ESCOLA BAHIANA DE MEDICINA E SAÚDE PUBLICA

PROGRAMA DE PÓS-GRADUAÇÃO EM MEDICINA E SAÚDE HUMANA

CAIAN LEAL DE AZEVEDO VINHAES

IMPACTO DA DISGLICEMIA NA ATIVAÇÃO INFLAMATÓRIA DE PACIENTES COM TUBERCULOSE

TESE DE DOUTORADO

SALVADOR – BAHIA

2024

CAIAN LEAL DE AZEVEDO VINHAES

IMPACTO DA DISGLICEMIA NA ATIVAÇÃO INFLAMATÓRIA DE PACIENTES COM TUBERCULOSE

Tese apresentada ao Programa de Pós-Graduação em Medicina e Saúde Humana da Escola de Medicina e Saúde Pública, como requisito parcial para obtenção do título de Doutor em Medicina e Saúde Humana.

Orientador: Prof. Dr. Bruno de Bezerril Andrade

Co-Orientador: Prof. Dr. Artur Trancoso Lopo de Queiroz

SALVADOR – BAHIA

2024

CAIAN LEAL DE AZEVEDO VINHAES

“IMPACTO DA DISGLICEMIA NA ATIVAÇÃO INFLAMATÓRIA DE PACIENTES COM TUBERCULOSE”

Tese apresentada à Escola Bahiana de Medicina e Saúde Pública, como requisito parcial para a obtenção do Título de Doutor em Medicina e Saúde Humana.

Salvador, 1 de dezembro de 2023.

BANCA EXAMINADORA

Dr. Moreno Magalhães de Souza Rodrigues
Doutor em Parasitologia
Fundação Oswaldo Cruz - Rondônia, FIOCRUZ-RO

Dr. Marcelo Nóbrega Litvoc
Doutor em Medicina
Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, HCFMUSP

Dr. Eduardo Martins Netto
Doutor em Medicina e Saúde
Universidade Federal da Bahia, UFBA

Prof.ª Dra. Maria Fernanda Rios Grassi
Doutora em Imunologia
Escola Bahiana de Medicina e Saúde Pública, EBMSP

Prof.ª Dra. Ana Marice Teixeira Ladeia
Doutora em Medicina e Saúde
Escola Bahiana de Medicina e Saúde Pública, EBMSP

Este documento foi assinado digitalmente por Alson Carlos De Souza Fernandes. Este documento foi assinado eletronicamente por Maria Fernanda Grassi, Moreno Magalhaes de Souza Rodrigues, Marcelo Nóbrega Litvoc, Ana Marice Teixeira Ladeia e Eduardo Martins Netto.
Para verificar as assinaturas vá ao site https://bahianaeducacao.portaldeassinaturas.com.br e utilize o código C5F7-22E0-7BC5-BD86.

Aos meus pais Tatiana e Sérgio

Às minhas avós Lúcia e Magali

Ao meu avô Tong

Aos meus tios e tias e

Toda minha família.

AGRADECIMENTOS

Ao meu orientador Bruno Bezerril e ao meu coorientador Artur Queiroz por todo apoio desde os anos de iniciação científica e por acreditarem no meu potencial e fomentarem o meu sonho de me tornar cientista.

Aos preceptores da residência médica que, de maneira indireta, me mostraram diariamente a importância da dedicação e aprimoramento científico.

A Elze Leite por toda ajuda e apoio administrativo.

Aos meus amigos, que desde sempre estiveram ao meu lado nas batalhas da vida.

A todo grupo MONSTER que mais me acolhe como casa.

Ao RePORT Brasil e RePORT internacional pela viabilização dos dados.

A todos aqueles que de alguma forma contribuíram para o meu desenvolvimento científico.

Muito obrigado.

RESUMO

A Tuberculose (TB) é uma doença infectocontagiosa de relevância global para saúde pública, principalmente em países subdesenvolvidos e em desenvolvimento. É uma das principais causas de morte por patógeno único, atrás apenas do SARS-CoV-2, somando 1.4 milhões de óbitos em 2021 em pessoas não portadoras de HIV e um total de 10.6 milhões de infectados. Inicialmente um patógeno pulmonar, o *Mycobacterium tuberculosis* (Mtb) pode se disseminar por outros órgãos e tecidos, além de apresentar diferentes níveis de gravidade clínica na dependência da interação entre a resposta imune do hospedeiro e o Mtb, sendo, portanto, uma imunopatogênese, onde a ativação inflamatória crônica do hospedeiro definirá o espectro clínico após o contato com o patógeno, e o dano tecidual determinado pela ativação inflamatória do paciente na tentativa de contenção bacilar. Com trabalhos iniciados ainda no período de iniciação científica, eu e o meu grupo desenvolvemos diversos projetos que visam o melhor entendimento de biomarcadores moleculares associados nesse processo que definirá o espectro da doença em seus portadores. Por tratar-se de imunopatogenia, processos que afetem os mecanismos homeostáticos de ativação inflamatória podem interferir na montagem de uma resposta adequada e capaz de conter a proliferação bacilar. Nesse contexto, ganha destaque o diabetes melitus, doença metabólica crônica com incidência crescente, principalmente em consequência dos maus hábitos de vida e obesidade, com aproximadamente 422 milhões portadores ao redor do mundo e, responsável por cerca de 1.5 milhões de mortes anualmente. Diversos estudos têm avaliado a influência do diabetes na TB, porém o impacto da disglicemia na ativação inflamatória de portadores de TB permanece incerto. Essa tese utiliza dados do RePORT Brasil e Índia, somados a dados públicos do consórcio TANDEM, de pacientes com TB seguidos por 6 meses durante a terapia anti-tubercular. Amostras de fluidos biológicos foram coletadas e analisadas. Os dados utilizados foram previamente coletados e obtidos através da plataforma Luminex e transcriptoma do sangue periférico e dosagem de eicosanoides na urina. Utilizando análises multidimensionais e de inteligência artificial do tipo aprendizado de máquina, encontramos que as condições clínicas TB, DM e TBDM compartilham poucos genes diferencialmente expressos na comparação entre os países do estudo, pacientes com TBDM apresentam maior grau de perturbação molecular, a influência dos genes diferencialmente expressos nas vias biológicas é distinta, não apresentando um mesmo padrão entre os países, exceto as vias de degranulação de neutrófilos, peptídeos antimicrobianos e a via de organização da matriz extracelular, além de identificar a correlação entre níveis de HbA1c e vias biológicas. Além disso, estudamos as alterações multi-ômicas no contexto TBDM, além de avaliar o impacto de RNA não codificadores no processo. Em conjunto, os manuscritos dessa tese identificam alguns aspectos do impacto da disglicemia na ativação inflamatória de pacientes com TBDM, que poderão futuramente ser utilizados como alvo para estratégias terapêuticas, além de trazer marcadores multi-ômicos que poderão embasar

novos estudos mecanísticos na busca de marcadores para definição de prognóstico e evolução da TB em portadores de DM.

Palavras-chave: Tuberculose, Diabetes mellitus, ativação inflamatória, transcriptomica, multi-ômica, RNA não codificante.

ABSTRACT

Tuberculosis (TB) is an infectious disease of global relevance to public health, mainly in low and middle incomes countries. It is one of the leading causes of death by single pathogen, second only to SARS-CoV-2, adding up to 1.4 million deaths in 2021 in people without HIV and a total of 10.6 million infected people. Initially a pulmonary pathogen, Mycobacterium tuberculosis (Mtb) can spread to other organs and tissues, presenting different levels of clinical severity depending on the interaction between the host's immune response and Mtb, thus being an immunopathogenesis, where the chronic inflammatory activation of the host will define the clinical spectrum after contact with the pathogen, and the tissue damage being determined by the inflammatory activation of the patient in an attempt to contain the bacillary. With work started still in the period of scientific initiation, my group and I developed several projects aimed at a better understanding of molecular biomarkers associated in this process that will define the spectrum of disease. Because it is an immunopathogenesis, processes that affect the homeostatic mechanisms of inflammatory activation can interfere with the assembly of an adequate response capable of containing bacillary proliferation. In this context, diabetes mellitus, a chronic metabolic disease with increasing incidence, stands out, mainly because of poor lifestyle habits and obesity, with approximately 422 million carriers around the world, and responsible for about 1.5 million deaths annually. Several studies have evaluated the influence of diabetes on TB, but the impact of dysglycemia on the inflammatory activation of TB patients remains uncertain. This thesis uses data from RePORT Brazil and India, added to public data from the TANDEM consortium, from TB patients followed for 6 months during anti-tubercular therapy. Samples of biological fluids were collected and analyzed. The data used were previously collected and obtained through Luminex and peripheral blood transcriptome and measurement of eicosanoids in the urine. Using multidimensional and artificial intelligence analyses, we found that the clinical conditions TB, DM and TBDM share few differentially expressed genes in the comparison between the countries of the study, patients with TBDM have a higher degree of molecular disturbance, the influence of differentially expressed genes on biological pathways is distinct, not showing the same pattern between countries, except for neutrophil degranulation pathways, antimicrobial peptides and the extracellular matrix organization pathway, in addition to identifying the correlation between HbA1c levels and biological pathways. Together, the manuscripts of this thesis identify some aspects of the impact of dysglycemia on the inflammatory activation of patients with TBDM, which may in the future be used as a target for therapeutic strategies, in addition to bringing multi-omic markers that may support new mechanistic studies in the search for markers to defining the prognosis and evolution of TB in patients with DM.

Keywords: Tuberculosis, diabetes mellitus, inflamatory activation, transcriptome, multi-omics, non-coding RNA.

LISTA DE TABELAS

Tabela 1**:** Função fisiológica das principais células envolvidas na imunidade inata contra o *Mycobacterium* e as consequências da disglicemia em sua ativação. Página 21.

LISTA DE ILUSTRAÇÕES

Figura 1: Efeito da disglicemia na dinâmica das células T helper e o seu impacto na fisiopatologia da TB. Página 22.

Figura 2: Entendendo o Grau de Perturbação Inflamatória (GPI) ou “Molecular Degree of Perturbation (MDP). Página 28.

LISTA DE ABREVIATURAS

ADA	American Diabetes Association
AI	Inteligência artificial
AMP	Peptídeos antimicrobianos
ATT	Terapia Anti-Tubercular
CD	Células dendríticas
DEG	Genes diferencialmente expressos
DM	Diabetes melitus
EPTB	Tuberculose extrapulmonar
FDR	False Discovery Ratio
HC	Controles saudáveis
HIV	Vírus da imunodeficiência humana
IFN	Interferon
IL	Interleucina
IMC	Índice de Massa Corporal
MDP	Grau de perturbação molecular
miRNA	MicroRNA
MMP	Metaloproteínase de matriz
MSTD	Molecular Signatures of Tuberculosis-Diabetes Interaction
Mtb	Mycobacterium tuberculosis
OMS	Organização Mundial de Saúde
PTB	Tuberculose pulmonar
RNAseq	Sequenciamento de RNA
TB	Tuberculose
TIMP	Inibidor tecidual de metaloproteinase
Th	Linfocitos T helper
HbA1c	Hemoglobina glicosilada

SUMÁRIO

1. **INTRODUÇÃO**...12
2. **OBJETIVO**...16
2.1 **Geral**...16
2.2 **Específicos**...16
3. **REVISÃO DA LITERATURA**...17
4. **METÓDOS**...23
4.1 **Popoulações dos estudos**...23
4.2 **Classificação e Amostragem Sanguínea**...24
4.3 **Sequenciamento de RNA**...24
4.4 **Análises estatísticas**...24
5. **ÍNDICE DOS ARTIGOS CIENTÍFICOS**...26
6. **ARTIGOS CIENTÍFICOS**...27
6.1 **Artigo I**... 27
6.2 **Artigo II**...52
6.3 **Artigo III**...74
6.4 **Artigo IV**...94
7. **DISCUSSÃO**...111
8. **CONCLUSÃO**...118
Referências...120
Perspectiva Histórica da Tese...127
Anexo I: Produção científica no Doutorado...128
Anexo II: Rede de colaboração do estudante...140

1 INTRODUÇÃO

A Tuberculose (TB) é uma doença infectocontagiosa milenar, que leva à impactos econômicos e sociais catastróficos em todo mundo, principalmente nos países subdesenvolvidos e em desenvolvimento, como Brasil, Índia, África do Sul e China. Trata-se de uma das principais causas de morte por patógeno único, atrás apenas do SARS-CoV-2, somando 1.4 milhões de óbitos em 2021 em pessoas não portadoras de HIV e um total de 10.6 milhões de infectados (1). Ademais, durante a pandemia de COVID-19, houve grande dificuldade no processo diagnóstico e, principalmente, falhas de notificação, podendo esse número ser ainda maior. Adicionalmente, emerge a preocupação com as cepas resistentes às principais drogas utilizadas no esquema terapêutico, colocando a TB dentre as doenças de preocupação global pela Organização Mundial de Saúde (OMS).

Inicialmente um patógeno pulmonar, *o Mycobacterium tuberculosis* (Mtb) pode se disseminar por outros órgãos e tecidos, resultando em diferentes níveis de gravidade clínica na dependência da interação entre a resposta imune do hospedeiro e o Mtb (2). Por essa razão, a TB é entendida e estudada como uma imunopatologia, um processo de ativação inflamatória crônica do hospedeiro que definirá o espectro clínico após o contato com o patógeno, sendo o dano tecidual determinado pela ativação inflamatória do paciente na tentativa de contenção bacilar. Nosso grupo vem se empenhado no entendimento das vias imunopatogênicas da TB, seja isoladamente ou em associação com HIV (3-9). Em trabalho publicado ainda como estudante de iniciação científica, pudemos avaliar diferenças na ativação inflamatória de acordo com o sítio de infecção por Mtb, TB pulmonar (PTB) ou extrapulmonar (EPTB) (10). Nossos achados revelaram, como esperado, um perfil de ativação similar, entre PTB e EPTB. No entanto, evidenciamos que a terapia anti-tubercular (ATT) induz alterações inflamatórias mais precoces em PTB quando comparado com EPTB (10). Digno de nota, concluímos que as alterações em concentrações plasmáticas de proteínas inflamatórias e mediadores lipídicos permanecem mesmo após o término da ATT (10). Em outros dois trabalhos, visamos avaliar se a nacionalidade (11) e a idade (12) influenciam tal ativação. Nossos achados revelaram perfis distintos comparando populações da China e Índia utilizando o grau de perturbação molecular, “molecular degree of perturbation (MDP)” (11). Ainda aplicando MDP, evidenciamos

que a idade se associa a aumento no grau de perturbação molecular (12). Isso sugere que o envelhecimento está intimamente associado a capacidade do hospedeiro infectado de induzir um grau de inflamação mais exacerbado, possivelmente resultando em maior dano tecidual.

Portanto, condições outras que afetem a resposta imunológica, e sua efetividade antimicrobiana assim como a sua regulação para minimizar dano tecidual colateral, podem influenciar na ativação das vias inflamatórias dos pacientes com TB. Nesse contexto destacam-se outras doenças infecciosas, como o HIV, determinantes genéticos (13), ambientais (14), nutricionais (15) e metabólicos como as alterações de glicemias no plasma, denominadas aqui disglicemias.

Disglicemia refere-se às alterações, para mais ou para menos, no considerado nível normal para glicemia plasmática, entre 80-100 mg/dL. A presente tese foca principalmente nas hiperglicemias e diabetes mellitus (DM). O DM é uma doença metabólica crônica de grande relevância mundial, com incidência crescente, principalmente em consequência de hábitos de vida e obesidade. De acordo com a OMS, atualmente 422 milhões de pessoas possuem DM ao redor do mundo, a maioria em países de baixa e média renda, com 1.5 milhões de mortes atribuídas à doença anualmente (16).

Usualmente, os níveis de glicose elevados a longo prazo geram lesões em órgãos alvo, como rins, olhos, coração, nervos periféricos e vasos sanguíneos. No entanto, hiperglicemias sustentadas levam a um processo de inflamação sistêmica desencadeada pela alteração metabólica, com efeitos na regulação imune do portador de DM, que por muitas vezes é caracterizado como imunossupresso ou desregulado (17). Exemplo atual dessa condição, é o DM como fator de risco para desenvolvimento de quadros mais graves de COVID-19 (18). Dessa forma, tratando-se de patologia que compromete os mecanismos de ativação inflamatória e com grande incidência mundial, principalmente nos países de baixa e média renda, que também são os que reúnem maiores incidências de TB, emerge no mundo uma superposição das condições e, a relação TBDM passou a ser amplamente estudada nos últimos anos.

Nesse contexto o DM surge como condição prevalente que contribui com a carga global da TB (19). A superposição dos processos inflamatórios, um mediado por condição infecciosa e, o outro por condição metabólica, vem sendo marcada por

apresentações mais graves de PTB (20), maior transmissibilidade à contatos próximos (21), alterações de mediadores inflamatórios derivados do ácido araquidônico na urina (22) e, a desfechos desfavoráveis em um estudo peruano (23) e, desfechos desfavoráveis e morte no Brasil (24).

No entanto, pouco ainda se sabe a respeito do impacto a nível molecular que a disglicemia exerce em portadores de TB, culminando nas consequências supra escritas. O entendimento dessas vias e mecanismos pode auxiliar tanto no diagnóstico de TB em pacientes com DM, tanto quanto no seguimento e monitoramento de tratamento e definição de prognóstico naqueles portadores das condições. Porém, até pouco tempo atrás, as ferramentas investigativas não eram capazes de extrapolar simultaneamente aspectos clínicos ou, no máximo, bioquímicos, para elucidação de pontos determinantes no manjo clínico das doenças.

Graças às inovações na coleta e armazenamento de dados, associados à Medicina de precisão, surgem inúmeras abordagens que unem determinantes genômicos e multiômicos com informações clínicas, a fim de aprimorar estratégias diagnósticas e prognósticas (25), com impactos diretos na assistência prestada a nível individual e global. Possibilitada pelos avanços descritos, a Inteligência Artificial (AI), surge trazendo entusiasmos ao meio científico, principalmente por viabilizar abordagens baseadas em Predição, Prevenção, Personalização e Participação, conhecido como modelo de Medicina 4P (26). Nesse contexto, diversos modelos vêm sendo construídos e testados utilizando AI, em medicina nuclear (27), oftalmologia (28), anestesiologia (29), neurocirurgia (30), cardiologia (31) e infectologia. Recentemente o nosso grupo pôde contribuir para o desenvolvimento de uma assinatura gênica para diagnóstico de TB em pacientes com HIV (32) e avaliação multi-ômica de marcadores de TB em HIV avançado (33).

Utilizando essas e outras técnicas estatísticas multidimensionais, a presente tese, composta por três manuscritos, visa identificar os fatores moleculares associados à disglicemia que influenciam na ativação inflamatória sistêmica de pessoas com TB. Utilizamos dados previamente coletados do RePORT internacional (34), no Brasil e Índia (35, 36), associados a dados públicos do consórcio TANDEM (37) nos sítios da África do Sul e Romênia, para comparar a expressão gênica de pacientes com TB pulmonar, quantificar e qualificar a resposta inflamatória utilizando grau molecular de

ativação inflamatória (10-12, 38-41) e vias de expressão biológicas, buscar um perfil de expressão gênica que caracterize a superposição TBDM, identificar o efeito das condições clinicas aqui estudadas na expressão das vias biológicas, além de estudar a expressão multiômica de TB e Diabetes, que ao nosso conhecimento é o primeiro estudo a avaliar múltiplas plataformas ômicas, coletadas nos mesmos pacientes e nos mesmos tempo de estudo, na relação TBDM. Adicionalmente, avaliamos o impacto dos RNAs não codificadores no contexto TBDM.

Os resultados obtidos nessa tese de doutorado poderão servir de base molecular para o desenvolvimento futuro de novas técnicas e métodos diagnósticos que utilizem assinaturas gênicas, ferramentas de seguimento terapêutico, além de identificar potenciais marcadores em plataformas multiômicas para desenvolvimento de novas vias de intervenção terapêutica, auxiliando assim o campo de estudos em TB afim de reduzir a carga global da doença.

2 OBJETIVOS

2.1 Geral:

Identificar os fatores moleculares associados à disglicemia que influenciam na ativação inflamatória sistêmica de pessoas com Tuberculose.

2.2 Específicos:

Comparar expressão gênica de pacientes com tuberculose pulmonar e diabetes mellitus (DM) entre os sítios do RePORT;

Quantificar e qualificar a resposta inflamatória de pessoas com TB e TBDM;

Avaliar se há um perfil de expressão genica que caracterize pessoas com Tuberculose e Diabetes mellitus;

Identificar o efeito das condições clínicas (TB, TBDM e DM) na expressão de vias biológicas;

Estabelecer o impacto da disglicemia na ativação das vias biológicas;

Avaliar o efeito da disglicemia nas múltiplas plataformas ômicas;

Estudar o impacto dos RNAs não codificadores na interação TBDM.

3 REVISÃO DE LITERATURA

A interação Tuberculose-Diabetes vem sendo amplamente estudada nos últimos anos, principalmente pela crescente incidência de ambas as doenças que têm gerado tal superposição de condições com relevantes impactos na saúde pública. É sabido que a presença de DM se impõe como fator de risco para TB e, que está associada a desfechos desfavoráveis. No entanto, as interações fisiopatológicas que justificam tais efeitos adversos na interação TBDM permanecem pouco elucidadas. Com os avanços na Medicina moderna, diversos estudos vêm sendo realizados a fim de delinear perfis inflamatórios e de expressão gênica que caracterizem as condições TB, DM e TBDM (35, 37, 42) a fim de buscar ferramentas alternativas para diagnóstico, seguimento terapêutico e definição de prognóstico.

Parte dos mecanismos imunológicos que contribuem para elevada susceptibilidade de pessoas diabéticas à TB se associam a deficiência no reconhecimento micobacteriano, atividade fagocítica e ativação de resposta imune celular, culminando em inadequada produção de citocinas e quimiocinas, sumarizados em uma revisão de literatura publicada em 2019 (43). O status de hiperglicemia crônica leva a diversas alterações na ativação imunológica que provavelmente estão associados à maior susceptibilidade e desfechos desfavoráveis, haja vista que para o controle da proliferação micobacteriana o hospedeiro precisa estruturar uma resposta eficaz. Estudos prévios demonstram diferenças na imunidade inata, primeira linha de defesa contra patógenos, entre DM e não DM. Parece haver um prejuízo na função de neutrófilos, macrófagos, células dendríticas (CD), natural killer e de outros aspectos da ativação inata (44-46).

As células dendríticas são uma das células apresentadoras de antígeno, que por sua vez tem a função de mediar a intercessão entre imunidade inata e adaptativa (47). Migrando para os linfonodos, as CD desempenham papel fundamental na ativação dos linfócitos T na TB (48). Estudos prévios demonstram redução significativa na frequência de CD em TBDM, acreditando-se haver correlação entre o número de CD e níveis de HbA1c (48, 49).

Outro tipo celular marcadamente afetado pela hiperglicemia são os neutrófilos. Neutrófilos se caracterizam não só por possuírem a função de migrar para o tecido infectado e matar patógenos, mas também, secretando citocinas e quimiocinas, estão

associados a uma série de reações que culminam na ativação e recrutamento de outros tipos celulares (49, 50). Indivíduos hiperglicêmicos apresentam neutrófilos com maior expressão de adesinas e integrinas, redução da quimiotaxia, defeitos em fagocitose, e atividade microbicida comprometida quando comparados a controles normoglicêmicos (49). As funções fisiológicas das células da imunidade inata e o consequente impacto da disglicemia nas suas atividades estão sumarizadas na Tabela 1.

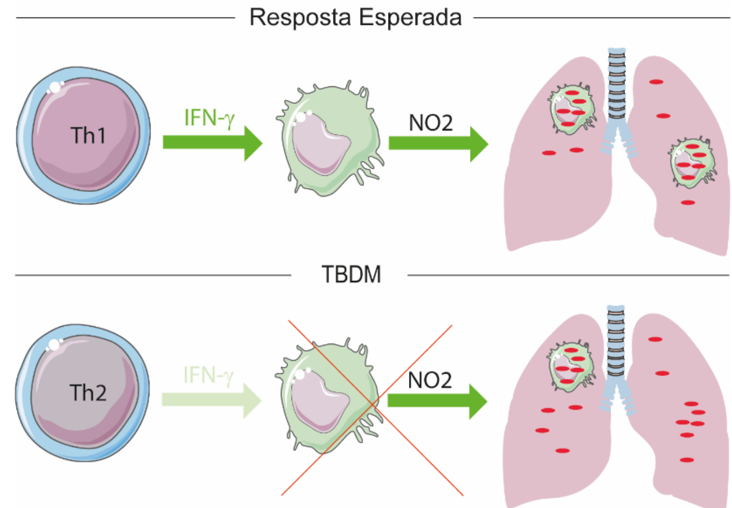
Tabela 1: Função fisiológica das principais células envolvidas na imunidade inata contra o *Mycobacterium* e as consequências da disglicemia em sua ativação.

Célula	Função Fisiológica	Impacto na disglicemia
Macrófagos	Fagocitam o Mtb; Ativam fusão do fagolisossomo; -Apoptose; -Redução da carga bacilar; -Ativação de outras vias imunológicas;	-Impacto no encontro dos macrófagos alveolares com o Mtb; -Atraso na transição inata -> adaptativa; -Aumento logarítimico na carga bacilar.
Células Dendríticas	- Apresentadoras de antígeno; -Intercessão imunidade Inata->adaptativa; -Papel na ativação dos linfócitos T.	-Redução na frequência de CD; -Correlação entre HbA1c X CD.
Neutrófilos	-Morte de patógenos; -Ativação e recrutamento de outras células.	-Maior expressão de adesinas e integrinas; -Quimiotaxia reduzida; -Defeitos na fagocitose.

Além das deficiências na imunidade inata que levam a uma ineficaz ativação da resposta adaptativa, a interação TBDM com elevados níveis glicêmicos afeta também diretamente os mecanismos adaptativos do hospedeiro. Tratando-se de defesa contra Mtb, a principal linha de imunidade adaptativa é a celular (43). Células T helper (Th) 1, estimulando a produção de interferon (IFN)- γ potencializam as vias dependentes de óxido nítrico nos macrófagos (51), além de estarem associadas a uma séria de outras cascatas de ativação celular. Postula-se que o prejuízo na ação das células Th no contexto da hiperglicemia, pode ser um dos maiores responsáveis pela maior susceptibilidade à TB (51). Por outro lado, postula-se ainda não só deficiência na qualidade da ativação, mas também na intensidade e frequência da resposta (52), com redução de Th1 e Th17, respostas marcadamente pró-inflamatórias, e aumento

no braço Th2, com característica menos inflamatória que antagoniza as duas primeiras (47, 51).

Figura 1. Efeito da disglicemia na dinâmica das células T helper e o seu impacto na fisiopatologia da TB.



No entanto, com os avanços no campo da imunologia, os estudos vêm evoluindo e ultrapassando o entendimento da ativação imunológica apenas no que tange imunidade inata e adaptativa. Nesse contexto, o estudo das ômicas, muitas vezes associados a aplicações de IA, emerge como chave para respostas ainda não obtidas no meio científico, contexto no qual se inserem os objetivos do presente trabalho. Apesar de algumas limitações no uso de transcriptomas de RNA mensageiro, diversas assinaturas genicas vêm sendo estudadas como ferramentas para diagnóstico, prognóstico e monitoramento do tratamento da infecção por Mtb (53-55).

Em um recente coorte indiano, foi realizada análise integrativa de citocinas plasmáticas e expressão genica de sangue total afim de comparar a ativação inflamatória de pacientes com TB, DM, TBDM (56). Os resultados revelam um aumento na expressão de biomarcadores inflamatórios no grupo TBDM. Adicionalmente, foram avaliados genes diferencialmente expressos (DEGs) e encontrados 993 DEGs comparando TBDM, TB e DM com o grupo controle (56). Entre esses, 455 genes foram comummente encontrados em TB e TBDM, alguns deles previamente reportados em assinaturas de TB. O estudo não encontrou assinatura genica no sangue periférico que distinguisse TBDM de TB.

Em um pequeno estudo Chines, incluindo 9 participantes (3 controles, 3 TB, 3 DM e 3 TBDM), foi utilizado sequenciamento de RNA (RNAseq) para identificar alterações em vias biológicas especificamente em TBDM (57). Em uma análise similar ao que foi avaliado em um dos trabalhos da tese, os autores encontraram um maior número de DEGs na comparação TBDM vs. Controle, sendo 582 dos 1516 positivamente regulados “up regulated” e 934 negativamente regulados “down regulated”. Além disso, uma análise de enriquecimento de vias foi realizada com os 952 DEGs identificados apenas nos pacientes com TBDM. Entre as vias biológicas com maior número de DEGs expressos estavam a via do ciclo celular, homeostase e as vias relacionadas a processos imunológicos. Algumas das análises realizadas pelo nosso grupo no presente projeto, incluem a identificação do efeito das condições TB, DM e TBDM na expressão de vias biológicas, além de tentar estabelecer o impacto da disglicemia na ativação dessas vias.

