

Um gene em movimento: o papel do *IS6110* no fenótipo, patogenicidade e diagnóstico do *Mycobacterium tuberculosis*



**João Guilherme Souza Oliveira¹, LÍlian Maria Lapa Montenegro²,
Willisses Henrique Veloso Carvalho-Silva³**

¹Bacharel em Biomedicina, mestrando pelo "Programa de Pós-graduação em Biotecnologia em Saúde (IAM/FIOCRUZ-PE)

²Graduação em Ciências Biológicas na Universidade Federal Rural de Pernambuco, mestrado em Biofísica pela Universidade Federal de Pernambuco e doutorado em Saúde Pública pela Fundação Oswaldo Cruz. Pesquisadora do Instituto Aggeu Magalhães/FIOCRUZ

³Bacharel em Ciências Biológicas (UFPE), Mestre em Genética (UFPE), Doutor em Biologia Aplicada à Saúde (Instituto Keizo Asami-UFPE). PDJ no Instituto Aggeu Magalhães (IAM/FIOCRUZ-PE) e professor de Medicina do Centro Universitário Facol (UNIFACOL)

Autor para correspondência - joao.gsoliveira@ufpe.br

Palavras-chave: tuberculose, transposon, sequência de inserção, teste molecular

O *Mycobacterium tuberculosis* (Mtb) é a bactéria que causa a tuberculose (TB), sendo a principal causa de morte por um único agente infeccioso no mundo. Uma parte do material genético dessa bactéria corresponde a uma sequência de nucleotídeos, a chamada sequência de inserção 6110 (IS6110), pertencente à família IS3. Essa sequência é um elemento genético móvel de classe 2 (transposon) sendo específico do *M. tuberculosis*. A partir dos seus mecanismos de transposição, o IS6110 pode ser capaz de induzir fenótipos associados à virulência e à vantagem adaptativa diante do sistema imunológico do hospedeiro. A compreensão da dinâmica do IS6110 é essencial para elucidar mecanismos de virulência do Mtb e aprimorar estratégias diagnósticas e de controle da TB.

Mycobacterium tuberculosis e a Tuberculose

O *Mycobacterium tuberculosis* (Mtb) é uma bactéria que é o agente etiológico da tuberculose (TB), uma doença infectocontagiosa responsável por altos índices de morbidade e mortalidade no mundo. Morfologicamente, o Mtb apresenta-se na forma de bacilo, e é transmitido pelo ar, alcançando os pulmões, onde tem como alvo inicial os **macrófagos** dos alvéolos pulmonares, etapa fundamental para o estabelecimento da infecção. Em 2023, ultrapassando a COVID-19, a TB voltou a ser a principal causa de morte por um único agente infeccioso, somando cerca de 1,25 milhões de óbitos e 10,8 milhões de casos. Atualmente, o Brasil está entre os 30 países mais afetados pela TB e considerados prioritários pela Organização Mundial de Saúde (OMS) no combate à doença, registrando, somente em 2024, cerca de 84 mil casos e cerca de 6 mil óbitos em 2024.

O Mtb faz parte de um grupo de sete organismos muito próximos geneticamente, denominado de **complexo** *Mycobacterium tuberculosis* (CMTB), incluindo *M. tuberculosis*, *M. bovis*, *M. africanum*, *M. pinnipedii*, *M. caprae*, *M. microti* e *M. canetti* (Figura 1). O CMTB é um grupo que compartilha, no seu genoma, sequências idênticas de DNA que codificam RNA ribossômico 16S

e com alta homogeneidade genética (99,9% de similaridade de nucleotídeos), mas distinguem-se em suas características fenotípicas e mecanismos patológicos no hospedeiro. Tal elevada similaridade sugere um **gargalo evolutivo** recente e uma subsequente **expansão clonal** a partir de um ancestral comum (do inglês, *most recent common ancestor* – MRCA), que surgiu há cerca de 40.000 anos na África Oriental, região considerada o local de sua origem. Entretanto, evidências históricas e moleculares indicam que a associação entre o Mtb e os seres humanos é muito mais antiga, com infecções presentes em homínidos na África Oriental há milhões de anos. Registros de tuberculose foram identificados em múmias do Egito Antigo, confirmados por análises de DNA e por alterações ósseas características, demonstrando a persistência da doença ao longo de diferentes civilizações até os dias atuais, quando permanece como um importante problema de saúde pública mundial.

No genoma do Mtb, certos elementos genéticos, como a sequência de inserção 6110 (do inglês, *Insertion Sequence 6110* – IS6110), apresentam a capacidade de transpor-se (“saltar”) para diferentes regiões genômicas, caracterizando um transposon. A compreensão do papel do IS6110 é fundamental para o entendimento dos processos de patogênese e evolução do patógeno, além da sua relevância clínica no diagnóstico da TB e, conseqüentemente, no controle da transmissibilidade da doença em escala global.

