Netherlands, the Russian Republic, Spain, Sweden, the United States, Venezuela and Zimbabwe, as verified in the SpolDB3 database, suggests the presence of RD<sup>Rio</sup> genotype in those countries. We also additionally identified them in Italy (16), Portugal (38) and in all Caribbean region (39) and because RD<sup>Rio</sup> is not restricted to these two LAM subtypes, its prevalence could even be wider. Finally, LAM 1 and LAM 2 alone have recently been estimated to make up 2.5% of all world spoligotypes (40). In summary, all these data strongly suggests that the RD<sup>Rio</sup> genotype may have some transmissible advantage over other strains, allowing them to disseminate worldwide. It is remarkable that anyone else has ever detected this deletion before. Moreover, Gagneux et al recently demonstrated that different Mtb strains may adapt better to different human populations (41) and that the Euro-American lineage of Mtb seems to be the best adapted and to more likely generate a secondary case when compared to other Mtb lineages.

It has been shown that particular strains are more prevalent in certain geographic regions, raising the question about the importance of genotype related bacterial factors in pathogenesis and presentation of disease and the importance of the interaction between parasite and host (42). Probably the best studied family in this regard in the widely distributed W/Beijing family, known to be involved in outbreaks (43), multi-drug resistance (44), febrile response during treatment (4) and to be more virulent and elicit a non-protective immune response in a mouse model (45). On the contrary, no



reports on clinical aspects, virulence or transmissibility of the LAM family exist, although this family is also widespread and considered one of the most important Mtb superfamilies (20). As our data showed a higher incidence of hemoptysis and probable weight loss in patients infected by RD<sup>Rio</sup> strains, suggesting a more severe infection, we may be facing another significant yet unrecognized Mtb population that deserves better characterization. Although we did not observe a higher cluster rate among RD<sup>Rio</sup> strains, they are undoubtedly infectious because of the demonstration of two pairs of strains, proven to be direct transmission by fingerprinting. One concerned a patient that was diagnosed for drug sensitive TB in 1996, became multidrug-resistant and transmitted the strain to his mother, who acquired multidrug TB in 1998. The second transmission case was between a zookeeper and a chimpanzee. These data demonstrate transmissibility of RD<sup>Rio</sup> strains but because our current sampling was not accompanied by data collection for definition of transmission chains in Rio de Janeiro, a prospective, population-based study is underway.

Several publications suggest that MIRU-VNTR, associated or not to spoligotyping, is the preferred method for Mtb genotyping and an adequate alternative to IS6110-RFLP (14,16-18). It is expected that with the discovery and improved combination of MIRU alleles, this technique will become the best tool for Mtb typing (16). MIRU-VNTR is potentially useful to study evolutionary genetics, as MIRU 24 is able to distinguish ancestral and modern Mtb (44). The present study adds another informative MIRUs loci for distinguishing what might be different Mtb phylogenetic lineages: all but two RD<sup>Rio</sup> strains were differentiated from the LAM-WT strains by the presence of two copies of MIRU 2 and one copy of MIRU 40. Amazingly, MIRU 40 is considered one of the most polymorphic loci (16,17,46), what makes the linkage of RD<sup>Rio</sup> with this MIRU allele even more significant. Although not always the case in genotypes other than LAM, the presence of two MIRU 2 copies and 1 MIRU 40 copy in a LAM strain is highly indicative for the presence of the deletion and a large scale study in different regions of Brazil and other countries from Latin America is underway. Remarkably, when analyzing literature data where MIRU-VNTR of LAM strains was reported, our proposed RD<sup>Rio</sup> associated MIRU-VNTR profile was found in all LAM 1 and LAM 2 and in none of the other non-LAM Mtb (16,48,49), suggesting that our results can be extrapolated to other places.

Another consideration is whether RD<sup>Rio</sup> strains represent a recent evolutionary branch or has been present for a long time. The finding that they had a chance to generate a considerable copy number diversity in different MIRU-VNTR loci and that our preliminary sampling data indicate the presence of these strains in several countries suggest that the deletion has been established and

spread quite some time ago.

One additional interesting finding was the documentation of simultaneous infection by two different strains in some patients (in our case, a mixed of WT and RD<sup>Rio</sup> strains), a situation already described between Beijing and non-Beijing isolates (50). Like in several other reports, our incidence of mixed infection was low (less than 1%) but these numbers may be underrepresenting the true mixed infection rates. Although uncommon, the presence of diverse populations of Mtb at the same time may have epidemiological and treatment implications as well as are source of studies of TB pathogenesis and host predisposition to disease. As described recently (51), MIRU-VNTR can be used to successfully identify and differentiate polyclonal infection.

Concluding, we uncovered here a new infectious Mtb genotype representing a significant proportion of clinical disease Mtb isolates from Rio that has an unprecedented large-scale chromosomal deletion resulting from homologous recombination between two genes that preserved the transmissibility and virulence of the organism. According to the population structure of the deleted and non-deleted strains, it is likely that the RD<sup>Rio</sup> deletion resulted from a single genetic event and subsequently diverged. Future studies might help to define the role of this population to global TB burden.

## **ACKNOWLEDGMENTS**

We are grateful to Solange Cavalcante, Coordinator of City TB Program, for allowing us to review the medical charts of patients from the health unit centers. We thank Dick van Soolingen for sending samples for analysis. We also thanks the medical students Thiago T. Mafort, Pedro G. Macedo, Rodrigo L. L. Fernandes, André P. Shecaira, Fernanda S. César, Edson da Silva S. Júnior e Cláudio J. Valença for helping in collecting clinical data.

Funding support was provided by NIH grants R0-1 AI39606 and R0-1 HL61960 (to J.L.H.), an NIH Fogarty International Center Training grant (FICTG) (D43 TW00018), a grant from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (Ministry of Education-Brazil), and a grant from the Laura Cook Hull Trust Fund (LCHTF) (Warren D. Johnson, principal investigator). LCOL was a FICTG and CAPES trainee and RCH was supported by the LCHTF. PNS and JRLP received funding

from CNPq, Milenium Institute, Pronex and PAPES/Fiocruz.

## REFERENCES

1. World Health Organization. 2000. Tuberculosis fact sheet no. 104. World Health Organization, Geneva, Switzerland.
2. Filliol, I., Driscoll, J.R., van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valétudie, V., Anh, D.D., Barlow, R., Banerjee, D., Bifani, P.J. *et al.* (2002). Global Distribution of *Mycobacterium tuberculosis* Spoligotypes. *Emerg. Infect. Dis.* **8**, 1347-49.
3. Manca, C., Tsenova, L., Barry, C.E. III., Bergtold, A., Freeman, S., Haslett, P.A., Musser, J.M., Freedman, V.H. & Kaplan G (1999). *Mycobacterium tuberculosis* CDC1551 induces a more vigorous host response in vivo and in vitro, but is not more virulent than other clinical isolates. *J. Immunol.* **162**, 6740-6.
4. van Crevel, R.V., Nelwan, R.H.H., Lenne, W., Veeraragu, Y., van der Zanden, A.G., Amin, Z., van der Meer, J.W.M. & van Soolingen, D. (2001). *Mycobacterium tuberculosis* Beijing Genotype Strains Associated with Febrile Response to Treatment. *Emerg. Infect. Dis.* **7**, 880-3.
5. Heym, B., Honore, N., Truffot-Pernot, C., Banerjee, A., Schurra, C., Jacobs, W.R. Jr, van Embden, J.D., Grosset, J.H. & Cole, S.T. (1994). Implications of multidrug resistance for the future of short-course chemotherapy of tuberculosis: a molecular study. *Lancet* **344**, 293-298.
6. Zhang, M., Gong, J.H., Yang, Z., Samten, B., Cave, M.D. & Barnes, P.F. (1999). Enhanced capacity of a widespread strain of *Mycobacterium tuberculosis* to grow in human macrophages. *J. Infect. Dis.* **179**, 1213-7)
7. Sreevatsan, S., Pan, X., Stockbauer, K.E., Connell, N.D., Kreiswirth, B.N., Whittam, T.S. & Musser, J.M. (1997). Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc. Natl. Acad. Sci. USA.* **94**, 9869-74.
8. Mahillon, J. & Chandler, M. (1998). Insertion sequences. *Microbiol. Mol. Biol. Rev.* **62**, 725-74.
9. Huard, R.C., Fabre, M., de Haas, P., Lazzarini, L.C.O., van Soolingen, D., Cousins, D., Ho J.L. (2006). Novel Genetic Polymorphisms that Further Delineate the Phylogeny of the *Mycobacterium tuberculosis* Complex (*J. Bacteriol.* accepted).

10. Tsolaki, A.G., Hirsh, A.E., DeRiemer, K., Enciso, J.A., Wong, M.Z., Hannan, M., de la Salmoniere, Y.O.L.G., Amam, K., Kato-Maeda, M. & Small, P.M. (2004). Functional and evolutionary genomics of *Mycobacterium tuberculosis*: Insights from the genomic deletion in 100 strains. *Proc. Natl. Acad. Sci. USA*. **101**, 4865-70.
11. Lewin, B. (1997) in *Genes VI*, eds. Oxford University Press, pp 569-572.
12. Parish, T., Smith, D.A., Kendall, S., Casali, N., Bancroft, G.J. & Stoker, N.G. (2003). Deletion of Two-Component Regulatory Systems Increases the Virulence of *Mycobacterium tuberculosis*. *Infect. Immun.* **71**, 1134-40.
13. Shimono, N., Morici, L., Casali, N., Cantrell, S., Sidders, B., Ehrt, S. & Riley, L.W. (2003). Hypervirulent mutant of *Mycobacterium tuberculosis* resulting from disruption of the *mce1* operon. *Proc. Natl. Acad. Sci. USA*. **100**, 15918-23.
14. Kanduma, E., McHugh, T.D. & Gillespie, S.H. (2003). Molecular methods for *Mycobacterium tuberculosis* strain typing: a users guide. *J. Appl. Microbiol.* **94**,781-91.
15. Suffys, P.N., de Araújo, M.E.I. & Degraeve, W.M. (1997). The Changing Face of the Epidemiology of Tuberculosis due to Molecular Strain Typing - A Review. *Mem. Inst. Oswaldo Cruz* **92**, 297-316.
16. Sola, C., Filliol, I., Legrand, E., Lesjean, S., Locht, C., Supply, P. & Rastogi, N. (2003). Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. *Infect. Genet. Evol.* **3**, 125-133.
17. Cowan, L.S., Mosher, L., Diem, L., Massey, J.P. & Crawford, J.T. (2002). Variable-Number Tandem Repeat Typing of *Mycobacterium tuberculosis* Isolates with Low Copy Numbers of IS6110 by Using Mycobacterial Interspersed Repetitive Units. *J. Clin. Microbiol.* **40**, 1592–1602.
18. Hawkey, P.M., Smith, E.G., Evans, J.T., Monk, P., Bryan, G., Mohamed, H.H., Bardhan, M. & Pugh, R.N. (2003). Mycobacterial Interspersed Repetitive Unit Typing of *Mycobacterium tuberculosis* Compared to IS6110-Based Restriction Fragment Length Polymorphism Analysis for Investigation of Apparently Clustered Cases of Tuberculosis. *J. Clin. Microbiol.* **41**, 3514–20.

19. Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D. & Locht, C. (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.
20. Filliol, I., Driscoll, J.R., van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valetudie, G. Anh D.D., Barlow, R., Banerjee, D. Bifani, P.J. *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**, 1963-70.
21. Huard, R.C., Lazzarini, L.C.O., Butler, W.R., van Soolingen, D. & Ho, J.L. (2003). A PCR-Based Method to Differentiate the Subspecies of the Mycobacterium tuberculosis Complex on the Basis of Genomic Deletions. *J. Clin. Microbiol.* **41**, 1637-50.
22. 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**, 907-14.
23. Marmiesse, M., Brodin, P., Buchrieser, C., Gutierrez, C., Simoes, N., Vincent, V., Glaser, P., Cole, S.T. & Brosch, R. (2004). Macro-array and bioinformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the *Mycobacterium tuberculosis* complex. *Microbiology* **150**, 483-96.
24. Camus, J.C., Pryor, M.J., Medigue, C. & Cole, S.T. (2002). Reannotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* **148**, 2967-73.
25. Kato-Maeda, M., Rhee, J.T., Gingeras, T.R., Slamon, H., Drenkow, J., Smittipat, N. & Small, P.M. (2001). Comparing Genomes within the Species *Mycobacterium tuberculosis*. *Genome* **11**, 547-54.
26. Ahmed, N., Alam, M., Rao, K.R., Kauser, F., Kumar, N.A., Qazi, N.N., Sangal, V., Sharma, V.D., Das, R., Katoch, V.M., Murthy, K.J. *et al.* (2004). Molecular Genotyping of a Large, Multicentric Collection of Tubercle Bacilli Indicates Geographical Partitioning of Strain Variation and Has Implications for Global Epidemiology of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **42**, 3240–47.
27. Brennan, M.J., Delogu, G., Chen, Y., Bardarov, S., Kriakov, J., Alavi, M. & Jacobs, W.R.

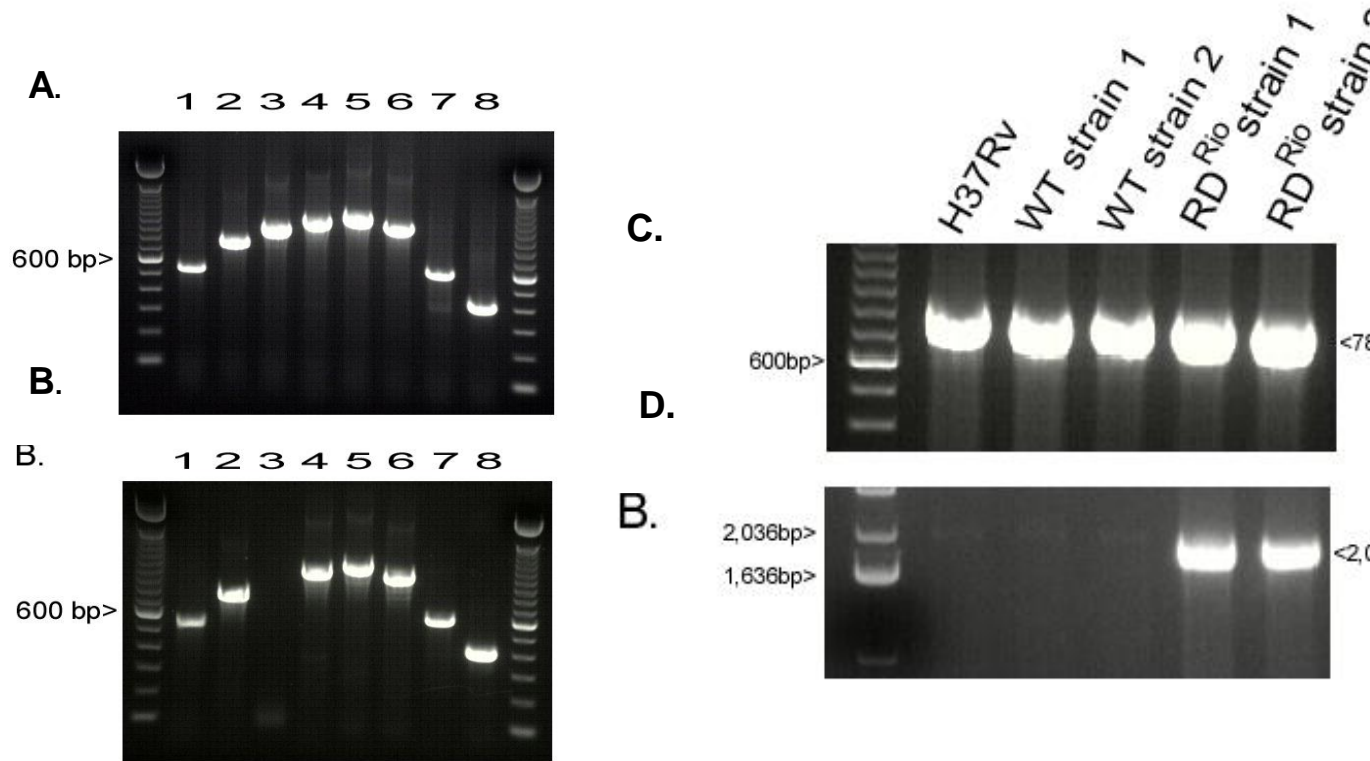
- (2001). Evidence that Mycobacterial PE-PGRS Proteins Are Cell Surface Constituents That Influence Interactions with Other Cells. *Infect. Immun.* **69**, 7326–7333.
28. Brennan, M.J. & Delogu, G. (2002). The PE multigene family: a ‘molecular mantra’ for mycobacteria. *Trends Microbiology* **10**, 246-9.
29. Ramakrishnan, L., Federspiel, N.A. & Falkow, S. (2000). Granuloma-Specific Expression of Mycobacterium Virulence Proteins from the Glycine-Rich PE-PGRS Family. *Science* **288**, 1436-9.
30. Voskuil, M.I., Schnappinger, D., Rutherford, R., Liu, Y., Schoolnik, G.K. (2004). Regulation of the Mycobacterium tuberculosis PE/PPE genes. *Tuberculosis*. **84**, 256-62.
31. Garnier, T., Eiglmeier, K., Camus, J.C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsempe, C. *et al.* (2003). The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci. USA* **100**, 7877–82.
32. Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, J., Venter, J.C., Fraser, C.M. *et al.* (2002). Whole-Genome Comparison of *Mycobacterium tuberculosis* Clinical and Laboratory Strains. *J. Bacteriol.* **184**, 5479–90.
33. Valway, S.E., Sanchez, M.P.C., Shinnick TM, Orme, I., Agerton, T., Hoy, D., Jones, J.S., Westmoreland, H., Onorato, I.M. (1998). An outbreak involving extensive transmission of a virulent strain of *Mycobacterium tuberculosis*. *N. Engl. J. Med.* **338**, 633-9.
34. Meyer, T.F., Gibbs, C.P. & Haas, R. (1990). Variation and control of protein expression in *Neisseria*. *Annu. Rev. Microbiol.* **44**, 451–77.
35. Gilsdorf, J.R. (1998). Antigenic diversity and gene polymorphisms in *Haemophilus influenzae*. *Infect. Immun.* **66**, 5053–9.
36. Singh, S.K., Dong, Y., Patibandla, S.A., McMurray, D.N., Arora, V.K., Laal, S. (2005). Immunogenicity of the *Mycobacterium tuberculosis* PPE55 (Rv3347c) Protein during Incipient and Clinical Tuberculosis. *Infect. Immun.* **73**, 5004-14.
37. Gutacker, M.M., Mathema, B., Soini, H., Shashkina, E., Kreiswirth, B.N., Graviss, E.A., Musser, J.M. (2006). Single-Nucleotide Polymorphism–Based Population Genetic Analysis of *Mycobacterium tuberculosis* Strains from 4 Geographic Sites. *J. Infect. Dis.* **193**, 121-8.
38. David, S., Portugal, C., Antunes, A., Cardoso, A., Caldo, A., Barros, V. & Sancho, L. (2004). Molecular identification using Spoligotyping of strains from the *Mycobacterium tuberculosis*