Em um recente coorte multinacional, com amostras da África do Sul, Romênia, Indonésia e Peru, foi novamente utilizada abordagem de RNAseq para identificar vias imunológicas alteradas em TBDM comparados com TB (42). Além disso, o trabalho do consórcio TANDEM avaliou o impacto de níveis intermediários de HbA1c na TB, a fim de entender como o DM aumenta a susceptibilidade à TB (42). O estudo conclui que indivíduos TBDM apresentam alterações transcriptômicas com maior número de DEGs positivamente regulados em comparação ao grupo controle. Dentre esses, genes associados a citocinas pró inflamatórias, como interleucina (IL)-1β, IL-15, IL-18 e IL-10 (42), marcador regulatório das vias inflamatórias. Curiosamente, foi encontrada redução nas vias de interferon do tipo I, sugerindo um inexplicado desacoplamento no fenótipo na associação TBDM (42). Achado similar foi encontrado em indivíduos com níveis intermediários de HbA1c, sugerindo potencial susceptibilidade à TB (42). No entanto, outros estudos se fazem necessário para melhora avaliação do impacto da disglicemia na ativação inflamatória de pacientes com TB.

Em outro trabalho utilizando transcriptoma, pesquisadores do consórcio TANDEM avaliaram amostras da África do Sul e Indonésia em um coorte que acompanhou pacientes com TB com e sem DM por 12 meses. Utilizando estratégia de inteligência artificial similar às aplicadas nos artigos que compõem essa tese, os autores buscaram uma assinatura transcrptômica que pudesse predizer desfecho ao tratamento da TB

(58). Como esperado, foi encontrada uma assinatura transcriptômica distinta no sangue periférico daqueles com TBDM (58). Além disso, duas assinaturas, uma composta por 8 genes no início do tratamento e, outra composta por 22 genes 2 semanas após o início da terapia, distinguiram pacientes considerados com desfecho favorável em relação aos que desenvolveram desfechos desfavoráveis. Os resultados obtidos utilizando análise de curva ROC revelaram acurácia de 81.5% e 83.4%, respectivamente, na predição do desfecho do tratamento de TB (58).

Apesar do grande e relevante aumento nos trabalhos utilizando ômicas no campo da TB, ao nosso conhecimento nenhum projeto avaliou mais de um estrato ômico coletado ao mesmo tempo e dos mesmos pacientes com TB e DM. Estudos multi-ômicos vem sendo publicados no campo de doenças parasitárias como a Leishmaniose (59, 60) e na coinfecção TB-HIV (33). Integrando citocinas plasmáticas, transcriptoma e lipidômica, nosso grupo participou de um projeto que visou uma assinatura de predição à de falha terapêutica em pessoas de área endêmica para Leishmaniose cutânea (59). Mais uma vez, foi aplicada inteligência artificial e foi encontrada uma assinatura biológica incluindo citocinas plasmáticas e lipídios para falha terapêutica (59).

Recentemente nosso grupo participou de um estudo inovador no campo da TB, conduzido com pacientes HIV severamente imunossupressos, com CD4<50 células/µL (33). Foram avaliados microRNAs (miRNAs), metabólitos e citocinas plasmáticas a fim de avaliar associação dos extratos ômicos estudados com novos diagnósticos de TB nos participantes HIV. Após aplicar uma estratégia de IA, árvore de decisão, o grupo encontrou dois marcadores com potencial para classificar TB incidente, com acurácia de 96% na análise de curva ROC.

Algumas barreiras são iminentes na produção científica envolvendo múltiplos extratos ômicos. O primeiro deles, financiamento. As estratégias de dosagem para cada um dos extratos, sejam eles citocinas, RNA, lipídios, microrganismos, é distinta e envolve etapas diferentes, muitas vezes em laboratórios alocados em cidades diferentes, o que onera muito o processo. Em segundo lugar, a coleta das amostras, que muitas vezes ocorre em fluídos corporais distintos (como sangue, urina, fezes) e o encaminhamento para análise que impõe grande dificuldade operacional ao processo. Aqui, utilizando o biorrepositório do RePORT Brasil, pudemos utilizar uma abordagem

inovadora no campo da imunologia e da TB e trazemos o que, ao nosso conhecimento, é o primeiro estudo multi-ômico em TB, a fim de avaliar o efeito da disglicemia em três plataformas ômicas, citocinas plasmáticas, trasncriptoma e eicosanoides dosados na urina de pacientes com PTB.

4 MÉTODOS

4.1 População dos estudos

A presente tese fundamenta-se em três manuscritos utilizando dados previamente coletados do RePORT Brasil (21, 22, 24) e RePORT Índia (35, 36) e dados públicos do consórcio TANDEM (42) dos sítios África do Sul e Romênia, para o primeiro manuscrito, nomeado Assinaturas Moleculares da Interação Tuberculose-Diabetes (Molecular Signatures of Tuberculosis-Diabetes Interaction – MSTDI). O RePORT trata-se de um consórcio internacional que visa erradicação da TB no mundo, com sítios de pesquisa clínica em 7 países. Termo de consentimento escrito foi obtido de todos os participantes do estudo e os protocolos de estudo aprovados pelos respectivos Comitês de Ética do RePORT Brasil e RePORT Índia.

Seguindo o protocolo do RePORT Internacional (34) foram incluídos pacientes maiores de 18 anos, diagnosticados com TB pulmonar (ou sem o diagnóstico para o grupo controle). Foram excluídos aqueles com TB resistente, retratamento para TB, gestantes, em uso de medicações imunossupressoras e HIV.

Os pacientes com TB pulmonar ativa confirmados por culturas incluídos no estudo foram seguidos no início do tratamento e nos meses 2 e 6, quando é finalizada a terapia de TB pulmonar. Durante as visitas foram coletadas amostras biológicas de múltiplos sítios. Para os estudos incluídos nessa tese utilizamos plasma com extração de RNA e dosagem de biomarcadores inflamatórios por metodologia LUMINEX e, urina, de onde foram dosados eicosanoides, mediadores inflamatórios lipídicos derivados do ácido araquidônico. Ao final dos seis meses de acompanhamento, todos os participantes foram adequadamente tratados conforme os critérios da OMS (61).

O grupo controle foi definido por pessoas sem diabetes e sem TB, pareados por sexo e idade nas amostras do RePORT, formando então 4 grupos de estudo que foram utilizados nos três trabalhos que compõem essa tese: HC, o grupo controle sem nenhuma das condições clínicas; DM, aqueles apenas com Diabetes Mellitus; TB, aqueles apenas com TB pulmonar e, por fim TBDM, os que possuíam ambas as condições clínicas estudadas. Ao todo, foram avaliadas amostras de 429 participantes.

4.2 Classificação e Amostragem Sanguínea

Para os três trabalhos que compõe essa tesa, o diagnóstico de TB foi baseado em cultura do escarro positiva para Mtb com um raio-X de tórax compatível no tempo zero. Culturas negativas e raio-X não compatíveis, definiram o grupo controle. A classificação de DM foi baseada na história médica pregressa e uso de hipoglicemiantes orais, ou ainda nos níveis de hemoglobina glicosilada (HbA1c) ≥ 6.5% no tempo zero, seguindo os critérios da Associação Americana de Diabetes (American Diabetes Association – ADA) (62).

4.3 Sequenciamento de RNA

As amostras de pacientes com TB pulmonar foram coletadas no tempo zero, dois e seis. Amostras de pacientes do grupo controle ou DM foram coletadas apenas no tempo zero. O RNA total do sangue venoso nos tubos PAXgenes RNA foi extraído usando o kit PAXgene blood miRNA (Qiagen) e quantificados usando sistema LabChip GX HiSens (PerkinElmer). Maiores detalhes do processamento das amostras de RNA estão contidos na sessão de métodos dos artigos apresentados na tese.

4.4 Análises estatísticas

A análise estatística foi feita utilizando mediana e intervalo interquartil como medida de tendencia central e dispersão, aplicando o teste de Kruskal-Wallis com o teste de Dunn para múltiplas comparações. Variáveis categóricas foram avaliadas com o teste de Qui-Quadrado. Correlações realizadas nos trabalhos da tese, foram feitas utilizando o modelo de Spearman. Utilizamos também análise de cluster hierárquico, com o método de Ward, representados por mapas de calor (heatmap), com bootstrap de 100x do score z normalizado.

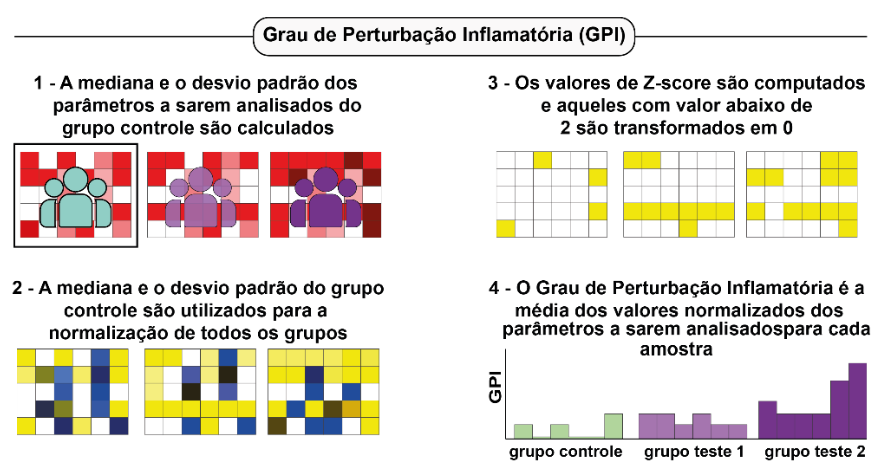
Genes diferencialmente expressos (DEGs) foram definidos como aqueles com Log_2 fold-change > ±1.4, a divisão do valor de expressão de um gene nos grupos DM, TB e TBDM pelo valor de expressão desse mesmo gene no grupo controle. Aqueles que apresentaram valores corrigidos de $p<0.05$, por False Discovery Ratio (FDR) foram considerados significantes.

Para análise utilizando inteligência artificial, escolhemos aplicar o modelo de Random Forrest (63). No primeiro artigo, aplicamos a técnica utilizando os valores dos DEGs na tentativa de identificar uma assinatura especifica que caracterizasse a população

TBDM. Já no segundo artigo, onde utilizamos múltiplas plataformas ômicas, a técnica foi aplicada para reduzir a dimensão dos dados.

Adicionalmente, empregamos a metodologia do grau de perturbação molecular (Molecular Degree of Perturbation – MDP), estratégia estatística adaptada de estudos em genoma, que vem sendo aplicada em diversos projetos do nosso grupo (10-12, 38-41). Utilizando o MDP somos capazes de mensurar o quão distante da normalidade (definida pelo grupo controle) está um determinado biomarcador. Inicialmente calculamos a mediana e o desvio padrão dos marcadores de interesse no grupo controle. Esses valores são utilizados para normalização de todos os outros grupos. Posteriormente, os valores de z-score são computados, e aqueles maiores que 2 são considerados perturbados, e serão incluídos, enquanto os menores que 2 são transformados em zero e, não serão considerados no cálculo do MDP. Então, calculamos a média dos valores normalizados para cada biomarcador em cada amostra, e chegamos ao valor final de MDP, como demonstrado na Figura 1.

Figura 2 – Entendendo o Grau de Perturbação Inflamatória (GPI) ou "Molecular Degree of Perturbation (MDP).



O GPI ou MDP é calculado usando mediana e o desvio padrão do parâmetro a ser analisado no grupo controle como ponto de partida. Em seguida, calcula-se o Z-score para todos os grupos, se estabelece um ponto de corte e, por fim, realiza-se o cálculo de perturbação média para cada amostra. Araújo-Pereira 2022.

Demais análises exclusivas de cada um dos trabalhos que compõem a presente tese estão detalhadas na sessão de apresentação dos artigos.

5 ÍNDICE DE ARTIGOS CIENTÍFICOS

Artigo I:

A Multi-center, Prospective Cohort Study of Whole Blood Gene Expression in the Tuberculosis-Diabetes Interaction

Artigo II:

An Integrative Multi-Omics Approach to Characterize Interactions Between Tuberculosis and Diabetes Mellitus

Artigo III:

The sound of silent RNA: The role of lncRNAs on TB infection in four different populations

Artigo IV:

Intersecting Epidemics: Deciphering the Complexities of Tuberculosis-Diabetes Comorbidity

6 ARTIGOS

6.1 Artigo I

www.nature.com/scientificreports

scientific reports

OPEN

A multi-center, prospective cohort study of whole blood gene expression in the tuberculosis-diabetes interaction

Artur T. L. Queiroz[1,2,3,25], Caian L. Vinhaes[2,3,4,25], Eduardo R. Fukutani[1], Akshay N. Gupte[5], Nathella Pavan Kumar[6], Kiyoshi F. Fukutani[2], María B. Arriaga[2,3], Timothy R. Sterling[7], Subash Babu[6], Sanjay Gaikwad[8], Rajesh Karyakarte[9], Vidya Mave[10,11], Mandar Paradhkar[10,11], Vijay Viswanathan[12], Amita Gupta[5], Bruno B. Andrade[2,3,4,13,26]✉, Hardy Kornfeld[14,15,26]✉, the RePORT Brazil* & RePORT India Consortia*

Diabetes mellitus (DM) increases tuberculosis (TB) severity. We compared blood gene expression in adults with pulmonary TB, with or without diabetes mellitus (DM) from sites in Brazil and India. RNA sequencing (RNAseq) performed at baseline and during TB treatment. Publicly available baseline RNAseq data from South Africa and Romania reported by the TANDEM Consortium were also analyzed. Across the sites, differentially expressed genes varied for each condition (DM, TB, and TBDM) and no pattern classified any one group across all sites. A concise signature of TB disease was identified but this was expressed equally in TB and TBDM. Pathway enrichment analysis failed to distinguish TB from TBDM, although there was a trend for greater neutrophil and innate immune pathway activation in TBDM participants. Pathways associated with insulin resistance, metabolic dysfunction, diabetic complications, and chromosomal instability were positively correlated with glycohemoglobin. The immune response to pulmonary TB as reflected by whole blood gene expression is substantially similar with or without comorbid DM. Gene expression pathways associated with the microvascular and macrovascular complications of DM are upregulated during TB, supporting a syndemic interaction between these coprevalent diseases.

Diabetes mellitus (DM) has been associated with increased risk for tuberculosis (TB) progression and adverse TB treatment outcomes in most clinical studies[1]. Mirroring the human data, animal models combining chronic

[1]Centro de Integração de Dados e Conhecimentos para Saúde, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil. [2]Laboratório de Inflamação e Biomarcadores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil. [3]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador 41810-710, Brazil. [4]Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador 40290-150, Brazil. [5]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [6]National Institutes of Health- NIRT - International Center for Excellence in Research, Chennai, India. [7]Division of Infectious Diseases, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [8]Department of Pulmonary Medicine, Byramjee-Jeejeebhoy Government Medical College and Sassoon General Hospitals, Pune, India. [9]Department of Microbiology, Byramjee-Jeejeebhoy Government Medical College and Sassoon General Hospitals, Pune, India. [10]Byramjee-Jeejeebhoy Government Medical College-Johns Hopkins University Clinical Research Site, Pune, India. [11]Johns Hopkins Center for Infectious Diseases in India, Pune, India. [12]Prof. M. Viswanathan Diabetes Research Centre, Chennai, India. [13]Faculdade de Tecnologia e Ciências, Instituto de Pesquisa Clínica e Translacional, Salvador 41741-590, Brazil. [14]Department of Medicine, University of Massachusetts Medical School, Worcester, MA, USA. [15]UMass Chan Medical School, Worcester, MA, USA. [25]These authors contributed equally: Artur T. L. Queiroz and Caian L. Vinhaes. [26]These authors jointly supervised this work: Bruno B. Andrade and Hardy Kornfeld. *List of authors and their affiliations appears at the end of the paper. ✉email: bruno.andrade@fiocruz.br; hardy.kornfeld@umassmed.edu

hyperglycemia with *Mycobacterium tuberculosis* challenge showed higher lung bacterial burden and more TB immune pathology[2]. The global population-attributable fraction of TB associated with DM is comparable to that of HIV/AIDS[3]. Despite its significance as a barrier to TB elimination[4], the mechanisms whereby DM impairs host defense against *M. tuberculosis* are not well understood[5].

Mechanistic studies of human immunity to TB are limited to accessible tissue samples. Blood transcriptomic studies have described a consensus TB signature of increased type I interferon (IFN) signaling[6]. In a prior blood transcriptome study using microarrays to evaluate a South Indian cohort, we found no pattern of immune gene expression that distinguished TB-DM comorbidity from TB in participants without DM[7]. The TANDEM Consortium later published a whole blood RNA sequencing (RNAseq) study from four national sites (South Africa, Peru, Indonesia, Romania)[8]. Their analysis revealed increased inflammatory pathway gene expression but reduced type 1 interferon signaling in TB combined with dysglycemia compared to TB in people with normoglycemia. To further explore the impact of DM on the host response to TB in diverse populations, we performed RNAseq on whole blood RNA sampled at TB diagnosis (baseline) and treatment months 2 and 6 from adults with pulmonary TB, with or without DM, at sites in India and Brazil.

The Molecular Signatures of Tuberculosis-Diabetes Interaction (MSTDI) study reported here leveraged unreported participant data and whole blood RNA samples from prospective observational pulmonary TB cohorts from two sites of the RePORT India and one from RePORT Brazil consortia. A single vendor performed RNAseq on baseline samples from all three sites and longitudinal samples during TB treatment from one Indian site and Brazil. To check the variability of the blood transcriptome in TB-DM interaction, comparison was made with published blood RNAseq data from two TANDEM consortium sites (Romania and South Africa). Our analysis was aimed to compare the intensity and quality of inflammatory activation between the clinical conditions and sites and evaluate the impact of HbA1c levels in the biological pathways. the This investigation did not reveal insights to the mechanisms of TB susceptibility in DM but the data support the existence of a syndemic interaction between TB and DM[9].

Results

Differential gene expression between clinical conditions and sites. The MSTDI cohort, comprising 290 participants, was recruited from two sites in India and one site in Brazil. Adults newly diagnosed with drug-sensitive pulmonary TB, with or without DM (TB and TBDM, respectively), and control group participants without TB, with or without DM (DM and HC, respectively) were enrolled. Characteristics of the population are shown in Supplementary Fig. S1 and Supplementary Tables S1 and S2. Additional comparison was made with publicly available RNAseq data from two sites (Romania and South Africa) of a TANDEM consortium gene expression study with similar group structure and where data were available for site-specific healthy control (HC) participants.

Gene expression within the disease condition groups (DM, TB, TBDM) was compared against site-specific HC participants. The raw data on DEGs are shown in Supplementary File S1. Across all four sites, the DM groups exhibited no DEGs in common, while the TB and TBDM groups shared six and twelve DEGs, respectively (Fig. 1A). A z-score normalized heatmap using the combined DEG gene expression values further demonstrated the variability within and between groups and sites, finding no pattern that classified any clinical group across all sites (Fig. 1B). A principal component analysis model (PCA) applied to DEGs differentiated between groups by the presence or absence of TB disease but did not discriminate between TB and TBDM (Supplementary Fig. S2A,B). Furthermore, we calculated the molecular degree of perturbation (MDP) score[10] to estimate the overall level of inflammation within disease groups at each site. The highest individual and median MDP values were present in the TB and TBDM groups and there was a non-significant trend for higher MDP in the TBDM than in the TB groups at all sites (Fig. 2A–D). Statistically significant differences in MDP scores across the four study sites were identified in the HC, DM, TB, and TBDM groups (Fig. 2E). To evaluate possible effects of clinical and epidemiological features in the differential gene expression, we performed a PCA, labelling the participants according to presence or absence of cavitation (Supplementary Fig. S3A,B). Furthermore, a Spearman correlation between BMI and the MDP values was performed (Supplementary Fig. S3C,D). No association between BMI values and gene expression variation was observed in the TB and TBDM groups in both sites. The results showed that we could not segregate the participants according to the presence or absence of cavitation, and that the BMI values were not associated with the degree of inflammatory activation, highlighting that the population-specific differences for all these clinical conditions were likely not associated with clinical and epidemiological features.

Discovery and validation of a TB-associated gene expression signature. As an alternative approach to identify gene expression patterns associated with specific groups across all sites, we inputted expression values of the combined total of 3427 DEGs for a random forest model (Supplementary Fig. S4). The South Africa and Romania sites were used as a discovery set for each of the four conditions (HC, DM, TB, TBDM) and the model was tested in a validation set composed of samples from India and Brazil (Supplementary Fig. S4). Model validation resulted in the accuracy of 58.97 (63.32–74.84); $p = 0.001658$). Gene expression data were used in a feature selection analysis with a random forest algorithm to rank the variables according to their model importance and those above the third quartile were selected, identifying SMARCD3, VAMP5, ANKRD22, and BATF2 as most informative genes to distinguish the clinical conditions (Supplementary Fig. S4A). An unsupervised cluster analysis of z-score normalized expression data identified higher baseline expression of these four genes in participants with TB at all four sites, irrespective of DM status (Fig. 3). Consistent with that observation, Spearman correlation analysis performed between HbA1c levels and expression of the most informative genes did not reveal a consistent association between the expression of these genes and HbA1c levels in participants from Brazil or India (Supplementary Fig. S4B,C). While the 4-gene signature did not discriminate between TB

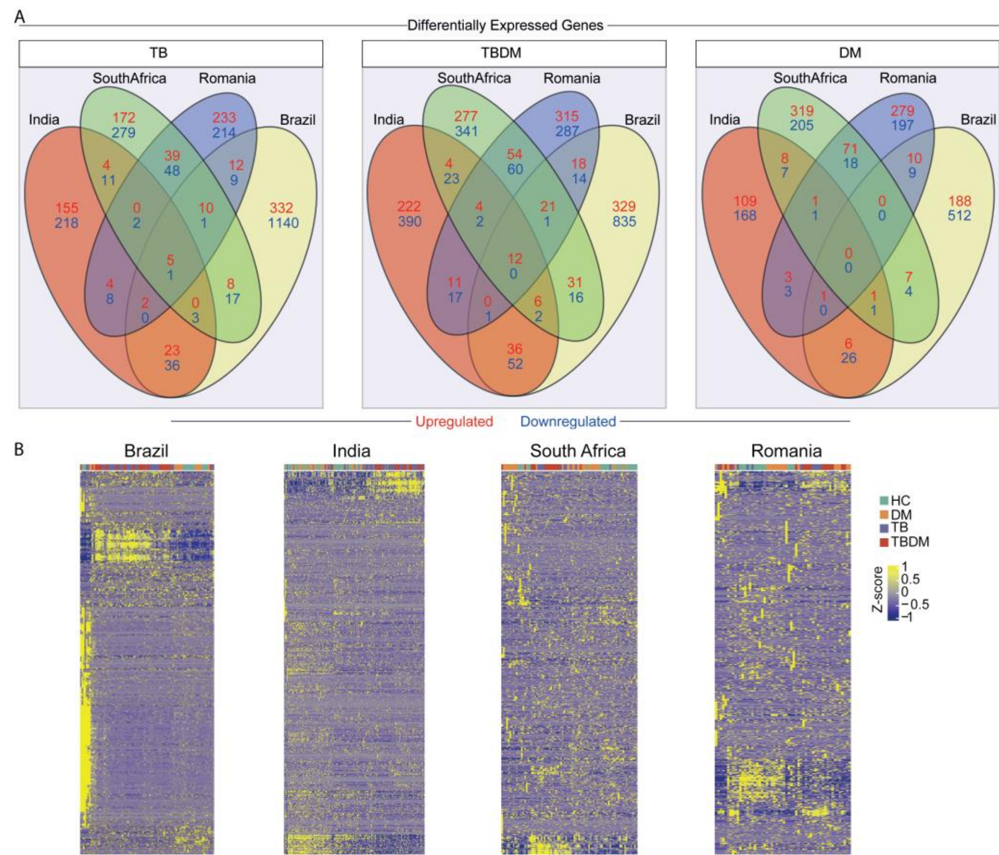


Figure 1. Distribution of differentially expressed genes (DEGs) between clinical groups in the sites from MSTDI and TANDEM cohorts. **(A)** Venn Diagrams show the DEGs defined by the threshold corrected p-value < 0.05 and log twofold change $> +/- 1.4$ from each comparison of TB vs, TBDM vs, and DM vs HC from the subjects within the cohort sites. **(B)** Heat maps of z-score normalized data of all DEGs from each cohort site. Each heat map depicts all DEGS from TB vs, TBDM vs, and DM vs HC comparison from subjects within each cohort site.

and TBDM, receiver operator characteristic (ROC) curve analysis demonstrated that the signature presented good accuracy to classify TB and TBDM from DM and HC at all four sites with ROC ≥ 0.85 (Fig. 4). Additionally, we tested the accuracy of previously published gene biosignatures in our study sites, with relatively high accuracy of most of signatures to identify either TB or TBDM participants (Supplementary Figs. S5, S6, respectively). The temporal expression of these signature genes over the course of TB treatment differed among the individual genes, the condition (TB or TBDM), and between the Indian and Brazilian cohorts (Fig. 5). Expression levels of all four signature genes tended to be higher in TBDM than TB in the Brazil cohort at month-6, rising from a nadir at month-3. In the India cohort, expression of BATF2, VAMP5, and ANKRD22 tended to be higher in TBDM than TB at baseline and month-2, with VAMP5 and ANKRD22 continuing that trend to month-6. These patterns might reflect persistent inflammation in TBDM, which was identified in prior study measuring plasma cytokines[11].

Pathway enrichment and interaction analysis across conditions and populations. While comparison of DEGs failed to discriminate between TB and TBDM, we questioned whether pathway enrichment had the potential to reveal condition-specific differences. Reactome pathways of interest were identified within significant ($p < 0.05$) and false discovery rate (FDR)-corrected DEGs from all conditions and sites. No pathway was uniquely enriched within the TB or TBDM groups at all sites, but interferon signaling that has been identi-

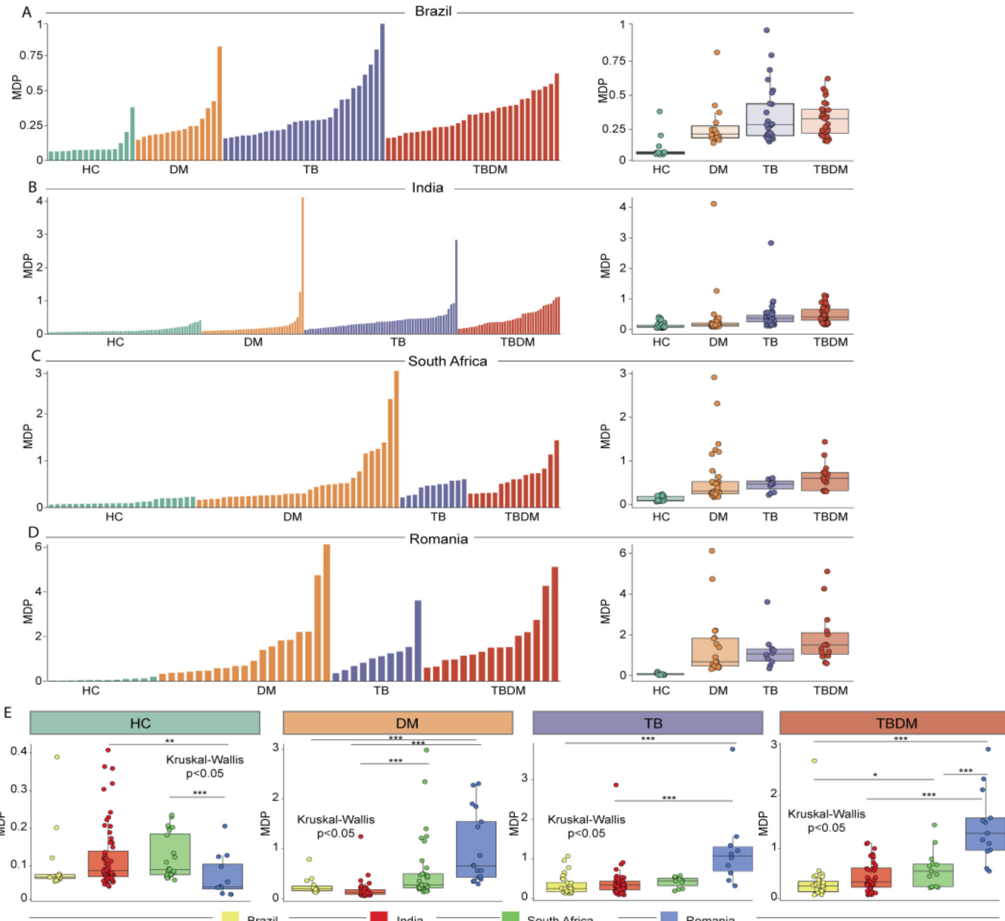


Figure 2. Molecular degree of perturbation stratified by clinical group and site. Histograms show molecular degree of perturbation (MDP) scores of each sample (left) and boxplots of each group (right) are shown for Brazil (A), India (B), South Africa (C) and Romania (D), using Kruskal–Wallis test. (E) Boxplots of MDP by of each clinical group (DM, TB, TBDM and HC) compared between sites. To estimate which groups are different from others, Tukey's post test was employed. Thus, *refers to $p \leq 0.05$, **refers to $p \leq 0.01$, ***refers to $p \leq 0.001$ and ****refers to $p \leq 0.0001$.

fied in many TB gene expression studies[12] was enriched among TB participants from Brazil, South Africa, and Romania (Fig. 6A). The India and Brazil TBDM groups shared enrichment of the Neutrophil.

Degranulation, Antimicrobial Peptides, and Extracellular Matrix Organization pathways. Interferon signaling was enriched only in the Romanian TBDM group but not in TBDM participants from the other sites (Fig. 6B). Mirroring the DEG analysis, the changes in pathway predominance during TB treatment were markedly different between sites (Supplementary Fig. S7). Although pathway enrichment did not consistently distinguish TBDM from TB across all sites, the trend for increased Neutrophil Degranulation, Antimicrobial Peptides, and Matrix Organization fits with prior reports of neutrophilic inflammation and elevated circulating levels of cathelicidin, human beta defensin- 2, human neutrophil peptides 1–3, and matrix metalloproteinases in TB-DM comorbidity[7,13,14].