Gargalo evolutivo - Trata-se de um evento evolutivo no qual o tamanho de uma população é drasticamente reduzido, decorrente de eventos ambientais ou isolamento geográfico.

Expansão clonal - Multiplicação de um grupo de bactérias geneticamente idênticas a partir de um ancestral comum.

Macrófagos - Células do sistema imunológico responsáveis por reconhecer, englobar e destruir microrganismos, além de participar da ativação da resposta imune por meio da apresentação de antígenos a outras células de defesa.

Complexo - Conjunto de microrganismos geneticamente muito semelhantes, pertencentes ao mesmo grupo, que compartilham características biológicas e patogênicas, mas que podem ser classificados em espécies diferentes.

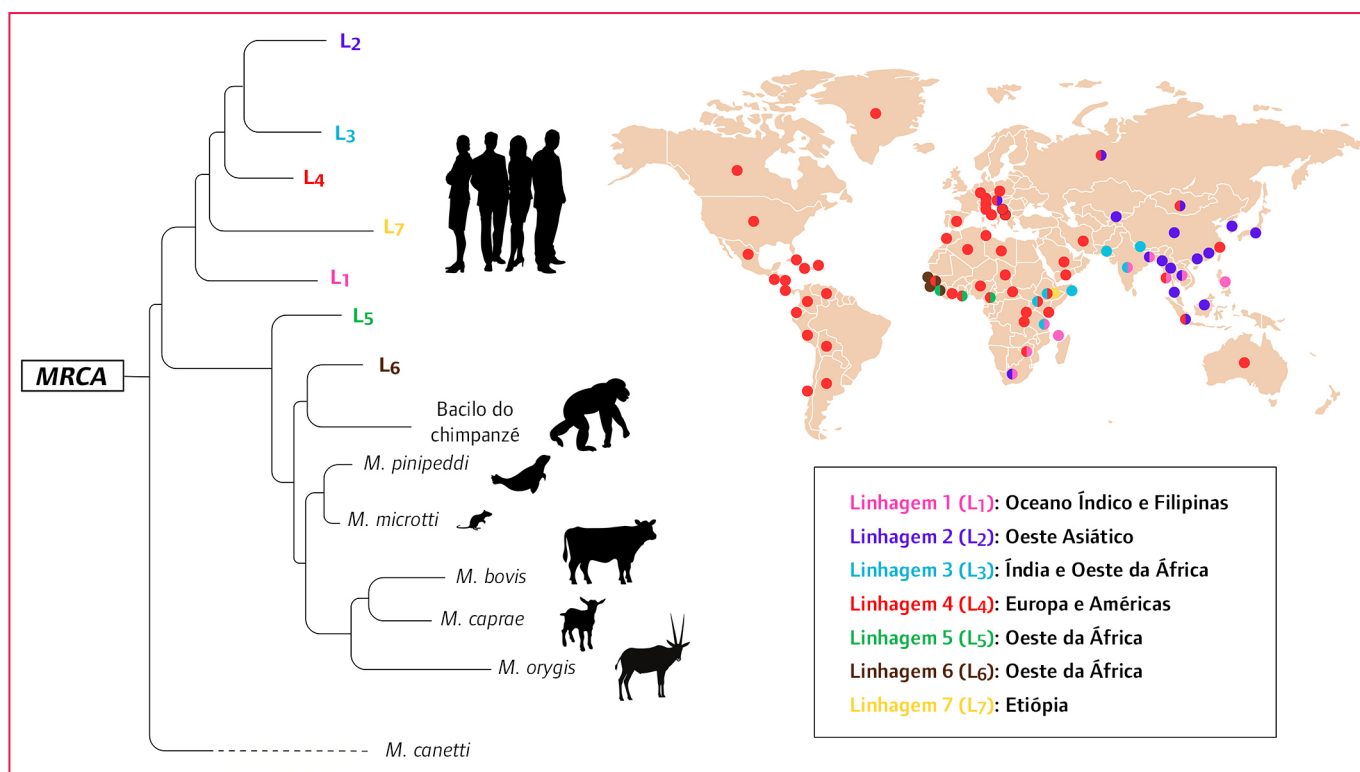


Figura 1.
Filogenia e distribuição geográfica do CMTB.

Sugere-se que as espécies do complexo *Mycobacterium tuberculosis* (CMTB) surgiram de um ancestral comum mais recente (MRCA), que teria surgido há cerca de 40.000 anos. As linhagens capazes de infectar exclusivamente os seres humanos estão representadas de forma colorida, enquanto as linhagens adaptadas aos animais estão na cor cinza. Além disso, destaca-se o mapa com a distribuição global das sete principais linhagens de CMTB adaptadas aos humanos.

ORF - Trecho de uma sequência do DNA que potencialmente pode ser traduzido na célula para produzir uma proteína. Ele começa em um ponto específico de início (côdon ATG), termina em um côdon de parada e possui vários códons correspondentes a aminoácidos, indicando que aquela sequência pode gerar uma proteína funcional.