- complex isolated from the Hospital Fernando Fonseca. *Rev. Port. Pneumol.* **10**, 195-204
39. Sola, C., Filliol, I., Gutierrez, M. C., Mokrousov, I., Vincent, V. & Rastogi N. (2001). *Emerg. Infect. Dis* **7**, 390-396.
40. Brudey, K., Driscoll, J.R., Rigouts, L., Prodinger, W.M., Gori, A., Al-Hajoj, S.A.M., Allix, C., Aristimuño, L., Arora, J., Baumanis, V., Binder, L., Cafrune, P., et al. (2006). *Mycobacterium tuberculosis* Complex genetic diversity: mining the fourth international Spoligotyping database (SpolDB4) for Classification, Population. Genetics and Epidemiology. *BMC microbial.* **6**, 23
41. Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., 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*. *Proc. Natl. Acad. Sci. USA.* **103**, 2869–2873
42. Malik, A.N.J.; Godfrey-Fausset, P. (2005). Effects of genetic variability of *Mycobacterium tuberculosis* strains on the presentation of disease. *Lancet Infect. Dis.* **5**: S174-183.
43. Bifani, P.J., Mathema, B., Kurepina, N.E., Kreiswirth, B.N. (2002). Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. *Trends Microbiol.* **10**, 45-52.
44. Kruuner, A., Hoffner, S.E., Sillastu, H., Danilovits, M., Levina, K., Svenson, S.B., Ghebremichael, S., Koivula, T., Kallenius, G. (2001). Spread of drug-resistant pulmonary tuberculosis in Estonia. *J. Clin. Microbiol.* **39**, 3339-45.
45. Dormans, J., Buerger, M., Aguilar, D., Hernandez-Pando, R., Kremer, K., Roholl, P., Arend, S.M., van Soolingen, D. (2004). Correlation of virulence, lung pathology, bacterial load and delayed type hypersensitivity responses after infection with different *M. tuberculosis* genotypes in a BALB/c mouse model. *Clin. Exp. Immunol.* **137**, 460-468.
46. Sun, Y.J., Bellamy, R., Lee, A.S.G., Ng, S.T., Ravindran, S., Wong, S.Y., Locht, C., Supply, P. & Paton, N.I. (2004). Use of *Mycobacterium tuberculosis* Interspersed Repetitive Unit–Variable-Number Tandem Repeat typing to Examine Genetic Diversity of *Mycobacterium tuberculosis* in Singapore. *J. Clin. Microbiol.* **42**, 1986–1993.
47. Supply, P., Warren, R.M., Bañuls, A.L., Lesjean, S., van der Spuy, G.D., Lewis, L.A., Tibayrenc, M., van Helden, P.D. & Locht, C. (2003). Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. *Mol. Microbiol.* **47**, 529–538.

48. Cowan, L. S., Diem, L., Monson, T., Wand, P., Temporado, D., Oemig, T.V., Crawford, J. T. (2005). Evaluation of a Two-Step Approach for Large-Scale, Prospective Genotyping of *Mycobacterium tuberculosis* Isolates in the United States. *J. Clin. Microbiol.* **43**, 688–695.
49. Kwara, A., Schiro, R., Cowan, L. S., Hyslop, N. E., Wiser, M. F., Harrison, S. R., Kissinger, P., Diem, L., Crawford J. T. (2003). Evaluation of the Epidemiologic Utility of Secondary Typing Methods for Differentiation of *Mycobacterium tuberculosis* Isolates. *J. Clin. Microbiol.* **41**, 2683–2685.
50. Warren, R.M., Victor, T.C., Streicher, E.M., Richardson, M., Beyers, N., van Pittius, N.C.G. & van Helden, P.D. (2004). Patients with Active Tuberculosis often Have Different Strains in the Same Sputum Specimen. *Am. J. Respir. Crit. Care Med.* **169**, 610–614.
51. de Viedma, D. G.; Rodriguez, N. A; Andrés, S.; Serrano, M. J. R.; Bouza, E. (2005). *J. Clin. Microbiol.* **43**, 5660-64.

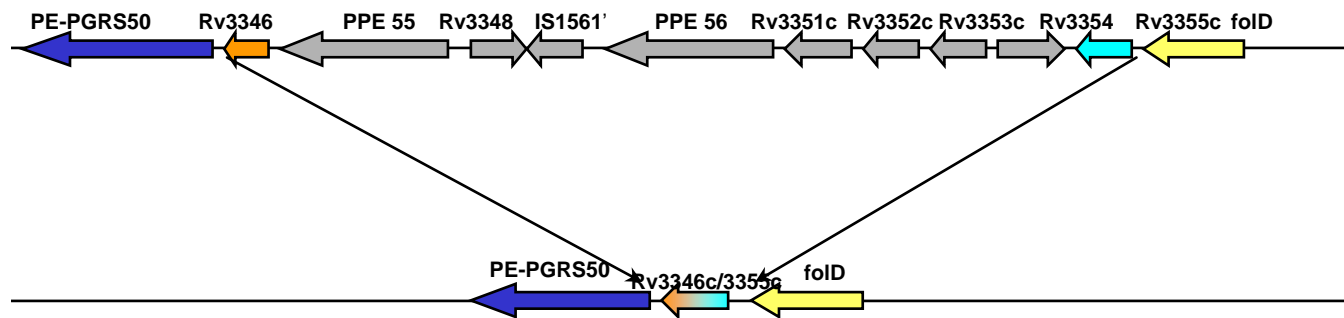


Figure 1.



**Figure 1. Panel A and B illustrated the composite *M. tuberculosis* complex (MTC) PCR typing panel.** The PCR panel is composed of eight primer pairs, which amplify within the specific target loci and run in separate but simultaneous reactions. Each primer pair either amplifies specifically for a DNA fragment of a unique size, or fails, depending upon the presence of the locus in the genome tested. **Figure 1A** is the typical amplification pattern in “Wildtype” (WT) Mtb strains and in **Figure 1B** is the typical amplification pattern in RD<sup>Rio</sup> Mtb strains. *Lanes:* 1, 16S rRNA; 2, Rv0577; 3, IS1561; 4, Rv1510 (RD4); 5, Rv1970 (RD7); 6, Rv3877/8 (RD1); 7, Rv2073c (RD9); 8, Rv3120 (RD12). Unlabeled lanes contain the 100-bp ladder. **Panel C and D illustrated the amplification bridging the RD<sup>Rio</sup> deletion.** **Figure 1C** shows the positive control amplification (Rv0577; reference 22). All strains amplified. *Lane 1*, Marker 100-bp ladder. **Figure 1D.** PCR amplification to bridge the RD<sup>Rio</sup> locus. Only RD<sup>Rio</sup> Mtb strains amplified; Mtb WT strains are negative because the amplicon is too large for the polymerase to complete the transcript under the PCR conditions and polymerase used. *Lane 1*, Marker 1-kb ladder; *Lane 2*, Mtb H37Rv; *Lanes 3 and 4*, two representative WT strains; *lanes 5 and 6*, two representative RD<sup>Rio</sup> Mtb strains.

A.



**B.**

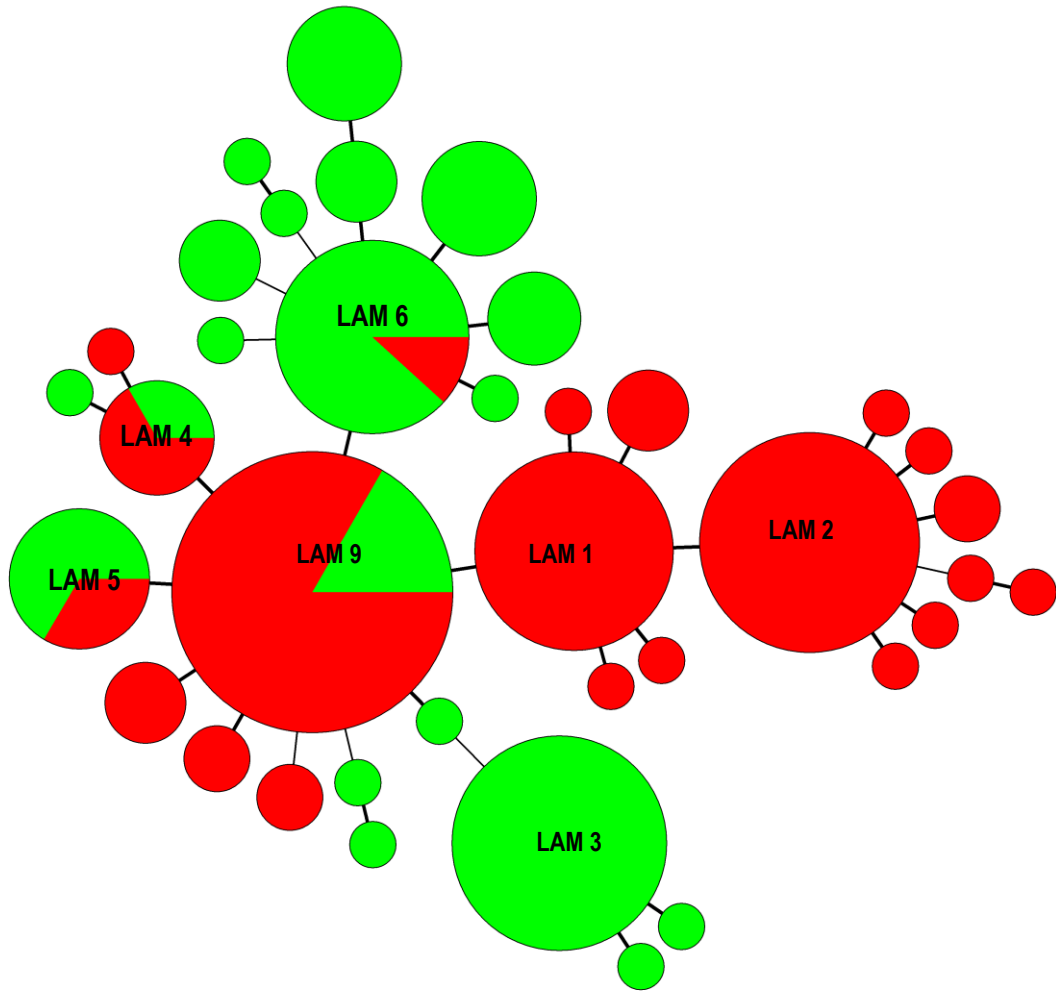
Position*	3.743.204	3.743.210	3.743.204	3.743.216	3.743.219	3.743.204	3.769.709	3.769.710	3.769.787
Rv3346/55	<b>C</b>	<b>C</b>	<b>G</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>A</b>	<b>G</b>	<b>G</b>
Rv3346	<b>C</b>	<b>C</b>	<b>G</b>	<b>A</b>	<b>G</b>	<b>A</b>	G	A	C
Rv3355	G	T	A	G	C	G	<b>A</b>	<b>G</b>	<b>G</b>

\* Position in relation to the Mtb strain H37Rv (available at <http://genolist.pasteur.fr/Tuberculist>)

**Figure 2. A.** Illustration of the deleted RD<sup>Rio</sup> locus in RD<sup>Rio</sup> *Mtb* strains. The recombination between 2 homologues genes Rv3346c and Rv3355c formed a new in-frame fusion sequence (Rv3346/55c). All elements between these 2 genes were excised and lost (comprising a length of 26,314 kb). **B.**

Comparison of the nucleotides in the chimeric gene (Rv 3346/55c) and the original Rv3346c and Rv3355c genes. Illustrated only the position where differences were found. The first five polymorphisms in the chimeric gene are identical to the Rv3346c whereas the last three are identical to the Rv3355c sequence.

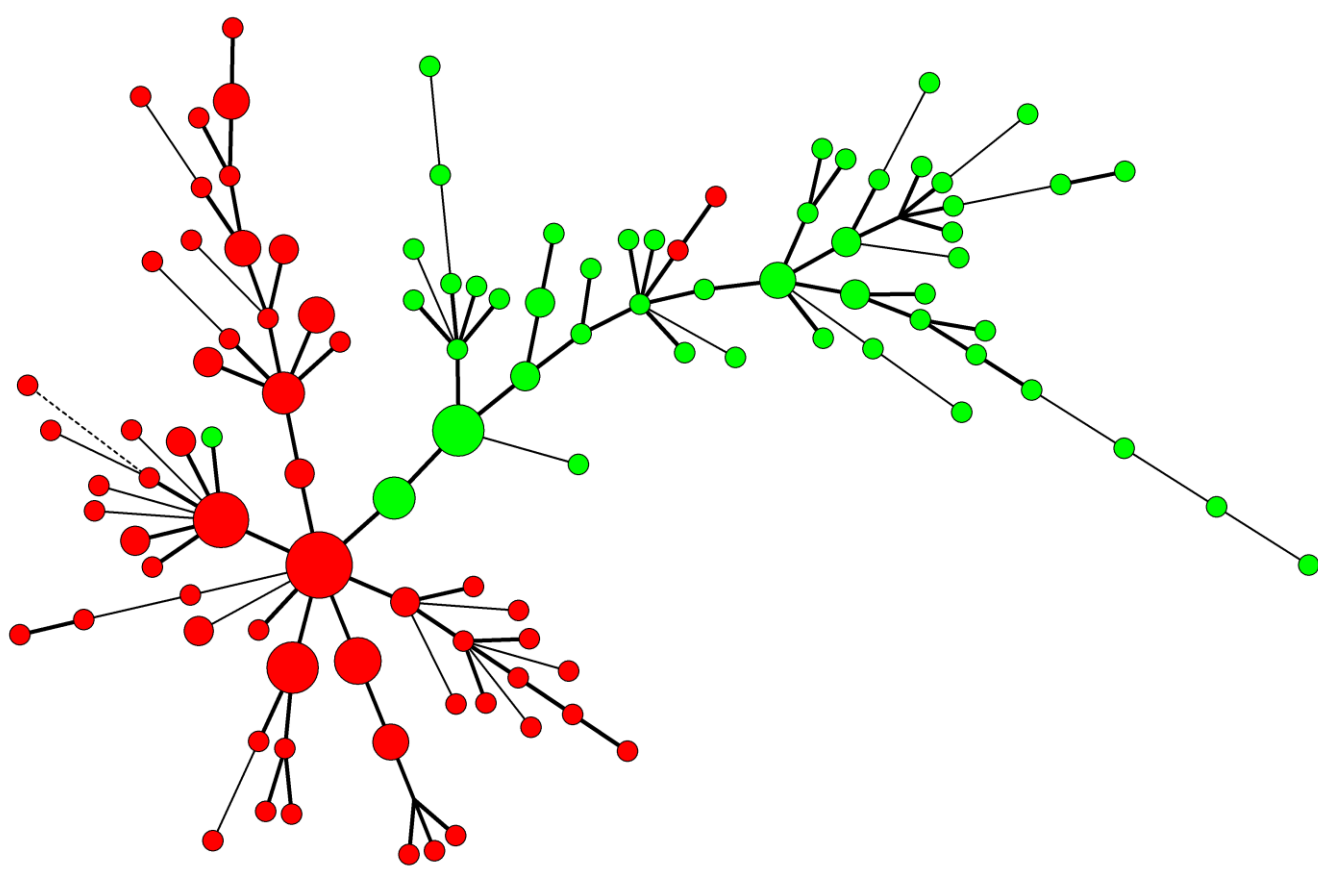
Figure 3.