The Neutrophil Degranulation, Antimicrobial Peptides, Interferon Signaling, Regulation of Complement Cascade, Complement Cascade, and Interferon alpha/beta Signaling pathways were evaluated using a hierarchical

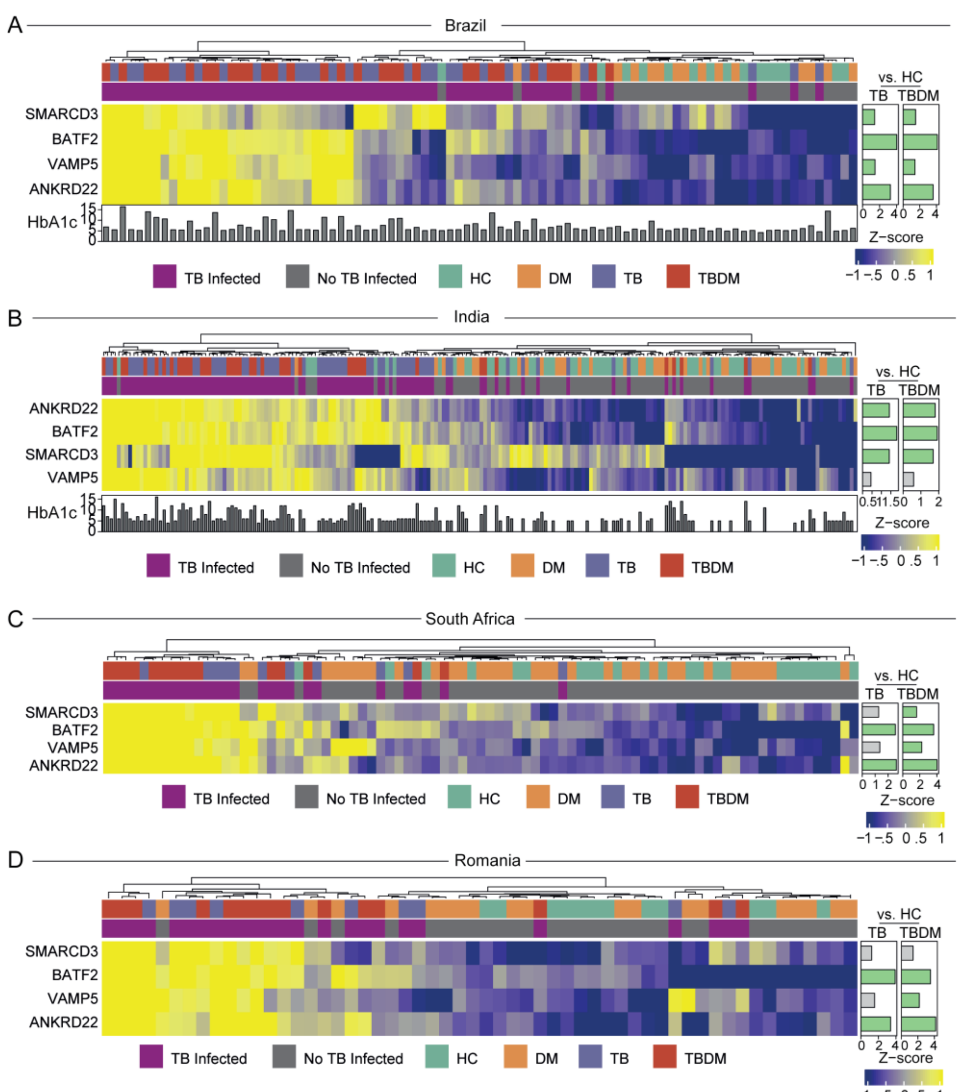


Figure 3. Relative expression of four TB signature genes at each site. Using a machine learning analysis, we defined four top genes among the clinical sites (as showed in Supplementary Material). A z-score normalized heatmap was employed to depict overall trends in gene expression among the clinical groups each study site, as indicated. Panels to the right of heatmaps show the average fold-difference between the signature gene expression in the HC group versus TB and TBDM (log-transformed values).

cluster analysis ordinating participants by MDP levels in TB and TBDM groups (Fig. 7A,B). This analysis was limited by the absence of some pathway enrichment data from South Africa and Romania, thus box plots from only Brazil and India are displayed. We observed a higher expression of the Interferon Signaling and Neutrophil Degranulation pathways in Brazilian TB participants when compared with the TB group from India (Fig. 7A).

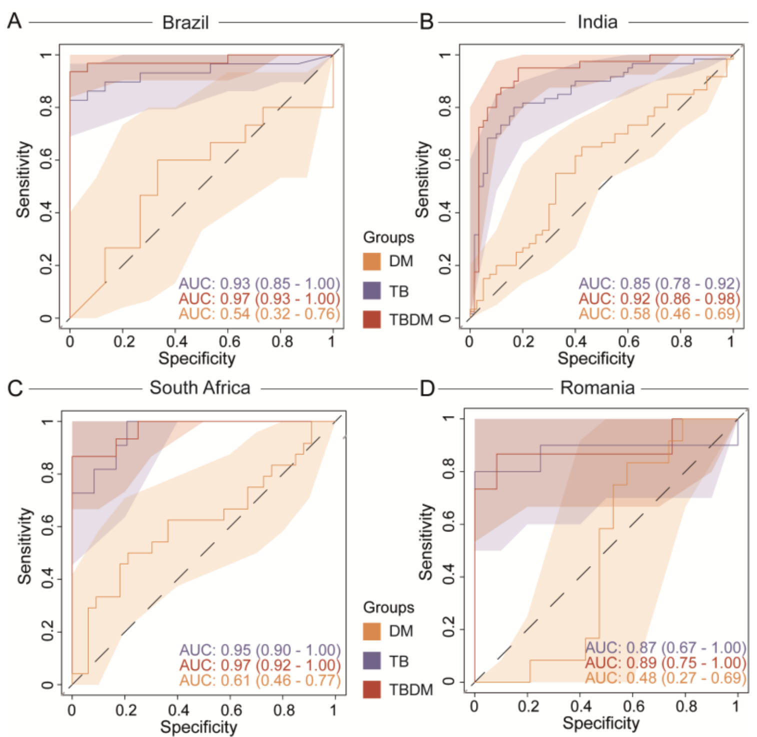


Figure 4. Accuracy of top the 4-gene signature to classify TB, with or without DM. Receiver operator curve (ROC) analysis was used to check the accuracy of the signature genes identified by the random forest model to classify the TB, TBDM, and DM groups in each clinical site as indicated with respect to TB disease.

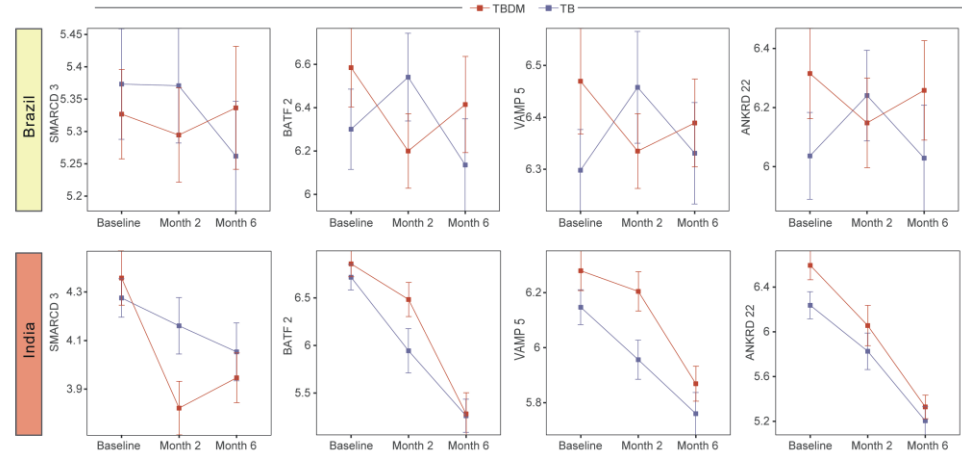


Figure 5. A distinctive pattern of gene expression during TB treatment. Expression of the four TB signature genes was analyzed at baseline, month-2 and month-6 in MSTDI participants in India and Brazil. Nemenyi's non-parametric all-pairs comparison was used.

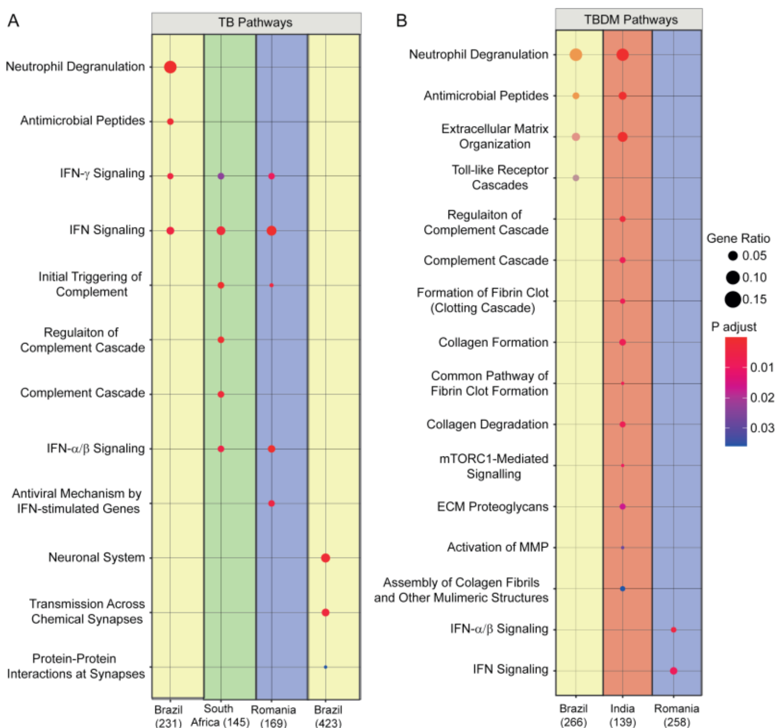


Figure 6. Pathway enrichment by condition and site. The colored spots indicate selected upregulated and downregulated Reactome pathways of interest identified by the combined significant ($p < 0.05$) and FDR-corrected DEGs from the TB groups (A) and TBDM groups (B) in Brazil (yellow), India (red), South Africa (green), and Romania (blue). The hue of spots corresponds to p-value and the area corresponds to gene ratio. The method did not detect any pathways in TB participants from India.

Among TBDM groups, the Interferon Signaling pathway was more highly expressed in Brazil while the Complement Cascade pathway was more highly expressed in India (Fig. 7B).

To further evaluate the interaction between DM and TB, we examined the interactions between pathways identified from DEGs compared to Reactome in the TBDM groups (Fig. 8) and TB groups (Supplementary Fig. S8) which were observed at least two sites. This approach differed from our initial pathway analysis by estimating enrichment scores computed in each sample instead of a set of DEGs, enabling further analyses such as correlation or fold change score. Correlation network analysis and fold-change analysis using single sample gene set enrichment analysis (ssGSEA) to generate normalized enrichment scores (NES) identified differences across sites and clinical groups, along with highly correlated pathways at some sites that were not identified by the prior enrichment analysis (Fig. 6), which was not suitable to investigate correlation. As shown in Fig. 8, higher density of network pathways in TBDM was found in South Africa (0.108), India (0.091), and Brazil (0.069). Network pathway density was markedly lower in the Romanian TBDM group (0.013) despite a higher number of vertices found in these correlations (Fig. 8). Excluding the Romanian cohort, the top correlated pathway in TBDM was Complement Cascade, with 22 connections in Brazil, 30 correlations in India and 25 in South Africa, however, the pathway was not uniformly up or downregulated in TBDM at all the three sites compared with the site-specific healthy control group. Correlation analysis for the normoglycemic TB groups using ssGSEA-NES values showed comparable diversity between sites (Supplementary Fig. S8).

Correlation of ssGSEA pathways with HbA1c. Finally, to assess the influence of average blood glucose levels with pathway engagement, we employed a correlation model using HbA1c levels and ssGSEA NES values with available data from Brazil and India (Fig. 9). Among TB group participants, the India cohort data showed a substantially higher number correlations (68 positive and 6 negative) than Brazil (9 positive and only one negative) (Fig. 9). The TBDM comparison showed a lower total number of correlations than was seen for TB, again

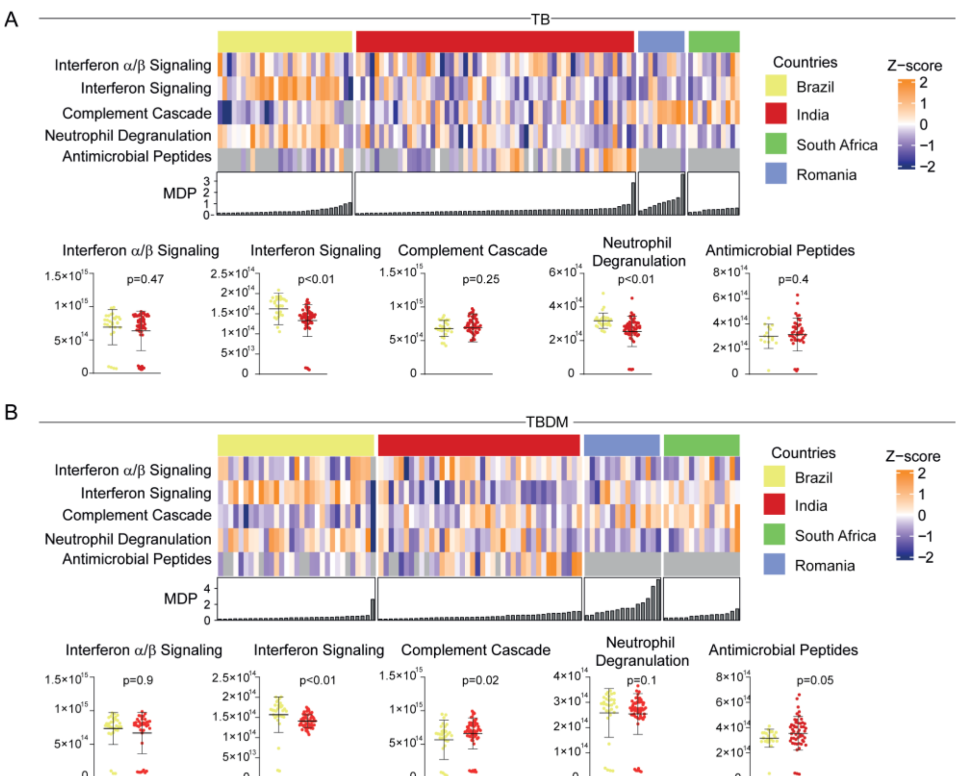


Figure 7. Changes in the pathways expression across clinical sites and diseases. A heatmap Z-score normalized was employed using the overlapped pathways identified in the enrichment analysis from DEGs (as described in Fig. 3). Box plots represents median and interquartile ranges, Mann Whitney test was used.

with more in the Indian cohort (7 positive and 7 negative correlations) compared to Brazil (6 positive and 2 negative correlations) (Fig. 9). There was no overlap in these correlated pathways between the sites or conditions. Notably, many of the positive correlations from the India TB group were for pathways associated with insulin resistance and metabolic syndrome (e.g. Passive Transport by Aquaporins, Yap1 and WWTR1 Taz Stimulated Gene Expression, Complex I Biogenesis, Mitochondrial Translation), diabetic complications (Axon Guidance, Regulation of Kit Signaling, Kinesins, MAPK Family Signaling Cascades, Mitochondrial Translation, VEGF Signaling, Factors Involved in Megakaryocyte Development and Platelet Production, Unblocking of NMDA Receptors Glutamate Binding and Activation), and pathways associated with chromosomal instability (Mitotic Prometaphase, Centrosome Maturation, G2 M DNA Damage, M Phase, Chromatin Organization, Mitotic Metaphase and Anaphase). Correlation with immune and inflammatory pathways were less prominent but included positive correlations with the RORA Activates Gene Expression and the Innate Immune System pathways in the India TB group, and negative correlations with B Cells Memory and Potassium Channels Pathways. The universally high level of HbA1c prevented correlation of these pathways in TBDM individuals.

Discussion

Whole blood RNAseq potentially offers a window into biological processes at the primary site of TB disease in the lung. In an earlier blood gene expression study at Chennai with a different group of participants, we found no differences in immune response pathway activation between TB and TBDM[7]. To address shortcomings of that study done at one site and using microarrays, the MSTDI study included participants from Chennai and Pune in India and Salvador in Brazil and used RNAseq performed on the same platform. Results of the MSTDI study supported our prior findings. There was no evident signature pattern of immune response gene expression or pathway activation that offered mechanistic insight to the basis of TB susceptibility in people living with DM. However, trends for increased neutrophil and innate immune pathway activation in TBDM were noted. Studies

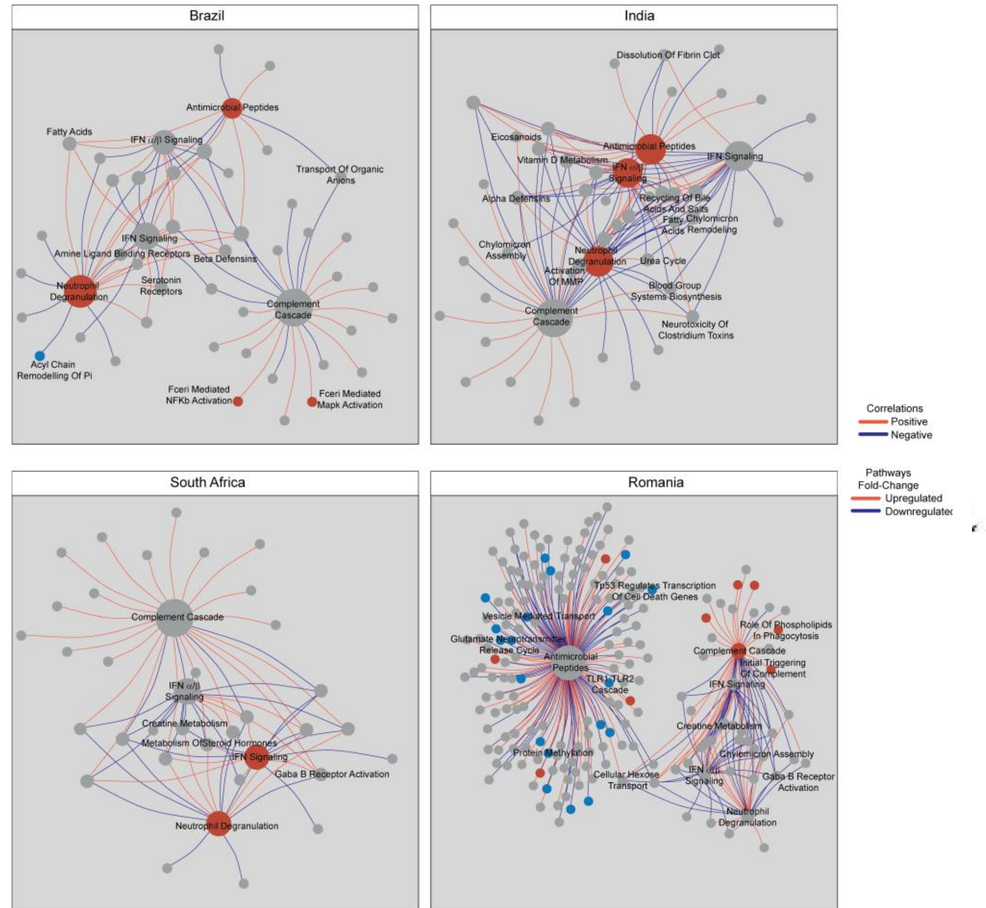


Figure 8. Pathway dynamicity across the sites and disease groups. A Spearman correlation analysis was performed using the pathways from TBDM participants in each clinical site, as indicated. Each node indicates a pathway, blue nodes indicate downregulation compared to the site-specific HC group, while red nodes indicate upregulation. Grey nodes represent pathways that were not statistically significant compared to the site-specific healthy control group. Red lines infer positive correlation and blue lines negative interaction.

in hyperglycemic mice identified a defect in alveolar macrophage sentinel function as the key immune mechanism of TB susceptibility in DM[15–17]. This impacts the initial encounter of alveolar macrophages with inhaled *M. tuberculosis*, delaying the transition from innate to adaptive immunity during the period of logarithmic increase in lung bacterial load. Once the adaptive immune response is expressed in diabetic mice, it effectively limits *M. tuberculosis* replication and is qualitatively indistinguishable from the response of normoglycemic control mice with TB. If a similar mechanism operates in human DM, then the key immunological events contributing to susceptibility will have occurred months before TB diagnosis. Gene expression at baseline in human studies corresponds to the later time points of mouse experiments when the cell-mediated immune response is fully activated.

Based on animal model data, we anticipated that immune pathology in human TBDM would be qualitatively similar to euglycemic TB but quantitatively more severe[18]. That prediction was supported by plasma cytokine and radiographic studies[7,19], but not so clearly by differences in whole blood gene expression[7]. The TANDEM study included four geographic sites and reported an overall increase of innate immune gene expression with decreased expression of genes linked to adaptive immunity in participants with TB combined with DM or pre-DM[20]. That result supports the general notion of increased immune pathology in TBDM but does not explain the broadly

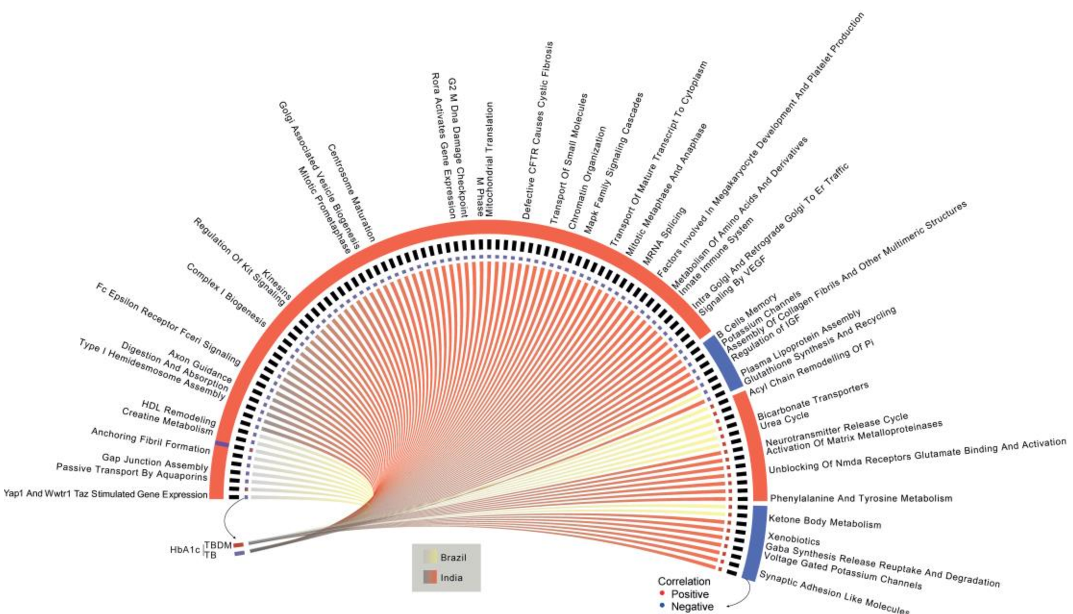


Figure 9. Integrative analysis of HbA1c and Pathways expression in Brazil and India. Levels of HbA1c were correlated with the pathways according to disease group as indicated in the left box (red refers to TBDM whereas purple to TB). The external-colored bars indicated the direction of correlation, as indicated in right circles (red for positive correlation, blue for negative correlation, and purple for positive and negative correlations). Lines in center of the figure refers to the country (red for India and yellow for Brazil).

elevated plasma levels of Th1, Th2, and Th17 cytokines reported in this condition[7,21]. The MSTDI data presented here showed a non-significant trend for higher MDP values in TBDM than TB participants, whereas the differences in MDP between sites were statistically significant. Pathway enrichment, interaction, and correlation analyses likewise failed to identify a universally consistent TBDM-specific pattern of engagement. Overall, the data suggest greater severity of immune pathology in TBDM, which more likely reflects a consequence rather than cause of TB susceptibility in people living with DM.

Results from the MSTDI study, notably the correlation of pathway engagement with HbA1c levels, further support the potential for TB to exacerbate non-communicable disease processes[7,9]. Bacterial pneumonia and COVID-19 are both associated with increased risk of cardiovascular disease events[22,23]. A similar association has been made with TB[24] where the longer duration of inflammation, particularly with comorbid DM[11], could be particularly damaging. Similar reasoning would apply to the microvascular complications of DM, but this has not been studied.

An unexpected finding in our analyses was the high degree of variability in baseline gene expression between all four populations evaluated. Of note, some differences in the clinical presentation and disease severity were found in our study, as in TB, with a higher percentage of cavitations in India when compared with Brazilian TBDM participants, as well in DM, with higher HbA1c levels detected in India. The heterogeneity in this proportion could affect the gene expression and to verify its influence we performed an additional analysis that demonstrated an apparent lack of association between such parameters and differences in gene expression profiles among the countries. Equally remarkable were the different temporal patterns of gene expression in the India and Brazil cohorts from baseline through TB treatment completion at month 6. This variability presumably reflects differences in host and microbe genetics, behaviors, and environmental exposures between sites. Similar issues have hampered the application of concise gene expression signatures for TB diagnosis and treatment response at different sites[25]. Further research will be required to identify the fundamental processes influencing the TB-DM interaction and their expression in diverse populations.

This work has some limitations. First, not all clinical and epidemiological data are available from all datasets used. The data from South Africa and Romania did not present other information such as HbA1c levels, Cavitation or alcohol, and smoking usage. For this reason, only age and sex were used to adjust the models in differential expression analysis. The second limitation was the differences observed in the BMI values, age, sex, smoking, and alcohol use between the population of Brazil and India (Supplementary Tables S1 and S2). Additionally, some patients were under treatment with metformin and statin, that could affect the inflammatory responses. We have applied a negative binomial model for the differential gene expression analysis. This modeling has limitations regarding multiple variable adjustments. Despite we have evaluated the Cavitary TB and BMI association with

gene expression, their systemic influence could not be fully corrected in the model. Moreover, use of metformin and/or statins, and frequency of cavitary TB were significantly different in TB-infected individuals between those countries. Cavitary TB and BMI values are known to affect immune activation and thus could influence the gene expression profile of those subjects and may explain the absence of overlapping pathways in Brazil and India sites. Another limitation was associated with the different sample sizes observed in the cohorts (Supplementary Fig. S1). South Africa data presented 11 TB and 15 TBDM samples, Romania 10 TB and 15 TBDM samples, Brazil 29 TB and 31 TBDM samples, and India 60 TB and 40 TBDM samples. This unbalanced sample size could have influenced both gene expression analysis and signature performance. DM is a metabolic disease, and it is reasonable to speculate that using other platforms, such as metabolomics, could increase the odds of detecting discrepancies between TB and TBDM. We are currently performing new investigations using multi-omics to portray the TB-DM interaction.

This work has some differences compared with the TANDEM work which makes the straightforward comparison of the results difficult. First, the TANDEM has used data from 4 different sites (South Africa, Romania, Peru, and Indonesia). However, all the analysis was performed with whole data, without performing the comparisons within the sites, as was performed in the MSTDI. Also, two TANDEM sites do not have healthy controls, thus all comparisons from the case subjects from Peru and Indonesia were performed with foreign control, which may have inserted variation. Second, the enrichment analysis was performed using different approaches. In TANDEM, the transcriptional module enrichment analysis was performed with tmod package while two different approaches were performed in the MSTDI, using clusterProfiles and Reactome database, and the single sample gene set enrichment analysis with ssGSEA package. This shows that the cohorts explored the disease dynamics in a different optic, with the first exploring the overall impact and the second one investigating the populational contribution on this impact.

In summary, we found substantial variability in whole blood gene expression between HC, DM, TB and TBDM participants across the two MSTDI study sites in Brazil and India and the TANDEM study South African and Romanian cohorts. No mechanistically informative signature of immune pathway gene expression distinguished TBDM vs TB in all populations, although increased innate immune and vascular complication pathway activation was common across some sites. Our findings lend evidence in support of adjunctive anti-inflammatory and antioxidant therapies during TB treatment. In that regard, retrospective evidence demonstrated that metformin added to antibiotic treatment reduces the mortality risk in TBDM independent of glycemic control[26] and a recently completed randomized controlled trial of TBDM patients in India showed that metformin reduced inflammation and radiographic severity in disease[27].

Materials and methods

Ethics statement and study population.