Remodelação Cromossômica - Alteração na estrutura, quantidade ou arranjo dos cromossomos, como deleção, duplicação e inserção.

IS6110: o gene saltador

Na década de 1950, a pesquisadora Barbara McClintock identificou a presença de elementos genéticos móveis (EGMs) e elucidou seu papel na **remodelação cromossômica** de espécies de milho. Os EGMs podem ser classificados em duas categorias principais: classe I e classe II. Os EGMs de classe I, também conhecidos como retrotransposons, possuem a capacidade de mover-se no genoma por meio de um mecanismo de “copiar e colar”. Em primeiro lugar, são transcritos em RNA e posteriormente convertidos novamente em DNA por ação de uma enzima transcriptase reversa (Figura 2A). No caso dos EGMs de classe II, geralmente denominados transposons, eles codificam outro tipo de enzima chamada transposase, que permite seu deslocamento no genoma por meio de um mecanismo de “cortar e colar” (Figura 2B). Em bactérias, esses transposons

são chamados de sequências de inserção (IS – *Insertion Sequences*).

O IS6110 é um elemento de inserção do tipo transposon, específico do CMTB. Trata-se de uma sequência com 1358 pares de bases (pb), que pertence à família IS3 de sequências de inserção. Estudos demonstram que esse elemento pode estar presente em múltiplas cópias – variando de aproximadamente 10 a 156 – em diferentes cepas de Mtb. O IS6110 possui duas regiões do DNA chamadas **ORFs** (*open reading frames* ou quadros de leitura abertos), denominadas orfA e orfB que, após serem transcritas, podem ser traduzidas em proteínas nas células. Essas duas ORFs apresentam sequências repetidas invertidas em suas extremidades e se sobrepõem parcialmente, ou seja, compartilham um par de bases (1 pb) da sua sequência no DNA (Figura 3). Cada uma dessas ORFs pode ser traduzida separadamente, originando proteínas individuais (A e B) que ajudam a controlar a movimentação do IS6110 dentro do genoma.