**Figure 3.** The figure shows a Minimum Spanning Tree (MST) built from 37 different spoligotypes obtained from 186 *M. tuberculosis* strains. For its construction, a binary similarity coefficient and the highest number of Single Locus Variants as a priority rule were used; creation of hypothetical links was allowed. Colors indicate spoligopatterns of RD<sup>Rio</sup> strains (red) or WT strains (green) while circle size is related to the number of strains with the particular genotype; circles that are linked by a thicker line consist of strains that have a one spacer different spoligopattern while those linked by the thinner line have a 2-spacer difference. Besides the main LAM types indicated directly on the figure, the following types are presented: new29 (1), Type95 (2), new30 (3), new26 (4), Type176 (5), new43 (6), new28 (7), new47 (8), new35 (9), Type828 (10), new49 (11), new16 (12), new2 (13), new20 (14), Type1241 (15), new39 (16), Type81 (17), new40 (18), new41 (19), new7 (20), new13 (21), new22 (22), new11 (23), new 12 (24), new19 (25), Type194 (26), new23 (27), new10 (28), T960 (29) and new14 (30). Three strains, including one with the WT LAM new26 pattern and

two strains with the RD<sup>Rio</sup> LAM new17 also belonged to the complex but were excluded from the figure because they did not match the expected WT to RD<sup>Rio</sup> evolution pattern, probably because of the lack of the intermediate patterns between LAM9 and LAM new26, as evidenced by the creation of a hypothetical pattern (data not shown).

Figure 4.



**Figure 4.** The figure shows a Minimum Spanning Tree (MST) built from 107 different MIRU patterns from 165 *M. tuberculosis* strains. Settings from the construction of the tree were as in figure 3, except for the use of a categorical similarity coefficient. Green dots contain WT strains and red dots RD<sup>Rio</sup> strains; again, circle size is related to the number of strains with the particular genotype and linkage is as explained in figure 3.

Table 1.

<b>Clinical characteristic</b>	<b>RD<sup>Rio</sup></b>	<b>WT</b>	<b>p value (stat test)</b>
Mean age ± SD (yrs)	38 ± 15	36 ± 15	0,36 (a)
Gender (M/F) %	36 (50%)/ 36(50%)	108 (61%)/ 69(39%)	0,11(b)
HIV+ (% , n)	24% (6/25)	28% (18/63)	0,66 (b)
PPD mm	11 ± 8	10 ± 11	0,54 (a)
Fever (% , n)	57% (25/44)	68% (66/96)	0,16 (b)
Nights sweats (% , n)	23% (10/44)	31% (30/96)	0,3 (b)
Cough (% , n)	73% (33/45)	78% (75/96)	0,53 (b)
Sputum (% , n)	50% (22/44)	51% (49/96)	0,9 (b)
Hemoptysis (% , n)	26% (12/45)	12% (12/96)	<b>0.03*</b> (b)
Weight loss (% , n)	73% (33/45)	57% (55/96)	<b>0.06</b> (b)
Duration of symptoms (days)	92 ± 113	89 ± 109	0,95 (a)
Extrapulmonary TB (% , n)	43% (21/49)	40% (41/103)	0,72 (b)
<1 lung lobe involved in CXR(% , n)	53% (22/41)	57% (52/90)	0,81
>1 lung lobe involved in CXR(% , n)	46% (19/41)	42% (38/90)	
Cavitation(% , n)	36% (15/41)	33% (30/90)	0,71
Interstitial infiltrates	27% (10/36)	15% (12/80)	0,1
AFB in smears (% , n)	52% (59/114)	57% (15/272)	0,28
<b>No. of colonies on plate:</b>			
up to 19 colonies(% , n)	32% (36/110)	29% (79/270)	<b>0.05*</b> (b)
20-100 colonies (% , n)	35% (39/110)	37% (10/270)	
101-200 colonies (% , n)	14% (15/110)	23% (63/270)	
201-500 colonies (% , n)	18% (20/110)	10% (28/270)	
<200/>200 colonies (% , n)	82% (90/110) / 18% (20/110)	90% (242/270) / 10% (28/270)	<b>0.04*</b> (b)
<b>Sensitivity to anti-TB drugs:</b>			
Rifampin (% , n)	90% (67/74)	94% (169/180)	0,34 (b)
Isoniazid (% , n)	87% (90/103)	91% (225/247)	0,29 (b)
Ethambutol (% , n)	94% (97/103)	95% (235/247)	0,7 (b)
Etionamide (% , n)	99% (84/85)	99% (206/208)	0,86 (b)
Streptomycin (% , n)	100% (79/79)	97% (185/191)	0,11 (b)

*Abbreviation:* BAL: broncoalveolar lavage, CXR: Chest radiography, AFB: Acid-fast bacilli, stat: Statistical test used: (a): two-sample Wilcoxon (Mann-Whitney), (b) Pearson Chi-square.

Clinical records were reviewed by a team of researches lacking knowledge of the genotype (WT or RD<sup>Rio</sup>) of the patient's Mtb isolate. Of the total 309 patients evaluated, 219 (70.8%) were WT, 89 (28.8%) were RD<sup>Rio</sup> and 02 (0.3%) were mixed WT and RD<sup>Rio</sup>. Of these, 146 patients (47%) had available medical records for review: 45% WT (100/219) and 51% RD<sup>Rio</sup> (46/89). Only 50% of the patients came from our hospital and the others were scattered in several health centers.

\* - means statistically significant

**Capítulo X. Manuscrito 2**

Distribution of *Mycobacterium tuberculosis* Strains in Brazil as Reflected by Spoligotyping

Gomes HM et al

## Distribution of *Mycobacterium tuberculosis* strains in Brazil as reflected by spoligotyping

Harrison Magdinier Gomes<sup>1</sup>, Atina Ribeiro Elias<sup>1</sup>, Maranibia Aparecida Cardoso Oelemann<sup>1</sup>, Márcia Aparecida da Silva Pereira<sup>1</sup>, Fátima Fandinho Onofre Montes<sup>1</sup>, Luís Cláudio de Oliveira Lazzarini<sup>2</sup>, Ana Grazia Marsico<sup>2</sup>, Afrânio Lineu Kritski<sup>2</sup>, Luciano dos Anjos Filho<sup>2</sup>, Paulo Caldas<sup>3</sup>, Ângela Werneck<sup>3</sup>, Lia Gonçalves Possuelo<sup>4</sup>, Patrícia Cafrune<sup>4</sup>, Maria Lúcia Rossetti<sup>4</sup>, Norma Lucena<sup>5</sup>, Marcílio Fernandes Baliza<sup>5</sup>, Maria Helena Feres Saad<sup>6</sup>, Hebe Rodrigues Cavalcanti<sup>6</sup>, Clarisse Queico Fujimura Leite<sup>7</sup>, Rossana Coimbra de Brito<sup>8</sup>, Maria Luiza Lopes<sup>9</sup>, Karla Lima<sup>9</sup>, Maisa Souza<sup>9</sup>, Rita de Cássia Trindade<sup>10</sup>, Júlia Salen<sup>11</sup>, Cristophe Sola<sup>12</sup>, Nalin Rastogi<sup>12</sup> & Philip Noel Suffys<sup>1</sup>

<sup>1</sup>Laboratory of Molecular Biology Applied to Mycobacteria, Dept. Mycobacteriosis, Oswaldo Cruz Institute, Rio de Janeiro, Brazil. <sup>2</sup>Institute of Thoracic Disease. Federal University of Rio de Janeiro, Brazil. <sup>3</sup>Reference Center Prof. Hélio Fraga, Ministry of Healthy, Rio de Janeiro, Brasil. <sup>4</sup>Federal University of Rio Grande do Sul, Brazil. <sup>5</sup>Centro de Pesquisa Aggeu Magalhães. <sup>6</sup> Dept. Mycobacteriosis, Oswaldo Cruz Institute, Rio de Janeiro, Brazil. <sup>7</sup> Dept. Pharmacy State University of São Paulo, Araraquara, Brazil. <sup>8</sup>State University of Rio de Janeiro, Rio de Janeiro, Brazil. <sup>9</sup>Evandro Chagas Institute, Pará, Brazil. <sup>10</sup>UFederal University of Sergipe, Brazil. <sup>11</sup>Instituto de Pesquisas da Amazônia., Brazil. <sup>12</sup> Institute Pasteur de Guadeloupe, França

Corresponding author: Fax: +55-21-2270.9997. Email: psuffys@ioc.fiocruz.br

Key words: Spoligotyping, Brazil, *Mycobacterium tuberculosis*

### Abstract

Tuberculosis (TB) is an important cause of death in Brazil and the countries highest incidence is in Rio de Janeiro state (160/100.000), but the disease are spread in all states with different frequency. Genotyping has been performed for better understanding of Mycobacterium population structure and TB transmission and techniques such as IS6110-RFLP, MIRU-VNTR and Spoligotyping have been used widely. The latter procedure has been used to visualize the global spread of *M. tuberculosis* genotypes (SpolDB) and following this approach, for better knowledge of the *M. tuberculosis* genotypes spread over the country, we performed spoligotyping from isolates obtained from residents of 11 states, including Amazonas, Ceará, Goiás, Minas Gerais, Pará, Paraná, Pernambuco, Rio Grande do Sul, Rio de Janeiro, São Paulo and Sergipe). Among the 2005 isolates, 181 shared



spoligotypes totalizing 1582 isolates, 118 orphan isolates and 302 new spoligotypes were observed (or 302 strains belonging to new types;). The three most frequently observed families were the Latin-American and Mediterranean family (LAM, 46.5%), the T family (19.7%) and the Haarlem family (13.3%). Within the LAM family, the classes LAM7, LAM8 and LAM10 were not detected while LAM9 (ST42) was the most frequent (21,11%); the recently described LAM11 was present with minor frequency. Some genotypes such as the East African Indian (EAI) were almost exclusively observed in the North of the country suggesting regions-specific genotype distribution while the Beijing family was present at very low frequency. The region associated genotype distribution and speculation of the influence of human population migration is presented.

## Introduction

About two billion people are infected with *Mycobacterium tuberculosis* and this suggests the existence of a long lasting and close relationship between the parasite and its host, as supported by the detection of *M. tuberculosis* in several thousand-years-old Egyptian (Zink et al, 2003) and Peruvian mummies (Salo et al, 1994). The oldest infection with *M. tuberculosis* has been observed in a bison in North America prairie around 17,000 years ago and probably *M. tuberculosis* reached the American continent when the bovines and humans crossed the Behring strait (Rothschild et al. 2001). In fact, one speculates that the geographic origin of *M. tuberculosis* is the African continent, origin of its principal host, *Homo sapiens sapiens sapiens*.

Analysis of genome sequences and of deletion distribution in organisms of the *M. tuberculosis* complex (MTBC) revealed that *M. bovis* and *M. tuberculosis* evolved separately from a common precursor and this species definition seems to have occurred about 20,000 years ago (Mostowy et al, 2005 & Sreevatsan et al, 1997). Upon sequencing of MTBC isolates from European and African patients, a better organization on genetic level of the complex was presented, including the definition of *M. prototuberculosis* as a general precursor and co-existence of humans with these organisms since about 3 million years ago (Gutierrez et al. 2006).

In spite of this long standing interaction between parasite and host, most severe tuberculosis (TB) epidemics were reported mainly in Europe between the 16<sup>th</sup> and 18<sup>th</sup> century, being responsible at that time for one in every four deaths (Fletcher et al, 2003). At present, TB continues to causes more adult deaths worldwide than any other infectious disease (Blower et al., 1996) and many factors contributed to TB transmission and disease dynamics, co-infection with HIV and development of multi-drug-resistant bacteria among the most important. Therefore, TB was considered a reemergence disease by the World Health Organization (WHO) with the highest incidence being concentrated in the Third-World. Many countries in the African and Asian continent and some in Latin-America present TB incidence rates between 100-300/100,000. Although Brazil presents a milder national incidence rate of 58/100,000, a considerable contrast in TB levels exist in regions of the country, ranging from 31/100,000 in the Southern state of Paraná to 160/100,000 in the state of Rio de Janeiro, Southeast Brazil (Kritski & Ruffino-Netto, 2000).

Brazilians are among the ethnically most heterogeneous populations and this is the result of five centuries of ethnic mixing of people from Europe, Africa and America (Parra et al., 2003). The influence of ethnicity on susceptibility to infection and disease development has been described before and of the most striking examples is infection with plasmodium, causing malaria (Kwiatkowsky, 2000). Very recently, it was suggested that different ethnic groups had different susceptibility for particular *M. tuberculosis* strains and that this could even be observed in complex population settings (Gagneux et al, 2006).

Besides host genetic background, human migration has also been demonstrated to influence the genetic composition and region associated genetic variability of parasites, as exemplified in leprosy (Monot et al 2005); this is also the case for *M. tuberculosis* (Mokrousov et al, 2006). Therefore, the study of *M. tuberculosis* genotype from different states and regions of Brazil should permit us to better understand the importance of migratory flux during five centuries to design the molecular epidemiological profile in Brazil. We therefore performed spoligotyping from over 2,000 MTBC strains from 11 Brazilian states

## Materials and Methods

### ***Strains and patients.***

Isolates of TB patients had been cultured in Löwenstein-Jensen medium and were obtained from several institutions among the country, after being submitted to conventional identification procedures (Canetti et al. 1963). Bacterial mass was submitted either to DNA extraction using a CTAB- based protocol as described by van Soolingen et al. (1991) or submitted to freeze-thaw lyses.

### ***Spoligotyping.***

This typing procedure was performed slightly modified from the protocol described by Kamerbeek et al. (1997). Ten nanograms of purified DNA or 5 ul of lysate was submitted to PCR amplification of the DR region using 20 pmol of the primers DRa (biotinilated) and DRb and the rest of the PCR mixture containing 1U of Taq Polimerase, 0.2uM of dNTPs and 1.5mM of Magnesium Chloride Amplification was performed by initial denaturizing at 96°C by 3 min followed by 35 cycles of 1 min at 96°C, 1 min at 55°C and 30 sec at 72°C, and a final extension of 5 min at 72°C. For reverse hybridization, 30 uL PCR products were diluted in 30 µL of 0.1% SDS/2x SSPE, denatured at 99°C by 10 min and put on ice bath until application in the MN45 miniblotter (Isogen, the Netherlands). The spoligotyping membrane (Isogen) was pre-washed at 60°C in 0.1% SDS /2x SSPE during 5 min, and exposed to the PCR products in the miniblotter at 60°C by 60 min. After this, the membrane was washed twice with 0,5%SDS/2xSSPE at 60°C and incubated at 42°C for 60 min with 10 ul streptavidin peroxidase (Amersham Bioscience, UK) in 10 ml of 0,5%SDS/2xSSPE, followed by washing twice in 0,5% SDS /2x SSPE at 42°C during 10 min. For SDS removal, the membrane was rinsed twice with 2x SSPE at room temperature and hybridization visualizes by incubating 1 min with 15 mL of ECL liquid detection solution and exposure to a Hyperfilm for 2 hours.

### ***Computer Analysis***

All spoligopatterns were introduced in an Excel spreadsheet by converting to a binary format. Spoligopatterns were compared with those present in the international database of spoligoypes (spolDB4.0 – <http://www.pasteur-guadeloupe.fr/tb/spolddb4>; organized and maintained by Tuberculosis Unit from Institute Pasteur de Guadeloupe, France) by Dr Cristophe Sola. The spoligotypes were classified into types, classes and families according to the nomenclature described by Filliol et al. (2002). Excel tables using spoligotypes were introduced into Bionumerics (Version 4.2, Applied Maths, Belgium) using an Open Database Connection Protocol (ODCP) and similarity matrix and phylogenetic trees constructed using the binary similarity coefficient and Neighbor Joining and Minimum Spanning Tree algorithms.

## **Results**

A total of 2005 isolates gave good quality spoligopatterns and these included strains from the states from the Northern states Amazonas (n=45, 2.24%) and Pará (n=186,9.27%), the Northeastern states Pernambuco (n=78, 3.89%), Sergipe (n=49, 2.49%) and Ceará (n=71, 3,54%), the Southern states Rio Grande do Sul (n=294, 14.66%) and Paraná (n=6, 0.299%), the Southeastern states Rio de Janeiro (n=1034, 51.57%, São Paulo (n=177, 8.82%) and Minas Gerais (n=55, 2.74%) and the Middle-Eastern state Goiás (n=7, 0.349%). The Southeastern region was the origin of 63% of iolates, 14.96% came from Southern region, 11.52% came from Northern region, 9.87% came from Northeastern region and 0.349% came from Middle-Eastern region.

Spoligotypes were grouped in 181 Shared-types, totalizing 1582 isolates, while were 118 (5.9%) the orphans isolates; 15.06% (n=302 spoligotypes) were not present on the database (didn't match with any entry in SpolDB4) it was denominated news. Among the isolates, 46.53 % were identified as belonging to the Latin-American and Mediterranean (LAM) family, 13.26% as Haarlem family, 19.70% as T family (T1, T2, T3, T4 and T5), 1.74% as S family, 0.15% as Beijing family (3 isolates), 1.047% (21) as East-African Indian (EAI) family, 0.199% (4) as Manu Family and 7.1% of the isolates were unclassifieds.

The high percentage of LAM family in Brazil isolates is distributed in the classes LAM1 (5.46%), LAM2 (9%), LAM3 (7.28%), LAM4 (1.5%), LAM5 (2.46%), LAM6 (7.71%), LAM9 (21.11%) and LAM11 (0.32%), totalizing 54.55.36% from all isolates of LAM Family and 10.71% of spoligotypes of LAM, weren't classified (unclassified) and are recognized only by the type number., 16.29% of the isolates from these family were news. The class LAM7 (ST41), LAM8 (ST290) and LAM10 (ST61) were absent in the spoligotypes from Brazil. Only one strain (ST1815) found in Rio Grande do Sul was classified how to an intermediate spoligotype between LAM1 and LAM8, called in SpolDB4 LAM1-LAM8.