The MSTDI study used unpublished data and samples from the RePORT India and the RePORT Brazil consortia. Samples from the RePORT India were enrolled under protocols approved by the Ethics Committee of the Prof. M. Viswanathan Diabetes Research Center and the Institutional Review Boards of Byramjee Jeejeebhoy Government Medical College, Pune and National Institute for Research in Tuberculosis and Johns Hopkins University. The samples enrolled from the RePORT Brazil had their protocols approved by the institutional review boards of the Instituto Gonçalo Moniz, Fundação Oswaldo Cruz and Vanderbilt University Medical Center. Written informed consent was obtained from all participants prospectively enrolled at two sites of the RePORT India and one site of the RePORT Brazil consortia, with organizational support from RePORT International[28]. The study was conducted according to the principles of the Declaration of Helsinki. The Indian sites were in Chennai (EDOTS study)[29] and Pune (CTRIUMPH study[30]), while the Brazilian site was in Salvador (RePORT International Common Protocol)[28]. The combined MSTDI cohort of 290 individuals comprised 120 participants from Chennai, 80 from Pune and 90 from Salvador. These participants were selected from a larger cohort of the RePORT protocols based of sample availability. Participant groups included pulmonary TB disease with or without DM (TBDM and TB groups, respectively) and two control groups without TB, with or without DM (DM and HC groups, respectively). Inclusion criteria were age 18–65 and new diagnosis of pulmonary TB (or absence of pulmonary TB for the control group participants). Drug-resistant TB, retreatment TB, treatment of incident TB for > 7 days prior to enrollment, pregnancy, immunosuppressive medications, and HIV infection were exclusions. All TB data used here were obtained from participants accompanied during the 6 months of anti-TB therapy according to the RePORT common protocol and who were successfully treated according to WHO criteria[31]. Participant characteristics are presented in Supplementary Tables S1 and S2 of the Supplementary Material. Secondary data were used from the TANDEM study[32] that explored the TB-DM interactions at clinical sites in Indonesia, Peru, Romania, and South Africa. Patients from TANDEM were also on TB treatment according to the respective local TB program. We used TANDEM data from South Africa and Romania where site-specific healthy control participant data were available. This data set comprised 83 participants from South Africa (24 HC, 33 DM, 11 TB, 15 TBDM) 56 from Romania (12 HC, 19 DM, 10 TB, 15 TBDM). The group sizes for all sites used in our analysis are summarized in Supplementary Fig. S1.

Classification and blood sampling.

TB diagnosis was based on positive sputum culture for *M. tuberculosis* with a compatible chest x-ray at enrollment, while negative culture and x-ray defined the control groups. Classification with DM was based on self-reported medical and anti-diabetic medication history or glycohemoglobin (HbA1c) ≥ 6.5%. Classification as euglycemic was based on self-reported medical history and HbA1c < 5.7% or 75-g oral glucose tolerance test 2-h blood glucose < 140 mg/dL. Baseline blood samples were collected in RNA storage tubes from all participants at the time of enrollment and no later than 7 days after the initiation of anti-tubercular treatment.

Library preparation and RNA sequencing. The samples from TB-infected individuals from both India and Brazil sites were collected at baseline, 2, and 6 months of treatment. Samples for the HC and DM groups were collected at baseline. Whole blood (5 mL) was collected PAXgene Blood RNA tubes (Qiagen, catalog #762165) and frozen at − 80 °C. RNA was extracted using the PAXgene Blood RNA kit (Qiagen, catalog #762174) and quantified using Qubit RNA assay HS (Invitrogen, Cat #Q32852). RNA purity was checked using QIAxpert, and RNA integrity was assessed on TapeStation using RNA HS ScreenTapes (Agilent, Cat #5067-5579). NEB Ultra II Directional RNA-Seq Library Prep kit protocol was used to prepare libraries for total RNA sequencing. Prepared libraries were quantified using Qubit High Sensitivity Assay (Invitrogen, Cat #Q32852), pooled and diluted to final optimal loading concentration before cluster amplification on Illumina flow cell. Once the cluster generation was completed, the cluster flow cell was loaded on Illumina HiSeq 2500 instrument to generate paired end reads at MedGenome in Bangalore, India.

RNA-seq data analysis. Raw RNA-seq data from the MSTDI cohort were retrieved from Illumina HiSeq 2500 platform. Sequence data from the TANDEM cohort[8] was retrieved from the SRA database using BioProject PRJNA470512 using the SRA tools. Sequence data from MSTDI and TANDEM were quality control processed by removing low-quality bases and adapters using *Trimmomatic V0.32*. After the quality check, sequences were pseudo-aligned against the human transcriptome (GRCh38 version) comprising both mRNA and miRNA with *salmon* v0.8.2[33] and presented a mean mapping rate of 68.03 ± 2.28. After mapping, the output was converted to a count table using *tximport* package[34] from *R 4.1.3*. Count gene expression matrix was examined using the *edgeR* package[35] from *R4.1.3* to identify differentially expressed genes (DEGs). MSTDI gene expression data are available at the GEO database (Accession number GSE181143, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE181143).

Differential expression analysis in the MSTDI cohort compared three conditions (TB, TBDM, DM) at two sites (Brazil and India) to the healthy control (HC) group at the respective site to determine the fold change and *p*-value of each gene. For instance, the TB, TBDM, and DM subjects from one country are compared with the HC subjects from the same country, and not from another country, to avoid inserting operational variation in the analysis. From the TANDEM cohort, only two sites (South Africa and Romania) were used in the differential expression analysis. The other sites (Peru and Indonesia) lack the HC group at the respective sites. Multiple testing correction was performed using the Benjamini & Hochberg false discovery rate (FDR) method[36]. Changes in gene expression were considered significant when corrected *p*-values remained < 0.05 after FDR adjustment and if the fold change differences were higher than ± 1.4. The complete list of genes is available in Supplementary File 1. The DEGs were visualized using the *VennDiagram* package[37] from *R.4.1.3*. The *compareCluster* package[38] from *R4.1.3* was used with the obtained DEGs to scan the *REACTOME* pathway database[39] to perform the pathway enrichment analyses and the the Benjamini & Hochberg false discovery rate (FDR) method[36] was used to correct the p-values for multiple testing.

Participant characteristics tables present the median and interquartile range or percentage for the nominal variables. All comparisons were performed with Kruskal–Wallis with Tukey's post test and Chi-squares tests using R*.4.1.3*.

Population heterogeneity evaluation and feature selection analysis. Sample variation within and between sites was evaluated using the molecular degree of perturbation (MDP) package[10] applied in the gene expression values after variance-stabilizing transformation (VST), for Brazil, South Africa and Romania data, and batch effect correction with sva package[40] in the case of India site data. A gene was classified as perturbed when its variation compared to HC was > 2 standard deviations.

To evaluate the sample clustering and classification across the sites we performed one-sided unsupervised hierarchical clustering Ward's method[41], Heatmaps[42] and the Principal component analysis (PCA) plot in the VST gene expression values from each cohort (Supplementary Figs. S2, S3A,B). This approach allows the visualization of sample dispersion across the groups in the sites. To maximize the performance in the TB and TBDM classification, we employed a dimensionality reduction approach to reduce the number of genes associated with TB. Thus, we retrieved the VST gene expression of all 3427 DEGs from the comparisons. The data from South Africa and Romania were used as a discovery set in a random forest algorithm with leave-one-out cross-validation with the *caret*[43] and *randomForest*[44] packages. The data from Brazil and India was used as a validation set (Supplementary Fig. S4). The minimal gene set exhibiting higher classification power to describe the groups was defined by the variable importance in the random forest model. Gene expression values from each were retrieved from each site and the Receiver Operator Characteristics (ROC)[45] were used to assess the accuracy of the gene set to distinguish between comparison groups specified in the TANDEM and MSTDI datasets.

Performance analysis with previously identified signatures. We conducted a performance comparison using 69 previously published gene expression signatures for TB diagnosis, progression, and treatment provided by the *TBSignatureProfiler* package (https://github.com/compbiomed/TBSignatureProfiler). In addition, we have included RISK6, RISK11 and BATF2 signatures for comparison (Supplementary Figs. S5 and S6). The signature classification performance was performed using data from each signature data from TB and HC (Supplementary Fig. S5), and TBDM and HC (Supplementary Fig. S5) within each country, similar to the analysis performed to identify the DEGs. We estimated the area under the curve (AUC) values of each signature with its confidence interval (CI), by applying a general linear model to gene expression values from each comparison (TB and HC or TBDM vs HC, within the countries). The detailed performance of each signature in each clinical conditions is shown in the Supplementary File S2. Moreover, linear modeling allows the comparison of performance from signatures composed by DEGs and composed by scores, as well as a fair comparison between

them[46,47]. The outcomes were binarized to measure the sensitivity and specificity of classification, allowing us to measure each group rate and plot each signature AUC and CI value, from each country. This allows the direct performance comparison of each signature in data from different countries and different conditions (TB or TBDM).

Single sample gene set enrichment analysis (ssGSEA). The normalized enrichment scores (NES) from each sample were calculated with ssGSEA[48] using the Reactome database[39]. Only significant NES values were used (FDR < 0.05 and 100 permutations) to perform correlation analysis with the enriched pathways, HbA1c values and gene expression levels. Correlation relationships were depicted as chord diagrams and networks, performed by *circlize*[49] and *igraph*[50] package.

Data availability

The dataset from the TANDEM cohort analyzed during the current study is available at the BioProject data repository, identified by the accession code PRJNA470512 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA470512). The MSTDI gene expression data is available at the geoNCBI data repository, identified by the accession number GSE181143 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE181143).

Received: 27 February 2023; Accepted: 9 May 2023
Published online: 12 May 2023

References

1. Critchley, J. A. *et al.* Defining a research agenda to address the converging epidemics of tuberculosis and diabetes: Part 1: Epidemiology and clinical management. *Chest* 152, 165–173 (2017).
2. Ronacher, K. *et al.* Defining a research agenda to address the converging epidemics of tuberculosis and diabetes: Part 2: Underlying biologic mechanisms. *Chest* 152, 174–180 (2017).
3. Jeon, C. Y. & Murray, M. B. Diabetes mellitus increases the risk of active tuberculosis: A systematic review of 13 observational studies. *PLoS Med.* 5, e152 (2008).
4. Odone, A., Houben, R. M., White, R. G. & Lonnroth, K. The effect of diabetes and undernutrition trends on reaching 2035 global tuberculosis targets. *Lancet Diab. Endocrinol.* 2, 754–764 (2014).
5. Martinez, N. & Kornfeld, H. Diabetes and immunity to tuberculosis. *Eur. J. Immunol.* 44, 617–626 (2014).
6. Blankley, S. *et al.* The application of transcriptional blood signatures to enhance our understanding of the host response to infection: the example of tuberculosis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130427 (2014).
7. Prada-Medina, C. A. *et al.* Systems immunology of diabetes-tuberculosis comorbidity reveals signatures of disease complications. *Sci. Rep.* 7, 1999 (2017).
8. Eckold, C. *et al.* Impact of intermediate hyperglycemia and diabetes on immune dysfunction in tuberculosis. *Clin. Infect. Dis.* 72, 69–78 (2021).
9. Magee, M. J. *et al.* Convergence of non-communicable diseases and tuberculosis: A two-way street?. *Int. J. Tuberc. Lung Dis.* 22, 1258–1268 (2018).
10. Oliveira-de-Souza, D. *et al.* Molecular degree of perturbation of plasma inflammatory markers associated with tuberculosis reveals distinct disease profiles between Indian and Chinese populations. *Sci. Rep.* 9, 8002 (2019).
11. Kumar, N. P. *et al.* Persistent inflammation during anti-tuberculosis treatment with diabetes comorbidity. *eLife* https://doi.org/10.7554/eLife.46477 (2019).
12. Singhania, A., Wilkinson, R. J., Rodrigue, M., Haldar, P. & O'Garra, A. The value of transcriptomics in advancing knowledge of the immune response and diagnosis in tuberculosis. *Nat. Immunol.* 19, 1159–1168 (2018).
13. Kumar, N. P. *et al.* Heightened circulating levels of antimicrobial peptides in tuberculosis-diabetes co-morbidity and reversal upon treatment. *PLoS One* 12, e0184753 (2017).
14. Kumar, N. P. *et al.* Association of plasma matrix metalloproteinase and tissue inhibitors of matrix metalloproteinase levels with adverse treatment outcomes among patients with pulmonary tuberculosis. *JAMA Netw. Open* 3, e2027754 (2020).
15. Martens, G. W. *et al.* Tuberculosis susceptibility of diabetic mice. *Am. J. Respir. Cell Mol. Biol.* 37, 518–524 (2007).
16. Vallerskog, T., Martens, G. W. & Kornfeld, H. Diabetic mice display a delayed adaptive immune response to *Mycobacterium tuberculosis*. *J. Immunol.* 184, 6275–6282 (2010).
17. Martinez, N., Ketheesan, N., West, K., Vallerskog, T. & Kornfeld, H. Impaired recognition of *Mycobacterium tuberculosis* by alveolar macrophages from diabetic mice. *J. Infect. Dis.* 214, 1629–1637 (2016).
18. Martinez, N. & Kornfeld, H. Diabetes and immunity to tuberculosis. *Eur. J. Immunol.* 44, 617–626 (2014).
19. Barreda, N. N. *et al.* Severe pulmonary radiological manifestations are associated with a distinct biochemical profile in blood of tuberculosis patients with dysglycemia. *BMC Infect. Dis.* 20, 139 (2020).
20. Eckold, C. *et al.* Impact of intermediate hyperglycaemia as well as diabetes on immune dysfunction in tuberculosis. *Clin. Infect. Dis.* 72, 69–78 (2020).
21. Kumar, N. P. *et al.* Type 2 diabetes mellitus coincident with pulmonary tuberculosis is associated with heightened systemic type 1, type 17 and other pro-inflammatory cytokines. *Ann. Am. Thorac. Soc.* 10, 441–449 (2013).
22. Corrales-Medina, V. F. *et al.* Association between hospitalization for pneumonia and subsequent risk of cardiovascular disease. *JAMA* 313, 264–274 (2015).
23. Xie, Y., Xu, E., Bowe, B. & Al-Aly, Z. Long-term cardiovascular outcomes of COVID-19. *Nat. Med.* 28, 583–590 (2022).
24. Basham, C. A., Smith, S. J., Romanowski, K. & Johnston, J. C. Cardiovascular morbidity and mortality among persons diagnosed with tuberculosis: A systematic review and meta-analysis. *PLoS One* 15, e0235821 (2020).
25. Penn-Nicholson, A. *et al.* RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629 (2020).
26. Degner, N. R., Wang, J. Y., Golub, J. E. & Karakousis, P. C. Metformin use reverses the increased mortality associated with diabetes mellitus during tuberculosis treatment. *Clin. Infect. Dis.* 66, 198–205 (2018).
27. Padmapriydarsini, C. *et al.* Randomized trial of metformin with anti-tuberculosis drugs for early sputum conversion in adults with pulmonary tuberculosis. *Clin. Infect. Dis.* https://doi.org/10.1093/cid/ciab964 (2021).
28. Hamilton, C. D. *et al.* RePORT international: Advancing tuberculosis biomarker research through global collaboration. *Clin. Infect. Dis.* 61(Suppl 3), S155-159 (2015).
29. Kornfeld, H. *et al.* High prevalence and heterogeneity of diabetes in patients with TB in South India: A report from the Effects of Diabetes on Tuberculosis Severity (EDOTS) study. *Chest* 149, 1501–1508 (2016).

30. Gupte, A. *et al.* Cohort for tuberculosis research by the Indo-US medical partnership (CTRIUMPH): Protocol for a multicentric prospective observational study. *BMJ Open* 6, e010542 (2016).
31. Linh, N. N. *et al.* World health organization treatment outcome definitions for tuberculosis: 2021 update. *Eur. Respir. J.* 58, 2100804 (2021).
32. van Crevel, R. & Dockrell, H. M. TANDEM: Understanding diabetes and tuberculosis. *Lancet Diab. Endocrinol.* 2, 270–272 (2014).
33. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419 (2017).
34. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research* 4, 1521 (2015).
35. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140 (2010).
36. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300 (1995).
37. Chen, H. & Boutros, P. C. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinform.* 12, 35 (2011).
38. Guangchuang, Y., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287 (2012).
39. Yu, G. & He, Q. Y. ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualization. *Mol. BioSyst.* 12, 477–479 (2016).
40. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883 (2012).
41. Ward, J. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58, 236–244 (1963).
42. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849 (2016).
43. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* 28, 1–26 (2008).
44. Breiman, L. Random forests. *Mach. Learn.* 45, 2–32 (2001).
45. Zou, K. H., O'Malley, A. J. & Mauri, L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* 115, 654–657 (2007).
46. Kulkarni, V. *et al.* A two-gene signature for tuberculosis diagnosis in persons with advanced HIV. *Front. Immunol.* 12, 631165 (2021).
47. Mathad, J. S. *et al.* Transcriptional analysis for tuberculosis in pregnant women from the PRACHITI study. *Clin. Infect. Dis.* 75, 2239–2242 (2022).
48. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550 (2005).
49. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. Circlize implements and enhances circular visualization in R. *Bioinformatics* 30, 2811–2812 (2014).
50. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJ. Complex Syst.* 1695, 1–9 (2006).

Acknowledgements
We thank Ms. Daphne Martin and Ms. Samyra Cox for outstanding administrative support.

Author contributions
K.F.F., A.T.Q.L., C.L.V., E.R.F. and M.B.A. performed the data curation, analysis, interpretation, and draft of the first version of the manuscript. A.N.G., S.G., R.K., V.M., M.P. and A.G. performed the data interpretation, revising manuscript critically for important intellectual content, final approval of the manuscript. N.P.K. performed sample preparation and curation for the Chennai cohort. T.R.S. supervised the Brazilian study and helped with data interpretation. S.B. and V.V. coordinated the clinical study in Chennai. B.B.A. performed the study conceptualization, data analysis and interpretation, and draft of the manuscript. H.K. performed the study conceptualization, data curation, analysis and interpretation, revising manuscript critically for important intellectual content and coordinated all sites studies.

Funding
This work was funded by: OISE-17-63459-1from the National Institutes of Health, administered by CRDF Global; DAA3-18-64718-1, formerly USB1-31149-XX-13 from the Indo-US Vaccine Action Initiative on TB Research, administered by CRDF Global. The Brazilian site was supported by the National Institutes of Health (NIH U01AI069923 and R01AI120790), CCASAnet, RePORT-Brazil Tennessee Center for AIDS Research (TNC-FAR). The study was also supported by the Intramural Research Program of the Fundação José Silveira, and the Intramural Research Program of the Oswaldo Cruz Foundation, Brazil. BBA is a senior investigator from the Brazilian Council for Science and Technology (CNPq) and MBA received research fellowship from the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-34847-9.

Correspondence and requests for materials should be addressed to B.B.A. or H.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023

---

the RePORT Brazil

Alice M. S. Andrade[2,3,13], Marina C. Figueiredo[7], Vanessa Nascimento[16], Juan Manuel Cubillos-Angulo[3], Hayna Malta-Santos[3], Jéssica Rebouças-Silva[2,17], Adriano Gomes-Silva[18], Saulo R. N. Santos[16], André Ramos[16], Pedro Brito[2,3], Carolina A. S. Schmaltz[5], Alysson G. Costa[19,20,21], Leandro Sousa Garcia[19,20], Brenda K. de Sousa Carvalho[19,20], Bruna P. de Loiola[19,20], Francine P. Ignácio[18], Maria C. Lourenço[18], Elisangela C. Silva[19], Mayla Mello[19], Alexandra B. Souza[19,20], Michael S. Rocha[3,16], Aline Benjamin[18], Adriana S. R. Moreira[20], Jamile G. de Oliveira[18], Solange Cavalcante[18], Betina Durovni[18], Marcelo Cordeiro-Santos[19,20], Afrânio L. Kristki[22], Valeria C. Rolla[23] & José R. Lapa-e-Silva[24]

[16]Instituto Brasileiro para Investigação da Tuberculose, Fundação José Silveira, Salvador, Brazil. [17]Faculdade de Medicina, Universidade Federal da Bahia, Salvador, Brazil. [18]Laboratório de Pesquisa Clínica em Micobacteriose, Instituto Nacional de Infectologia Evandro Chagas, Fiocruz, Rio de Janeiro, Brazil. [19]Fundação Medicina Tropical Dr Heitor Vieira Dourado, Manaus, Brazil. [20]Programa de Pós-Graduação em Medicina Tropical, Universidade do Estado do Amazonas, Manaus, Brazil. [21]Universidade Federal do Amazonas, Manaus, Brazil. [22]Programa Acadêmico de Tuberculose da Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. [23]Instituto Nacional de Infectologia Evandro Chagas, Fiocruz, Rio de Janeiro, Brazil. [24]Faculdade de Medicina, Programa Acadêmico de Tuberculose, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

RePORT India Consortia

Kim West[15], Vandana Kulkami[9] & Nikhil Gupte[5]

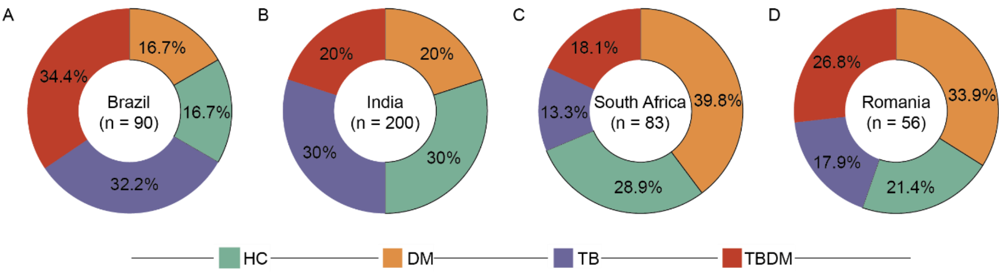
Supplementary Material

Country	Healthy controls	DM	TB	TBDM	P-value
Brazil	15	15	29	31	
India	60	40	60	40	
Total (N)	75	55	89	71	
Age – y	35 (30-42)	54 (44-59)	36 (28-46.5)	48 (40-55)	**<0.001**
Female – no. (%)	47 (62%)	24 (43%)	25 (28%)	15 (21%)	**<0.001**
BMI (Kg/m^2)	20.9 (15.3-24.5)	27.9 (24.8-30)	18.7 (17.3-21)	22.2 (19-27.3)	**<0.001**
Smoking (current)	7 (9%)	13 (23%)	23 (25%)	18 (25%)	**0.03**
Alcohol (current)	24 (32%)	23 (42%)	42 (47%)	33 (46%)	0.19
Metformin	N/A	26 (47%)	N/A	33 (46%)	0.9
Statin	N/A	3 (5%)	N/A	9 (12%)	0.22
Cavitary TB	N/A	N/A	37 (41.5%)	25 (35.2%)	0.42
HbA1c (%)	5.1 (5-5.5)	9 (7.7-11)	5.4 (5-5.7)	10.9 (8-12.1)	**<0.001**

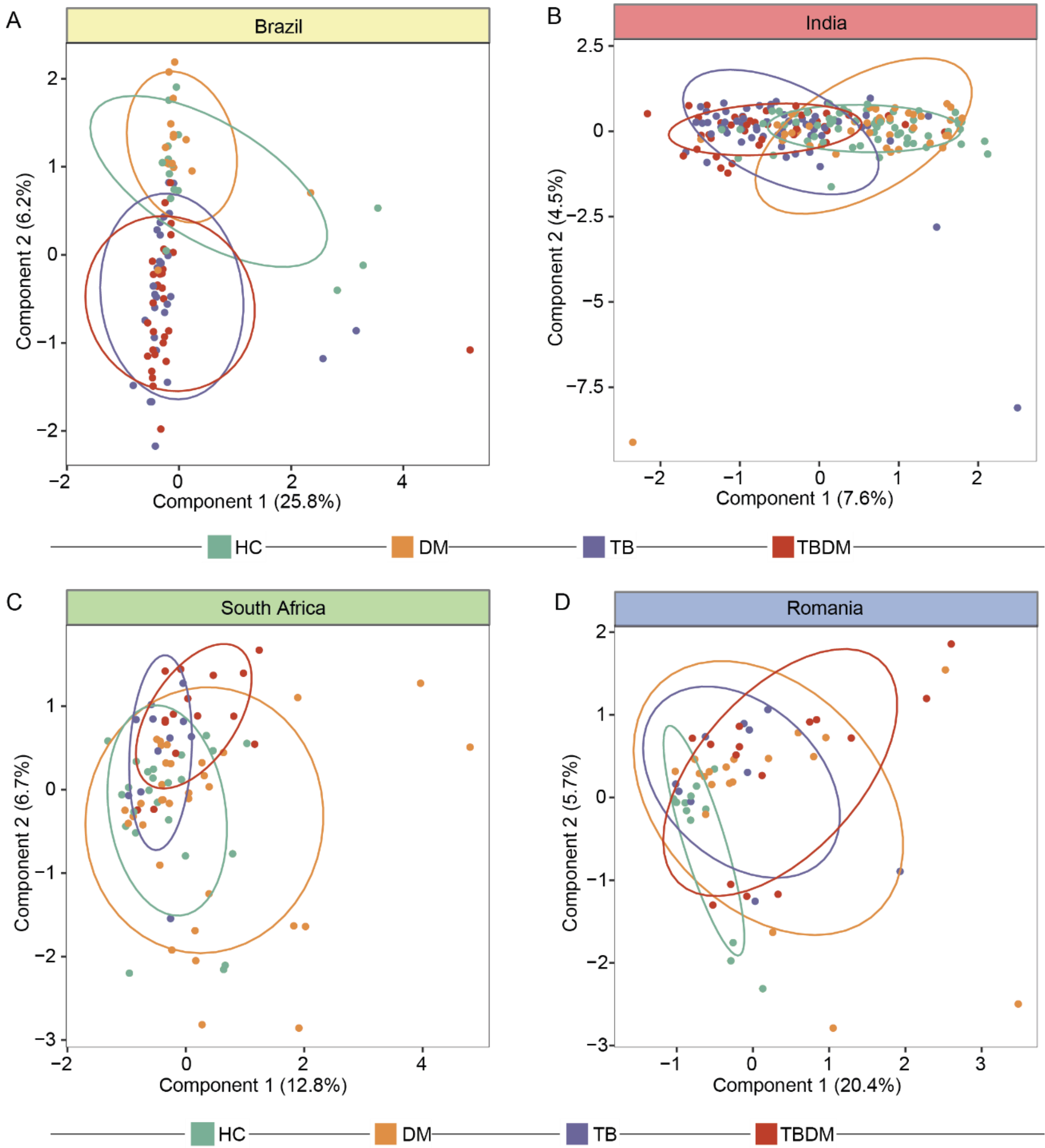
Supplementary Table S1: Characteristics of study population according to clinical groups among the sites of study. Data represent medians and interquartile ranges (age, BMI and HbA1c) and frequencies (female gender, smoking current use, alcohol current use, metformin and statin use and cavitary TB). The Kruskal-Wallis test was used to compare distributions of age, BMI and HbA1c while the Chi-square test was used to compare frequencies of gender, smoking current use, alcohol current use and cavitary, whereas Fisher Exact test used to compare frequencies of metformin and statin use TB. P-values in bold font are statistically significant. Abbreviations: BMI, Body Mass Index; HbA1c, glycosylated hemoglobin, N/A, Non-Applicable.