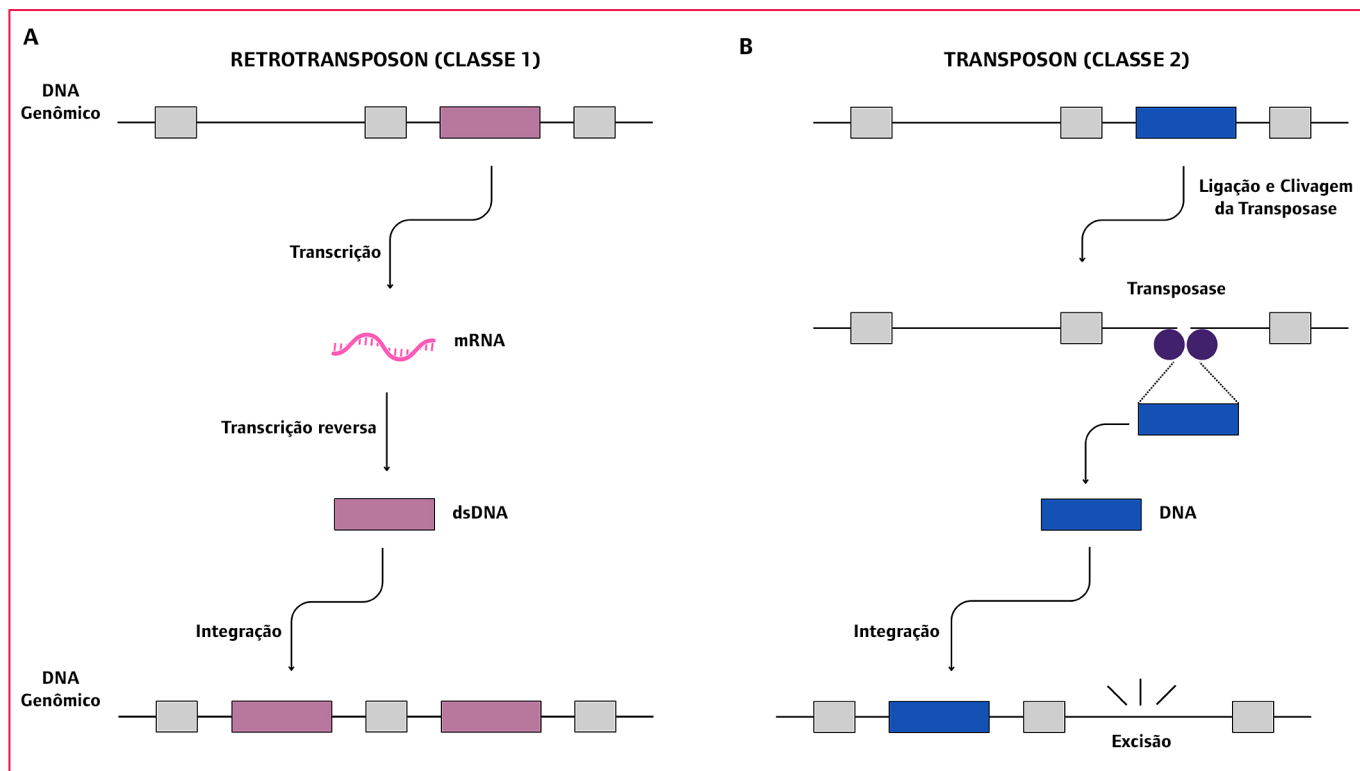


Figura 2. As duas principais classes de elementos genéticos móveis (EGMs). (A) Os retrotransposons (Classe 1), utilizam o método de “copiar e colar”, a partir da transcrição dos transposons em RNA, seguida de uma transcrição reversa e integração a um novo local do DNA genômico. (B) Os transposons (Classe 2) executam um processo de “cortar e colar”, no qual são removidos do local de origem a partir da ação das transposases e integrados em outra região alvo do DNA genômico, enquanto o local de origem sofre o processo de excisão e reparo.

Frameshift translacional - Mecanismo no qual o ribossomo muda a forma de leitura do RNA (deslocando a matriz de leituras das trinças) durante a produção de uma proteína. Essa mudança permite que dois quadros de leitura (ORFs) diferentes do DNA sejam traduzidos juntos, formando uma única proteína maior e funcional.

Além disso, em algumas situações, ocorre um mecanismo especial na tradução do RNA, chamado de mudança do quadro de leitura (**frameshift**) translacional, no qual o ribossomo (a estrutura responsável por produzir proteínas) muda a forma de leitura do RNA e passa a ler as trinças de modo diferente e deslocado. Esse “deslizamento” permite que as duas ORFs sejam traduzidas em conjunto, formando uma **proteína de fusão**, chamada orfAB. No caso do *IS6110*, essa proteína de fusão corresponde à transposase, a enzima responsável por permitir que essa sequência de inserção se mova e se insira em diferentes regiões do DNA bacteriano.

Proteína de fusão - É uma única proteína formada pela junção de duas sequências nucleotídicas diferentes, originárias de sequências que codificam duas proteínas distintas.

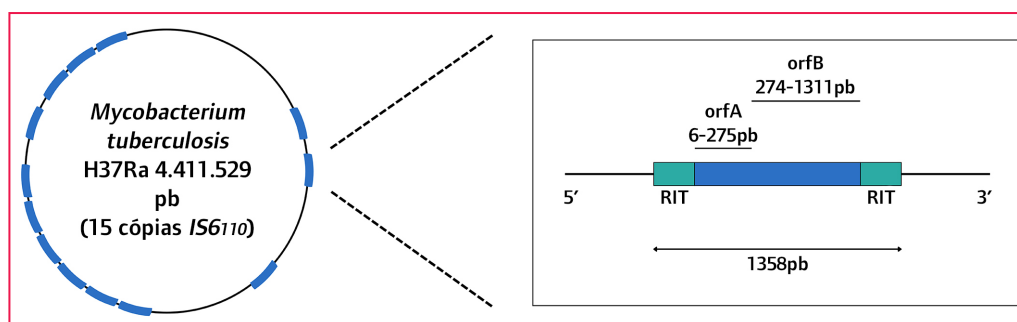


Figura 3. Ilustração esquemática da sequência *IS6110*. O gene *IS6110*, comumente presente em 10 a 15 cópias nas cepas de *Mycobacterium tuberculosis* (Mtb), é constituído por dois quadros de leitura abertos sobrepostos (ORFs, orfA e orfB) e possui repetições invertidas terminais (RITs) em suas extremidades.

IS6110: uma faca de dois gumes para o Mtb?

Atualmente, sabe-se que a frequência de transposição do *IS6110* pode influenciar diretamente o fenótipo do Mtb, uma vez que esse elemento pode atuar como sequência promotora no local em que ele se insere, modulando a expressão de genes vizinhos e, potencialmente, aumentando a virulência do patógeno. Isso ocorre porque o *IS6110* contém, em sua própria sequência, regiões semelhantes à sequência dos **promotores bacterianos**, incluindo sequências consenso reconhecidas pela RNA polimerase, especialmente em seu extremo 3'. Assim, quando o *IS6110* se insere próximo a um gene e na orientação adequada, ele pode aumentar ou alterar o nível de expressão desse gene. Entretanto, ainda não há consenso na literatura acerca do impacto do número de cópias, da redução ou do aumento da atividade de transposição sobre o aumento ou diminuição da patogenicidade do Mtb.

Já foi relatado, por exemplo, que o *IS6110* pode atuar como um promotor móvel, uma vez que, após sua transposição e integração próxima a regiões gênicas, pode iniciar ou aumentar a transcrição de genes adjacentes. Dentre os genes que sofreram influência, destaca-se o *phoP*, um gene associado ao crescimento prejudicado de Mtb em macrófagos de camundongos. Em um **isolado clínico** responsável por um surto de *Mycobacterium tuberculosis* resistente a drogas (Mtb-DR), observou-se que a transcrição de *phoP* foi aumentada cerca de dez vezes após a inserção do *IS6110* a montante do códon de início da tradução (ATG). Portanto, acredita-se que essa regulação positiva provavelmente esteja associada ao aumento da virulência durante o processo infeccioso.