The 266 starins from Haarlem family are spliced in 4 classes and 20.3% of this total belongs to the Haarlem1 Class (ST47), 41.72% to the Haarlem3 (ST50) and 3.0% to the Haarlem2 (ST2) and 0.75% from Haarlem4 (ST777) these percentages are completed by 7.52% of new spoligotypes.

,The EAI family are represented by 21 isolates distributed in 3 Brazilian states, 15 of these isolates came from Pará state, North Region, (EAI5 ST48 n=4; EAI5 ST702 n=4; EAI5 2543 n=2; EAI5 ST129 n=1; EAI5 ST924 n=1 and 3 news ), one isolate (ST1983) from Recife, capital of Pernambuco (Northeast Region) and 5 isolates from Rio de Janeiro ( four new and one EAI3\_IND ST11).

The Manu family is represented by Manu2 class (ST54) in Rio de Janeiro (n=3) and in São Paulo (n=1) and another isolate by ST59 (LAM11\_ZWE) could be found in a low frequency in Rio de Janeiro (n=1), São Paulo (n=1) and Ceará (n=1).

The class Haarlem2 is represented by 8 strains distributed in states of, Amazonas (n=4), of Rio de Janeiro (n=2),of São Paulo (n=1) and of Rio Grande do Sul (n=1). The Haarlem3 class was found hole states analyzed except in Goiás state and the Haarlem1 class wasn't observed in 4 states (Ceará, Minas Gerais, Sergipe, and Goiás). The members of Haarlem family that were classifieds only by ST weren't found in 2 states (Goiás and Paraná).

The LAM family is present in all states of Brazil, but the same doesn't occur with LAMs Classes. The LAM1 (ST20) wasn't found in 5 states (Minas Gerais, Amazonas, Segipe, Goiás and Paraná), LAM2 (ST17) wasn't found in São Paulo e Paraná, LAM3 (ST33) wasn't found in Minas Gerais, Amazonas, and Goiás, LAM4 (ST60) wasn't found in 7 states, LAM5 (ST93) wasn't found in 4 states (Minas Gerais, Amazonas, Pernambuco and Ceará), LAM6 (ST64) wasn't found in Minas Gerais, Sergipe, Goiás e Paraná, LAM9 (ST42) WAS present in all states, except in Paraná, the new types and isolates classifieds only by ST wasn't present in Goiás and Paraná state.

The total number of isolates from S family, was 35 and these isolates were concentrate in three states, Rio de Janeiro (n=29), São Paulo (n=2) and Rio Grande do Sul (n=4). The ST34 is the prototype of S family and was the most frequent strain from this family found in Brazil in a total of 30 strains.

. The T family represent 19.7% from the total of Brazilian spoligotypes and the ST53 is the most frequent class of this family in Brazil with 38.22% from the total of T family. The frequency of class T2 (ST52), T3 (ST37), T4, T5 and six new ST is approximately 7% and others ST from this family are composed by

different spoligotypes that belong to T1 to T5 class. We noted the absence of this family in Sergipe, Goiás and Paraná, and 65.6% of these isolates came from Rio de Janeiro.

Other ST were found specifically or preferentially in one region of Brazil, specifically in Rio de Janeiro were found the types: ST18 (n=15), ST37 (n=5), ST156 (n=4), ST370 (n=6) ST1163 (n=5), ST2495 (n=5) not described in spolDB4, ST746 (n=4), ST176 (n=7), ST534 (n=4), ST770 (n=8), ST1106 (n=4).. In the state of São Paulo we found the ST481 (n=7), in Pará we found the ST2517 (n=8), in Pernambuco the strain specifically was 2553 (n=5) and other STs were shared between one, two or three states: ST65 (Rio Grande do Sul n=21; Rio de Janeiro n=3; São Paulo n=1), ST119 (Rio Grande do Sul n=9; Pará n=1; Pernambuco n=1), ST157 (Amazonas n=8; Rio de Janeiro n=7; São Paulo n=1), ST240 (São Paulo n=5; Rio Grande do Sul n=1), ST291 (Rio de Janeiro n=7; São Paulo n=1), ST462 (Rio de Janeiro n=5; Rio Grande do Sul n=1); ST498 (Rio Grande do Sul n=10; Pará n=1), ST535 (Rio de Janeiro n=13; São Paulo n=5).

The total frequencies of LAM family in Brazil obey the same behavior to the states, where the LAM frequency to the major of the states is around 50%..

The cluster tax observed in this study was approximately 80%. The minor cluster is composed by two isolates (164 isolates) and the major cluster is composed by the type 42, LAM 9 (197 isolates)..

### **Discussion**

The Brazil is a Federative Republic composed by 26 states and 1 Federal District, the last demographic statistic point to a population of around 170 million of habitants and 70,78% of this people live in the eleven states that are responsible to supply the hole samples used in our study. The origin of principal group of isolates (42%) came from the Southeastern region and due the fact that is the richest and industrialized Brazilian region and one of the most important region from South America, the Southeastern receive migrants from hole parts of Brazil and probably is representative in relation to culture, people and other aspects inherent to human been.

The diversity genetic of tuberculosis in Brazil is designed by politic, economic and social aspects of his people and these aspects were building by five centuries of interethnic relationship from people of the whole world. [In the first decade of the 19<sup>th</sup> century, the Brazilian ports were opened to all friendly nations and this attitude was responsible to increase the numbers of immigrants from all parts of the world.](#) Between 1500 (Brazil discovery) and 1972, 58% of the immigrants who arrived in Brazil were Europeans, 40% Africans and 2% were Asians, between 1551 and 1850 (when the slave trade was abolished), approximately 3,5 million Africans arrived in Brazil (Alves-Silva et al., 2000).

The global study of spoligotypes receive an important improve after the Institute Pasteur of Guadeloupe start to publish the first version of the international database of spoligotypes in 1999. Since the presentation of the first version until the last version of these databases, SpolDB4 (2006) (Sola et al. 1999, Sola et al. 2001, Filliol et al. 2003, Filliol et al 2006), the number of spoligotypes patterns increase from 53 (SpolDB1) to 1939 (SpolDB4) and this last version include 122 countries. The spolDB offer a panoramic view of *M. tuberculosis* strains distribution in the world and each edition change this vision. One example is the detection of 13 isolates of Strain Type 18 collected during the years 1997, 1998 and 1999, related to Brazil in SpolDB4, before this, the strain type 18 was related, in low frequency, to GBR, Nederland and United States, in despite of the occurrence of the ST18 can be documented in Brazil database after 1997 its frequency in the world continue to be low and in Brazil is specific of Rio de Janeiro. The ST18 belong to the X family that may be linked to Indian or Anglo-Saxon ancestry (Brudey et al 2006), the frequency of X family in Brazil is 4% and the isolates are distributed in

Pará, Pernambuco, Ceará, Rio Grande do Sul, Rio de Janeiro and São Paulo, even though the low frequency of this family in Brazil its spread in 6 from the 11 states included in this study suggesting a good capacity disseminates.

From the Spoligotypes patterns, approximately 51% was isolated in Rio de Janeiro, this fact could be explain the great number of ST specifically of Rio de Janeiro, however ST found only in this state can be consider rare, because the distribution of these spoligotypes are limited to a few countries and in a low frequency, one example is the ST1163 present in Brazil (5 isolates) and United States (4 isolates), another example is the ST176 distributed in Brazil (7 isolates), French Guiana (2 isolate), Central African Republic (1 isolate) and United States of America (12 isolates).

In the analyze of SpolDB4, the LAM family in South America maintain the frequency of approximately 50%, and these frequency is too observed in Brazil that present 46,53% of Isolates belonged to LAM family, but the frequency of LAM in Brazil states shows some difference. The highest frequency of Brazilian LAM, was observed in Sergipe state with 81.6% followed by 70% from Ceará state and the low frequency of LAM family in Brazilian states was observed in São Paulo state with 35% (show table frequency). The behavior of Brazilian spoligotypes patterns, obey the distribution of LAM family in South America, and the same difference can be observed between Venezuela with 65% of Isolates belong to the LAM family and 17.4% in French Guiana and this difference point to a strong relationship between colonizers and Tuberculosis.

The origin of *M. tuberculosis* in Brazil can be related to the exclusive relationship between Portugal and Brazil, during approximately 300 hundred years (1500 – 1808). The 50% of the STs more frequent in Portugal are LAM1 ST20 (11,1%), LAM9 ST42 (9,8%), T1 (6,6%), Haarlem1 (4%) and 3 (3,1) and LAM6 ST64 (2,9%) and the ST244 (6%), ST150 (4,2%) and ST389 (3,1%), the Brazilian frequency present many similarities with these results, although there are an important difference in frequency of LAM1 (2,5%) and ST244 (0,38%), ST150 (0,4%) and ST389 (0,05%) (Suzana Davi, 2006).

Other important aspect in spoligotypes patterns from Brazil and Portugal is the frequency of LAM family, which is high and so close in Brazil and in Portugal, 46.53% and 51.19% respectively is about 0.47%. Many countries (Angola, Mozambique, Guinea-Bissau, Cabo Verde, São Tomé e Príncipe e Timor East) and regions, such as Goa, Diu, Damão, Galiza e Macau were colonized by Portuguese and all of them are in the area, where LAM family is highly prevalent, but there are spoligotypes in SpolDB4 only from Mozambique (n=28), Angola (n=4) and Guinea-Bissau (n=217) and the frequency of LAM family in these countries is 28.57% in Mozambique and 16,13% in Guinea-Bissau. The other spoligotypes **that are in destac in Portugal is ST150, ST244 and ST389. The ST389 is the most conspicuous strain of this group because there are four countries (Brazil, Finland, Portugal and Venezuela) where this strain could be found, and Portugal and Brazil concentrates 78,57% of the ST389 (11 from 14)**

In the account of SpolDB3, the most significant intercontinental match was observed between North America and Europe, one explanation to this match is the historical relationship of these regions since the begin of North America colonization, and an other alternative is the recent transmission between identical high prevalence countries (Filliol snapshot). The migratory flux between countries or region can be explaining the specific presence of the East-African Indian family (EAI) in Belém (PA). The EAI is spliced in 5 main subclades (Filliol et al.,2002 global distribution) and the subclades I and V (ST48, ST129, ST702, ST924 and ST2543) were found, in a high frequency in Pará state, from a total of 21 isolates, characterized by the absence of spacers 29-32 and 34, 16 of them are found in Pará, 4 in Rio de Janeiro and 1 in Pernambuco In South America these types were found in a low frequency in the French Guiana and two isolates from ST702 was found in Cuba,

Considering the presence of EAI isolates in French Guiana and its geographic proximity with the Pará state, there are a strong possibility to exist a epidemiological link between French Guiana and Belém (PA).

The presence of spoligotypes apparently uncommon in Brazil could be related to low frequency of isolates from Brazil in SpolDB4 (spoligotypes brasileiros no SpolDB4 %) or by the absence of spoligotypes from at least a half of Brazilian states in our database or by the strict biogeographic preference of Families, clades or sub-clades from *M. tuberculosis* strains and these biogeographic preference can be realize clearly in SpolDB4 (Brudey et al, 2006). The Manu family, recently characterized how to a family from India, and probably the ancestor of EAI, CAS and Beijing families. This family is represented in Brazil by 4 isolates of sub-clade Manu2 (ST54) genotyped in Rio de Janeiro (n=3), São Paulo (n=1) and due the low frequency and its state origin there is a possibility to this strains get a epidemiological link. The 4 isolates from Manu2 represent the total of this family in South America.

The same possibility is considered to the presence in Brazil of LAM11\_ZWE (ST59) with 57,8% of isolates from Zimbabwe and with 4 isolates (ST54) in Brasil, in states of Rio de Janeiro (n=3) and Minas Gerais (n=1), this is the total number of Manu2 in South America. In the situation of T1\_Russia2 ST280 (n=1) and T1\_Tuskany Variant ST 159 (n=1), both isolates from Rio de Janeiro, is possible to consider many possibilities, such as: patient origin, foreign contact, touristy or busy travel. An other example is the ST 58, T5\_Madrid2 highly prevalent genotype that is potentially specific Spain -related genotypes. In Brazil the Madrid2 genotype are spread in four states: Pará (ST58, n=1), Rio de Janeiro (ST58 n=4; ST1227 n=3), Rio Grande do Sul (n=3) and São Paulo (n=1).

The absence of LAM 7 in Brazil, apparently show that this class has an other biogeographically preference and this could be observed in the SpolDB4 where 26% of all isolates came from Turkey . The similar event occur with LAM 10 that is absent in Brazil and prevalent in Cameroon (more than 50% of all isolates) and LAM8 ST290 (Brudey 2006)

## References

Salo, WL, Aufderheid AC, Jane buikstra and Holcomb. 1994. Identification of *M. tuberculosis* in a pre-Columbian peruvian mummy, 91:2091-2094. PNAS

Rothschild, BM, Martin LD, Lev G, Bercovier H, Bar-Gal GK, Greenblat C, Donoghue H, Spigelman M, Britain D.2001. *M. tuberculosis* Complex DNA from na extinct bison dated 17,000 years before the present. Clin. Infect. Dis. 33:305-311.

Zink, AR, Sola, C, Reischl, U, Grabner, W, Rastogi, N, Wolf, H, Nerlich AG. 2003. Characterization of *M. tuberculosis* Complex DNAs from Egiptian Mummies by Spoligotyping. J. Clin. Microbiol. 41(1):359-367.

Sreevatsan, S, Pan, X, Stockbauer, KE, Connell, ND, Kreiswirth, BN, Whittam, TS, and Muser, JM. Restricted structural gene polymorphism in the *M. tuberculosis* complex indicates evolutionarily recent global dissemination.1997. Proc. Natl. Acad. Sci. USA. 94:9869-9874.

- Mostowy, S, Behrm, MA. 2005. The origin and evolution of *M. tuberculosis*. Clin. Chest. Med. 26:207-216.
- Konomi, N, Lebwahl, E, Mowbray, K, et al. 2002. Detection of mycobactireal DNA in Andean mummies. J. Clin. Microbiol. 40:4738-4740.
- Blower, SM, Small, PM, Hopewell, PC. 1996. Control strategies for tuberculosis epidemics: new models for world problems. Science 273: 497-500.
- Parra, FC, Amado, RC, Lambertucci, JR, Rocha, J, Antunes, CM, Pena, SDJ. 2003. Color and genomic ancestry in Brazilians. Proc. Natl. Acad. Sci. USA. (100) 1:177-182.
- Alves-Silva, J, Santos, MS, Guimarães, PEM, Ferreira, ACS, Bandelt, H-J, Pena, SDJ, Prado, VF. 2000. The ancestry of Brazilian mtDNA lineages. Am. J. Hum. Genet. 67: 444-461.
- Kritski, AL, Ruffino-Netto, A, 2000, Health sector reform in Brazil: impact on tuberculosis control. Int. J. Tuberc. Lung. Dis. 4(7):1-5.
- van Soolingen, D., Hermans, PWM, de Haas, PEW, Sool, DR, van Embden, JDA. 1991. The occurrence and stability of insertion sequences in *M. tuberculosis* complex strains: evaluation of a insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J. Clin. Microbiol. 29: 2578-2586.
- Kamerbeek, J, Schouls, L, kolk, A, van Agterveld, M, van Soolingen, D., kuijper, S, Bunschoten, A, Molhuizen, H, Shaw, R, Goyal, M, van Embden, JDA. 1997, Simultaneous detection and strain differentiation of *M. tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. 35: 907-914.
- Filliol, I, Driscoll, JR, van Soolingen, D, Kreioswirth, BN, Kremer, K, Valétudie, G, Anh, DD et al. 2002. Global distribution of *M. tuberculosis* spoligotypes. Emerg. Infect. Dis. 8:1341-1343.
- Gagneux S, DeRiemer K, Van T, Kato-Maeda M, Jong, BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, Hilty M, Hopewell PC, Small PM. 2006. Variable host-pathogen compatibility in *M. tuberculosis*. PNAS. 103(8): 2869-2873.
- Monot M, Honore N, Garnier T, Araoz R, Coppee JA, Lacroix C, Sow S, Spencer JS, Truman RW, Williams DL, Gelber R, Virmon M, Flageul B, Sho SN, Ji B, Paniz-Moldolfi A, Convit J, Yaung S, Fine PS, Rasolofo V, Brennan PJ, Cole S. 2005. On the origin of leprosy. Science. 308:1040-1042.
- Kwiatkowski D. 2000. Genetic susceptibility to malaria getting complex. Curr. Opin Genet. Dev. 10:320-324.



## Tables

State	Isolates	Habitants	Tax. Incid.
Amazonas	45	2,812.557	75-97
Pará	186	6,192.307	40-50
Ceará	71	7,430.661	40-50
Pernambuco	78	7,918.344	50-75
Sergipe	49	1,784.475	25-40
Minas Gerais	55	17,891.494	25-40
Rio de Janeiro	1034	14,391.282	75-97
São Paulo	177	37,032.403	50-75
Paraná	6	9,563.458	25-40
Rio Grande do Sul	294	10,187.798	50-75
Goiás	7	5,003.228	19-25

Table; Showing the number of isolates obtained from and incidence tax and habitants.

The habitant's data is from IBGE (Brazilian Institute of Geography and Statistic), 2000. The incidence tax is from SINAN (Sistema de Informação de Agravo de Notificação), 2000.