	Healthy controls			DM			TB			TBDM		
	Brazil	India		Brazil	India		Brazil	India		Brazil	India	
N	15	60	P-value	15	40	P-value	29	60	P-value	31	40	P-value
Age – y	35 (28-51)	35 (31-39.7)	0.64	56 (51-57)	52.5 (42.5-78)	0.46	29 (25-43.5)	39.5 (31-67)	0.002	48 (38-67)	48 (40.2-65)	0.7
Female – no. (%)	9 (60%)	32 (53%)	0.6	8 (57%)	22 (55%)	0.88	10 (34%)	15 (24%)	0.35	9 (29%)	10 (25%)	0.7
BMI (Kg/m^2)	24.9 (20.5-28.9)	16.9 (16-20)	0.2	30.3 (26-32)	25.4 (23.6-27.7)	<0.001	20.6 (18.6-22.1)	16.9 (16-20)	<0.001	22.5 (20-25.7)	21.9 (18-28.9)	0.79
Smoking (current)	5 (33.3%)	2 (3.3%)	0.004	3 (20%)	10 (25%)	0.6	8 (27.6%)	15 (25%)	0.7	12 (38.7%)	6 (15%)	0.02
Alcohol (current)	13 (86.7%)	11 (18.4%)	<0.001	14 (93.4%)	9 (22.5%)	<0.001	25 (86.2%)	17 (28.4%)	<0.001	28 (90.3%)	5 (12.5%)	<0.001
Metformin	N/A	N/A	N/A	Not assessed	26 (6%)	Not assessed	N/A	N/A	N/A	6 (19.4%)	27 (87%)	<0.001
Statin	N/A	N/A	N/A	Not assessed	3 (7.5%)	Not assessed	N/A	N/A	N/A	Not assessed	9 (22.5%)	<0.001
Cavitary TB	N/A	N/A	N/A	N/A	N/A	N/A	15 (51.7%)	22 (36.7%)	0.17	9 (29%)	26 (65%)	<0.001
HbA1c (%)	5.1 (4.9-5.2)	5 (5-5.5)	0.25	6.1 (5.9-7.4)	9.4 (8.4-11.1)	<0.001	5.5 (5.2-5.6)	5.3 (5-5.9)	0.84	8.5 (6.8-11.4)	11.7 (10-12.5)	0.001

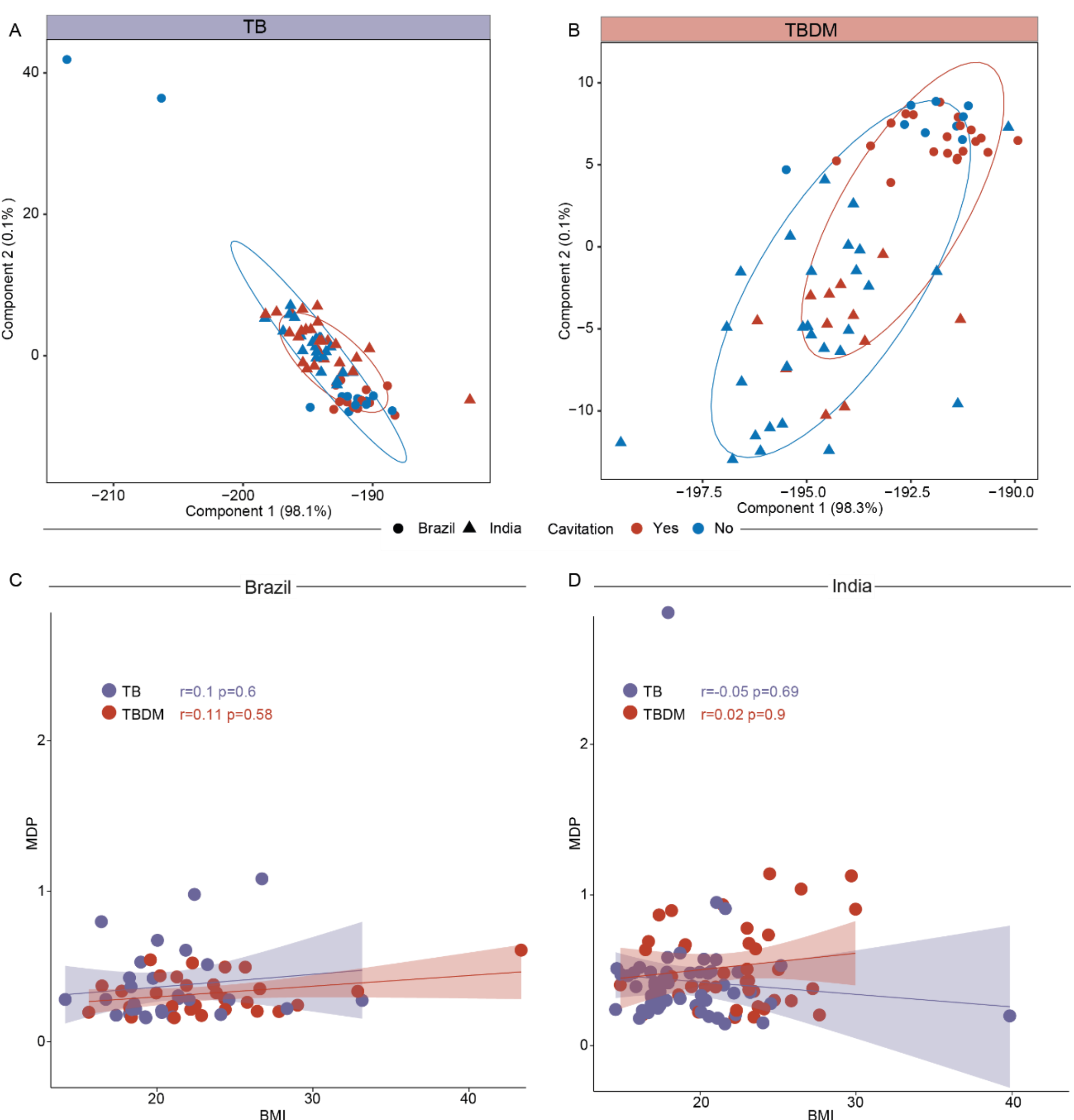
Supplementary Table S2: Characteristics of study population according to clinical groups among the sites of study. Data represent medians and interquartile ranges (age, BMI and HbA1c) and frequencies (female gender, smoking current use, alcohol current use, metformin and statin use and cavitary TB). The Kruskal-Wallis test was used to compare distributions of age, BMI and HbA1c while the Chi-square test was used to compare frequencies. P-values in bold font are statistically significant. Abbreviations: BMI, Body Mass Index; HbA1c, glycosylated hemoglobin, N/A, Non-Applicable.



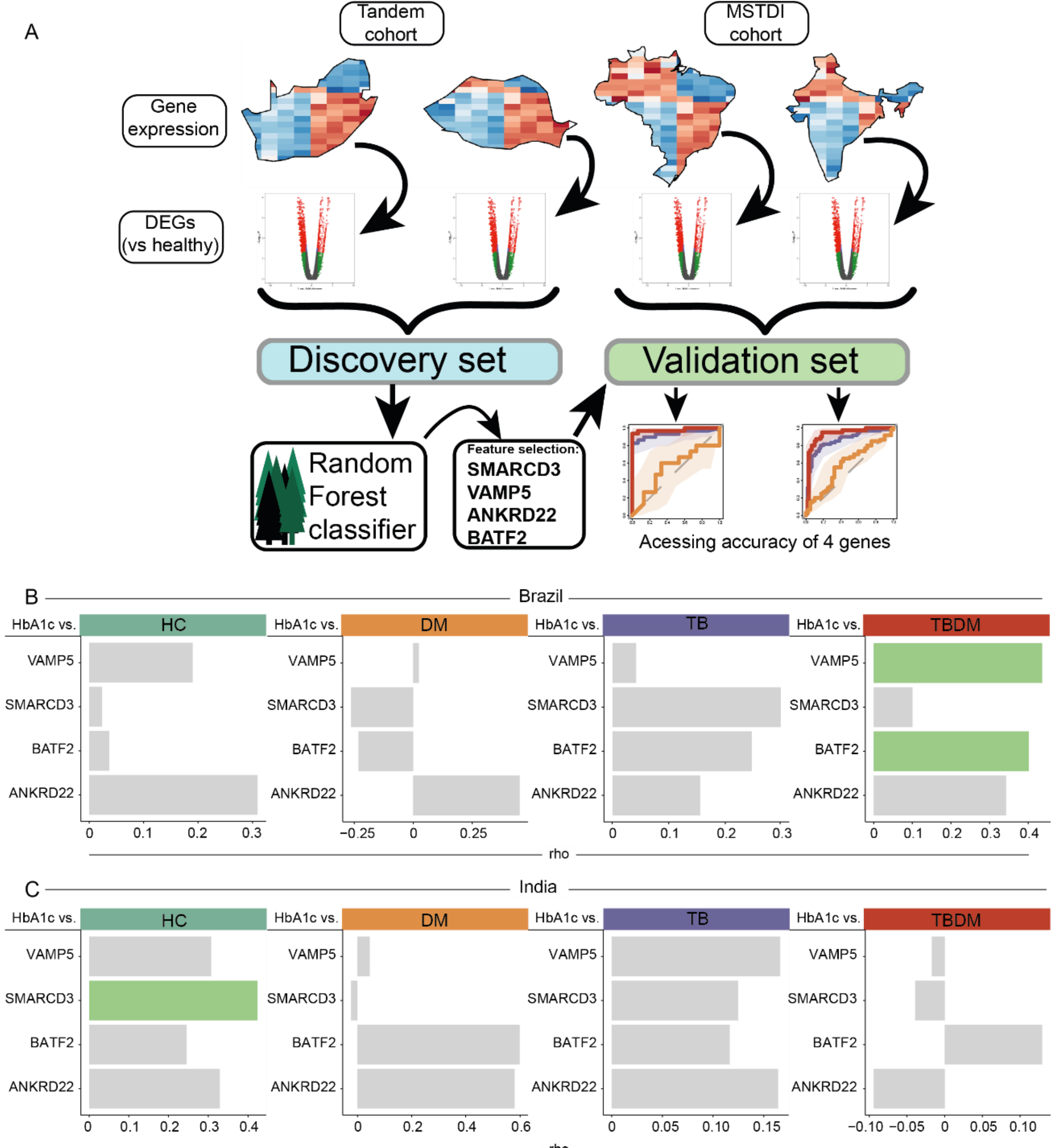
Supplementary Fig. S1. Sample distribution in the MSTDI and TANDEN cohort. Colored sections show the percentage of individuals at (A) Brazil, (B) India, and (C) South Africa and (D) Romania in the healthy control group (HC), diabetic control group (DM), and the non-diabetic and diabetic pulmonary TB groups (TB and TBDM, respectively).



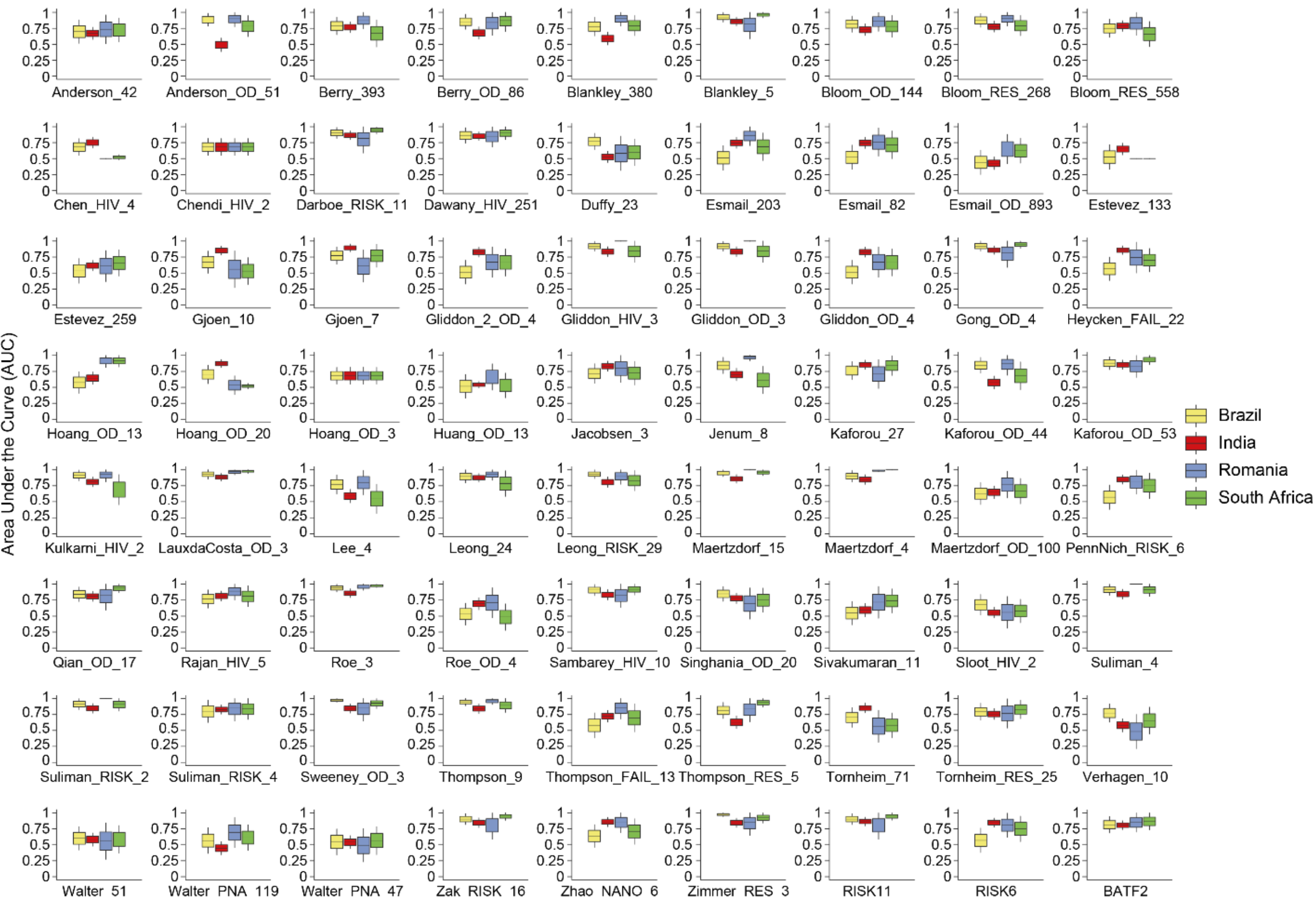
Supplementary Fig. S2. Identifying TB participants using Differential Expression Genes (DEGs). A principal component model was employed to test whether DEGs could cluster the patients in each clinical site, as indicated.



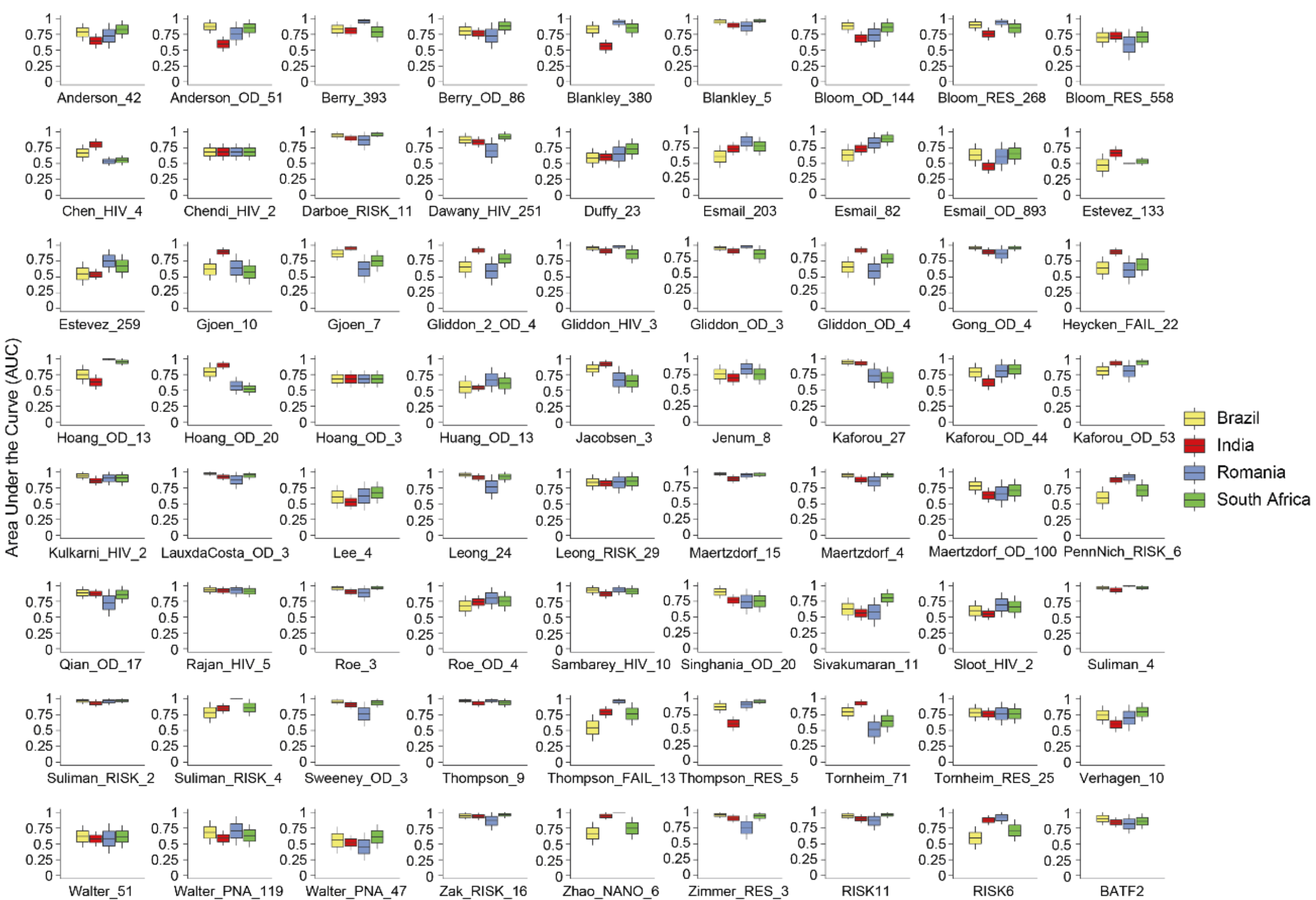
Supplementary Fig. S3. Impact of clinical and epidemiological features in the gene expression variability. (A-B) A principal component model was employed to test whether the presence or absence of cavitation could explain the differences in the gene expression profile between the countries. (C-D) A Spearman correlation analysis was used to evaluate if changes in BMI values is associated with the degree of the molecular degree of perturbation (MDP) in each clinical group.



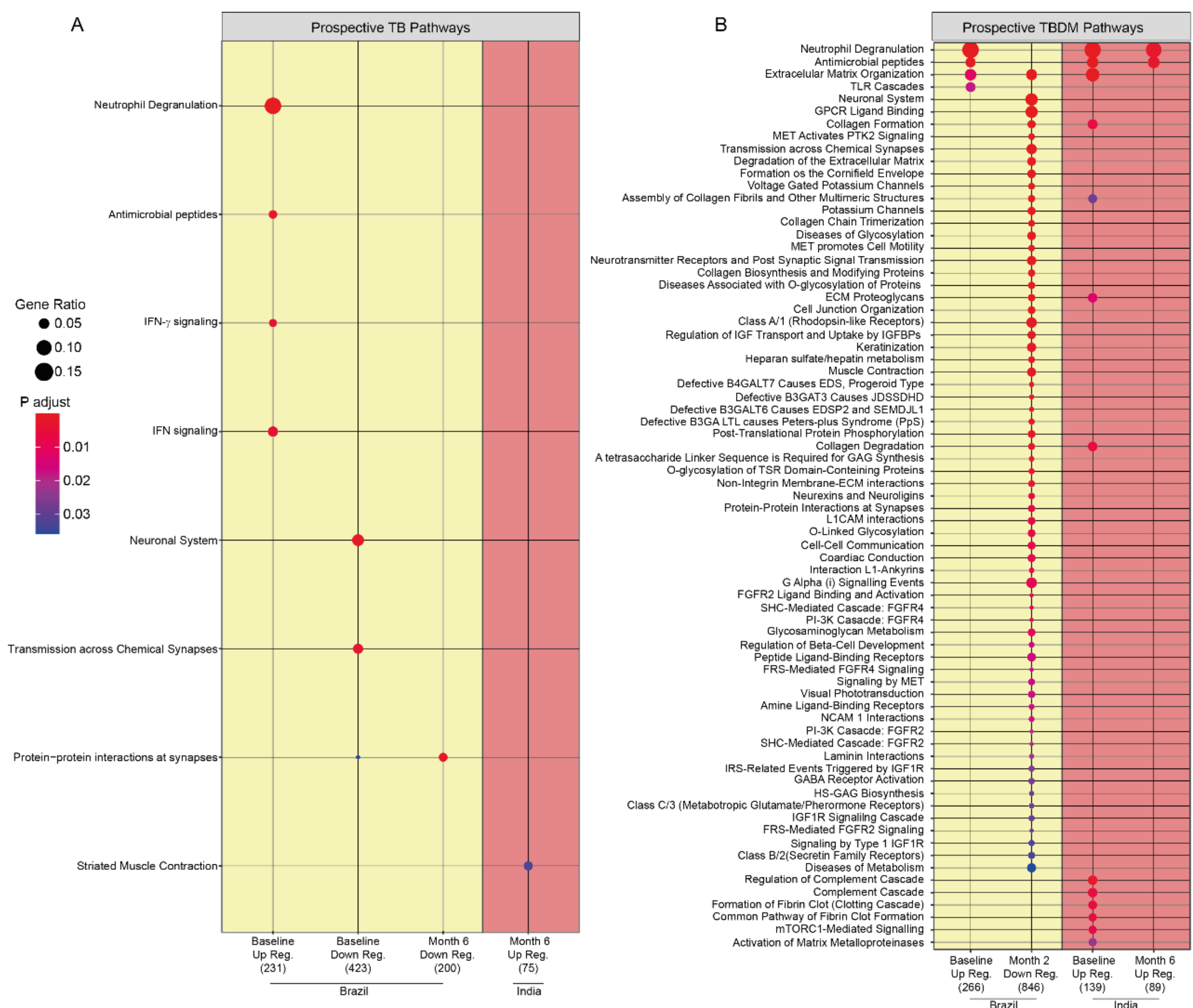
Supplementary Fig. S4. Identifying the top genes expression associated with TBDM. (A)DEGs values were inputted for a Random Forest model. South Africa and Romania sites were used as discovery set and the accuracy of the model was tested in the validation set composed of samples from the Brazil and India site. (B) A Spearman correlation analysis were performed between top genes expression and levels of HbA1c in each group from Brazil and India. Correlations with p-value < 0.05 were indicated as green bars.



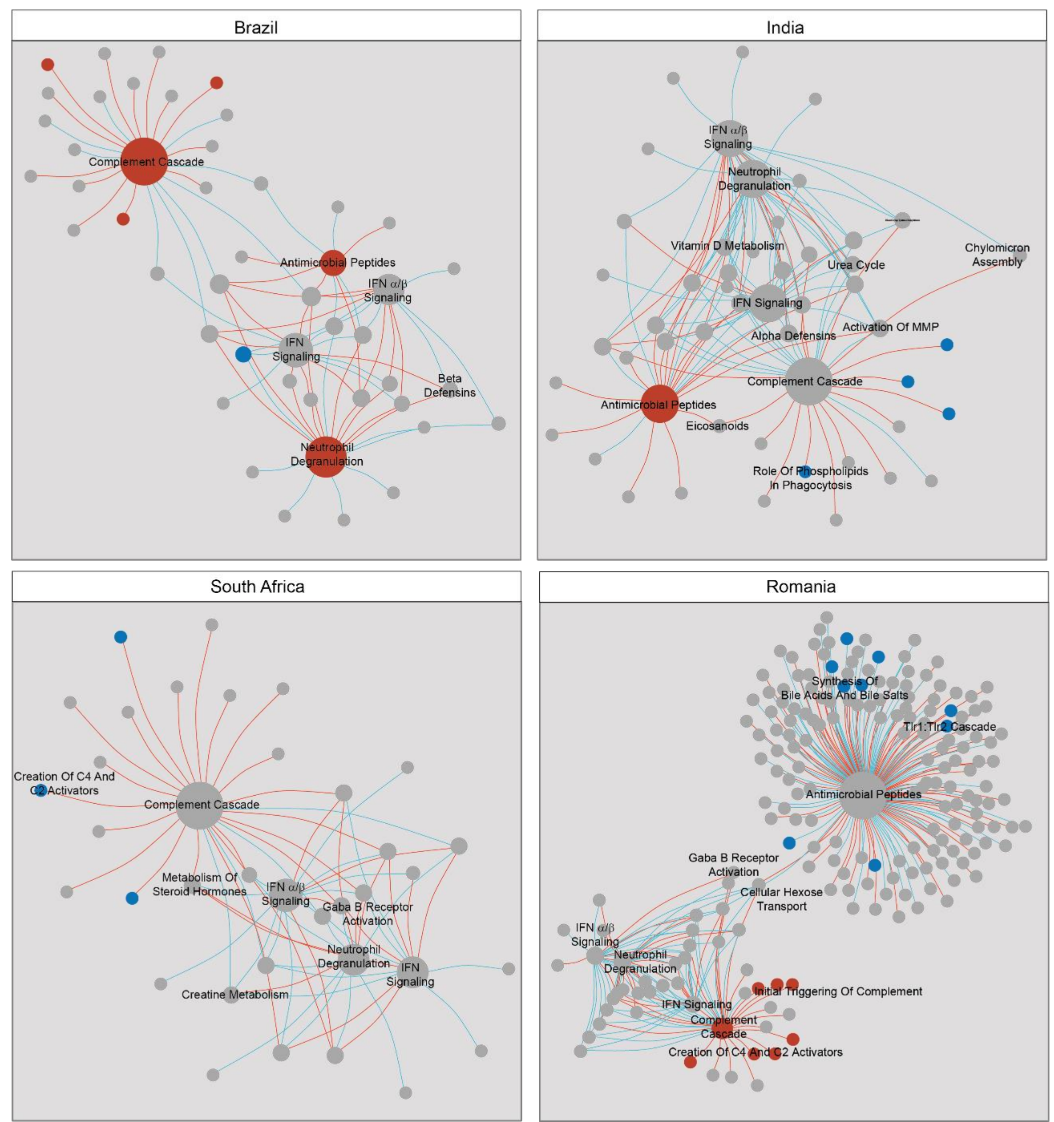
Supplementary Fig. S5. Assessing performance of previously published gene biosignatures to identify TB cases in each country. Receiver Operator Characteristics (ROC) analysis were performed to test the accuracy of previous reported signatures in our TB population. The y-axis shows the area under curve (AUC) values.



Supplementary Fig. S6. Assessing performance of previously published gene biosignatures to identify TB-diabetes cases in each country. Receiver Operator Characteristics (ROC) analysis were performed to test the accuracy of previous reported signatures in our TB population. The y-axis shows the area under curve (AUC) values.



Supplementary Fig. S7. Changes in pathways expression after started antitubercular therapy. The colored spots indicate the enriched pathways identified from the DEGs of the comparisons of (A) TB and (B) TBDM in each time point with the HC. The sites are highlighted by colors: pathways from Brazi are colored yellow and India are colored red. The color gradient of spots corresponds to the FDR-corrected p-value and the size corresponds to the gene ratio.



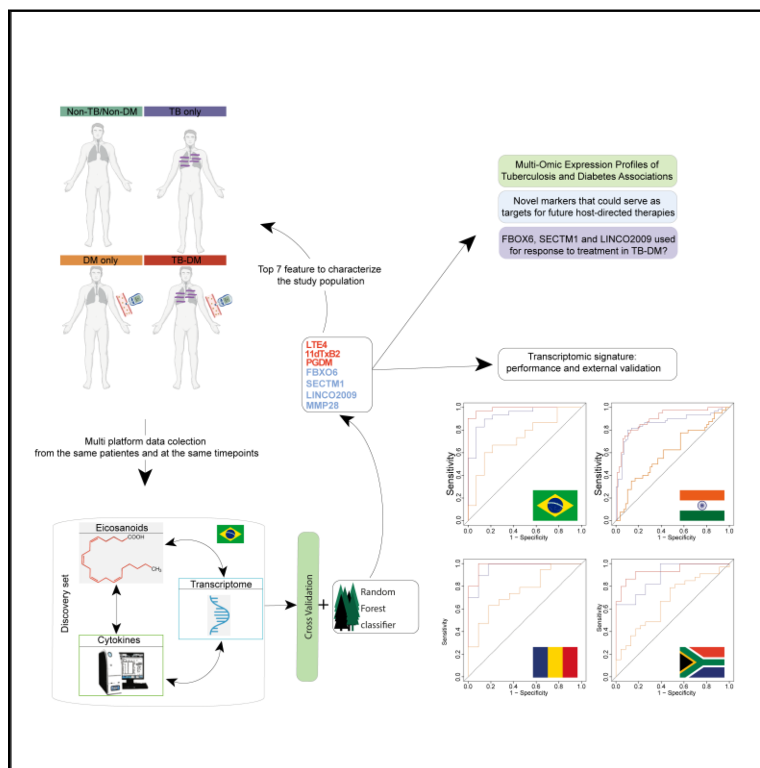
Supplementary Fig. S8. Changes in the dynamicity of pathways across the clinical sites in TB. A Spearman correlation analysis was performed using the pathways from TB participants in each clinical site, as indicated. Each node indicated a pathway, blue node infers downregulation when compared with control group, whereas red nodes

6.2 Artigo II

iScience

Article

An integrative multi-omics approach to characterize interactions between tuberculosis and diabetes mellitus



Caian L. Vinhaes, Eduardo R. Fukutani, Gabriel C. Santana, ..., Bruno B. Andrade, Artur T.L. Queiroz, for the RePORT Brazil Consortium

bruno.andrade@fiocruz.br

Highlights

A distinct multi-omic signature characterizes TB regardless of DM status

Anti-tubercular therapy leads to decreases in expression of the multi-omic signature

The biomarkers identified may serve as candidates for host-directed therapies

Vinhaes et al., iScience 27, 109135
March 15, 2024 © 2024 The Author(s).
https://doi.org/10.1016/j.isci.2024.109135

iScience

Article

An integrative multi-omics approach to characterize interactions between tuberculosis and diabetes mellitus

Caian L. Vinhaes,[1,2,3,4,5,16] Eduardo R. Fukutani,[2,5,6,16] Gabriel C. Santana,[1,2,7] María B. Arriaga,[8] Beatriz Barreto-Duarte,[1,2,5,7,9] Mariana Araújo-Pereira,[1,2,5,10] Mateus Maggiti-Bezerril,[2,5] Alice M.S. Andrade,[1,2,5] Marina C. Figueiredo,[8] Ginger L. Milne,[11] Valeria C. Rolla,[12] Afrânio L. Kristki,[9] Marcelo Cordeiro-Santos,[13,14,15] Timothy R. Sterling,[8] Bruno B. Andrade,[1,2,3,5,7,8,10,17,18,*] Artur T.L. Queiroz,[2,5,6,17] and for the RePORT Brazil Consortium

SUMMARY

Tuberculosis-diabetes mellitus (TB-DM) is linked to a distinct inflammatory profile, which can be assessed using multi-omics analyses. Here, a machine learning algorithm was applied to multi-platform data, including cytokines and gene expression in peripheral blood and eicosanoids in urine, in a Brazilian multi-center TB cohort. There were four clinical groups: TB-DM(n = 24), TB only(n = 28), DM(HbA1c ≥ 6.5%) only(n = 11), and a control group of close TB contacts who did not have TB or DM(n = 13). After cross-validation, baseline expression or abundance of MMP-28, LTE-4, 11-dTxB2, PGDM, FBXO6, SECTM1, and LINC02009 differentiated the four patient groups. A distinct multi-omic-derived, dimensionally reduced, signature was associated with TB, regardless of glycemic status. SECTM1 and FBXO6 mRNA levels were positively correlated with sputum acid-fast bacilli grade in TB-DM. Values of the biomarkers decreased during the course of anti-TB therapy. Our study identified several markers associated with the pathophysiology of TB-DM that could be evaluated in future mechanistic investigations.