Nesse contexto, surgiu a hipótese de que o maior número de cópias do *IS6110* estaria

relacionado ao aumento da virulência do Mtb durante a infecção. Essa suposição baseou-se em um estudo realizado em 2004, no qual se observou que a linhagem *Mycobacterium tuberculosis* Beijing/W apresentava 21 cópias do *IS6110* – o maior número entre todas as cepas analisadas –, sendo reconhecida por sua elevada patogenicidade e resistência a múltiplos fármacos.

No entanto, cepas com baixo número de cópias desse transposon também são capazes de causar surtos e desenvolver resistência a medicamentos. Espécies como *M. bovis*, que apresentam apenas uma cópia do *IS6110* em seu genoma, demonstram elevada capacidade de transmissão e infecção. Além disso, já foi descrito na literatura que a presença de um número excessivo de cópias desse elemento em diversas regiões do genoma do Mtb pode resultar na inativação ou exclusão de genes essenciais, e isso pode ser potencialmente prejudicial ao patógeno. Sendo assim, os efeitos da inserção do *IS6110* no genoma micobacteriano podem ser deletérios ou vantajosos do ponto de vista evolutivo, a depender da região afetada. Esses achados sugerem que diferentes espécies micobacterianas adotam distintos mecanismos de sobrevivência e adaptação, independentemente do conteúdo genômico relacionado ao *IS6110*.

Diante disso, surge o questionamento: qual é, de fato, o papel do *IS6110*? Trata-se de um gene egoísta ou de um fator vantajoso para o patógeno? Na primeira hipótese, sua presença não teria impacto significativo para o Mtb, atuando apenas como um elemento genético passivo, com potencial de gerar desequilíbrios quando presente em elevado número de cópias. Já na segunda hipótese, o *IS6110* poderia representar uma estratégia adaptativa utilizada pelo patógeno como uma “válvula de escape” frente a ambientes hostis ou condições de estresse. Essa característica possibilitaria eventos de transposição que culminariam na indução de fenótipos associados à virulência e à vantagem adaptativa diante do sistema imunológico do hospedeiro.

Promotor bacteriano -

Conjunto de sequências consenso (conservadas evolutivamente) que são reconhecidas pelas enzimas RNA polimerases bacterianas, permitindo o início da transcrição de um RNA.

Isolado clínico -

Corresponde a uma amostra bacteriana obtida diretamente de um paciente e posteriormente cultivada em laboratório para fins de identificação e análise.

Importância diagnóstica do *IS6110*

O *IS6110* possui significativa importância laboratorial, sendo amplamente utilizado como alvo em métodos moleculares para a detecção do Mtb. A presença de múltiplas cópias no genoma aumenta o número de alvos disponíveis, contribuindo para a elevação da sensibilidade dos testes diagnósticos. Atualmente, diversas metodologias, incluindo algumas em desenvolvimento ou aprimoramento, utilizam o *IS6110* como gene-alvo principal para diagnóstico da TB.

Entre essas metodologias, destaca-se a PCR (do inglês, *Polymerase Chain Reaction*), técnica versátil, específica, econômica e de rápida execução, que permite utilizar diferentes tipos de amostras biológicas. Para detalhes sobre a PCR, ver *Genética na Escola*. V 14, n.2 (2019). Neste caso, o gene *IS6110* é o alvo da reação, sendo multiplicado de forma exponencial e detectado em amostras biológicas de pacientes com TB. A PCR requer ciclos térmicos distintos para a desnaturação, hibridação dos iniciadores (*primers*) e extensão do DNA. Existem variações dessa técnica. Além da PCR convencional, de natureza qualitativa, com análise do produto por eletroforese e/ou sequenciamento, existe também a PCR em tempo real (qPCR ou PCR quantitativa), que permite a avaliação dinâmica da quantidade do DNA amplificado durante a reação (Figura 4A).

O **LAMP** (do inglês, *Loop-mediated isothermal amplification*, veja o BOX 1) representa uma alternativa simples de diagnóstico, também específica e de baixo custo. No método LAMP, estruturas em alça se formam devido ao modo como a síntese do DNA ocorre. Durante o processo, as fitas recém-formadas passam a conter sequências complementares em sua própria estrutura. Quando essas fitas são liberadas da molécula original, elas tendem a se dobrar sobre si mesmas, pois suas extremidades complementares se pareiam