Figure 1. Political map of Brazil

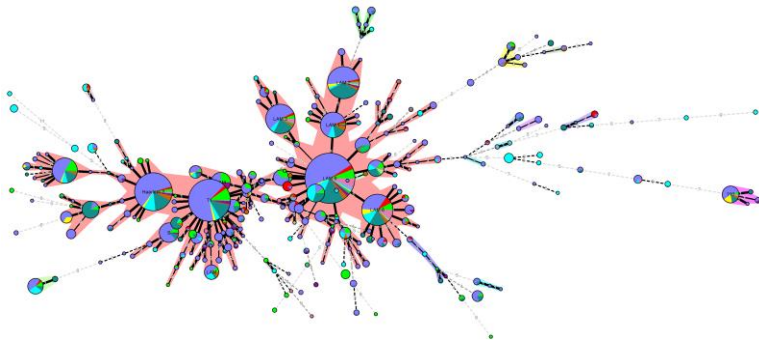


Figure 2. Minimal Spanning Tree constructed with the spoligotyping from 2,005 isolates from Brazil.

The Minimum spanning tree used to define the phylogeny and the distribution of spoligotypes in Brazil, showing the effect of the high prevalence of LAM family (~50%), its presence in the Brazilian states analyzed and its connection with the T1 family. The major of the out-groups, represent the new types (18%) that offspring from different families and types and are responsible to generate the diversity of spoligotypes. The sphere size represent the number of the strains, the colors represent the Brazilian states, the number near the line represent the difference of spacers between the groups.

## Capítulo XI – Discussão e Conclusão

### IV.1. Discussão

A relação entre as espécies do Complexo *M. tuberculosis* e a espécie humana está bem documentada através da presença de DNA identificado em um bisão datado de 17.000 anos atrás (Rothschild *et al.*, 2001), em múmias egípcias do período de 5000 a 1650 AC (Zink *et al.*, 2003), em uma múmia peruana pré-Colombiana datada de 1000 anos (Wilmar *et al.*, 1994); em restos humanos medievais e em corpos naturalmente mumificados

de uma família Européia do Século 18 (Fletcher *et al.*, 2003). Apesar destes registros destacarem esta longa convivência, a falta de uma política internacional de combate a TB, principalmente para os países pobres e em desenvolvimento, faz com que esta doença continue ameaçando esta e as futuras gerações.

Para o controle global da TB, destacam-se três importantes abordagens: i) detecção dos casos (multibacilar, paucibacilar e extrapulmonar), ii) teste de suscetibilidade a drogas (tratamento e vigilância de TB-MDR) e iii) a identificação da TB latente. Destes três pilares do combate a TB, a detecção de resistência apresenta uma variedade de alternativas, tanto em relação ao tempo necessário para chegar ao resultado, quanto à tecnologia utilizada. A detecção rápida e precisa de cepas resistentes permite a administração da droga anti-TB no tempo certo, o que contribui para a redução do custo do tratamento e principalmente reduz o sofrimento, melhorando a qualidade de vida do paciente durante o longo tratamento.

Muitos métodos eficientes e capazes de diagnosticar com rapidez cepas sensíveis e resistentes a rifampicina estão disponíveis. No entanto, o seu alto custo assim como o custo dos equipamentos, sua manutenção, o suprimento de insumos necessários para a manutenção do sistema e a dependência, das empresas que os fornecem, torna a sua utilização urgente, principalmente em países com altas taxas de TB, que em sua maioria são países pobres e desinteressados do destino de seu povo.

A técnica de SSCP parece ser promissora para a detecção de mutações no gene *rpoB* relacionadas a resistência a rifampicina, cuja relação custo-efetividade permita a sua utilização nas mais diferentes realidades econômicas. Em nosso trabalho (artigo 1), utilizamos uma versão não radioativa da técnica de SSCP para determinar a suscetibilidade de cepas de *M. tuberculosis* a rifampicina. Em um estudo multi-cêntrico, analisamos 16 cepas brasileiras de *M. tuberculosis* resistentes a rifampicina. O SSCP mostrou que das cepas brasileiras identificadas pelo método das proporções e pelo seqüenciamento (considerada a

técnica adequada para detectar cepas de *M. tuberculosis* resistentes a rifampicina), somente nove (43%) apresentaram o genótipo resistente pelo SSCP, uma frequência muito abaixo daquela esperada quando comparada com a análise das 32 cepas (Seul, Coreia) analisadas pela PCR radioativa. A baixa sensibilidade encontrada em nosso estudo foi semelhante à observada por Bobadilla-del-Valle *et al.* (2001) que, das 35 cepas resistentes a rifampicina, analisadas por SSCP não radioativo, somente 24 (68,6%) apresentaram o perfil de resistência. O fator em comum às duas abordagens foi a eletroforese em gel de poliacrilamida; no entanto o melhor resultado obtido por Bobadilla-del-Vale *et al.* (2001) (68,6%) pode ser atribuído à diferença na concentração de acrilamida, enquanto usamos 8% Bobadilla del Valle *et al.* (2001), utilizou 12%.

A baixa capacidade de detecção de cepas resistentes a rifampicina em cepas brasileiras, pode ser atribuída as matrizes utilizadas na eletroforese, enquanto os grupos da “Yonsei University College of Medicine” (Coreia do Sul), e da “University of Stellenbosch”, (África do Sul), utilizando MDE detectaram 76 e 80% respectivamente do total das cepas resistentes, nós utilizamos a poliacrilamida e detectamos somente 43% de resistência em um total de 16 cepas resistentes. A análise multicêntrica não definiu um protocolo rígido para a técnica de PCR-SSCP, no entanto esta mesma flexibilidade em relação às condições metodológicas possibilitou uma avaliação das diferentes realidades dos laboratórios participantes.

A melhor experiência da utilização da matriz especial MDE para detecção de mutações em *M. tuberculosis* relacionadas à resistência a rifampicina foi descrita por Lee *et al.* (2003), no artigo 1, Bobadilla del Valle *et al.* (2001) e Kim *et al.* (1997), no entanto um excelente resultado pode ser visto utilizando a poliacrilamida (SSCP) para detecção de mutações em *M. tuberculosis* relacionadas à resistência a isoniazida. Neste caso SSCP foi capaz de detectar mudança na mobilidade eletroforética diferente do controle sensível nas 60 cepas com mutação no códon 315 de *katG* e diferentes mobilidades foram observadas para seis mutações (AGC → ACC, AAC, CGC, ATC, GGC e AGG), o mesmo resultado foi obtido para a detecção de mutações no gene *kasA* em 157 cepas (Cardoso *et al.*, 2004).

A adequação do SSCP para obtermos resultados como os mostrados pela detecção radioativa, utilizando a poliacrilamida e a coloração pela prata, deve ser um objetivo do diagnóstico de resistência a rifampicina e a outras drogas. Estas adequações já se mostraram

úteis para uso em amostras clínicas, tais como: escarro (Kim *et al.*, 2001), biopsia ou esfregaços corados por Ziehl-Neelsen. Uma importante característica do SSCP é a sua capacidade de detectar mutações em um gene inteiro, através de sua mobilidade eletroforética. No entanto, algumas limitações metodológicas precisam ser consideradas, como por exemplo: mutações silenciosas em cepas sensíveis a drogas que podem gerar uma migração semelhante ao perfil de uma cepa resistente (Kim *et al.*, 1997). Além disto, (como pode ser observado na figura 1 do artigo 1), a diversidade de perfis obtidos requer a utilização de controles sensíveis e resistentes que permitam uma boa análise dos resultados do gel, assim como condições de eletroforese rigorosamente padronizadas, de modo que se possa estabelecer com clareza a relação entre perfil resistente ou sensível com a migração observada na eletroforese.

Em comum, todos os trabalhos de detecção de mutações em *rpoB* tem como alvo a seqüência de 157 pb do gene *rpoB*, onde mais de 95% das mutações associadas a resistência a rifampicina estão localizadas. Embora seja a mesma seqüência alvo, pequenas modificações, na seqüência dos “primers”, da PCR podem ser observadas.

A busca de técnicas de detecção de mutações relacionadas à resistência em *M. tuberculosis* tem gerado um grande número de métodos aplicáveis a diferentes situações. No entanto, estas metodologias também são desenvolvidas de modo a serem aplicadas nas mais diferentes realidades econômicas, como é o caso dos países que concentram o maior número de casos de TB e resistência. A hibridação reversa como um destes métodos se baseia na imobilização em linha de sondas específicas para determinado alvo. As suas principais vantagens são: i) está baseada na reação de PCR, portanto pode ser aplicada a qualquer amostra de onde se possa extrair DNA em pequenas concentrações (da ordem de nanogramas), o que reduz o tempo da obtenção do resultado, ii) pode ser aplicada simultaneamente a um grande número de amostras, iii) pode ser adequada a diferentes versões de detecção: radioativa, quimioluminescente e colorimétrica, iv) as sondas nas membranas podem ser adequadas às diferentes especificidades biogeográficas, v) o sistema pode ser automatizado, aumentando sua reprodutibilidade,

Avaliamos uma metodologia baseada na hibridação reversa em um segundo estudo multicêntrico. Nesta versão do estudo (artigo 2), foram imobilizadas sondas para detecção de alelos selvagens e/ou contendo mutações nos genes *rpoB*, *inhA*, *ahpC*, *rpsL*, *rrs* e *embB*. Diferente da maioria dos genes associados à

resistência a drogas, aproximadamente 95% das cepas resistentes a rifampicina possuem mutações concentradas em uma região de 81pb do

gene *rpoB*, nas demais cepas, a resistência pode estar associada a mutações em outras regiões do genoma ou a outros mecanismos. Durante a avaliação da RLBH, mutações no gene *rpoB* também foram observadas em cepas brasileiros, com 58% das mutações no códon 531-TTG. Esta alta frequência foi também observada em cepas de pacientes da Coreia do Sul, Nova Delhi (Índia) e Rússia, enquanto que em cepas de Mumbai (Índia) e da Malásia, apresentaram uma baixa frequência. A natureza e frequência de mutações no gene *rpoB* em cepas resistentes a rifampicina do Brasil foi avaliada através de sequenciamento por Valim *et al.* (2000) e a predominância de mutações no códon 531 varia entre 52 e 55% para os estados do RJ, RS e SP.

Morcillo *et al.* (2002) descreveram a hibridação reversa para detecção específica de mutações relacionadas à resistência em *rpoB*, denominada de “rifoligotyping” contendo cinco sondas correspondentes ao tipo selvagem e 6 sondas indicativas dos alelos mutantes. Trata-se da adaptação do teste comercial, Inno-LiPA (Innogenetics, Bélgica) que também tem como alvo o gene *rpoB*. Antes de avaliarmos uma versão “in house” do “rifoligotyping”, avaliamos a sensibilidade, especificidade e outras características do INNO-LiPA em cepas brasileiros de *M. tuberculosis* (de Oliveira *et al.*, 2003; artigo 3).

No artigo 4 (Senna *et al.*, 2006) apresentamos os resultados, obtidos com a avaliação de uma versão modificada do “rifoligotyping” (adição do códon mutante 522) de cepas brasileiras com suscetibilidade a rifampicina e com a sequência do gene *rpoB* conhecida. Embora os resultados sejam promissores, algumas modificações para tornar a técnica mais reprodutível precisam ser consideradas. O fato de termos trabalhado tanto com o teste comercial, quanto com a versão “in house” nos permite avaliar com maior segurança a diferença entre as duas abordagens. No estudo de Senna *et al.* (2006), artigo 4 foram testadas cepas provenientes de quatro estados brasileiros, Rio Grande do Sul (Região Sul) e 3 Estados da Região Sudeste (Espírito Santo, Rio de Janeiro e São Paulo). Os resultados foram comparáveis ao estudo anterior, onde as mutações foram determinadas por sequenciamento. Em relação à frequência de mutações, o trabalho de Oliveira *et al.* (2003), já indicava esta tendência, utilizando cepas do Rio de Janeiro, Rio Grande do Sul, Goiás, Mato Grosso e Mato Grosso do Sul. Os resultados mostraram que 49% das cepas resistentes apresentavam mutação no códon 531, com a substituição da Serina pela Leucina.

Alguns estudos tem mostrado uma relação entre uma mutação no gene *rpoB* com um

genótipo de *M. tuberculosis*, no entanto a maioria dos registros mostram que a natureza e a frequência das mutações em *rpoB* independem de seus perfis genotípicos estabelecidos por RFLP e “spoligotyping” (Van der Zanden *et al.*, 2003).

Em nosso trabalho (Mokrousov *et al.*, 2004), artigo 2, observamos que a frequência do códon 531 sofreu uma significativa diferença, quando comparamos duas regiões da Índia (Nova Delhi - 48% e Mumbai - 7,7%). Suresh *et al.*, (2006), analisando cepas provenientes de diferentes regiões da Índia com uma RLBH “in house” obteve as seguintes frequências para mutações no códon 531 (Delhi - 72%, Trivandrum - 50%, Chennai - 41,4%, Pune - 69,5% e Lucknow - 64%), estes resultados estão em acordo com a afirmação de que não existe

uma significativa diferença na distribuição de mutações entre cepas resistentes a rifampicina entre diferentes países, o que poderíamos estender para regiões de um mesmo país, como mostra o exemplo de Suresh *et al.* (2006) da Turquia, nas regiões de Anatólia, Marmara e região do Mediterrâneo, com uma distribuição média de 58% para mutações no códon 531 (Yang *et al.*, 2005). Em todos estes casos a maior frequência de mutações no códon 531 ocorreu na transição TCG (Serina) - TTG (Leucina) e em segundo lugar a transição de TCG (Serina) - TGG (Tryptofano) em uma frequência significativamente menor.

A relativa uniformidade na distribuição das mutações do códon 531, não significa dizer que as demais mutações neste gene estão distribuídas em todas as regiões ou países. Este fato justificou a modificação da RLBH (Senna *et al.*, 2006) com a adição do alelo mutante 522-GTT. Este alelo está associado à resistência a rifampicina, além de ter sido encontrado em cepas brasileiras foi também observada na Ásia, Austrália, Rússia e nas Américas, mas ainda não foi descrito na Europa.

A resistência em *M. tuberculosis* as diferentes drogas utilizadas no tratamento da TB, não esta somente relacionada ao acúmulo de determinadas mutações em um pequeno número de genes, mas também a outros mecanismos muito bem descritos para uma variedade de bactérias. Estes mecanismos incluem a permeabilidade da membrana e a ação de mecanismos de transporte ativo que previnem o acesso das drogas ao alvo intracelular, portanto constituem mecanismos gerais de resistência a drogas capazes de atuarem sobre uma variedade de drogas estruturalmente não relacionadas e outros compostos tóxicos (Silva *et al.*, 2001; Nikaido, 1998). Estes mecanismos gerais poderiam completar a resistência de cepas de *M. tuberculosis* que não apresentam as mutações esperadas nos genes envolvidos na resistência a isoniazida, rifampicina, pirazinamida etc.

Na resistência a isoniazida, 42-58% (Musser *et al.*, 1995) dos mutantes apresentam mutação no gene *katG*, 21- 34% no gene *inhA*, 10 - 15% em *ahpC* e para o gene *kasA* esta relação ainda não esta estabelecida. Mutações no gene *pncA*, relacionada a resistência a pirazinamida, estão presentes em 72 a 97% dos mutantes. A resistência a etambutol pode ser observada em 47 a 65% dos mutantes e para estreptomicina, 52-59% das mutações estão no gene *rpsL* e 8 a 21% no gene *rrs* (16S RNA) (Rossetti *et al.*, 2002).

Ao compararmos os nossos dois estudos (Mokrousov *et al.*, 2004, artigo 2 e Senna *et al.*, 2006, artigo 4) utilizando RLBH para detecção de mutações relacionadas a resistência a rifampicina em *M tuberculosis* com o de Suresh *et al.*, 2006 que utilizam também uma versão “in house” de RLBH e com o trabalho de Oliveira *et al.*, (2003), artigo 3, observamos que 3 oligos mutados foram comuns às quatro abordagens, são eles: 531 TTG, 526 TAC e 526 GAC (Tabela IV.1).

Tabela IV.1. Oligos mutados presentes nos quatro estudos

Oligo	Mukrousov <sup>1</sup>	Senna <sup>2</sup>	Suresh <sup>3</sup>	de Oliveira <sup>4</sup>
531 TTG	X	X	X	X
531 TGG	X	X	X	
533CCG	X	X		
529 CAA		X		
526 TAC	X	X	X	X



526 GAC	X	X	X	X
526 CGC	X			
526 CTC	X		X	
526 TGC	X			
526 CCC	X			
526 AAC	X			
526 ACC	X			
526 CAG	X			
522 TTG		X		
513 TTC		X		
513 CCA	X	X		
511 CCG	X	X	X	
511 CGG	X	X		
514 Ittc	X			
516 GAC				X
516 GTG		X		
516 GTC	X			
516 TAC	X		X	
Δ516-517	X			
Δ518	X			

<sup>1</sup>Mokrousov *et al.* 2004, <sup>2</sup>Senna *et al.* 2006, <sup>3</sup>Suresh *et al.* 2006, <sup>4</sup>de Oliveira *et al.* 2003.