INTRODUCTION

Tuberculosis (TB) affects approximately 10 million people annually, and despite being a preventable and curable disease, 1.6 million people died from TB in 2021.[1] About one-quarter of the world's population is estimated to be infected with *Mycobacterium tuberculosis* (Mtb).[2]

After infecting persons via the lungs, Mtb can disseminate and affect several other organ systems, resulting in a large potential range of clinical manifestations and varying degrees of severity depending on the interaction between Mtb and host immune responses. Factors that affect immune responses can influence the inflammatory reactions against Mtb; these include infection with HIV,[3] and genetic,[4] environmental,[5] nutritional,[6] and metabolic features such as dysglycemia and diabetes mellitus (DM).

DM is one of the most common metabolic diseases worldwide. Currently, 422 million people are living with DM,[7] most of whom reside in low and middle-income countries. Approximately 1.5 million deaths have been attributed to DM annually.[7] Sustained and uncontrolled hyperglycemia leads to changes in immune regulation[8]; therefore, people living with DM are considered immunosuppressed hosts.[9] The

[1]Laboratório de Pesquisa Clínica e Translacional, Instituto Gonçalo Moniz, FUndação Oswaldo Cruz, Salvador 40296-710, Brazil
[2]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador 41810-710, Brazil
[3]Programa de Pós-Graduação em Medicina e Saúde Humana, Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador 40290-150, Brazil
[4]Departamento de Infectologia, Hospital Português da Bahia, Salvador 40140-901, Brazil
[5]Instituto de Pesquisa Clínica e Translacional, Faculdade de Tecnologia e Ciências, Salvador 41741-590, Brazil
[6]Centro de Integração de Dados e Conhecimentos para Saúde, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil
[7]Curso de Medicina, Universidade Salvador, Salvador, Brazil
[8]Division of Infectious Diseases, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
[9]Programa Acadêmico de Tuberculose. Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[10]Faculdade de Medicina, Univerdidade Federal da Bahia, Salvador, Brazil
[11]Division of Clinical Pharmacology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA
[12]Instituto Nacional de Infectologia Evandro Chagas, Fiocruz, Rio de Janeiro, Brazil
[13]Fundação Medicina Tropical Doutor Heitor Vieira Dourado, Manaus, Brazil
[14]Programa de Pós-Graduação em Medicina Tropical, Universidade do Estado do Amazonas, Manaus, Brazil
[15]Universidade Nilton Lins, Manaus, Brazil
[16]These authors contributed equally
[17]These authors contributed equally
[18]Lead contact
*Correspondence: bruno.andrade@fiocruz.br
https://doi.org/10.1016/j.isci.2024.109135

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

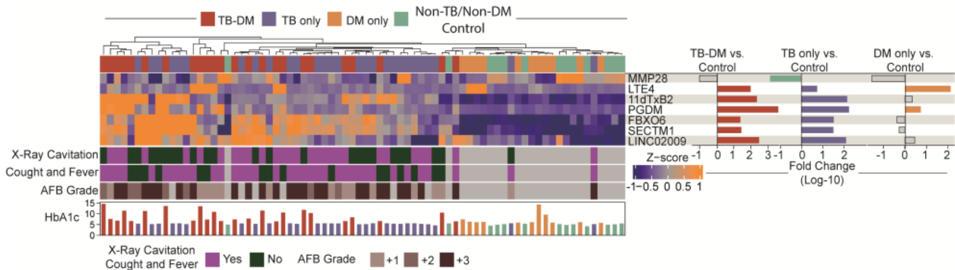


Figure 1. Distinct multi-omic expression profiles identified tuberculosis regardless the glycemic status

Right panel. A hierarchical cluster analysis (Ward method with 100 × bootstrap) was employed to test the overall expression of plasma cytokines, gene expression and urinary eicosanoids in the study population. Dendrograms represent Canberra distance. *Left panel.* Differential expression analysis was used to calculate the fold-changes and show differences in biomarkers levels for each clinical group (TB, TB-DM, and DM) versus control. Differences that reached statistical significance after adjustment for multiple comparisons (adjusted $p < 0.05$) are represented in colored bars.

association between TB and DM has important health consequences. Previous studies have revealed that DM increases the risk of incident TB 3-fold and increases the risk of TB unfavorable outcomes 2-fold.[10,11]

Our group has studied TB-DM interactions, including activation of transcriptional pathways in disease complications,[12] persistent inflammation during anti-TB treatment (ATT),[13] increased Mtb transmission to close contacts,[14] more severe TB clinical presentation,[15] and increased risk of unfavorable ATT outcomes.[16–18] However, while most of the studies used different assays, which were analyzed separately to characterize TB-DM, the use of multiple omics platforms remains poorly explored in the context of TB-DM interactions. Multi-omics provides the opportunity to gain insights into disease pathogenesis, given that it simultaneously explores several components of immune responses through multiple assay platforms. This is a powerful tool to generate hypotheses pertaining to the intricate molecular relationships that may be driving disease. To our knowledge, no previous investigation has examined whether multi-omics analyses can help characterize TB-DM pathogenesis. To fill this knowledge gap, we leveraged samples and data collected in the multi-center prospective cohort of the Regional Prospective Observational Research in Tuberculosis (RePORT)-Brazil (www.reportbrazil.org). We used transcriptomic and cytokine data from peripheral blood and eicosanoids measured in urine, collected at the same time points from persons with TB-DM, TB only, DM only, and non-TB/non-DM close contacts (control group). We applied a random forest model to investigate a possible multi-omic signature that could characterize TB-DM. Furthermore, we tracked the expression of the components of the multi-omic signature in the study groups during the course of ATT. Our results identified novel pathways that may drive disease pathogenesis and serve as targets for future host-directed therapies for TB-DM and as markers of progression to cure after antitubercular therapy initiation. Additionally, the transcriptomic markers identified here have shown a robust accuracy to detect TB-DM in an external validation study using cohorts from India, South Africa, and Romania.

RESULTS

Characteristics of study population

All patients included in our analysis were classified according to their status regarding TB disease and DM. The groups were composed of 24 with TB-DM, 28 TB only, 11 DM only, and 13 non-TB/non-DM (controls). TB-DM and DM only participants were older, with a median age of 46.5 (37–55.7) and 56 (IQR:51–59) years, respectively ($p < 0.001$; Table S1). In addition, controls and participants with DM were more frequently female than those with TB-DM and TB only. Higher BMI was found in the controls and in DM only groups, with a median (IQR) of 29 (24–33) and 29 (27–32), respectively, compared to 20 (18–22) and 22 (18–26) in TB only and TB-DM patients ($p < 0.001$) (Table S1). The levels of HbA1c were higher in those with TB-DM (median [IQR] 8.1 [6.6–11.3], followed by DM only 7.1 [6.9–8.4] $p < 0.001$). Of note, TB only patients were not significantly different regarding AFB (acid-fast bacilli) smear grade and the proportion with cavities on chest X-ray (Table S1).

Multi-omic expression profiles of tuberculosis and diabetes associations

To address the aim of finding a distinct profile that distinguishes TB regardless of DM, our first step was to study the multi-omic alterations in TB and DM using a well-established dimensionality reduction approach (Figure S1). This approach selected 7 parameters that could most contribute to discrimination between all the study groups (Figure S1). The Gini score and mean accuracy for each marker used in our model to discriminate between the clinical groups were provided in Table S2 and Figure S2. To further evaluate the baseline multi-omic profiles of the clinical groups, we first examined the expression of the seven selected features (Figure 1), as described in the STAR methods section. Using a heatmap with log-10 transformed values to assess the overall sample expression, we discovered that TB patients (TB-DM and TB only) had a distinct expression profile compared to non-TB groups, including controls and DM only, with a tendency toward higher expression

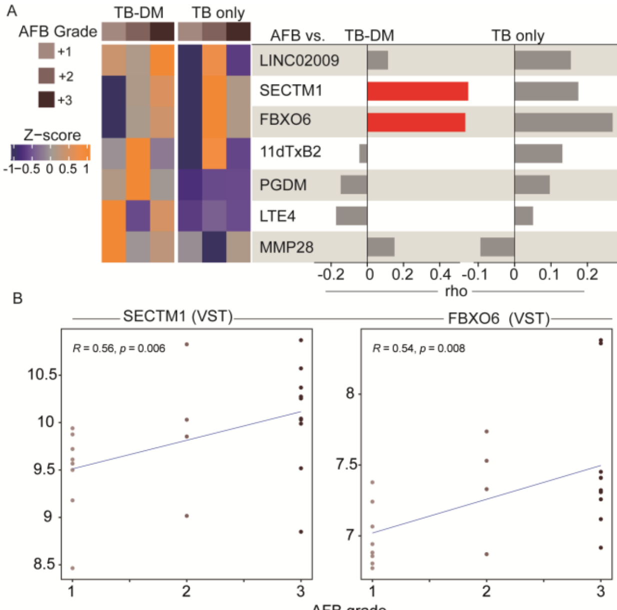


Figure 2. A distinct multi-omic expression and correlation profile between AFB grade and markers among TB groups
(A) *Left panel.* A hierarchical cluster analysis (Ward method with 100 × bootstrap) was employed to evaluate multi-omic marker expression according to the AFB smear grade in TB and TB-DM, as indicated. *Right panel.* A Spearman correlation analysis was used to study the influence of the AFB smear grade on multi-omic marker expression. The rho values are shown. Red bars indicate correlation with p value < 0.05.
(B) Spearman correlation plots demonstrating the associations between the expression levels of the indicated genes and the AFB smear grade. Line represent linear regression. R: Spearman rho value.

of 11dTxB2, PGDM, FBXO6, SECTM1, and LINC02009 in TB patients (Figure 1, left panel; Table S3). Of note, the heatmap demonstrated that the impact of TB disease on the biosignature expression values was able to cluster TB-disease participants independent of DM status (Figure 1, left panel; Table S2). The fold-difference of the multi-omic markers between the clinical groups versus control group is summarized in Figure 1, right panel. When the TB only group was compared with controls, the profile was similar to that observed between TB-DM vs. control group, with higher expressions observed in the majority of the markers in both TB groups, except for MMP28, which was higher in the control group when compared with the TB only group. (Figure 1, right panel). Extending the approach to DM only and control differences, we found higher levels of LTE4 (leukotriene E4) and PGDM (prostaglandin D metabolite) in DM only participants (Figure 1, right panel). These findings reveal a higher multi-omic expression of biomarkers capable of clustering TB patients, regardless of DM status, which suggests differential expression triggered by TB disease.

Variation in the multi-omic activation according to AFB grade and clinical group

After identifying a multi-omic profile that characterized the TB only and TB-DM groups, our next step was to evaluate whether glycemic status affected the expression of markers according to AFB smear grade (Figure 2). Using hierarchical clustering on the *Z* score normalized data at baseline, we observed that patients with TB (with and without DM) and AFB smear grade equal to 2+ tended to express higher levels of SECTM1, FBXO6, and 11dTxB2 compared with other AFB grades (Figure 2A). Those marker levels were lower among those with TB with AFB grade 1+ (Figure 2A). On the other hand, the TB-DM group showed higher expression of LTE4 and MMP28 in those with AFB grade 1+, PGDM and 11dTxB2 in AFB grade 2+, and finally FBXO6, SECTM1, and LINC02009 in AFB grade 3+ (Figure 2A). Next, using Spearman correlation analysis we found that the AFB smear grade was correlated with SECTM1 and FBXO6 expression in TB-DM participants. No significant correlation was observed in the TB only group (Figures 2A and 2B). No significant correlation was observed in the TB only group (Figure 2A). This finding suggests that the presence of TB was associated with changes in multi-omic expression according to AFB smear grade.

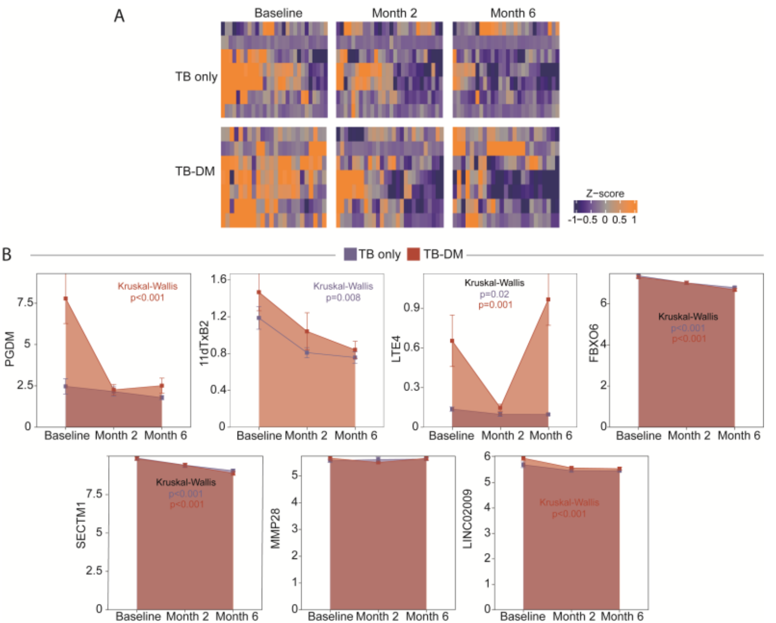


Figure 3. Changes in multi-omic expression after anti-tuberculosis therapy initiation

(A) A hierarchical cluster analysis (Ward method with 100 × bootstrap) was employed to evaluate multi-omic marker expression in TB and TB-DM after anti-tuberculosis therapy initiation.

(B) A boxplot was used to test the changes in multi-omic levels in months 2 and 6 after ATT. In the graphs, dots represent median and whiskers represent interquartile range values.

Changes in multi-omic expression after anti-tuberculosis therapy initiation

We next evaluated changes in the biomarker expressions or concentrations after ATT initiation (Figure 3). Using a heatmap approach, we were able to show an overall decreasing tendency in marker values, in both the TB and TB-DM groups (Figure 3A). Next, we prospectively assessed marker expressions to evaluate the dynamics of multi-omic expression during ATT (Figure 3B). We discerned a reduced level of 11 dTxB2 in the TB group, as well as a reduced level of FBXO6 and SECTM1 expression in both the TB and TB-DM groups, and reduced LINC02009 expression in the TB-DM group.

The multi-omic approach resulted in a transcriptomic signature to detect TB-DM

After evaluating the multi-omic profile of TB and TB-DM interaction, we extended our analysis to verify the expression of the transcriptomic markers identified here. Using data from a multi-center, prospective cohort study of our group,[19] which assessed transcriptomic data from 290 participants, recruited from two sites in India, one in Brazil and publicity available RNA-seq data from two sites of a TANDEM consortium (Romania and South Africa),[20] we test the accuracy of the transcriptomic markers identified here by our machine learning analysis (Figures 4 and S3). First, we used a heatmap approach applied to the four countries, verifying the expression of MMP28, SECTM1, FBOX6, and LINC02009 (Figure S3). Our results identified that these genes were capable of detecting TB regardless of the glycemic status in Brazil, India, and Romania (Figure S3), marked by higher expression of the markers in TB-DM and TB only (Figure S3). Additionally, we used a fold-change Log2 to study the expression of the markers in each clinical group and in each country (Figure S3). Of interest, our results revealed a similar expression profile in TB-DM and TB only in Brazil and India, with higher expression of SECTM1, FBOX6, and LINC02009 when compared with non-TB/non-DM controls, whereas MMP28 was higher in the controls (Figure S3). Next, we used the receiver operating characteristic (ROC) curve, employing the genes MMP28, SECTM1, FBOX6, and LINC02009 to verify the accuracy of these markers, together, to identify TB only, TB-DM, and DM only (Figure 4), shown a robust accuracy in detecting TB-DM, area under curve (AUC) 0.99, 0.91, 0.98, and 0.94 in Brazil, India, Romania, and South Africa, respectively (Figure 4). Our result identifies a transcriptomic signature, composed by expression levels of MMP28, SECTM1, FBOX6, and LINC02009, and identified using machine learning analysis applied to multi-platform data, with a great accuracy to detect TB-DM.

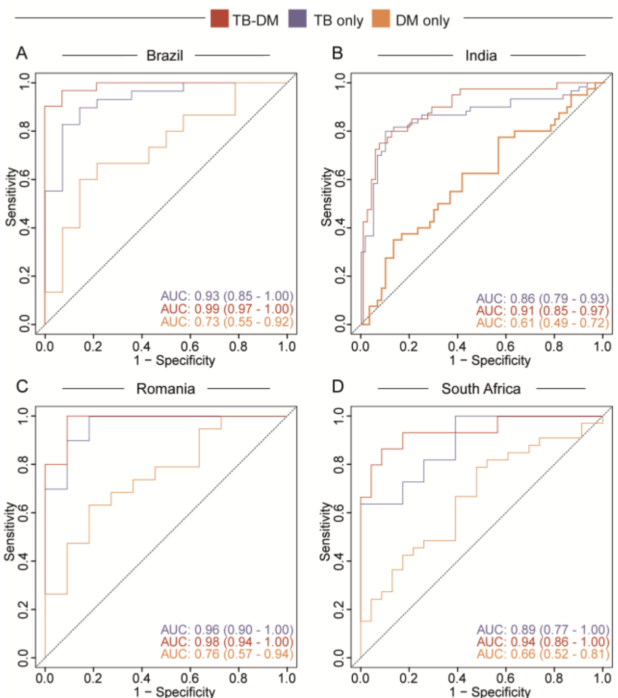


Figure 4. Accuracy of a new transcriptomic signature to detect TB-DM
Accuracy of the gene signature detected in the random forest analysis to classify TB, with or without DM in Brazil (A), India (B), Romania (C) and South Africa (D). Receiver operator characteristics (ROC) curve analysis was used to check the accuracy of the signature genes identified by the random forest model to classify the TB, TB-DM, and DM groups in each clinical site as indicated with respect to TB disease.

DISCUSSION

To achieve the World Health Organization targets for TB elimination, new strategies in the TB field are required. The study of omics has been extensively applied to provide insights into the pathophysiological changes induced by Mtb infection. Several mRNA transcriptomic signatures have been proposed in recent years as new tools for the diagnosis, prognosis, and treatment monitoring of Mtb infection and TB disease.[21–24] Nevertheless, advances have been made using isolated omics, and the employment of an integrative analysis could better dissect the changes triggered by Mtb infection. Here, we performed an integrated omics analysis, using transcriptomic data; eicosanoid measurements in urine; and soluble inflammatory biomarkers in blood, to identify a multi-omic profile that characterized TB and DM.

Our first step was to perform dimensionality data reduction. A random forest model was applied to the integrated multi-omic data (Luminex and RNAseq from peripheral blood and eicosanoids from urine) to select the most informative features. The random forest algorithm works based on summarizing multiple decision trees, evaluating the importance of markers in classifying the samples within the data. The variables were evaluated by using the mean decrease of accuracy and Gini indexes. This algorithm has been used by our group in TB and HIV coinfection studies.[25,26] This approach was also applied to the conceptualization of predictive models in cardiovascular diseases[27] and to predict outcomes in neurosurgery.[28] After cross-validation, MMP-28, LTE-4, 11- dTxB2, PGDM, FBXO6, SECTM1, and LINC02009 were selected. Importantly, despite the crucial role of cytokines in TB inflammatory activation, none of these markers were selected by this approach.

MMP-28 is a protein coding gene that encodes a secreted enzyme associated with casein degradation and is involved in the breakdown of extracellular matrix.[29] The role of MMPs in TB pathogenesis has been largely studied and linked with tissue remodeling.[30] After Mtb infection, inflammatory activation leads to a host response that increases tissue destruction involving MMPs, and disrupts homeostasis, which is required for Mtb dissemination. Some evidence has shown the expression of MMP28 in normal tissues, indicating its role in the preservation

or maintenance of homeostatic conditions.[29,31] In accordance, our results identified higher expression of MMP-28 in the control group when compared with TB only patients, which reinforces the marker activity in homeostasis. If confirmed in other cohorts and future studies, the MMP-28 pathway could be used in the future as a target for host-directed therapies, to improve tissue healing or recovery or limiting the damage caused by the responses against the Mtb.

The role of eicosanoids in TB pathogenesis has been largely explored,[32,33] and the changes in the arachidonic acid pathways persist after ATT.[34] Here, we identified three eicosanoids which, in TANDEM, could help future studies that develop a host-based therapy. LTE4 is a lipid inflammatory mediator and the unique cysteinyl leukotriene stable and abundant in vivo.[35] LTE4 has been described as having a critical role in pulmonary inflammation,[35] through leukocyte activation and triggering cytokine production and macrophage necrosis.[36] Their action in TB pathogenesis remains unclear, but LTE-4 was associated with mucosal eosinophilia and airway hyperresponsiveness in asthma diseases[35] and seems increased in persons with DM.[37] In accordance, our results demonstrated higher levels of LTE4 in TB only, TB-DM and DM only when compared with non-TB/non-DM control group. Of note, we showed a decrease in the levels of LTE4 after ATT initiation, mainly in those with TB only, suggesting an indirect attenuation of lung damage after the induction phase of ATT, the first two months. However, in TB-DM participants, the levels of LTE4 returned to higher expression in month 6, corroborating a hypothesis from one of our previous publications that showed persistent inflammation during ATT in TB-DM participants[13] (11dTxB2 is a urinary metabolite produced from the breakdown of the thromboxane A2). Their levels, measured in urine, can be used to evaluate the response to aspirin therapy in heart disease,[38] asthma,[39] and in diseases associated with higher platelet activation[40] as an indirect measure of platelet activity. The role of 11dTxB2 in TB pathogenesis remains unclear, but TB patients usually present with higher platelet counts that correlate with disease severity and hypercoagulability.[41] Platelets contribute to MMP-mediated tissue damage through their effects on monocytes, leading to upregulation of activation markers, increases in MMP secretion and enhanced phagocytosis.[41] Our results reveal higher levels of 11dTxB2 associated with TB, independent of glycemic status, with decreases in the expression after ATT that could be associated with platelet activation and tissue injury consequences. Of note, considering the effects of platelets on TB pathogenesis, some anti-platelet drugs such as aspirin could be a potential target for host-directed therapy in pulmonary TB.[42] Finally, PGDM is a urinary metabolite from PGD2, and little is known about its function in TB and in infected alveolar macrophages. A recent study in *Histoplasma capsulatum* infection demonstrated an immunostimulatory effect of PGD2, contributing to fungicidal mechanisms and controlling the inflammatory damage.[43] Our results showed higher levels of PGD2 in all three groups when compared with controls, suggesting lung damage triggered by TB disease, but the real role of the molecule in the pathophysiology of the disease needs to be better evaluated. Our finding highlights the potential role of eicosanoids as candidates for host-directed therapy in TB and TB-DM. The blockage of LTE4 pathways could attenuate the cytokine storm, inflammatory activation,[44] and tissue damage. Leukotriene blockers are being used in asthma treatment, and 11dTxB2 has the potential to be used in the future as a marker of platelet activation in TB, as in disease progression, as well as to measure response to the host-directed therapies using anti-platelet agents.

FBXO6 is a protein code gene associated with phosphorylation-dependent ubiquitination, largely studied in cancer,[45] without robust evidence in infectious diseases. In a murine study of Influenza A infection, deficiencies in FBXO6 expression were associated with decreased pulmonary viral replication, inflammatory responses, and mortality.[46] Despite the absence of robust publications involving the gene in the TB field, we found an interesting correlation between the gene expression levels of FBXO6 and the AFB smear grade in TB-DM patients. Additionally, after ATT initiation, we identified a decrease in the expression in both TB and TB-DM, highlighting the possibilities of a new molecule associated with bacilli proliferation and target for host directed therapies. To test this hypothesis, future studies are required to delineate the mechanisms of a potential role for FBXO6 in TB.

SECTM1 is a gene that encodes a transmembrane and secreted protein. In humans, SECTM1 works as a T/NK cell co-stimulatory molecule that has been associated with CD4 and CD8 T cell proliferation and interferon-gamma (IFN-γ) production[47] and reported as an IFN early response gene.[48] In TB, IFN- γ displays a pivotal role in macrophage activation and immune protection against the bacilli.[49,50] Importantly, SECTM1 was identified in a recent study that evaluated host protein features in persons with HIV, as a marker to evaluate Mtb infection prior to TB diagnosis.[51] In accordance with the literature, our findings revealed increased expression of SECTM1 in TB participants, independent of the glycemic status. Additionally, we showed a direct correlation between SECTM1 expression level and AFB smear grade in the TB-DM group, highlighting it as a possible marker to evaluate disease severity in TB-DM. Finally, during ATT the levels of SECTM1 decreased, in both the TB and TB-DM groups. Our findings emphasize the role of SECTM1 as a marker of TB, disease progression in TB-DM, and prognosis, as recently proposed.[52] The last selected marker by our random forest model was the LINC02009, a non-coding RNA with unknown function in TB. The non-coding RNA displays a high potential to modulate biochemical pathways, and our results could represent a potential topic of interest for further investigations in TB-DM pathology. Future studies are required to better evaluate the role and potential of LINC02009 in TB and TB-DM.

The final step of our study was to analyze the accuracy of the transcriptomic markers to detect TB-DM. Several studies around the world have been conducted to identify a transcriptomic profile that characterizes TB-DM interaction, but none of these obtained robust results.[12,20] We recently published a multi-center, prospective cohort study of whole blood gene expression in TB-DM (the MSTDI study).[19] We applied a random forest model to a whole RNA-seq from two sites in India[53,54] and one site in Brazil, as a discovery set, and used public RNA-seq data from two sites of the TANDEM consortium[20] as a validation set (Romania and South Africa), to verify the dynamicity across population. We identified four genes as more informative features to detect TB-DM, SMARCD3, BATF2, VAMP5, and ANKRD22, with an accuracy of 0.97 in Brazil and South Africa, 0.92 in India and 0.89 in Romania.[19] Here, using a similar machine learning analysis applied to multi-platform data from Brazil, we identified another four genes as more informative features to detect TB-DM, LINC02009, SECTM1, FBXO6, and MMP28, with a concise and robust accuracy in discrimination TB-DM. We extended our approach and tested this new transcriptomic signature in the sites

previously studied in MSTDI and found a better performance in all four sites using the transcriptomic markers identified here, with an accuracy of 0.99, 0.91, 0.98, and 0.94 in Brazil, India, Romania, and South Africa, respectively. This finding highlights the potential applicability of multi-omic platforms and machine learning approaches in the study of inflammatory and/or infectious diseases since the findings obtained through the application of the decision tree on the three platforms were better and more concise than what was found when we applied the model only to transcriptomic data.

Despite some limitations, to the best of our knowledge, this is the first study to evaluate the multi-omic analysis of TB and TB-DM, with data retrieved from samples collected from the same patients and time points. By identifying novel markers that could serve as targets for future host-directed therapies, our results expand the current knowledge regarding TB-DM pathogenesis and that can be useful to improve anti-TB treatment outcomes in this population.

Limitations of the study

This study had some limitations. First, the populations were different regarding age and BMI, factors that could influence inflammatory activation. A validation in an independent cohort is necessary to ensure the robustness of our findings. In the pre-analytical phase of the current study, we defined that all RePORT Brazil participants who had data on all the omics platforms investigated here would be included, and no pre-analytical matching of the study sample was performed. This approach may have introduced bias, which will require further investigation in additional cohorts to clarify. Second, the measurement of HbA1c levels was performed only at baseline, and changes in glycemia could also affect the immune responses after therapy initiation. We are currently developing a prospective investigation to assess HbA1c levels in a novel cohort from the RePORT Brazil attended population.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Ethics statement
  - Study design and population
- METHOD DETAILS
  - RNA sequencing
  - Collection, processing, and analysis of eicosanoid metabolites
  - Immunoassays
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Gene expression analysis
  - Feature selection analysis using machine learning
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.109135.

ACKNOWLEDGMENTS

We thank the study participants. We recognize the important contributions of the RePORT-Brazil Consortium, who participated at different stages of the consortium and helped to build the databases explored in the present study. Vanessa Nascimento and Michael S. Rocha (Instituto Brasileiro para Investigação da Tuberculose, Fundação José Silveira, Salvador, Brazil, Multinational Organization Network Sponsoring Translational and Epidemiological Research Initiative, Salvador, Brazil); Saulo R. N. Santos and André Ramos (Instituto Brasileiro para Investigação da Tuberculose, Fundação José Silveira, Salvador, Brazil); Juan Manuel Cubillos-Ángulo (Multinational Organization Network Sponsoring Translational and Epidemiological Research Initiative, Salvador, Brazil); Jéssica Rebouças-Silva (Faculdade de Medicina, Universidade Federal da Bahia, Salvador, Brazil, Laboratório de Inflamação e Biomarcadores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil); Pedro Brito (Multinational Organization Network Sponsoring Translational and Epidemiological Research Initiative, Salvador, Brazil, Laboratório de Inflamação e Biomarcadores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil); Carolina A. S. Schmaltz, Adriano Gomes-Silva, Francine P. Ignácio, Maria C. Lourenço, Aline Benjamin, Jamile G. de Oliveira, Solange Cavalcante, and Betina Durovni (Laboratório de Pesquisa Clínica em Micobacteriose, Instituto Nacional de Infectologia Evandro Chagas, Fiocruz, Rio de Janeiro, Brazil); Alysson G. Costa (Fundação Medicina Tropical Dr Heitor Vieira Dourado, Manaus, Brazil, Programa de Pós-Graduação em Medicina Tropical, Universidade do Estado do Amazonas, Manaus, Brazil, Universidade Federal do Amazonas, Manaus, Brazil); Leandro Sousa Garcia, Brenda

K. de Sousa Carvalho, Bruna P. de Loiola, and Alexandra B. Souza (Fundação Medicina Tropical Dr Heitor Vieira Dourado, Manaus, Brazil, Programa de Pós-Graduação em Medicina Tropical, Universidade do Estado do Amazonas, Manaus, Brazil); Elisangela C. Silva, Mayla Mello, and Adriana S. R. Moreira (Programa de Pós-Graduação em Medicina Tropical, Universidade do Estado do Amazonas, Manaus, Brazil); José R. Lapa-e-Silva (Programa Acadêmico de Tuberculose da Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; Megan Turner, John R Koethe, and Carlos Henrique Serezani (Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, United States). A special thanks to Elze Leite (FIOCRUZ, Salvador, Brazil), Eduardo Gama (FIOCRUZ, Rio de Janeiro, Brazil), Eldimar Junior (FMT-HVD, Manaus, Brazil), and Hilary Vansell (VUMC, Nashville, USA) for administrative and logistical support. Some data used here were originated from publicly available databases previously reported by the RePORT India and the TANDEM consortium, and thus we thank the research teams that generated the primary data. This work was supported by the Departamento de Ciência e Tecnologia (DECIT)—Secretaria de Ciência e Tecnologia (SCTIE)—Ministério da Saúde, Brazil [25029.000507/2013-07 to V.R], the National Institute of Allergy and Infectious Diseases [U01-AI069923] at NIH, and by The Civilian Research and Development Foundation (CRDF) Global #DAA3-17-63144. The study was partially supported by the Intramural Research Program of the Fundacaao Oswaldo Cruz and the Intramural Research Program of the Fundação José Silveira. B.B.A., V.C.R., A.K., and M.C.-S. are senior scientists from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). M.B.A. is Researcher Level I from The Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC) and is supported by Project Number 5 R01-AI147765-04 from NIH. G.C.S. is a scientific initiation fellow from CNPq. M.A.-P. and B.B.-D. received a research fellowship from the Coordenacaão de Aperfeicoamento de Pessoal de Nível Superior (CAPES, finance code: 001). The funders had no role in study design, data collection, and analysis, the decision to publish, or the preparation of the manuscript.