espontaneamente, resultando na formação de estruturas em alça (*loops*). Essas estruturas em alça facilitam a amplificação porque funcionam como moldes altamente eficientes para novas sínteses de DNA. Cada alça cria múltiplos pontos de início para a atividade da DNA polimerase, permitindo que a síntese ocorra de forma contínua e simultânea. Como resultado, o DNA-alvo é amplificado de maneira rápida e eficiente, mesmo em temperatura constante (65°C), o que explica a alta velocidade e sensibilidade do método LAMP. Essa característica elimina a necessidade de equipamentos complexos, como os termocicladores, geralmente utilizados na PCR. Por sua simplicidade e rapidez, o LAMP é especialmente indicado para o diagnóstico da tuberculose em locais com poucos recursos e é atualmente recomendado pela Organização Mundial da Saúde (OMS) (Figura 4B). A diferença principal entre o LAMP e a PCR consiste no fato de que a PCR requer variação de temperatura para desnaturação, hibridação de iniciadores (*primers*) e extensão do DNA, enquanto o LAMP promove a amplificação em temperatura constante, utilizando *primers* que induzem a formação de estruturas em alça, permitindo síntese contínua do DNA sem necessidade de variação de temperatura.

Já o TRM-TB (teste rápido molecular para tuberculose, ou Xpert MTB/RIF) é uma plataforma automatizada baseada em **PCR multiplex** em tempo real. O sistema possibilita tanto a extração do DNA na preparação da amostra biológica quanto a amplificação e detecção do Mtb em um único processo, fornecendo resultados em cerca de 90 minutos (Figura 4C). No Xpert MTB/RIF, duas sequências são o alvo para o diagnóstico de TB, ambas de transposons: o *IS6110* e o *IS1081*. Além da identificação do patógeno, o teste também detecta mutações associadas à resistência à rifampicina, principal fármaco utilizado no tratamento da TB. Apesar das limitações relacionadas ao custo elevado e à restrição de uso a certos tipos de amostras, excluindo o sangue, o TRM-TB é considerado, pela OMS, o padrão-ouro entre os testes moleculares para o diagnóstico de TB.

LAMP - É uma técnica de amplificação do DNA, que ocorre em uma temperatura constante (~65°C), sem a necessidade de um termociclador como na PCR. O DNA amplificado no processo pode ser detectado a partir da mudança de coloração, turbidez ou fluorescência.

PCR multiplex - Trata-se de um tipo de PCR que permite, em uma única reação, a amplificação simultânea de diversas sequências alvo, porque na reação são utilizados mais de um par de iniciadores (*primers*).