A escolha dos oligos mutantes a serem ligados a membrana pode ter determinado a frequência das cepas que perderam o sinal de hibridação com o oligo tipo selvagem e não ganharam no oligo mutante correspondente. No presente estudo, 32,2% (Senna *et al.*, 2006; artigo 4), 23% (de Oliveira *et al.*, 2003; artigo 3) e 16% (Mokrousov *et al.*, 2004; artigo 2) das cepas resistentes não apresentaram sinal com a sonda mutante. Esta diferença pode ser atribuída tanto a sondas presentes nas diferentes versões dos ensaios quanto à seleção de cepas avaliadas. A diferença observada entre Suresh *et al.* (2006) (9,58%) e Senna *et al.* (2006), artigo 4, que utilizam sistemas de hibridação muito semelhantes (inclusive com o mesmo número de oligos, sendo que quatro deles são comuns a ambos os sistemas) pode ser explicada por uma escolha de maior critério por parte de Suresh *et al.* (2006) com base na frequência das mutações que ocorrem nas diferentes regiões da Índia estudadas. Em Mokrousov *et al.*, (2004), artigo 2, este critério não foi rígido, mas incluía um significativo número de oligos mutados, capazes de se adequar as diferentes frequências regionais.

Em *rpoB*, a frequência considerável (32,2%) de cepas resistentes a rifampicina serem reconhecidos somente pela perda da hibridação com o oligo tipo selvagem sem a conseqüente hibridação com o oligo mutante não inviabilizou a determinação da resistência como pode ser comprovado pelo sequenciamento da região correspondente (Senna *et al.*, 2006; artigo 4). No entanto como observado, a detecção precisa do oligo mutado também pode ser uma importante informação para o início de um tratamento adequado, pois alguns trabalhos mostram uma estreita relação entre a concentração inibitória mínima e o códon mutado, como por exemplo,

mutações pontuais dentro dos códons: Ser-531, His-526, Gly-513 ou Asp-516 estão relacionadas a altos níveis de resistência em *E. coli* e *M. tuberculosis*. (Spindola *et al.*, 2001). Mais estudos associando natureza de mutações e efeitos no grau de resistência são necessários.

Algumas mutações mantêm-se distribuídas em uma determinada população ou região, isto pode ocorrer entre as diferentes regiões de um mesmo país ou entre diferentes países ou continentes. No Brasil, Valim *et al.*, (2000) observaram algumas mutações relacionadas à resistência a rifampicina somente no Rio de Janeiro, dos estudos realizados no Brasil foram encontrados um amplo número de mutações relacionadas a resistência: 511-CCG, 511-CGG, 512-TTT, 512-CAA, 513-CAC, 513-CCA, 514-GTC, 515-ATA, **516-GTC**, 516-CCG, 516-TAC, **522-TTG**, 524-TGG, 525-CCC, **526- GAC**, **526-TAC**, 526-CAG, 526-AAC, 526-CGC, 526-TGC, 526-CTC, 526-CTG, 530-CTA, **531-TTG**, **531-TGG**, 531-TCA, 532-TAC, 533-CCT (os códons em negrito são aqueles presentes na membrana elaborada em Senna *et al.*, 2006); destes códons, somente o 529-CAA não hibridou com nenhuma das amostras testadas. Dependendo da frequência, estas mutações devem ser consideradas na elaboração de uma nova membrana para detecção de resistência a rifampicina.

Além da sensibilidade dos métodos, o custo é um dos parâmetros importantes para utilização dos mesmos em países pobres e em desenvolvimento Segundo Suresh *et al.*, (2006), a metodologia de hibridação reversa “in house” parece ser uma metodologia que atende a estes requisitos porque o custo de um teste é 9 vezes inferior ao mesmo teste realizado por um teste comercialmente disponível, como é o caso do INNO-LiPa (Suresh *et al.*, 2006)

O desenvolvimento da hibridação reversa para detecção de “rifoligotypes” de *M. bovis* (Parreiras *et al.*, 2004, artigo 5) mostra que o método pode ser aplicado para detecção de resistência em outras espécies pertencentes ao Complexo *M. tuberculosis*; o mesmo mecanismo de resistência a rifampicina parece estar presente nas demais micobactérias do Complexo e inclusive em *M. leprae*, agente causador da hanseníase (Musser, 1995). O tratamento da hanseníase e da TB compartilham a rifampicina e sistemas preliminares baseadas no princípio de “rifoligotyping” para *M. leprae* já foram descritos.

A ocorrência de animais infectados por *M. bovis* resistente a drogas não tem sido relatados com frequência, embora tenhamos alguns casos de *M. bovis* MDR (Zanini *et al.*, 2005). Por outro lado, a resistência natural de *M. bovis* a pirazinamida é o principal critério utilizado para diferenciação entre *M. bovis* e *M. tuberculosis*. O nosso trabalho mostra que as cepas de *M. bovis*, analisadas por “rifoligotyping” apresentam um padrão de hibridação tipo selvagem. A ausência de cepas de *M. bovis* provenientes de animais com resistência a

rifampicina também foi observada através do Método das Proporções (Zanini *et al.*, 2000), no entanto, surtos de *M. bovis* MDR foram relatados em pacientes HIV-positivos (Blázquez *et al.*, 1997, Zanini *et al.*, 2005).

A resistência de *M. bovis* a rifampicina relacionada à mutação no gene *rpoB* está bem documentada em cepas provenientes de gado bovino da região da Sardenha, na Itália, onde em 10 cepas analisadas por seqüenciamento da região de 157pb do gene *rpoB*, 6 apresentavam mutações no códon 521, com mudança de Leucina para Prolina (4 cepas de CTG - CCG e dois cepas de CTG - CTT) duas cepas com mutações no códon

526 (CAC - GAC) e finalmente dois cepas com mutações no códon 513 (CAA - AAA) (Sechi *et al.*, 2001). As mutações relatadas em *M. bovis*, também estão relacionadas à resistência a rifampicina em *M. tuberculosis*.

A efetiva utilização da hibridação reversa “in house” na detecção de resistência em *M. bovis*, como descrita por Senna *et al.*, (2006), artigo 4, está diretamente relacionada à importância da disseminação de cepas de *M. bovis* resistentes a drogas provenientes de gado bovino em humanos.

Esclarecer a origem do *M. tuberculosis* poderá nos levar a uma melhor compreensão da patogenicidade e da epidemiologia global da tuberculose, além de nos auxiliar a perceber a tendência de distribuição deste microorganismo (Brosch *et al.*, 2002).

Os organismos do Complexo *M. tuberculosis* contêm um significativo número de elementos polimórficos repetitivos usados como marcadores em ferramentas de epidemiologia molecular e para estudos de filogenia. Os elementos repetitivos mais estudados são os “Direct Repeats” (DR) da região DR (“spoligotyping”), a seqüência de inserção IS6110 (IS6110-RFLP) e os MIRUs (MIRU-VNTR “typing”).

Através da construção de um banco de dados de “spoligotyping”, desde a primeira versão (SpolD) até a sua quarta e mais recente versão divulgada em 2006 (SpolDB4; Brudey *et al.*, 2006, artigo 6), adquiriu-se gradualmente um conhecimento mais profundo sobre a distribuição de genótipos de *M. tuberculosis* no mundo. Um exemplo da nomenclatura e composição do SpolDB4 encontra-se no anexo 1.

O “spoligotyping” pode ser aplicado em grande escala e gera informação: sobre a variabilidade genética da população bacteriana, sobre a ocorrência de surtos ou situações de risco para transmissão e sobre relações filogenéticas entre os parasitos, em função disto avaliamos a composição da região DR de isolados de *M. tuberculosis* provenientes de diferentes origens. Até o momento avaliamos em torno de 3.000 cepas de *M. tuberculosis* e o manuscrito em preparação descreve a análise de isolados de 2.005 pacientes de diferentes estados brasileiros, no entanto, na publicação apresentada SpolDB4 (Brudey *et al.*, 2006, artigo 6), relata a entrada de somente 842 cepas do Brasil, correspondendo a pouco mais de ¼ de todos os “spoligotypes” da América Latina. Isto significa que a contribuição brasileira para este banco de dados ainda está em fase inicial.

O principal objetivo desta parte do trabalho é visualizar a variabilidade genética de *M. Tuberculosis* entre os estados brasileiros e construir um cenário da distribuição destas. Verificamos que 80% dos isolados estão distribuídos entre as famílias LAM, Haarlem e T. O grupo dos genótipos mais abundantes no Brasil são aquelas que caracterizam a família LAM, observado em 48% das cepas. Esta família está distribuída principalmente entre os países da América Latina e da região do Mediterrâneo (Filliol *et al.*, 2002 ).

No Brasil encontramos dentro da família LAM as classes 1 (tipo 20), dois (tipo 17), 3 (tipo 33), 4 (tipo), 5 (tipo 93), 6 (tipo 64) e 9 (tipo 42). As classes LAM7 (tipo 41), 8 (tipo 290) e 10 (tipo 61) no continente americano só estão presentes na Guiana Francesa (LAM7, n=1 e LAM10, n=1) e em Guadeloupe (LAM8, n=1 e

LAM10, n=1), ambos os territórios que pertencem a Europa. Este continente concentra 61,4% dos isolados da classe LAM7 e 26,6% da classe LAM10; aproximadamente 50% de cepas LAM10 são provenientes da França. A distribuição da classe LAM8 é diferente das últimas duas porque não se concentra no continente europeu e possui baixa frequência (n=17) sem distribuição específica. A classe LAM7 apresenta uma alta frequência na Turquia, um país com uma importante relação histórica com a Europa e a Ásia; das 245 cepas provenientes deste país, 25% pertencem a esta classe e no SpolDB4 (LAM7 n=127), esta cepa passou a ser denominada de classe LAM7-TUR (Zozio *et al.*, 2005). A classe LAM10, com 331 cepas no SpolDB4, foi denominada de LAM10-CAM, em função da alta frequência (31 %) de casos de TB relacionados a esta classe na República dos Camarões. Parte deste país esteve sob domínio francês do início do século 20 até 1960; esta ligação histórica poderia explicar a presença desta classe em ambos os países.

A família LAM está amplamente distribuída na Europa, em boa parte do continente americano e na África, aonde pode ser vista principalmente no sul do continente. Na África do Sul, 21,4% dos casos de TB estão relacionados à classe LAM3 (ST11). Na América Latina, a maior frequência da família LAM está na Venezuela com 65% dos “spoligotypes” pertencendo a esta família enquanto a menor (17,4%) se encontra na Guiana Francesa. Sugere-se que a alta frequência da família LAM na América Latina esteja relacionada à colonização Espanhola e Portuguesa. A busca da família LAM no SpolDB4 nas ex-colônias portuguesas na África: Angola, Moçambique, Guiné-Bissau, Cabo Verde, São Tomé e Príncipe e Timor Leste, mostrou que somente os três primeiros países depositaram “spoligotypes” no SpolDB, sendo Angola com 4 cepas, Moçambique com 28 cepas e Guiné-Bissau com 217 cepas. A frequência de LAM nestes países foi de 28,6% em Moçambique e 16,1% em Guiné-Bissau. No SpolDB4 estão depositadas 336 cepas oriundas de Portugal e a metade pertencem à família LAM, frequência similar à observada no Brasil. As diferenças podem ser observadas entre as frequências das classes desta família.

Na tentativa de se estabelecer a origem da família LAM na América Latina, não podemos deixar de considerar a origem autóctone da família LAM em nosso continente, “co-evoluindo” com os primeiros habitantes da América. A identificação de DNA micobacteriano em restos de um bisão de 17.000 anos e em uma múmia de 1.000 anos mostra que este microorganismo já circulava entre as populações ameríndias muito antes da chegada dos primeiros colonizadores. Devido a essas observações, podemos propor duas hipóteses: (i) a família LAM já existia na América pré-colombiana e se espalhou pelo mundo após o contato com os colonizadores que continuaram navegando para outras terras e (ii) a família LAM foi introduzida no continente americano pelos primeiros colonizadores e aos poucos foi substituindo as cepas de *M. tuberculosis* que aqui circulavam como está ocorrendo com o *M. africanum*, substituído pelo *M. tuberculosis* em alguns países da África.

Algumas particularidades podem ser relacionadas à distribuição de “spoligotypes” no Brasil. Exemplo é a detecção de 21 isolados da família EAI distribuídos entre três dos onze estados brasileiros analisados, sendo que a grande maioria (n=16) foi observada no Estado do Pará. Na América do Sul, a sua presença só é relatada em um pequeno número na Guiana Francesa e em Cuba. A Guiana Francesa é território da França e imigrantes brasileiros em busca de melhores condições de vida atravessam a fronteira com relativa facilidade, estabelecendo assim uma forte ligação epidemiológica entre aquele território e o território

brasileiro. Devido à raridade de isolados EAI em outros lugares do Brasil, suspeitamos que estas cepas sejam originárias da Guiana. Um fator que reforça esta hipótese é a alta frequência desta família na Europa, inclusive na França, que mantém uma comunicação regular com a Guiana. Também é importante notar a baixa frequência desta família em Portugal e Espanha, que mantém uma relação histórica com o Brasil, portanto não podemos relacionar EAI com estes países.

A concentração de isolados da família EAI no Pará não significa que esta família não possa ser encontrada em outros estados, pois não obtivemos isolados de pacientes de todos os estados brasileiros. Entretanto, 51% das cepas são provenientes do Estado do Rio de Janeiro onde a presença da família EAI foi rara.

O fluxo migratório moderno pode ser o responsável pela presença de outros genótipos no Brasil, como por exemplo, cepas pertencentes à família “Manu”, a classe LAM11-ZWE (tipo 59), responsável por 57,8% dos casos de TB em Zimbábue, a classe Madrid2 (Viedma *et al.*, 2005), a classe Haarlem 2.e a família S. Estas cepas foram encontradas no Brasil em baixa frequência, e até o momento, somente em grandes capitais. O pequeno número destes casos, ocorrendo especialmente em uma região, indica que pode tratar-se de casos importados, pois geralmente são as grandes cidades, que, por razões variadas, recebem frequentes ondas migratórias.

A família Haarlem, que recebeu este nome porque estes genótipos foram observados inicialmente em isolados de pacientes provenientes da cidade holandesa de mesmo nome. A presença desta família nas Américas pode estar relacionada à colonização deste continente pelos europeus (Kremer *et al.*, 1999). O tipo 51 (T1) em um total de 11 amostras encontradas no Rio de Janeiro, São Paulo e Rio Grande do Sul, é um tipo prevalente na Sicília e Sardenha, ambas com forte herança espanhola (Soini *et al.*, 2000). A família X, com alta prevalência nos Estados Unidos da América e no Reino Unido (Sebban *et al.*, 2002), está presente no Brasil em quase todos os estados, mas com baixa frequência. A família T, com terceira maior frequência no Brasil, tem alta prevalência em alguns países africanos, o que pode sugerir uma origem africana para este alelo (Duchêne *et al.*, 2004).

Sugerimos que a presença dos diferentes genótipos de *M. Tuberculosis* nos estados brasileiros é consequência de fatores, tais como (i) o caráter multirracial da formação do povo brasileiro, (ii) o fluxo migratório moderno e (iii) o processo de globalização evidenciado na última década do Século 20.

Não identificamos no Brasil família tipo ou classe de “spoligotype” que possa estar diretamente associada com determinado estado ou região. Embora a família LAM se apresente em uma alta frequência no país, não se trata de um aspecto exclusivo do Brasil devido a sua ampla distribuição na América Latina, na região do mediterrâneo e também em grande parte da Europa. Entretanto, a existência de grandes diferenças entre a distribuição das classes da família LAM no mundo podem estar relacionadas com a associação destas cepas com a população hospedeira local (Hirsh *et al.*, 2004). No Brasil, a maior frequência foi observada para a classe LAM9, enquanto em Portugal a maior frequência é observada na classe LAM1.

Os trezentos e sessenta e quatro isolados incluídas manuscrito em preparação (anexo 1) foram classificadas de acordo com a divisão em Grupo Genético Principal (PGG, Sreevatsan *et al.*, 1997) e 99,3% pertencem aos grupos genéticos 2 e 3. Esta distribuição de isolados brasileiros assemelha-se a observado por Duchêne *et al.*, (2004), onde na região do Caribe (Cuba, Haiti, Guadeloupe e Martinica) todos os 321 isolados

pertencem ao PGG 2 e 3. Estes resultados reforçam a hipótese da introdução recente da TB no Brasil (após a sua descoberta em 1500)

A presença da TB na América pré-Colombiana está bem documentada, porém, ao que parece, “a TB não representava um grave problema para os ameríndios já que viviam em comunidades com baixa densidade populacional, um fator importante que manteve a TB sobre controle naqueles tempos”. Esta afirmação pode ser considerada razoável, já que os alelos no Brasil pertencem todos ao grupo genético recente. Por outro lado, estudos sugerem que uma comunidade com uma população em torno de 180-440 pessoas seria suficiente para estabelecer uma situação endêmica e transmissão vertical (Duchêne *et al*, 2004) então se acredita que a importância de *M. tuberculosis* na América pré-Colombiana ainda esta por ser desvendada.

Outra espécie de Complexo *M. tuberculosis* importante para TB humana é *M. africanum*, subdividida por sua origem geográfica e propriedades bioquímicas em dois grupos principais: *M. africanum* subtipo I, originário do Oeste da África e com características semelhantes ao *M. bovis* enquanto que *M. africanum* subtipo II, tem sua origem no leste da África e apresenta características semelhantes ao *M. tuberculosis*. (Mostowy *et al.*, 2004).

Aproximadamente 95% das cepas de *M. africanum* no SpolDB4 são caracterizadas pela ausência dos espaçadores 8, 9 e 39 e entre os “spoligotypes” dos isolados brasileiros, não observamos perfis com estas características, concluindo que *M. africanum* não é responsável por causar TB no Brasil, a semelhança de cepas da família CAS que podem ser encontradas na Ásia, sudeste africano, Europa e Estados Unidos (Brudey *et al.*, 2006, artigo 6).