AUTHOR CONTRIBUTIONS

A.T.L.Q., B.B.A., T.S., M.C.-S., A.K., V.R., and M.F. designed the study and mentored the work; M.B.A., B.B.-D., M.A.-P., A.A., and B.B.A. performed the experiments and data collections; C.L.V., E.R.F., G.C.S., B.B.A., and A.T.L.Q. analyzed the data; M.B.A., M.M.B., B.B.-D., M.A.-P., V.C.R., A.L.K., M.C.-S., and T.R.S. helped with data interpretation; C.L.V., G.C.S., B.B.A., and A.T.L.Q. wrote the manuscript. All authors contributed to the article and approved the submitted version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 29, 2023
Revised: January 2, 2024
Accepted: February 1, 2024
Published: February 5, 2024

REFERENCES

1. World Health Organization (2022). Global Tuberculosis Report.
2. Houben, R.M.G.J., and Dodd, P.J. (2016). The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. PLoS Med. 13, e1002152. https://doi.org/10.1371/journal.pmed.1002152.
3. Bell, L.C.K., and Noursadeghi, M. (2018). Pathogenesis of HIV-1 and Mycobacterium tuberculosis co-infection. Nat. Rev. Microbiol. 16, 80–90. https://doi.org/10.1038/nrmicro.2017.128.
4. Kathirvel, M., and Mahadevan, S. (2016). The role of epigenetics in tuberculosis infection. Epigenomics 8, 537–549. https://doi.org/10.2217/epi.16.1.
5. Murray, M., Oxlade, O., and Lin, H.H. (2011). Modeling social, environmental and biological determinants of tuberculosis. Int. J. Tuberc. Lung Dis. 15, 64–70. https://doi.org/10.5588/ijtld.10.0535.
6. Gupta, K.B., Gupta, R., Atreja, A., Verma, M., and Vishvkarma, S. (2009). Tuberculosis and nutrition. Lung India 26, 9–16. https://doi.org/10.4103/0970-2113.45198.
7. World Health Organization (2016). Global Report on Diabetes.
8. Mantovani, A., and Garlanda, C. (2023). Humoral Innate Immunity and Acute-Phase Proteins. N. Engl. J. Med. 388, 439–452. https://doi.org/10.1056/NEJMra2206346.
9. Berbudi, A., Rahmadika, N., Tjahjadi, A.I., and Ruslami, R. (2020). Type 2 Diabetes and its Impact on the Immune System. Curr. Diabetes Rev. 16, 442–449. https://doi.org/10.2174/1573399815666191024085838.
10. Restrepo, B.I. (2016). Diabetes and Tuberculosis. Microbiol. Spectr. 4, 1–11. https://doi.org/10.1128/microbiolspec.TNMI7-0023-2016.
11. Jeon, C.Y., and Murray, M.B. (2008). Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. PLoS Med. 5, e152. https://doi.org/10.1371/journal.pmed.0050152.
12. Prada-Medina, C.A., Fukutani, K.F., Pavan Kumar, N., Gil-Santana, L., Babu, S., Lichtenstein, F., West, K., Sivakumar, S., Menon, P.A., Viswanathan, V., et al. (2017). Systems Immunology of Diabetes-Tuberculosis Comorbidity Reveals Signatures of Disease Complications. Sci. Rep. 7, 1999. https://doi.org/10.1038/s41598-017-01767-4.
13. Kumar, N.P., Fukutani, K.F., Shruthi, B.S., Alves, T., Silveira-Mattos, P.S., Rocha, M.S., West, K., Natarajan, M., Viswanathan, V., Babu, S., et al. (2019). Persistent inflammation during anti-tuberculosis treatment with diabetes comorbidity. Elife 8, e46477. https://doi.org/10.7554/eLife.46477.
14. Arriaga, M.B., Rocha, M.S., Nogueira, B.M.F., Nascimento, V., Araújo-Pereira, M., Souza, A.B., Andrade, A.M.S., Costa, A.G., Gomes-Silva, A., Silva, E.C., et al. (2021). The Effect of Diabetes and Prediabetes on Mycobacterium tuberculosis Transmission to Close Contacts. J. Infect. Dis. 224, 2064–2072. https://doi.org/10.1093/infdis/jiab264.
15. Gil-Santana, L., Almeida-Junior, J.L., Oliveira, C.A.M., Hickson, L.S., Daltro, C., Castro, S., Kornfeld, H., Netto, E.M., and Andrade, B.B. (2016). Diabetes Is Associated with Worse Clinical Presentation in Tuberculosis Patients from Brazil: A Retrospective Cohort Study. PLoS One 11, e0146876. https://doi.org/10.1371/journal.pone.0146876.
16. Arriaga, M.B., Araújo-Pereira, M., Barreto-Duarte, B., Nogueira, B., Freire, M.V.C.N.S., Queiroz, A.T.L., Rodrigues, M.M.S., Rocha, M.S., Souza, A.B., Spener-Gomes, R., et al. (2022). The Effect of Diabetes and Prediabetes on Antituberculosis Treatment Outcomes: A Multicenter Prospective Cohort Study. J. Infect. Dis. 225, 617–626. https://doi.org/10.1093/infdis/jiab427.
17. Barreda, N.N., Arriaga, M.B., Aliaga, J.G., Lopez, K., Sanabria, O.M., Carmo, T.A., Fróes Neto, J.F., Lecca, L., Andrade, B.B., and

Calderon, R.I. (2020). Severe pulmonary radiological manifestations are associated with a distinct biochemical profile in blood of tuberculosis patients with dysglycemia. BMC Infect. Dis. 20, 139. https://doi.org/10.1186/s12879-020-4843-0.

18. Calderon, R.I., Arriaga, M.B., Aliaga, J.G., Barreda, N.N., Sanabria, O.M., Barreto-Duarte, B., Franco, J.P.D., Lecca, L., Andrade, B.B., Carvalho, A.C.C., and Kritski, A.L. (2022). Persistent dysglycemia is associated with unfavorable treatment outcomes in patients with pulmonary tuberculosis from Peru. Int. J. Infect. Dis. 116, 293–301. https://doi.org/10.1016/j.ijid.2022.01.012.
19. Queiroz, A.T.L., Vinhaes, C.L., Fukutani, E.R., Gupte, A.N., Kumar, N.P., Fukutani, K.F., Arriaga, M.B., Sterling, T.R., Babu, S., Gaikwad, S., et al. (2023). A multi-center, prospective cohort study of whole blood gene expression in the tuberculosis-diabetes interaction. Sci. Rep. 13, 7769. https://doi.org/10.1038/s41598-023-34847-9.
20. Eckold, C., Kumar, V., Weiner, J., Alisjahbana, B., Riza, A.L., Ronacher, K., Coronel, J., Kerry-Barnard, S., Malherbe, S.T., Kleynhans, L., et al. (2021). Impact of Intermediate Hyperglycemia and Diabetes on Immune Dysfunction in Tuberculosis. Clin. Infect. Dis. 72, 69–78. https://doi.org/10.1093/cid/ciaa751.
21. Hamada, Y., Penn-Nicholson, A., Krishnan, S., Cirillo, D.M., Matteelli, A., Wyss, R., Denkinger, C.M., Rangaka, M.X., Ruhwald, M., and Schumacher, S.G. (2022). Are mRNA based transcriptomic signatures ready for diagnosing tuberculosis in the clinic? - A review of evidence and the technological landscape. EBioMedicine 82, 104174. https://doi.org/10.1016/j.ebiom.2022.104174.
22. Singhania, A., Wilkinson, R.J., Rodrigue, M., Haldar, P., and O'Garra, A. (2018). The value of transcriptomics in advancing knowledge of the immune response and diagnosis in tuberculosis. Nat. Immunol. 19, 1159–1168. https://doi.org/10.1038/s41590-018-0225-9.
23. Burel, J.G., Babor, M., Pomaznoy, M., Lindestam Arlehamn, C.S., Khan, N., Sette, A., and Peters, B. (2019). Host Transcriptomics as a Tool to Identify Diagnostic and Mechanistic Immune Signatures of Tuberculosis. Front. Immunol. 10, 221. https://doi.org/10.3389/fimmu.2019.00221.
24. Warsinske, H., Vashisht, R., and Khatri, P. (2019). Host-response-based gene signatures for tuberculosis diagnosis: A systematic comparison of 16 signatures. PLoS Med. 16, e1002786. https://doi.org/10.1371/journal.pmed.1002786.
25. Krishnan, S., Queiroz, A.T.L., Gupta, A., Gupte, N., Bisson, G.P., Kumwenda, J., Naidoo, K., Mohapi, L., Mave, V., Mngqibisa, R., et al. (2021). Integrative Multi-Omics Reveals Serum Markers of Tuberculosis in Advanced HIV. Front. Immunol. 12, 676980. https://doi.org/10.3389/fimmu.2021.676980.
26. Kulkarni, V., Queiroz, A.T.L., Sangle, S., Kagal, A., Salvi, S., Gupta, A., Ellner, J., Kadam, D., Rolla, V.C., Andrade, B.B., et al. (2021). A Two-Gene Signature for Tuberculosis Diagnosis in Persons With Advanced HIV. Front. Immunol. 12, 631165. https://doi.org/10.3389/fimmu.2021.631165.
27. Yang, L., Wu, H., Jin, X., Zheng, P., Hu, S., Xu, X., Yu, W., and Yan, J. (2020). Study of cardiovascular disease prediction model based on random forest in eastern China. Sci. Rep. 10, 5245. https://doi.org/10.1038/s41598-020-62133-5.
28. Hanko, M., Grendár, M., Snopko, P., Opšenák, R., Šutovský, J., Benčo, M., Soršák, J., Zeleňák, K., and Kolarovszki, B. (2021). Random Forest-Based Prediction of Outcome and Mortality in Patients with Traumatic Brain Injury Undergoing Primary Decompressive Craniectomy. World Neurosurg. 148, e450–e458. https://doi.org/10.1016/j.wneu.2021.01.002.
29. Ra, H.J., and Parks, W.C. (2007). Control of matrix metalloproteinase catalytic activity. Matrix Biol. 26, 587–596. https://doi.org/10.1016/j.matbio.2007.07.001.
30. Amaral, E.P., Vinhaes, C.L., Oliveira-de-Souza, D., Nogueira, B., Akrami, K.M., and Andrade, B.B. (2021). The Interplay Between Systemic Inflammation, Oxidative Stress, and Tissue Remodeling in Tuberculosis. Antioxid. Redox Signal. 34, 471–485. https://doi.org/10.1089/ars.2020.8124.
31. Sabir, N., Hussain, T., Mangi, M.H., Zhao, D., and Zhou, X. (2019). Matrix metalloproteinases: Expression, regulation and role in the immunopathology of tuberculosis. Cell Prolif. 52, e12649. https://doi.org/10.1111/cpr.12649.
32. Arriaga, M.B., Karim, F., Queiroz, A.T.L., Araújo-Pereira, M., Barreto-Duarte, B., Sales, C., Moosa, M.Y.S., Mazibuko, M., Milne, G.L., Maruri, F., et al. (2022). Effect of Dysglycemia on Urinary Lipid Mediator Profiles in Persons With Pulmonary Tuberculosis. Front. Immunol. 13, 919802. https://doi.org/10.3389/fimmu.2022.919802.
33. Chen, M., Divangahi, M., Gan, H., Shin, D.S.J., Hong, S., Lee, D.M., Serhan, C.N., Behar, S.M., and Remold, H.G. (2008). Lipid mediators in innate immunity against tuberculosis: opposing roles of PGE2 and LXA4 in the induction of macrophage death. J. Exp. Med. 205, 2791–2801. https://doi.org/10.1084/jem.20080767.
34. Vinhaes, C.L., Oliveira-de-Souza, D., Silveira-Mattos, P.S., Nogueira, B., Shi, R., Wei, W., Yuan, X., Zhang, G., Cai, Y., Barry, C.E., 3rd, et al. (2019). Changes in inflammatory protein and lipid mediator profiles persist after antitubercular treatment of pulmonary and extrapulmonary tuberculosis: A prospective cohort study. Cytokine 123, 154759. https://doi.org/10.1016/j.cyto.2019.154759.
35. Paruchuri, S., Tashimo, H., Feng, C., Maekawa, A., Xing, W., Jiang, Y., Kanaoka, Y., Conley, P., and Boyce, J.A. (2009). Leukotriene E4-induced pulmonary inflammation is mediated by the P2Y12 receptor. J. Exp. Med. 206, 2543–2555. https://doi.org/10.1084/jem.20091240.
36. Salimi, M., Stöger, L., Liu, W., Go, S., Pavord, I., Klenerman, P., Ogg, G., and Xue, L. (2017). Cysteinyl leukotriene E(4) activates human group 2 innate lymphoid cells and enhances the effect of prostaglandin D(2) and epithelial cytokines. J. Allergy Clin. Immunol. 140, 1090–1100.e11. https://doi.org/10.1016/j.jaci.2016.12.958.
37. Rafnsson, A., and Bäck, M. (2013). Urinary leukotriene E4 is associated with renal function but not with endothelial function in type 2 diabetes. Dis. Markers 35, 475–480. https://doi.org/10.1155/2013/370461.
38. Lordkipanidzé, M., Pharand, C., Schampaert, E., Turgeon, J., Palisaitis, D.A., and Diodati, J.G. (2007). A comparison of six major platelet function tests to determine the prevalence of aspirin resistance in patients with stable coronary artery disease. Eur. Heart J. 28, 1702–1708. https://doi.org/10.1093/eurheartj/ehm226.
39. Oosaki, R., Mizushima, Y., Mita, H., Shida, T., Akiyama, K., and Kobayashi, M. (1997). Urinary leukotriene E4 and 11-dehydrothromboxane B2 in patients with aspirin-sensitive asthma. Allergy 52, 470–473. https://doi.org/10.1111/j.1398-9995.1997.tb01032.x.
40. Catella, F., Healy, D., Lawson, J.A., and FitzGerald, G.A. (1986). 11-Dehydrothromboxane B2: a quantitative index of thromboxane A2 formation in the human circulation. Proc. Natl. Acad. Sci. USA 83, 5861–5865. https://doi.org/10.1073/pnas.83.16.5861.
41. Kirwan, D.E., Chong, D.L.W., and Friedland, J.S. (2021). Platelet Activation and the Immune Response to Tuberculosis. Front. Immunol. 12, 631696. https://doi.org/10.3389/fimmu.2021.631696.
42. Cubillos-Angulo, J.M., Nogueira, B.M.F., Arriaga, M.B., Barreto-Duarte, B., Araújo-Pereira, M., Fernandes, C.D., Vinhaes, C.L., Villalva-Serra, K., Nunes, V.M., Miguez-Pinto, J.P., et al. (2022). Host-directed therapies in pulmonary tuberculosis: Updates on anti-inflammatory drugs. Front. Med. 9, 970408. https://doi.org/10.3389/fmed.2022.970408.
43. Pereira, P.A.T., Assis, P.A., Prado, M.K.B., Ramos, S.G., Aronoff, D.M., de Paula-Silva, F.W.G., Sorgi, C.A., and Faccioli, L.H. (2018). Prostaglandins D(2) and E(2) have opposite effects on alveolar macrophages infected with Histoplasma capsulatum. J. Lipid Res. 59, 195–206. https://doi.org/10.1194/jlr.M078162.
44. Tobin, D.M., Roca, F.J., Oh, S.F., McFarland, R., Vickery, T.W., Ray, J.P., Ko, D.C., Zou, Y., Bang, N.D., Chau, T.T.H., et al. (2012). Host genotype-specific therapies can optimize the inflammatory response to mycobacterial infections. Cell 148, 434–446. https://doi.org/10.1016/j.cell.2011.12.023.
45. Merry, C., Fu, K., Wang, J., Yeh, I.J., and Zhang, Y. (2010). Targeting the checkpoint kinase Chk1 in cancer therapy. Cell Cycle 9, 279–283. https://doi.org/10.4161/cc.9.2.10445.
46. Cen, M., Ouyang, W., Lin, X., Du, X., Hu, H., Lu, H., Zhang, W., Xia, J., Qin, X., and Xu, F. (2023). FBXO6 regulates the antiviral immune responses via mediating alveolar macrophages survival. J. Med. Virol. 95, e28203. https://doi.org/10.1002/jmv.28203.
47. Wang, T., Huang, C., Lopez-Coral, A., Slentz-Kesler, K.A., Xiao, M., Wherry, E.J., and Kaufman, R.E. (2012). K12/SECTM1, an interferon-gamma regulated molecule, synergizes with CD28 to costimulate human T cell proliferation. J. Leukoc. Biol. 91, 449–459. https://doi.org/10.1189/jlb.1011498.
48. Huyton, T., Göttmann, W., Bade-Döding, C., Paine, A., and Blasczyk, R. (2011). The T/NK cell co-stimulatory molecule SECTM1 is an IFN "early response gene" that is negatively regulated by LPS in human monocytic cells. Biochim. Biophys. Acta 1810, 1294–1301. https://doi.org/10.1016/j.bbagen.2011.06.020.
49. Khan, T.A., Mazhar, H., Saleha, S., Tipu, H.N., Muhammad, N., and Abbas, M.N. (2016). Interferon-Gamma Improves Macrophages Function against M. tuberculosis in Multidrug-Resistant Tuberculosis Patients. Chemother. Res. Pract. 2016, 7295390. https://doi.org/10.1155/2016/7295390.

50. Flynn, J.L., Chan, J., Triebold, K.J., Dalton, D.K., Stewart, T.A., and Bloom, B.R. (1993). An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection. J. Exp. Med. 178, 2249–2254. https://doi.org/10.1084/jem.178.6.2249.
51. Singer, S.N., Ndumnego, O.C., Kim, R.S., Ndung'u, T., Anastos, K., French, A., Churchyard, G., Paramithiothis, E., Kasprowicz, V.O., and Achkar, J.M. (2022). Plasma host protein biomarkers correlating with increasing Mycobacterium tuberculosis infection activity prior to tuberculosis diagnosis in people living with HIV. EBioMedicine 75, 103787. https://doi.org/10.1016/j.ebiom.2021.103787.
52. Young, B.L., Mlamla, Z., Gqamana, P.P., Smit, S., Roberts, T., Peter, J., Theron, G., Govender, U., Dheda, K., and Blackburn, J. (2014). The identification of tuberculosis biomarkers in human urine samples. Eur. Respir. J. 43, 1719–1729. https://doi.org/10.1183/09031936.00175113.
53. Kornfeld, H., West, K., Kane, K., Kumpatla, S., Zacharias, R.R., Martinez-Balzano, C., Li, W., and Viswanathan, V. (2016). High Prevalence and Heterogeneity of Diabetes in Patients With TB in South India: A Report from the Effects of Diabetes on Tuberculosis Severity (EDOTS) Study. Chest 149, 1501–1508. https://doi.org/10.1016/j.chest.2016.02.675.
54. Gupte, A., Padmapriyadarsini, C., Mave, V., Kadam, D., Suryavanshi, N., Shivakumar, S.V.B.Y., Kohli, R., Gupte, N., Thiruvengadam, K., Kagal, A., et al. (2016). Cohort for Tuberculosis Research by the Indo-US Medical Partnership (CTRIUMPH): protocol for a multicentric prospective observational study. BMJ Open 6, e010542. https://doi.org/10.1136/bmjopen-2015-010542.
55. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
56. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8.
57. Breiman, L. (2001). Random Forest. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.
58. Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. J. Stat. Softw. 28, 26. https://doi.org/10.18637/jss.v028.i05.
59. Arriaga, M.B., Amorim, G., Queiroz, A.T.L., Rodrigues, M.M.S., Araújo-Pereira, M., Nogueira, B.M.F., Souza, A.B., Rocha, M.S., Benjamin, A., Moreira, A.S.R., et al. (2021). Novel stepwise approach to assess representativeness of a large multicenter observational cohort of tuberculosis patients: The example of RePORT Brazil. Int. J. Infect. Dis. 103, 110–118. https://doi.org/10.1016/j.ijid.2020.11.140.
60. Geadas, C., Stoszek, S.K., Sherman, D., Andrade, B.B., Srinivasan, S., Hamilton, C.D., and Ellner, J. (2017). Advances in basic and translational tuberculosis research: Proceedings of the first meeting of RePORT international. Tuberculosis (Edinb) 102, 55–67. https://doi.org/10.1016/j.tube.2016.11.006.
61. van der Heijden, Y.F., Abdullah, F., Andrade, B.B., Andrews, J.R., Christopher, D.J., Croda, J., Ewing, H., Haas, D.W., Hatherill, M., Horsburgh, C.R., Jr., et al. (2018). Building capacity for advances in tuberculosis research; proceedings of the third RePORT international meeting. Tuberculosis (Edinb) 113, 153–162. https://doi.org/10.1016/j.tube.2018.09.009.
62. Arriaga, M.B., Araújo-Pereira, M., Barreto-Duarte, B., Sales, C., Miguez-Pinto, J.P., Nogueira, E.B., Nogueira, B.M.F., Rocha, M.S., Souza, A.B., Benjamin, A., et al. (2021). Prevalence and Clinical Profiling of Dysglycemia and HIV Infection in Persons With Pulmonary Tuberculosis in Brazil. Front. Med. 8, 804173. https://doi.org/10.3389/fmed.2021.804173.
63. Broekmans, J.F., Migliori, G.B., Rieder, H.L., Lees, J., Ruutu, P., Loddenkemper, R., Raviglione, M.C.; World Health Organization International Union Against Tuberculosis and Lung Disease and Royal Netherlands Tuberculosis Association Working Group, and Lung, D.; Royal Netherlands Tuberculosis Association Working Group (2002). European framework for tuberculosis control and elimination in countries with a low incidence. Recommendations of the World Health Organization (WHO), International Union Against Tuberculosis and Lung Disease (IUATLD) and Royal Netherlands Tuberculosis Association (KNCV) Working Group. Eur. Respir. J. 19, 765–775. https://doi.org/10.1183/09031936.02.00261402.
64. American Diabetes Association Professional Practice Committee, Draznin, B., Aroda, V.R., Bakris, G., Benson, G., Brown, F.M., Freeman, R., Green, J., Huang, E., Isaacs, D., et al. (2022). 6. Glycemic Targets: Standards of Medical Care in Diabetes-2022. Diabetes Care 45, S83–S96. https://doi.org/10.2337/dc22-S006.
65. Song, W.L., Wang, M., Ricciotti, E., Fries, S., Yu, Y., Grosser, T., Reilly, M., Lawson, J.A., and FitzGerald, G.A. (2008). Tetranor PGDM, an abundant urinary metabolite reflects biosynthesis of prostaglandin D2 in mice and humans. J. Biol. Chem. 283, 1179–1188. https://doi.org/10.1074/jbc.M706839200.
66. Murphey, L.J., Williams, M.K., Sanchez, S.C., Byrne, L.M., Csiki, I., Oates, J.A., Johnson, D.H., and Morrow, J.D. (2004). Quantification of the major urinary metabolite of PGE2 by a liquid chromatographic/mass spectrometric assay: determination of cyclooxygenase-specific PGE2 synthesis in healthy humans and those with lung cancer. Anal. Biochem. 334, 266–275. https://doi.org/10.1016/j.ab.2004.08.019.
67. Oliveira-de-Souza, D., Vinhaes, C.L., Arriaga, M.B., Kumar, N.P., Cubillos-Angulo, J.M., Shi, R., Wei, W., Yuan, X., Zhang, G., Cai, Y., et al. (2019). Molecular degree of perturbation of plasma inflammatory markers associated with tuberculosis reveals distinct disease profiles between Indian and Chinese populations. Sci. Rep. 9, 8002. https://doi.org/10.1038/s41598-019-44513-8.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw RNAseq data from TANDEM	Eckold et al.[20]	Bioproject ID: PRJNA470512
Raw RNAseq from RePORT	Kornfeld et al.[53]; Gupta et al.[54]; Queiroz et al.[19]	GEOncbi ID: GSE181143
Software and algorithms		
R version 4.2.2.	R Core Team	https://cran.r-project.org/
STAR version 2.7.10	Dobin et al.[55]	https://code.google.com/archive/p/rna-star/
DESeq2 version 1.40.2	Love et al.[56]	https://www.bioconductor.org/packages/release/bioc/html/DESeq2.html
randomForest version 4.7–1.1	Breiman et al.[57]	https://cran.r-project.org/web/packages/randomForest/index.html
caret version 6.0–94	Kuhn et al.[58]	https://cran.r-project.org/web/packages/caret/index.html
complexheatmap	https://doi.org/10.1093/bioinformatics/btw313	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
Hmisc version 5.1–1	https://cran.rproject.org/web/packages/Hmisc/Hmisc.pdf	https://hbiostat.org/r/hmisc/

RESOURCE AVAILABILITY

Lead contact

Further information and requests regarding the packages employed for the analysis performed in this study should be directed to and will be fulfilled, by the lead contact, Bruno B. Andrade (bruno.andrade@fiocruz.br).

Materials availability

- This study did not generate new unique reagents.

Data and code availability

- Data used here were from two previous works. The gene expression from the main population, Salvador site from RePORT Brazil, and Indian population (used only in the validation of the transcriptomic signature) were from the MSTDI study and have been deposited at the GEO ncbi database and is publicly available as of the date of publication (Accession number GSE181143, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE181143).

The validation of the transcriptomic signature identified here was performed using an external cohort, with population from South Africa and Romania, previously published by the TANDEM consortium and have been previously deposited at the SRA database and are publicly available as of the date of publication (BioProject ID PRJNA470512).

- All employed packages' references are available at the key resources tables and methodology.
- This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Ethics statement

The study was conducted according to the Declaration of Helsinki. Data and specimens were obtained from the RePORT-Brazil observational cohort study. The study was approved by the Institutional Review Boards of the Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Brazil (CAAE: 25102412.3.1001.5262). Written informed consents were obtained from all voluntary participants.

Study design and population

This was a retrospective analysis of a prospective observational study with data and specimens collected within the scope of RePORT-Brazil,[59–61] at the Salvador site. All participants were enrolled between June 2015 and June 2019 and followed for up to 24 months. Several studies from RePORT-Brazil have explored different aspects of TB-DM.[14,16,32,62] However, the evaluation of multi-platform data has not been previously explored. Additionally, to evaluate the accuracy of the transcriptomic signature identified in this study we used public data from a previously published work from the RePORT-India[53,54] and TANDEM consortium.[20]

Participants from the RePORT-Brazil were at least 18 years old, had an Mtb culture-positive sputum at enrollment (for those in the TB groups) and were evaluated at three in-person visits: (i) ATT initiation (baseline), (ii) two months after initiating ATT, and (iii) after completing treatment (approximately month 6). The sputum smears were prepared by the Ziel–Nielsen method using 1% carbol-fuchsin and scored using the International Union Against Tuberculosis and Lung Disease (IUATLD) scale.[63] Non-TB/Non-DM controls and DM patients (without TB) were selected from close TB contacts who agreed to participate in the study. These patients were evaluated at two visits: baseline and 6 months after enrollment. We selected those participants who tested negative, at both visits, for Mtb infection by QuantiFERON-TB Gold. The selection of the participants included in the sub-groups was based on availability of databases on each assay platform assessed in the present study. As a result, only participants who had data on all the platforms evaluated were included. This was pre-specified in the analysis plan to allow data integration and multi-omic analyses. Sociodemographic, clinical, and epidemiologic data such as age, sex, race/ethnicity (self-reported), and body mass index (BMI) were collected at baseline and shown in Table S1. Of note, gander variable was not collected. Regarding the biological sex and race/ethnicity (self-reported) our analysis found no differences between the groups. Glycated hemoglobin (HbA1C) was measured in all participants at enrollment.[32]

Diabetes was defined according to baseline HbA1c, following American Diabetes Association (ADA) guidelines.[64] Individuals were classified as having DM (HbA1c $\geq$ 6.5%) or normoglycemia (HbA1c<5.7%).