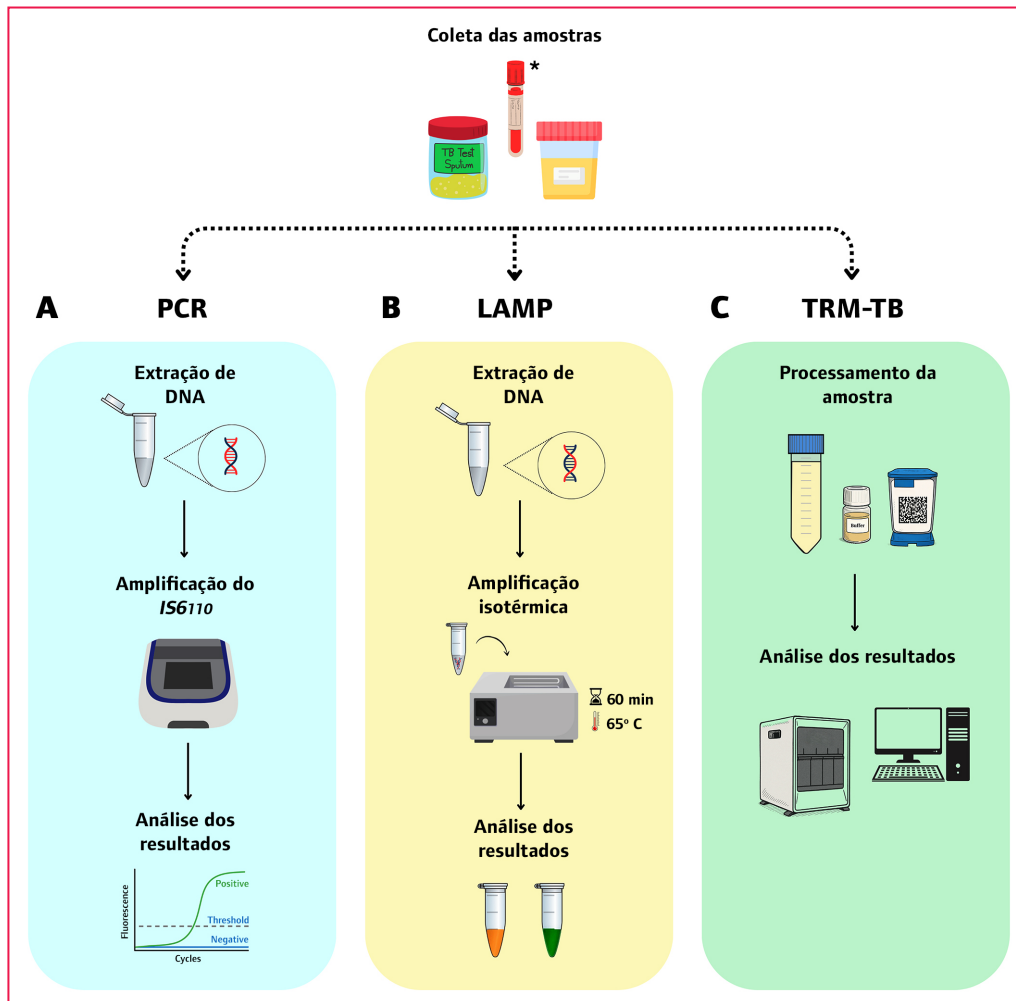


Figura 4. Principais técnicas moleculares utilizadas para a detecção do *Mycobacterium tuberculosis* (Mtb) a partir da amplificação do *IS6110*. (A) Para o diagnóstico por PCR, inicialmente realiza-se a extração do DNA genômico. Em seguida, ocorre a amplificação do *IS6110* e, por fim, análise de resultados a partir da quantificação dinâmica do DNA amplificado durante a reação. (B) O LAMP também é iniciado com a extração de DNA e o material genômico é submetido à **amplificação isotérmica** em banho-maria, por 60 minutos, a 65°C. A análise dos resultados baseia-se na colorimetria, sendo a cor verde indicativa de resultado positivo e a cor laranja, de resultado negativo. (C) No caso do TRM-TB, as amostras biológicas de urina e escarro são processadas, misturadas com um reagente (*buffer*) e posteriormente adicionadas ao cartucho, que é inserido no aparelho para as reações de PCR e análise dos resultados. O asterisco (*) na amostra de sangue indica que esse tipo de amostra biológica não é processada pelo TRM-TB.

O que aprendemos?

Diante do exposto, foi possível compreender de forma mais aprofundada aspectos conceituais, estruturais e mecânicos do gene *IS6110*. Sua classificação como EGM do tipo transposon possibilita mudanças fenotípicas significativas relacionadas à evolução, patogenicidade e sobrevivência do Mtb, agente etiológico da TB. Ademais, ficou evidenciada sua importância laboratorial para o diagnóstico molecular da doença, sendo o *IS6110* a principal sequência alvo dos métodos moleculares atualmente utilizados. Estudos que elucidam os mecanismos e as consequências da ação desse gene são indispensáveis para compreender seu papel no ciclo de vida do patógeno e como ele pode influenciar a infecção no hospedeiro – especialmente os seres humanos.

Para saber mais

GAGNEUX, S. Ecology and evolution of *Mycobacterium tuberculosis*. *Nature Reviews Microbiology*, v. 16, n. 4, p. 202–213, 19 fev. 2018.

GONZALO-ASENSIO, J. et al. New insights into the transposition mechanisms of *IS6110* and its dynamic distribution between *Mycobacterium tuberculosis* Complex lineages. *PLOS Genetics*, v. 14, n. 4, p. e1007282, 12 abr. 2018.

MCEVOY, C. R. E. et al. The role of *IS6110* in the evolution of *Mycobacterium tuberculosis*. *Tuberculosis*, v. 87, n. 5, p. 393–404, 1 set. 2007.

MUKHERJEE, S. et al. Evolution of tuberculosis diagnostics: From molecular strategies to nanodiagnosics. *Tuberculosis*, v. 140, p. 102340, 1 maio 2023.

WHO. *GLOBAL TUBERCULOSIS REPORT 2024*. Geneva: World Health Organization, 2024.

Amplificação isotérmica - É um método de multiplicação de moléculas do DNA que ocorre em temperatura constante, sem a necessidade de aquecer e resfriar a amostra várias vezes.

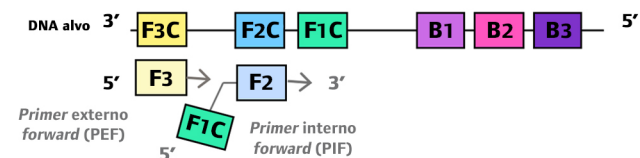
Box 1.

Amplificação isotérmica mediada por *loop*.

A amplificação isotérmica mediada por *loop* (do inglês, *Loop-mediated isothermal amplification - LAMP*) é uma técnica molecular de amplificação de DNA realizada em temperatura constante. O método utiliza um conjunto de *primers* específicos - geralmente quatro principais (F3, B3, PIF e PIB) - que reconhecem múltiplas regiões do DNA-alvo, conferindo alta especificidade. Durante a reação, são formadas estruturas em alça (*loops*) que permitem a amplificação exponencial do material genético. O DNA amplificado no LAMP pode ser detectado por eletroforese em gel, pela visualização de bandas de DNA; por fluorescência, utilizando corantes intercalantes que se ligam ao DNA dupla fita e emitem fluorescência em contato com a luz visível ou ultravioleta ou por métodos colorimétricos, nos quais indicadores químicos mudam de cor durante a reação, permitindo a identificação do produto amplificado a olho nu. Abaixo, as etapas do LAMP:

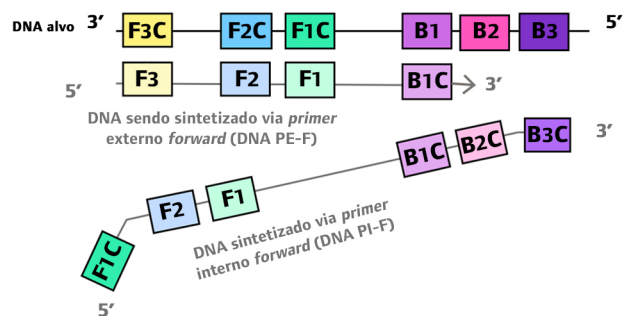
1. Hibridação do primer interno forward

O *primer interno forward* (PIF), composto por sequências das regiões F2 e F1C, se liga à região complementar F2C da fita molde e inicia a síntese de uma fita complementar pela DNA polimerase.



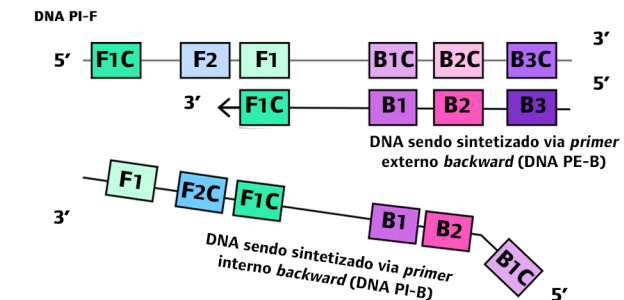
2. Extensão e deslocamento via primer externo

O *primer externo* F3 se hibrida à região F3C e promove a síntese de DNA (DNA PE-F), deslocando a fita recém-sintetizada iniciada pelo *primer interno forward*.



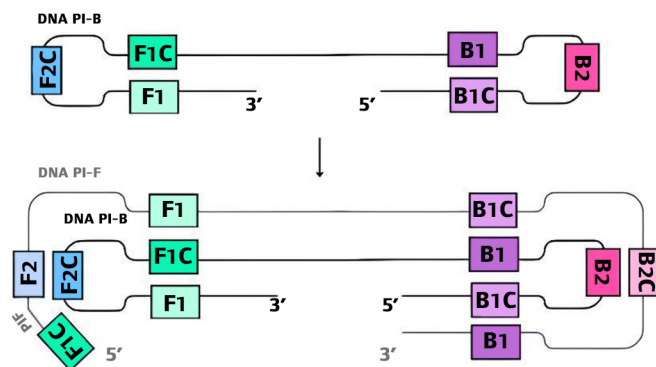
3. Ação do primer interno reverse

A fita deslocada forma uma estrutura complementar (DNA PI-F) contendo as regiões B1C, B2C e B3C. Então, ocorre o mesmo processo da etapa 2, porém no sentido *backward*: o *primer externo backward* (PEB) desloca a fita sintetizada pelo PIF. Dessa forma, ocorre a formação da fita DNA PI-B com as extremidades complementares.



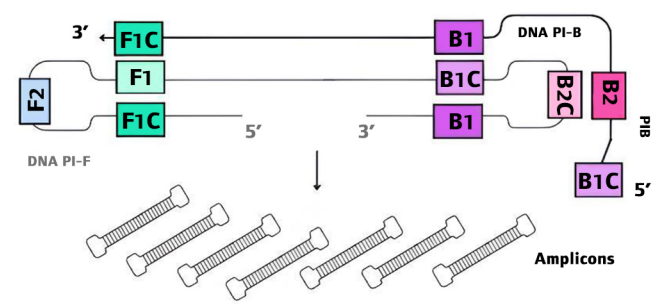
4. Formação do loop (alça) e hibridação PIF

Devido à complementaridade entre as regiões F1 e F1C (B1 e B1C), a molécula de DNA dobra-se, formando uma estrutura em alça (*loop*) nas extremidades 3' e 5', que servem como sítio para novas reações de amplificação a partir da hibridação dos *primers* internos. Nessa etapa, a partir da ligação do PIF, ocorre a formação de uma fita DNA PI-F.



5. Sítio de ligação permite a amplificação exponencial

A partir da complementaridade entre B1 e B1C ocorre a hibridação na extremidade 3' do PIB, ocorrendo o mesmo processo da etapa 4 e a formação de uma fita DNA PI-B. Como ambas as extremidades são complementares, isso permite que os *primers* PIF e PIB liguem-se a esses sítios de ligação, permitindo uma reação de amplificação exponencial.



6. Métodos de detecção

O DNA amplificado pode ser identificado tanto em gel de eletroforese (a), emissão de fluorescência (b) ou detecção colorimétrica (c), permitindo diferentes estratégias de visualização do resultado.