Perfis de 339 isolados de *M. africanum* estão presentes no SpolDB4 e distribuídos do seguinte modo: 203 (59,88%) na África, 86 (25,36%) na Europa, 44 (12,98%) nos E.U.A, 1 (0,29%) na Ásia, 3 (0,88%) na Guiana Francesa, 1 (0,29%) na América do Sul em uma cepa de origem desconhecida (Brudey *et al.*, 2006, artigo 6). Estes números mostram que esta espécie é geralmente associada a casos de TB na África Sub-Saariana, embora muitos casos possam ser observados na Europa em pacientes que não possuem relação epidemiológica com a África (Viana-Niero *et al.*, 2001). Somente uma foi encontrada na América do Sul e 3 na América Central em territórios que pertencem a Comunidade Européia e que, portanto mantém uma relação próxima com o continente europeu. A ausência do *M. africanum* no Brasil surpreende pelo fato de ser uma espécie de alta prevalência no continente africano com o qual o Brasil mantém uma forte relação histórica, estabelecida principalmente pelo tráfico de africanos durante os regimes colonial e imperial no Brasil.

Outro importante grupo de cepas de *M. tuberculosis* amplamente distribuída no mundo e observada em baixa frequência no Brasil são as pertencentes ao genótipo Beijing. Uma característica deste genótipo é que a maioria das unidades DR foi perdida, provavelmente em um único evento mediado por recombinação pela IS6110. As três cepas da família Beijing encontradas no Brasil pertencem às amostras do Rio de Janeiro (uma isolada em 1997 e duas em 2003), e embora ainda não tenha sido relatada esta família no SpolDB4; recentemente outros estados como São Paulo e Minas Gerais também têm apresentado alguns casos do genótipo Beijing entre seus isolados. É interessante observar que a análise dos genótipos destes isolados por MIRU-VNTR, apresenta

uma grande variabilidade, sugerindo diferentes origens desta linhagem (dados não mostrados).

Certas famílias definidas através de “spoligotype”, inclusive LAM, Haarlem e T estão distribuídas em todos os continentes; no entanto, para a família LAM a maior frequência pode estar na América do Sul e no continente africano. As cepas da família Beijing e EAI são principalmente encontradas no Sudeste asiático, confirmando que cepas do complexo *M.*

*tuberculosis* apresentam uma tendência a uma distribuição geográfica específica (Sola *et al.*, 2001). Isto suporta uma nova tendência de pensamento e de linha de pesquisa, sugerindo que há certa associação entre composição genética do parasito e do hospedeiro. Procura-se identificar o componente genômico das diferentes famílias que as relacione com uma população específica (hospedeiro) e como exemplo pode ser citado que algumas cepas estão associadas com alguns grupos de imigrantes e periodicamente são re-introduzidas em determinada região acompanhando o fluxo migratório destas populações (Gagneux *et al.*, 2006).

Outros “spoligotypes” encontrados tanto no Brasil quanto no SpolDB4 são o tipo 18 (X2) com 13 amostras exclusivas do Rio de Janeiro e 9 amostras em 3 países (5 na Grã-Bretanha, dois na Holanda e dois nos Estados Unidos da América). O tipo 176 (LAM5) com 7 no Brasil, localizadas no Rio de Janeiro e também encontrada na República Centro Africana (1), Guiana Francesa (2) e Estados Unidos da América (12). O tipo 498 (T1) com 11 amostras no Brasil (PA 1 e RJ 10) e 5 no mundo (1 em cada país: Áustria, República Checa, Grã-Bretanha, Estados Unidos da América e Holanda) e como último exemplo cito o tipo 770 (LAM9) com 8 amostras no Brasil (RJ) e 4 amostras no mundo (Argentina, Espanha, França e Itália). A baixa frequência destas cepas no mundo e a sua frequência relativamente alta no Brasil sugerem uma associação a uma população humana específica.

A definição de genótipo implica o uso de determinados marcadores genéticos até agora, enfocamos em genótipos definidos pela composição da região DR, no entanto existem “famílias” ou genótipos predominantes em certas regiões definidos por outros marcadores, não necessariamente detectados por “spoligotyping”, como por exemplo, a cepa Quebec (Nguyen *et al.*, 2004).

A construção de uma versão estendida do SpolDB, SITVIT (Spoligotyping International Types MIRU-VNTR International Types), combinando “spoligotypes” com perfis MIRU-VNTR utilizando 12 ou 15 VNTRs oferecerá informações sobre a distribuição de genótipos em nível mundial que não foram observadas através do “spoligotyping”, porém, a análise combinada de “spoligotyping” com outros marcadores, inclusive com as deleções, poderá revelar com mais clareza associações entre as populações brasileiras e certos genótipos.

Está descrita a prevalência de cepas de *M. tuberculosis* com deleções, como por exemplo, a cepa Quebec com a deleção DS6<sup>Quebec</sup> (Nguyen *et al.*, 2004). No Brasil, também esta sendo evidenciada a presença de uma linhagem de cepas com uma deleção, denominada de RD<sup>Rio</sup>. Isolados com esta deleção compõem em torno de um terço da população bacteriana, causando TB no Rio de Janeiro e “curiosamente”, a presença da deleção esta associada com “spoligotypes” particulares: 56% da população RD<sup>Rio</sup> concentra-se nas classes LAM1, LAM2

e LAM9.

A deleção foi denominada RD<sup>Rio</sup> porque inicialmente foi detectada no RJ; entretanto, isto não significa que sua origem seja (desta região) do Brasil. Estudos estão sendo conduzidos para identificação de cepas RD<sup>Rio</sup> em outros estados do Brasil, assim como, em outros países e uma avaliação preliminar mostrou que cepas RD<sup>Rio</sup> foram também identificadas em um pequeno número de amostras dos Estados Unidos da América, Holanda e África, ou seja, regiões que apresentam uma alta frequência da família LAM. As cepas LAM da classe 1 e 2 são exclusivamente RD<sup>Rio</sup> e no SpolDB4 estão depositadas 426 cepas da classe LAM1 (tipo 20) e 8 países reúnem 74,13% destas cepas (Alemanha, Brasil, Haiti, Holanda, Namíbia, Portugal e Venezuela). Para a classe LAM2 estão depositadas no SpolDB4 322 cepas e 89% delas se concentram em 6 países (Brasil, Espanha, Guadeloupe, Haiti, Estados Unidos da América e Venezuela). Isto sugere que as linhagens de *M. tuberculosis* RD<sup>Rio</sup> poderiam ser encontradas fora do Brasil e que esta e outras deleções poderiam ser, devido ao caráter irreversível da deleção (e tudo indica, evento único), marcadores tanto para avaliar migrações populacionais e a associação parasito-hospedeiro, quanto a evolução da população bacteriana.

Uma hipótese é que a capacidade de dispersão das cepas RD<sup>Rio</sup>, verificada através de sua alta frequência (30%) na população estudada pode estar relacionada a perda das proteínas da família PE/PPE, principal fonte de variabilidade antigênica nas micobactérias. Semelhante situação pode ocorrer na inserção ou excisão de alguns elementos móveis alterando a expressão de genes envolvidos na interação parasito-hospedeiro ou contribuindo para o “fitness” bacteriano (Simone *et al.*, 1996 e Hammerschmidt *et al.*, 1996).

O fenótipo atenuado de *M. microti* e de “dassye bacillus” está relacionado a perda de dois antígenos secretados CFP10 e ESAT6. Estas proteínas são simultaneamente perdidas nas cepas BCG Pasteur e BCG Tóquio através da RD1, provavelmente esta deleção confere algum

tipo de vantagem para o BCG (Mostowy e Behr, 2005), do mesmo modo que a perda das proteínas da família PE/PPE, através da deleção RD<sup>Rio</sup>.

As deleções evolutivamente estabelecem uma relação entre virulência e dispersão do organismo nas comunidades hospedeiras, pois embora virulência possa facilitar a transmissão, a excessiva virulência será desvantajosa se conduzir a uma grande mortalidade ou a uma forma, em geral, menos transmissível de TB. A transmissibilidade ótima requer um equilíbrio entre algum grau de virulência capaz de gerar os agentes necessários para a dispersão e excreção durante tempo suficiente para garantir esta transmissão para um número suficiente de contatos (Mostowy e Behr, 2005).

Em nosso estudo onde estão incluídas as cepas RD<sup>Rio</sup>, a relação entre o grupo genético principal (PGG) 1, 2 e 3 está em acordo com a hipótese de que as cepas do grupo genético 3 tem reduzida capacidade de se estabelecer em grandes “clusters” (Sreevatsan *et al.*, 1997 e Brosch *et al.*, 2002). Neste caso, encontramos cepas RD<sup>Rio</sup> exclusivamente no grupo genético 2. Cepas de outras famílias como EAI e Beijing pertencem ao grupo genético 1 e sugerem uma proximidade com grupos ancestrais do Complexo *M. tuberculosis*.

A maioria dos genótipos estabelecidos por determinados marcadores se distribui dentro dos três grupos



genéticos principais (PGG) com exceção para as cepas da família T (tipo 53) com 38% no grupo genético 2 e 62% no grupo genético 3, mostrando que este grupo ainda não está suficientemente caracterizado.

A RD<sup>Rio</sup> com um número de bases superior a 26 Kb é, se comparada as RDs (2 - 12,7 Kb) descritas por Brosch *et al.*, (2002), a maior RD já descrita, portanto se esta deleção se enquadrar em uma RD deverá seguir o conceito de que as RDs podem diferenciar cepas de *M. tuberculosis* independente de sua origem geográfica, como é o caso de nossas observações nas cepas RD<sup>Rio</sup> internacionais.

A frequência de 30% de cepas RD<sup>Rio</sup> pode indicar que esta deleção foi responsável por conferir alguma vantagem seletiva para *M. tuberculosis* do grupo genético 2, no entanto outras circunstâncias que podem estar relacionadas a este “cluster” devem ser investigadas.

#### IV.1. Conclusão

O PCR-SSCP (“Polymerase Chain Reaction”) – (“Single Strand Conformation Polymorphism”) e a RBLH (“Reverse Line Blot Hybridization”) são metodologias adequadas para detecção rápida de resistência a rifampicina de cepas de *M. tuberculosis* e *M. bovis* e que podem ser desenvolvidas e implementadas a um baixo custo, embora exijam uma criteriosa preparação técnica. Para que o SSCP possa ser uma alternativa realmente econômica, é necessário que possa ser viabilizada a sua utilização em géis de poliacrilamida. No Brasil, a utilização da RBLH, para análise do gene *rpoB*, necessita que se tenha um melhor conhecimento das principais mutações mais frequentes que ocorrem em cepas brasileiras. As principais mutações já começam a ser esclarecidas através deste e de outros trabalhos brasileiros ou com cepas brasileiras.

A detecção de resistência de *M. tuberculosis* a estreptomicina, isoniazida e ethambutol mostraram resultados promissores, no entanto ainda não está adequada para responder as necessidades deste diagnóstico. Novas mutações e genes precisam ser estudados de modo que se possa estabelecer a sua relação com o fenótipo de resistência a estas drogas. Por conta disto vislumbramos a análise de resistência a drogas de segunda linha utilizadas no tratamento da TB, as quais já estão relacionadas a casos de resistência como as fluoroquinolonas, aminoglicosídeos e outras.

No caso específico da resistência em *M. bovis*, a RLBH será uma importante ferramenta na monitoração da resistência nesta espécie, no entanto nossos resultados estabeleceram duas possibilidades, (i) a frequência de cepas de *M. bovis* resistentes no gado bovino brasileiro é extremamente baixa, não tendo importância para a saúde animal e humana, o que é muito pouco provável ou (ii) RLBH precisa ser drasticamente modificada, no que diz respeito aos códons mutados e presentes na membrana para detecção destes casos.

Um segundo objetivo deste trabalho foi à construção do perfil genotípico, por “spoligotyping” das cepas do Complexo *M. tuberculosis* do Brasil. Para responder a este desafio, lançamos mão de antigas colaborações e estabelecemos novas, o que resultou em cepas provenientes de 11 estados brasileiros, que nos permitiram alcançar este objetivo, que está muito bem documentado no SpoIDB4 e em nosso manuscrito que descreve os “spoligotypes” brasileiros. Este trabalho, ainda não está completo, já que almejamos estabelecer um perfil

genotípico de *M. tuberculosis* provenientes de todos os estados brasileiros de modo a compreendermos a importância desta ou daquela família no Brasil, além de construirmos um cenário de “spoligotyping” com as duas mais importantes (até o

momento) metodologias de genotipagem de *M. tuberculosis*, que são o IS6110-RFLP e o MIRU-VNTR. Incluída também nesta visão de futuro, está à análise e expansão dos estudos da deleção RD<sup>RIO</sup>.

À espécie humana é atribuído o caráter inalienável do “ser político”, portanto não podemos nos esquecer que as mazelas que a acometem também têm este caráter, e isto lamentavelmente fica mais evidenciado nos países do terceiro mundo, como a Tuberculose e *Mycobacterium tuberculosis* não nos deixam esquecer.

## Cpítulo V – Bibliografía

Alland D, Kalkut GE, Moss AR, McAdam RA, Hahn JA, Bosworth W, *et al.* Transmission of tuberculosis in New York. *N Engl J Med.* 1994; 24: 1710-1716.

Anti-tuberculosis drug resistance in the world. The WHO/IUATLD Global Project on Anti-tuberculosis Drug Resistance Surveillance. WHO/TB/97.229. Geneva: World Health Organization. 1997.

Baptista, IMFD, Oelemann MAC, Opromola DVA, Suffys PN. Drug resistance and genotypes of strains of *Mycobacterium tuberculosis* isolated from human immunodeficiency virus-infected and non-infected tuberculosis patients in Bauru, São Paulo, Brazil. *Mem Inst Oswaldo Cruz*. 2002; 97: 1147-1152.

Barnes PF, Cave MD. Molecular Epidemiology of Tuberculosis. *N Engl J Med* 2003; 27: 1149-1156.

Bergmann JS, Woods GL. Mycobacterial growth indicator tube for susceptibility testing of *M. tuberculosis* to isoniazid and rifampin. *Diagn Microbiol Infect Dis* 1997; 27: 153-156.

Bifani PJ, Mathema B, Kuperina NE, Kreiswirth BN. Global dissemination of the *M. tuberculosis* W-Beijing family strains. *Trends in Microbiol*. 2002; 10: 45-52.

Bifani PJ, Plykaytis BB, Kapur V, Stockbauer K, Pan X, Lutfey ML, *et al*. Origin and interstate spread of New York City multidrug-resistant *M. tuberculosis* clone family. *JAMA*. 1996; 275: 452-457.

Blázquez J, de Los Monteros LEE, Samper S, Martin C, Guerrero A, Cobo J, *et al*. Genetic characterization of multidrug-resistant *M. bovis* strains from a hospital outbreak involving human immunodeficiency virus-positive patients. *J Clin Microbiol*. 1997; 35: 1390-1393.

Bobadilla-del-Vale M, Ponce-de-Leon A, Arenas-Huertero C, Vargas-Alarcon G, Kato-Maeda M, Small PM, *et al*. *rpoB* gene mutation in rifampin-resistant *M. tuberculosis* identified by PCR-SSCP. *Emerg Infect. Dis*. 2001; 7: 1010-1013.

Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K. *et al*. A new evolutionary scenario for the *M. tuberculosis* complex. *Proc Natl Acad Sci U S A*. 2002; 99: 3684-9.

Burger M, Raskin S, Brockelt SR, Amthor B, Geiss HK, Haas WH. DNA Fingerprinting of *M. tuberculosis* complex culture isolates collected in Brazil and spotted onto filter paper. *J Clin Microbiol*. 1998; 36: 573-576.

Cardoso RF, Cooksey RC, Morlock GP, Barco P, Cecon L, Forestiero F, *et al*. Screening and characterization of mutations in isoniazid-resistant *M. tuberculosis* isolates obtained in Brazil. *Antimicrob Agents Chemother*. 2004; 48: 3373-3381.

Cohn DL, Bustreo F, Raviglione MC. Drug-resistant tuberculosis: review of the worldwide situation and the WHO/IUATLD global surveillance project. *Clin Infect Dis* 1997; 24 (Suppl 1): S121-30

.

Cousins D, Willians S, Liébana E, Aranaz A, Bunschoten A, van Embden J, *et al*. Evaluation of four DNA techniques in epidemiological investigation of bovine tuberculosis. *J Clin Microbiol*. 1998; 36: 168-178.

[De Beenhouwer H, Lhiang Z, Jannes G, Mijs W, Machtelinckx L, Rossau R, et al.](#) Rapid detection of rifampicin resistance in sputum and biopsy specimens from tuberculosis patients by PCR and line probe assay. *Tuber Lung Dis*. 1995; 76:425-30.

de la Salmonière GY-O, Li HM, Torrea G, Bunschotem A, van Embden A, Giquel B. Evaluation of spoligotyping in a study of the transmission of *M. tuberculosis*. *J Clin Microbiol*. 1997; 35: 2210-2214.

Duchêne V, Ferdinand S, Filliol I, Guégan JF, Rastogi N, Sola C. Phylogenetic reconstruction of *M. tuberculosis* within four settings of the Caribbean region: tree comparative analysis and first appraisal on their phylogeography. *Infect Genetic Evolut*. 2004; 4: 5-14.

Fandinho FCO, Kritski AL, Hoefler C, Conde H, Ferreira RMC, Saad MHF, LW, *et al*. RFLP patterns and risk factors for recent tuberculosis transmission among hospitalized tuberculosis patients in Rio de Janeiro, Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2000; 94: 271-275.

[Ferrazoli L, Palaci M, Marques LR, Jamal LF, Afiune JB, Chimara E, et al.](#) Transmission of tuberculosis in an endemic urban setting in Brazil. *Int J Tuberc Lung Dis*. 2000; 4: 18-25.

[Filliol I, Driscoll JR, Van Soolingen D, Kreiswirth BN, Kremer K, Valetudie G, et al.](#) Global distribution of *M.*

*tuberculosis* spoligotypes Emerg Infect. Dis. 2002; 8: 1347-9.

[Filliol I, Driscoll JR, van Soolingen D, Kreiswirth BN, Kremer K, Valetudie G, et al.](#) Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study. J Clin Microbiol. 2003; 41: 1963-70.

[Filliol I, Ferdinand S, Sola C, Thonnon J, Rastogi N.](#) 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. 2002; 153: 81-8.

[Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbon MH, Bobadilla del Valle M, et al.](#) 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. J Bacteriol. 2006; 188: 759-72.

Fletcher HA, Donoghue D, Taylor GM, van der Zanden AGM, Spigelman M. Molecular analysis of *M. Tuberculosis* DNA from a family of 18th century Hungarians. Microbiology. 2003; 149: 143-151.

Gagneux S, DeRiemer K, Van T, Kato-Maeda M, Jong BC, Narayanan S, et al. Variable host-pathogen compatibility in *M. tuberculosis*. Proc Natl Acad Sci. 2006; 103: 2869-2873.

Gutacker MM, Mathema B, Soini H, Shashikina E, Kreiswirth BN, Gravis EA, et al. Single Nucleotide Polymorphism-based population genetic analysis of *M. tuberculosis* strains from 4 geographic sites. J Infect Dis. 2006; 193: 121-128.

Gutierrez MC, Galan JC, Blazquez J, Bouvet E, Vincent V. Molecular markers demonstrate that the first described multidrug-resistant *Mycobacterium bovis* outbreak was due to *Mycobacterium tuberculosis*. J Clin Microbiol. 1999; 37(4):971-5

Hammerschmidt S, Hilse R, van Putten JPM, Gerardy-Schahn R, Unkmeir A, Frosch M. EMBO J. 1996; 15: 192-198.

Hermans PWM, van Soolingen D, Bik EM, de Haas PEW, Dale JW, van Embden JDA. The insertion element IS987 from *M. bovis* BCG is located in a hot spot integration region for insertion elements in *M. tuberculosis* complex strains. Infect Immun. 1991; 59: 2695-2705.

Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small P. Stable association between strains of *M. tuberculosis* and their human host populations. Proc Natl Acad Sci USA. 2004; 101: 4871-4876.

Horgen L, Sola C, Devallois A, Goh KS, Rastogi N. Follow up of *M. tuberculosis* transmission in the French West Indies By IS6110-DNA fingerprinting and DR-based Spoligotyping. Immunol Med Microbiol. 1998; 21: 203-212.

II Diretrizes Brasileiras para Tuberculose. J Bras Pneumo. 2004; (30) supl. 1

[Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al.](#) Simultaneous detection and strain differentiation of *M. tuberculosis* for diagnosis and epidemiology. J Clin Microbiol. 1997; 35: 907-914.

Kim BJ, Kim SY, Park BH, Lyu MA, Park IK, Bai GH, et al. Mutations in the *rpoB* gene of *M. tuberculosis* that interfere with PCR-SSCP analysis for rifampin susceptibility testing. J Clin Microbiol. 1997; 35: 492-494.

Kim BJ, Lee KH, Park BN, Kim SJ, Park EM, Park YG, Bai GH, et al. Detection of rifampin-resistant *M. tuberculosis* in sputa by nested PCR-linked SSCP and DNA sequencing. J Clin Microbiol. 2001; 39: 2610-2617.

[Kremer K, van Soolingen D, Frothingham R, Haas WH, Hermans PW, Martin C, et al. Related Articles,](#)

Comparison of methods based on different molecular epidemiological markers for typing of *M. tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol.* 1999; 37:2607-18.

Kremer K, Au BKY, Yip PCW, Skuce R, Supply P, Kam KM, *et al.* Use of variable-number tandem-repeat typing to differentiate *M. tuberculosis* Beijing family isolates from Hong Kong and comparison with IS6110 restriction fragment length polymorphism typing and spoligotyping. *J Clin Microbiol.* 2005; 43: 314-320.

Kremer K, van Soolingen D, Frothingham R, Haas WH, Hermans PWM, Martín C, *et al.* Comparison of methods based on different molecular epidemiological markers for typing of *M. tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol.* 1999; 37: 2607-2618.

Kritski AL, Ruffino-Netto. A. Health sector reform in Brazil: impact on tuberculosis control. *Int J Tuberc Lung Dis.* 2000; 4: 1-5.

Legrand E, Filliol I, Sola C, Rastogi N. Use of spoligotyping to study the evolution of the direct repeat locus by IS6110 transposition in *M. tuberculosis*. *J Clin Microbiol.* 2001; 39: 1595-1599.

Ministério da Saúde. 1994; Manual de Bacteriologia da Tuberculose.

Morcillo N, Zumarraga M, Alito A, Dolmann A, Schouls L, Cataldi A, *et al.* Low cost home-made, reverse line blot hybridization assay for rapid detection of rifampicin resistance in *M. tuberculosis*. *Int J Tuberc Lung Dis.* 2002; 6: 959-965.

Mostowy S, Behr MA. The origin and evolution of *M. tuberculosis*. *Clin Chest Med.* 2005; 26: 207-216.

Mostowy S, Cousins D, Brinkman J, Aranaz A, Behr MA. Genomic deletions suggest a Phylogeny for the *M. tuberculosis* complex. *J Infect Dis.* 2002; 186: 74-80.

Mostowy S, Onipede A, Gagneaux S, Niemann S, Kremer K, Desmond EP, *et al.* Genomic Analysis Distinguishes *M. tuberculosis*. *J Clin Microbiol.* 2004; 42: 3594-3599.

Musser JM, Ramaswamy S. Molecular genetic basis of antimicrobial agent resistance in *M. tuberculosis*: 1998 update. *Tubercle and Lung Dis.* 1998; 79: 3-29.

Musser JM. Antimicrobial agent resistance in mycobacteria: molecular genetic insights. *Clin Microbiol Rev.* 1995; 84: 496-514.

[Narayanan S.](#) Molecular epidemiology of tuberculosis. *Indian J Med Res.* 2004; 120:233-47.

Nguyen D, Brassard P, Menzies D, Thibert L, Warren R, Mostowy S, *et al.* Genomic characterization of an endemic *M. tuberculosis* strain: evolutionary and epidemiologic implications. *J Clin Microbiol.* 2004; 42: 2573-2580.

Niemann S, Kubika T, Bange FC, Adjei O, Browne EN, Chimbuah MA. The species *Mycobacterium africanum* in the light of new molecular markers, *et al.* *Clin Microbiol.* 2004; 42: 3958-3962.

Niemann S, Rüsç-Gerdes S, Joloba ML, Whalen CC, Guwatudde D, Ellner JJ, *et al.* *M. africanum* subtype II is associated with two distinct genotypes and is a major cause of human tuberculosis in Kampala, Uganda. *J Clin Microbiol.* 2002; 40: 3398-3405.

Nikaido H. Multiple Antibiotic Resistance and Efflux. *Curr Opin Microbiol.* 1998; 1: 516-523.

Niobe-Eyangoh SN, Kuaban C, Sorlin P, Thonnon J, Vincent V, Gutierrez MC. Molecular characteristics of strains of the cameroon family, the major group of *M. tuberculosis* in a country with a high prevalence of tuberculosis. *J Clin Microbiol.* 2004; 42: 5029-5035.

Oelemann MAC, Fontes ANB, Pereira MAS, Bravin Y, Barreira G, Degrave WM, *et al.* Typing *M. tuberculosis* strains isolated in 11 communities health centers of Rio de Janeiro City, Brazil. *Submetido*.

Pfyffer GE, Strassle A, Rose N, Wirth R, Brandi O, Shang H. Transmission of Tuberculosis in the metropolitan area of Zurich: a 3 year survey based on DNA fingerprint. *Eur Respir J.* 1998; 11: 804-808.

Qian L, van Embden J, van der Zandem A, Weltevreden EF, Duanmu H, Douglas JT. Retrospective analysis of the Beijing Family of *M. tuberculosis* in preserved lung tissues. *J Clin Microbiol.* 1999; 37: 471-474.

Roberts CA, Boylston A, Buckley L, Chamberlain AC, Murphy EM. Rib lesions and tuberculosis: the palaeopathological evidence. *Tuber Lung Dis.* 1998;79(1):55-60.

Rossetti MLRA, [Valin ARM](#), [Silva MSN](#), Rodrigues VS. Resistant tuberculosis: a molecular review. *Rev Saúde Pública.* 2002; 36: 525-532.

Rothschild BM, Martin LD, Lev G, Bercovier H, Bar-Gal GK, Greenblatt C. *et al.* *M. tuberculosis* Complex DNA from an extinct bison dated 17,000 years before the present. *Clin Infect Dis.* 2001; 33: 305-311.

Sechi LA, Zanetti S, Sanguinetti M, Mollicotti P, Romano L, Leori G, *et al.* Molecular basis of rifampin and Isoniazid resistance in *M. bovis* strains isolated in Sardinia, Italy. *Antimicrob Agents Chemother.* 2001; 45: 1645-1648.

Silva, PEA, Bigi F, de la Paz Santangelo M, Romano MI, Martin C, Cataldi A, *et al.* Characterization of P55, a multidrug efflux pump in *M. bovis* and *M. tuberculosis*. *Antimicrob Agents Chemother.* 2001; 45: 800-804.

Simone M, Riot B, Fortineau N, Berche P. Invasin production by *Yersinia pestis* is abolished by insertion of an IS200-like element within the *inv* gene. *Infect Immun.* 1996; 64: 375-379.

Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, *et al.* 1994; The epidemiology of tuberculosis in San Francisco. *N Engl J Med.* 330: 1703-1709.

Sola C, Devallois A, Horgen L, Maïset J, Filliol I, Legrand I, *et al.* Tuberculosis in the Caribbean: using spacer oligonucleotide typing to understand strain origin and transmission. *Emerg Infect Dis.* 1999; 5: 404-414.

Sola C, Ferdinand S, Mammina C, Nastasi A, Rastogi N. Genetic diversity of *M. tuberculosis* in Sicily based on spoligotyping and variable number of tandem DNA repeats and comparison with a Spoligotyping database for population-based analysis. *J Clin Microbiol.* 2001; 39: 1559-1565.

Sola C, Filliol I, Gutierrez MC, Mokrousov I, Vincent V, Rastogi N. Spoligotype Database of *M. tuberculosis*: Biogeographic Distribution of shared types and epidemiologic and phylogenetic perspectives. *Emerg Infect Dis.* 2001; 7: 390-396.

Sola C, Filliol I, Legrand E, Lesjean S, Locht C, Supply P, *et al.* Genotyping of *M. tuberculosis* complex using MIRUs- association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. *Infect Genet Evol.* 2003; 3: 125-133.

Sola C, Filliol I, Legrand E, Mokrousov I, Rastogi N. *M. tuberculosis* Phylogeny based 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.* 2001; 53: 680-689.

Sola C, Horgen L, Maïsetti J, Devallois A, Goh KS, Rastogi N. Spoligotyping followed by double-repetitive-element PCR as a rapid alternative to IS6110 fingerprinting for epidemiological studies of tuberculosis. *J Clin Microbiol.* 1998; 36: 1122-1124.

[Sharma SK](#), [Mohan A](#). Multidrug-resistant tuberculosis. *Indian J Med Res.* 2004; 120:354-76.

- Spindola de Miranda S, Kritski AL, Filliol I, Mabilat C, Panteix G, Drouet E. Mutations in the *rpoB* gene of rifampicin-resistant *M. tuberculosis* strains isolated in Brazil and France. *Mem Inst Oswaldo Cruz.* 2001; 96: 247-250.
- Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whitam TS, *et al.* Restricted structural gene polymorphism in the *M. tuberculosis* complex indicates evolutionary recent global dissemination. *Proc Natl Acad Sci USA.* 1997; 94: 9869-9874.
- Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C. Automated high-throughput genotyping for study of global epidemiology of *M. tuberculosis* based on MIRUs. *J Clin Microbiol.* 2001; 39: 3563-3571.
- Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Locht C. Variable human minisatellite-like regions in the *M. tuberculosis* genome. *Mol Microbiol.* 2000; 36(3): 762-771.
- [Supply P, Magdalena J, Himpens S, Locht C.](#) Identification of novel intergenic repetitive units in a mycobacterial two-component system operon. *Mol Microbiol.* [1997.](#) 26:991-1003.
- Suresh N, Singh UB, Arora J, Pande JN, Seth P, Samantaray JC. Rapid detection of rifampicin-resistant *M. tuberculosis* by in-house, reverse line blot assay. *Diagn Microbiol. Infect Dis.* 2006; *In press.*
- [Telenti A, Imboden P, Marchesi F, Schmidheini T, Bodmer T.](#) Direct, automated detection of rifampin-resistant *Mycobacterium tuberculosis* by polymerase chain reaction and single-strand conformation polymorphism analysis. *Antimicrob Agents Chemother.* 1993; 37:2054-8.
- Thierry D, Cave MD, Eisenach KD, Crawford JT, Bates JH, Gicquel B, *et al.* IS6110, an IS-like element of *Mycobacterium tuberculosis* complex. *Nucleic Acids Res.* 1990; 11(18): 188.
- Valim ARM, Rossetti MLR, Ribeiro MO, Zaha A. Mutations in the *rpoB* gene of multidrug-resistant *M. tuberculosis* isolates from Brasil. *J Clin Microbiol.* 2000; 38: 3119-3122.
- Van der Zandem AGM, Te Koppele-Vije EM, Bhanu NV, Van Soolingen D, Schouls LM. Use of DNA extracts from Ziehl-Neelsen-stained slides for molecular detection of rifampin resistance and spoligotyping of *M. tuberculosis*. *J Clin Microbiol.* 2003; 41: 1101-1108.
- van Deutekon H, Gerritsen JJJ, van Soolingen D, van Ameijden EJC, van Embden JD, Coutinho RA. A molecular epidemiological Approach to studying the transmission of tuberculosis in Amsterdam. *Clin Infect Dis.* 1997; 25: 1071-1077.
- van Embden JDA, Cave MD, Crawford JT, Strain Identification of *M. tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol.* 1993; 31: 406-409.
- Van Soolingen D, de Haas PEW, Haagsma J, Eger T, Hermans PWM, Ritacco V, *et al.* Use of various genetic markers in differentiation of *M. bovis* strains from animals and humans and for studying epidemiology of bovine tuberculosis. *J Clin Microbiol.* 1994; 32: 2425-2433.
- Viana-Niero C, Gutierrez C, Sola C, Filliol I, Boulahbal, Vincent V, *et al.* Genetic diversity of *M. africanum* clinical isolates base don IS6110-Restriction Fragment Length Polymorphism analysis, Spoligotyping, and Variable Number of Tandem DNA Repeats. 2001; *J Clin Microbiol.* 39: 57 – 65.
- Victor TC, de Haas PEW, Jordaan AM, van der Spuy GD, Richardson M, van Soolingen D, *et al.*, Molecular Characteristics and global spread of *M. tuberculosis* with a Western Cape F11 genotype. *J Clin Microbiol.* 2004; 42: 769-772.
- Viedma DG, Bouza E, Rastogi N, Sola C. Analysis of *M. tuberculosis* genotypes in Madrid and identification of two new families specific to Spain-related settings. *J Clin Microbiol.* 2005; 43: 1797-1806.
- Wanger A, Mills K. Testing of *M. tuberculosis* susceptibility to ethambutol, isoniazid, rifampin and streptomycin by using Etest. *J Clin Microbiol.* 1996; 34: 1672-1676.



Wilmar LS, Aufderheide AC, Buikstra J, Holcomb TA. Identification of *M. tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proc Natl Acad Sci.* 1994; 91: 2091-2094.

Yang Z, Durmaz R, Yang D, Gunal S, Zhang L, Foxman B, *et al.* Simultaneous detection of isoniazid, rifampin, and ethambutol resistance of *M. tuberculosis* by a single multiplex allele-specific polymerase chain reaction (PCR) assay. *Diag Microbiol Infect. Dis.* 2005; 53: 201-208.

Zanini MS, Moreira EC, Salas CE, Lopes MTP, Barouni AS, Roxo E, *et al.* Molecular typing of *M. bovis* isolates from South-east Brazil by Spoligotyping and RFLP. *J Vet Med B.* 2005; 52: 129-133.

Zink AR, Sola C, Reischl U, Grabner W, Rastogi N, Wolf H, *et al.* Characterization of *M. tuberculosis* complex DNAs from Egyptian mummies by spoligotyping. *J Clin Microbiol.* 2003; 41: 359-367.

Zozio T, Allix C, Gunal S, Saribas Z, Alp A, Durmaz R, Fauville M, *et al.* Genotyping of *M. tuberculosis* clinical isolates in two cities of Turkey: Description of a new family of genotypes that is phylogeographically specific for Asia Minor. *BMC Microbiol.* 2005; 26: 5:44.