METHOD DETAILS

RNA sequencing

Samples from TB patients (with and without DM) were collected at baseline, 2, and 6 months of treatment. Samples for the non-TB/non-DM control and DM only groups were collected at baseline. Total RNA from venous blood (2.5 mL) was collected in PAXgene Blood RNA Tubes (PreAnalytiX), extracted using a PAXgene blood miRNA kit (Qiagen) with the semi-automated QIAcube (Qiagen) and quantified using the LabChip GX HiSens RNA system (PerkinElmer). RNA-seq libraries were prepared using the Bioscientific NEXTflex-Rapid-Directional mRNA-seq sample preparation with the Caliper SciClone. Samples were sequenced using the NextSeq500 High Output kit V2 (Illumina) for 75 cycles. For all samples, RNA was sequenced by Illumina HiSeq 2500 at MedGenome in Bangalore, India.

Collection, processing, and analysis of eicosanoid metabolites

Urine samples from patients were collected at baseline, 2 months, and 6 months of treatment. Urine was stored at −80°C within 1 h of sample collection.[65,66] PGE-M and PGD-M are unstable if urine is not stored at −80°C within 90 min of collection. Concentrations of the major urinary prostaglandin (PG)E_2 metabolite, PGE-M; the tetranor-PGE_1, TN-E, the major urinary metabolite of PGD_2, PGD-M; 11-dehydro-thromboxane-B_2, $dTxB_2$; the metabolite of PGI_2, PGI-M; and leukotriene (LT)E_4 were measured in urine at each time point in all study participant. The assays were performed at the Eicosanoid Core Laboratory at Vanderbilt University Medical Center, in the USA. Samples were shipped on dry ice and stored at −80°C until analysis.

[$^2H_4$]-PGE-M and [$^2H_{11}$]-TN-E were synthesized as previously described.[32] [$^2H_4$]-PGI-M and [$^2H_4$]-11dTxB_2 were purchased from Cayman Chemicals (Ann Arbor, MI USA), and [20,20,20-$^2H_3$]-LTE_4 from Enzo Life Sciences (Farmingdale, NY USA). Sep-Pak C18 and Oasis HLB (3cc/60mg) extraction cartridges were obtained from Waters Corporation (Milford, MA USA). The organic reagents were of high-performance Liquid Chromatograph (LC) quality and purchased from Sigma Aldrich (St. Louis, MO USA).

Immunoassays

Plasma was stored at −80°C for immunology assays. Using the Bio-Plex Pro Human Cytokine Standard 27-Plex kit (Group I) on the Biorad Bioplex 200 Luminex platform the following analytes were measured: interleukin (IL)-1β, IL-1 receptor antagonist (IL-1Ra), IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12p70, IL-13, IL-15 IL-17A, eotaxin, granulocyte colony stimulating factor (G-CSF), granulocyte-macrophage colony stimulating factor (GM-CSF/CSF2), interferon gamma (IFN-γ), monocyte chemoattractant protein-1 (MCP-1)/C-C motif chemokine ligand 2 (CCL2), macrophage inflammatory protein-1 alpha (MIP-1α/CCL3), MIP-1 beta (MIP-1β/CCL4), platelet-derived growth factor-BB (PDGF), regulated on activation, normal T cell expressed and secreted (RANTES/CCL5), tumor necrosis factor-alpha (TNF-α), and vascular endothelial growth factor (VEGF). Biomarkers were chosen based on commercially available pre-mixed kits and included parameters previously described to be involved in TB pathogenesis.[34,67]

QUANTIFICATION AND STATISTICAL ANALYSIS

Gene expression analysis

After the quality check, sequences were aligned to the human transcriptome (GRCh38 version 100), comprising mRNA and ncRNA, using STAR.[55] Count gene expression matrix was examined using the DESeq2 R package,[56] version 4.2.2, to identify differentially expressed genes

(DEG) across the four patient groups. Changes in gene expression with false discovery rate (FDR)-adjusted p value <0.05 and log2 fold-difference +1.4 were considered significant.

Feature selection analysis using machine learning

We employed a feature selection approach to identify the most relevant features from the multi-omic factors used here associated with TB-DM. The integrated multi-omic data, composed by all three platforms measured in the study, transcriptomic data, cytokines, and eicosanoids, were used to trained the model and cross validation, applying the random forest algorithm, using the randomForest package.[57] 1000 trees were used in the model and at each split the number of features was 41.88, the square root of the total of variables in the dataset plus one. Furthermore, a leave-one-out cross-validation with 50-folds, and 5 repetitions was performed to assess the markers' accuracy, using the caret package.[58] The cross-validation allowed the estimation of the model's accuracy and the No information rate, which was 1(1-0.95) and 0.36 respectively. The p value for accuracy > No information rate was 0.0000000000000022. With this method, we aim to reduce the dimension of data and obtain the most accurate model. Seven features were selected after applied the random forest model and included and the analysis presented here. Figure S1 illustrates the model design. Table S2 and Figure S2 show the variable importance of the top 20 features.

Statistical analysis

The median values with interquartile ranges (IQR) were used as measures of central tendency and dispersion. Features were compared between the study groups using the Kruskal-Wallis test with Dunn's multiple comparisons. Categorical variables were compared using Chi-square test. Hierarchical cluster analyses (Ward's method), with 100X bootstrap of *Z* score normalized data were employed to depict the overall multi-omic expression profile in the study groups. Dendrograms represent Manhattan distance.

Spearman correlation analysis was used to evaluate the interaction between sputum smear AFB grade and multi-omic expression data, and between HbA1c levels and the multi-omic expression data. All univariate comparisons with p values <0.05 after adjustments were considered statistically significant. The statistical analyses were performed using *R* 4.2.2. The R package used to perform the analyses here is described in Table S4.

Supplemental Tables and Figures

Table S1. Characteristics of the study population

Characteristic	TB-DM	TB only	DM only	Non-TB/Non-DM controls	P-value
N	24	28	11	13	
Age – years	46.5 (37-55.7)	28.5 (25-40.5)	56 (51-59)	34 (27-50)	**<0.001**
Female – no. (%)	9 (37)	10 (35)	7 (63)	9 (69)	0.10
Race – no. (%)					0.6
White	3 (12)	3 (10)	0	1 (7)	
Black	14 (58)	11 (39)	6 (54)	7 (53)	
Pardo	7 (29)	14 (50)	5 (45)	5 (38)	
BMI (Kg/m²)	22 (18-26)	20 (18-22)	29 (27-32)	29 (24-33)	**<0.001**
HbA1c (%)	8.1 (6.6-11.3)	5.5 (5.2-5.6)	7.1 (6.9-8.4)	5 (4.8-5.3)	**<0.001**
AFB smear grade – no. (%)					0.36
0	5 (17)	1 (4)			
1+	8 (28)	6 (25)			
2+	4 (14)	3 (12)			
3+	11 (39)	14 (58)			
Cavities on chest X-ray – no. (%)	15 (62)	13 (46)			0.2

Data represent medians and interquartile ranges (age, BMI and HbA1c) and frequencies (female sex, race, AFB smear grade and cavities on chest X-ray). The Kruskal-Wallis test was used to compare continuous variables between the groups and the distributions while the Chi-square test was used to compare frequencies. AFB, acid-fast bacilli; BMI, Body Mass Index; HbA1c, Glycated hemoglobin. P-values in bold font are statistically significant.

Table S2. Top 20 variables from the multi-omic database that were identified by the Gini score and mean accuracy to discriminate between all the clinical groups.

Variable	Gini score	Accuracy
AKR1C5P	2,0	0,5
CASKIN1	1,6	0,3
CDC42P3	2,0	0,5
CSMD2	2,4	0,5
DUSP3	2,5	0,2
ESYT3	1,2	0,3
FBXO6	3,5	1,0
HNRNPA1P61	1,2	0,2
IL15	1,6	0,2
IL6	2,7	0,3
KCNH7	1,7	0,25
LINC02009	2,6	1,2
LTE4	3,7	4,8
MIR1256	1,4	0,3
MMP28	2,3	1,3
PGD/LTE4	2,6	0,3
PGDM	8,5	12,6
PGD/TNE	1,7	0,2
SECTM1	3,4	1,6
11dTxB2	4,9	7,3

PGDM, prostaglandin D Metabolite; 11dTxB2, 11-dehydrothromboxane B2; LTE4, Leukotriene E4; MMP28, Matrix Metalloproteinase 28; FBOX6, F-Box Protein 6; PGEM, Prostaglandin E Metabolite; IL- Interleukin; PGIM, Prostaglandin I Metabolite.

Table S3. Expression of multiomic markers used in the study according to the clinical groups.

Parameter	Unit	Non-TB/Non-DM controls	DM	TB	TBDM	P-value
FBXO6	VST	6.5 (6.4-6.7)	6.4 (6.2-6.5)	7.2 (6.9-7.7)	7.3 (6.9-7.4)	**<0.001**
SECTM1	VST	8.5 (8.3-8.8)	8.4 (8.2-8.7)	9.8 (9.3-10.5)	9.9 (9.5-10.2)	**<0.001**
MMP28	VST	5.8 (5.7-6.0)	5.5 (5.4-5.6)	5.5 (5.4-5.6)	5.6 (5.5-5.7)	**0.005**
LINC02009	VST	5.4 (5.1-5.6)	5.4 (5.2-5.5)	5.5 (5.4-5.9)	5.9 (5.7-6.1)	**0.002**
PGDM	ng_mg_Cr	0.4 (0.3-0.5)	0.7 (0.5-0.9)	1.8 (1.4-2.5)	4.4 (1.9-13.4)	**<0.001**
11dTxB2	ng_mgCr	0.22 (0.20-0.27)	0.3 (0.2-0.4)	0.9 (0.7-1.6)	0.9 (0.7-2.4)	**<0.001**
LTE4	ng_mgCr1_A	0.07 (0.05-0.09)	0.2 (0.1-1.1)	0.1 (0.08-0.2)	0.2 (0.1-0.9)	**<0.001**

Data represent medians and interquartile ranges. The Kruskal-Wallis test was used to compare the distributions of the multiomic markers between the study groups. P values in bold font are statistically significant. VST, variance stabilizing transformation gene expression; FBOX6, F-Box Protein 6; MMP28, Matrix Metalloproteinase 28; PGDM, Prostaglandin D Metabolite; 11dTxB2, 11-dehydrothromboxane B2; LTE4, Leukotriene E4.

Table S4: Packages used for statistical analyses.

Objective	R package	Version	Reference
Identify differentially expressed genes	DESeq2	R 4.2.2	1
Feature selection analysis using machine learning	Random Forest	R 4.2.2	2
Cross-validation	Caret	R 4.2.2	3
Heatmap	ComplexHeatmap	R 4.2.2	4
Spearman correlations	Hmisc	R 4.2.2	5

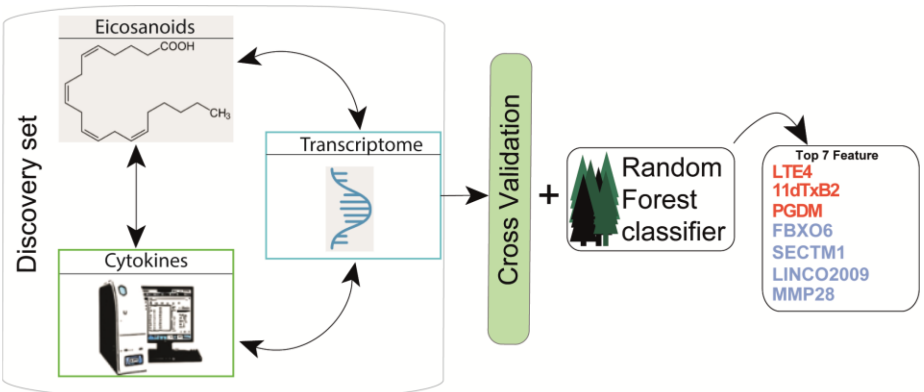


Figure S1. Dimensional data reduction. A graphical abstract from the dimensionality reduction approach: A random forest model was applied to the multiplatform data (Luminex and RNAseq from peripheral blood and eicosanoids from urine) for feature selection based on clinical groups.

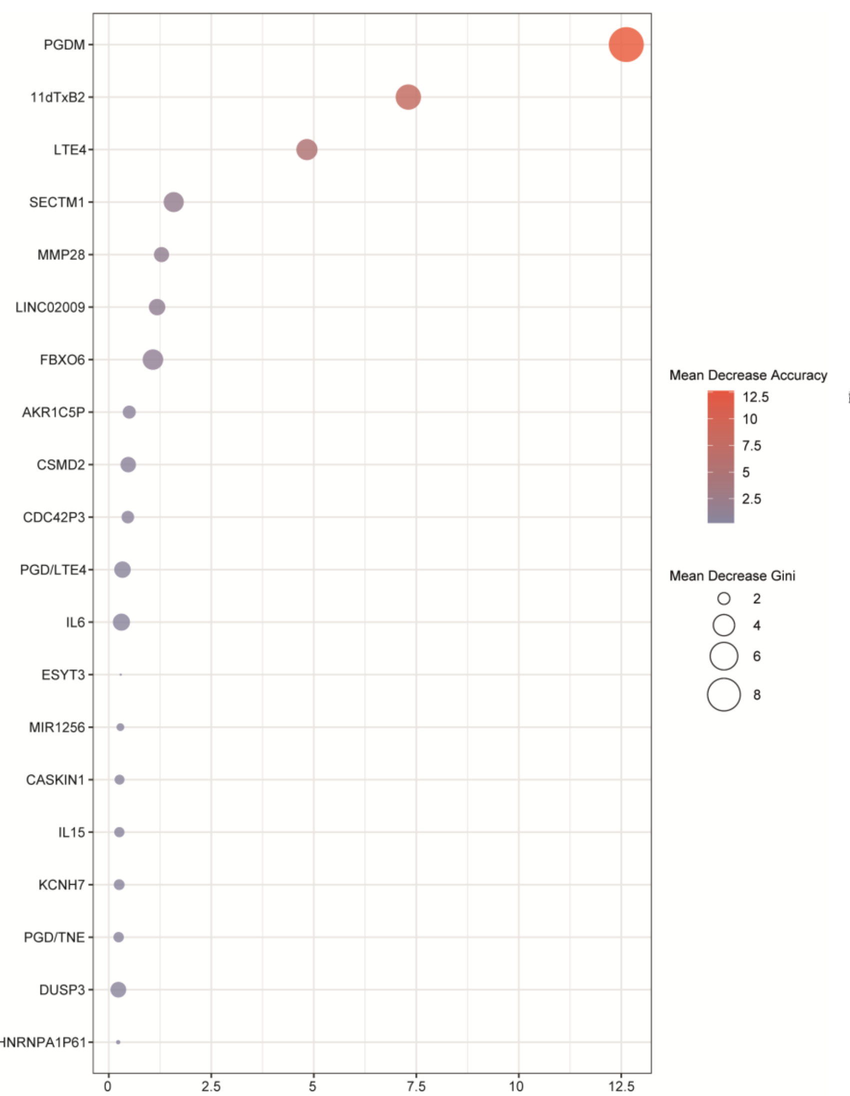


Figure S2. A dot plot to show the feature importance of the top 20 markers used in the randomForest model.

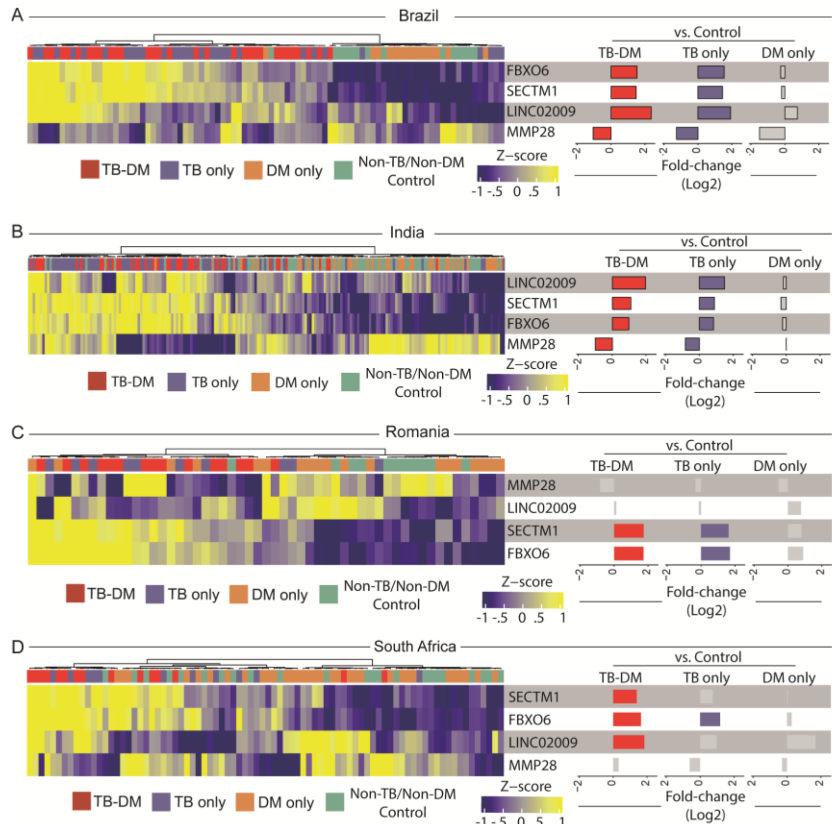


Figure S3. A partial transcriptomic signature to detect TB-DM. A z-score normalized heatmap was employed to depict overall trends in gene expression among the clinical groups on each study site, as indicated. Panels to the right of heatmaps show the average fold-difference between the signature gene expression in the HC group versus TB-DM, TB only and DM only (log-transformed values)

References

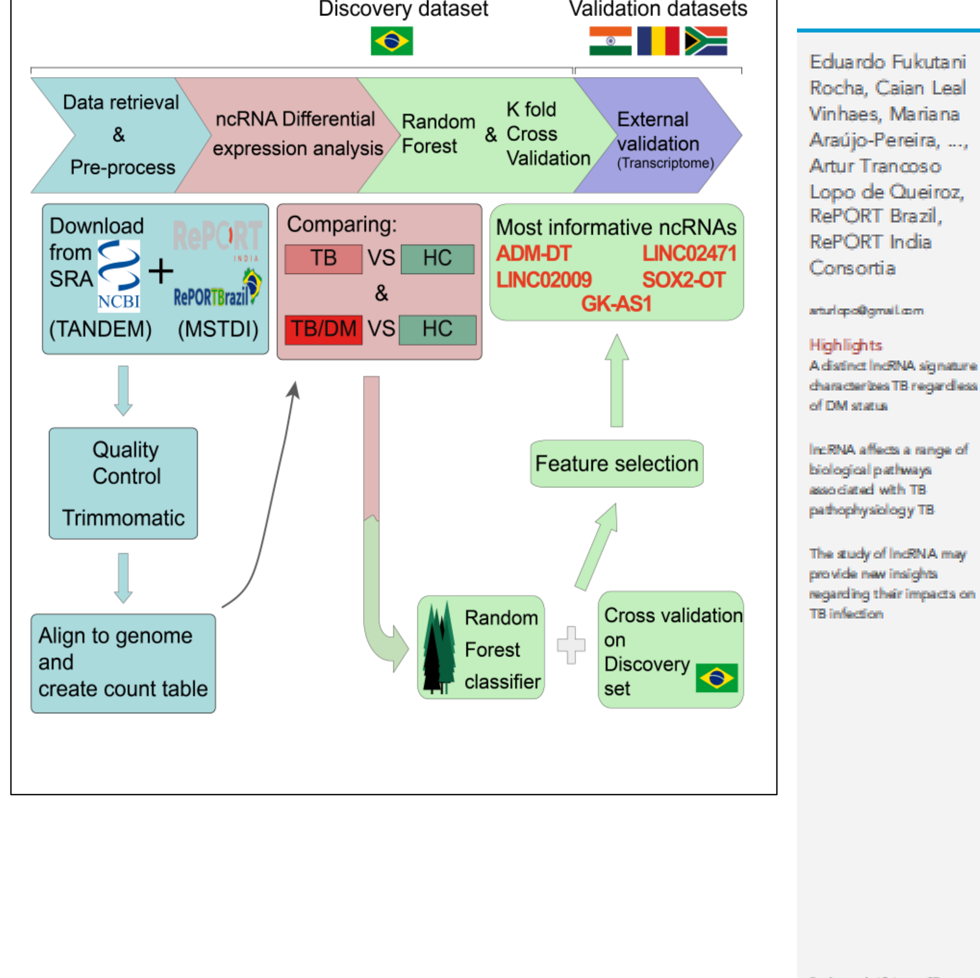
1. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. 10.1186/s13059-014-0550-8.
2. Breiman, L. (2001). Random Forest. Machine Learning. 10.1023/A:1010933404324.
3. Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software *28*, 26. 10.18637/jss.v028.i05.
4. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847-2849. 10.1093/bioinformatics/btw313.
5. Harrell, Frank E. Hmisc: A package of miscellaneous R functions. [Internet]. Available from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://cran.rproject.org/web/packages/Hmisc/Hmisc.pdf

Artigo III

iScience

Article

The sound of silent RNA in tuberculosis and the lncRNA role on infection



Eduardo Fukutani Rocha, Caian Leal Vinhaes, Mariana Araújo-Pereira, ..., Artur Trancoso Lopo de Queiroz, RePORT Brazil, RePORT India Consortia

arturlopo@gmail.com

Highlights

A distinct lncRNA signature characterizes TB regardless of DM status

lncRNA affects a range of biological pathways associated with TB pathophysiology TB

The study of lncRNA may provide new insights regarding their impacts on TB infection

Rocha et al., iScience 27, 108662
January 19, 2024 © 2023 The Authors.
https://doi.org/10.1016/j.isci.2023.108662

iScience

Article

The sound of silent RNA in tuberculosis and the lncRNA role on infection

Eduardo Fukutani Rocha,[1,3,17] Caian Leal Vinhaes,[2,3,4,17] Mariana Araújo-Pereira,[2,3,4,13] Tiago Feitosa Mota,[1,3] Akshay N. Gupte,[5] Nathella Pavan Kumar,[16] Maria Belen Arriaga,[2,3] Timothy R. Sterling,[7] Subash Babu,[6] Sanjay Gaikwad,[8] Rajesh Karyakarte,[9] Vidya Mave,[10,11] Vandana Kulkarni,[10,11] Mandar Paradkar,[10,11] Vijay Viswanathan,[12] Hardy Kornfeld,[14,15] Amita Gupta,[10] Bruno Bezerril Andrade,[2,3,4,13,18] Artur Trancoso Lopo de Queiroz,[1,2,3,18,19,*] and RePORT Brazil, RePORT India Consortia

SUMMARY

Tuberculosis (TB) is one of the leading causes of death worldwide, and Diabetes Mellitus is one of the major comorbidities (TB/DM) associated with the disease. A total of 103 differentially expressed ncRNAs have been identified in the TB and TB/DM comparisons. A machine learning algorithm was employed to identify the most informative lncRNAs: ADM-DT, LINC02009, LINC02471, SOX2-OT, and GK-AS1. These lncRNAs presented substantial accuracy in classifying TB from HC (AUCs >0.85) and TB/DM from HC (AUCs >0.90) in the other three countries. Genes with significant correlations with the five lncRNAs enriched common pathways in Brazil and India for both TB and TB/DM. This suggests that lncRNAs play an important role in the regulation of genes related to the TB immune response.

INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis* (Mtb) and is one of the leading causes of death worldwide due to a single pathogen.[1] It is estimated that ¼ of the global population have been infected by the Mtb[2] and about 10% of the infected people develop the active form of TB during their lifetime.[3] The clinical presentation of active TB varies depending on the site of infection and the host inflammatory response. An aggravating condition to TB is the comorbidity with diabetes mellitus (DM), a potentially devastating medical condition with an alarming increase in its prevalence since the beginning of this century.[4] DM is characterized as a metabolic disease with pathologically high blood glucose level due to insulin action failure.[5]

The main TB aggravating factor for DM is the immunological dysfunction caused by the hyperglycemia, as it impairs both the innate and adaptive immune responses toward infections, increasing the host susceptibility to develop active TB.[6,7] The comorbidity of TB and DM (TB/DM) also worsens the treatment for TB, often leading to prolonged sputum culture conversion, and unfavorable anti-TB treatment (ATT) outcomes, such as death, treatment failure, and TB recurrence after treatment.[8] Furthermore, TB/DM is associated with an altered transcriptome and perturbations in biological pathways,[9] which may also contribute to the dysregulation of non-coding RNAs (ncRNAs).

Around 60% of the transcriptional output in human cells is represented by ncRNAs[10] and its largest type is the long non-coding (lncRNA). These are composed of 200 or more nucleotides, and despite their functions being little understood, they have major importance in regulating a large range of biological processes.[11] At genetic level, lncRNAs participate in gene expression regulation by controlling access or

[1]Centro de Integração de Dados e Conhecimentos para Saúde, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil
[2]Laboratório de Inflamação e Biomarcadores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil
[3]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Brazil
[4]Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador 40290-150, Brazil
[5]Boston University School of Public Health, Boston, MA USA
[6]National Institutes of Health- NIRT - International Center for Excellence in Research, Chennai, India
[7]Division of Infectious Diseases, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN USA
[8]Department of Pulmonary Medicine, Byramjee-Jeejeebhoy Government Medical College and Sassoon General Hospitals, Pune, India
[9]Department of Microbiology, Byramjee-Jeejeebhoy Government Medical College and Sassoon General Hospitals, Pune, India
[10]Byramjee-Jeejeebhoy Government Medical College-Johns Hopkins University Clinical Research Site, Pune, India
[11]Johns Hopkins Center for Infectious Diseases in India, Pune, India
[12]Prof. M. Viswanathan Diabetes Research Centre, Chennai, India
[13]Faculdade de Tecnologia e Ciências, Instituto de Pesquisa Clínica e Translacional, Salvador, Brazil
[14]Department of Medicine, University of Massachusetts Medical School, Worcester, MA USA
[15]UMass Chan Medical School, Worcester, MA USA
[16]ICMR-National Institute for Research in Tuberculosis, Chennai, India
[17]These authors contributed equally to this work and share the first authorship.
[18]These authors contributed equally to this work and share the last authorship.
[19]Lead contact
*Correspondence: arturlopo@gmail.com
https://doi.org/10.1016/j.isci.2023.108662

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

dismissal of regulatory proteins from chromatin.[12] They can also regulate other ncRNAs[13] and microRNAs activities, by acting as microRNAs sponges as well as affecting the mRNAs translation.[14] Additionally, it is also known that lncRNAs can modify the mRNA expression by regulating the pre-mRNA splicing, editing and even stabilizing mRNAs.[15–17]

In the TB scenario, differentially expressed lncRNAs are important molecules with the role of regulating immune response pathways against Mtb. Such regulation is done on essential molecules and biological processes, such as TGF-β, IFN-γ, T- and B- cells differentiation and adaptive immune responses.[18–20] While for DM, differentially expressed lncRNAs have been mainly associated with insulin secretion by the pancreatic beta cells and insulin resistance.[21] Despite being less explored, lncRNAs emerge as potential biomarkers to evaluate the dynamics of TB infection and prognosis, as some lncRNA signatures have been previously proposed.[22,23] However, more studies are required to further validate the previously proposed TB biomarkers, including other populations.

Our group has recently described the patterns of coding gene expression in response toward TB and TB/DM in four different populations (Brazil, India, Romania, and South Africa).[24] The previous results depicted highly different patterns of gene expression, suggesting influence of population-specific differences on TB and TB/DM gene expression. In the present study, we used a robust bioinformatic approach to propose a lncRNA based biomarker for TB, which is consistently expressed through different regions and maintains its accuracy even with the TB/DM comorbidity. This biomarker was identified in RNA-seq data from patients from Brazil, enrolled by the Report Brazil[25] and had its accuracy validated in data from patients enrolled from India,[26,27] Romania, and South Africa.[28]

RESULTS

Identifying lncRNA that characterize tuberculosis and tuberculosis/diabetes mellitus

We used previously published and public data to identify differentially expressed ncRNAs and evaluate their expression. A detailed population description can be found in our previous work. In Brazil, our discovery set, a total of 189 DEGs between TB and HC groups were identified, from which 120 were upregulated and 69 were downregulated (Figure S1). Regarding the TB/DM vs. HC comparison, a total of 1128 DEGs were identified, from which 182 DEGs were upregulated and 946 were downregulated (Figure S1). Following, the lncRNAs and microRNAs were filtered from the DEGs on each comparison, i.e., TB vs. HC and TB/DM vs. HC. A total of 25 differentially expressed ncRNAs (DEncRNAs) were identified in the TB comparison, being 15 upregulated and 10 downregulated. The TB/DM comparison identified 95 DEncRNAs (28 upregulated and 67 downregulated). A summary of all identified DEGs and DEncRNAs, as well as the statistical values are available in supplemental material S1. The overall study procedure and downstream analyses are resumed in a flowchart (Figure 1).

Discriming tuberculosis and tuberculosis/diabetes mellitus using differentially expressed non-coding RNAs signature from machine learning application

After identifying 103 DEncRNAs, we applied the Random Forest (RF) machine learning algorithm to their expression data, aiming to detect the five best classifying DEncRNAs to characterize TB and TB/DM. The lncRNAs *ADM-DT*, *LINC02009*, *LINC02471*, *SOX2-OT* and *GK-AS1* were the top five features in terms of variable importance according to the RF model (Table S1). The five most informative lncRNAs' fold change was shown in Figure S2. To evaluate these genes' expression in each clinical group in Brazil, the discovery set, we employed a heatmap, displaying the z-scores of VST normalized expression data (Figure 2A). Our analysis revealed a total of two major clusters: the first was predominantly composed of patients with TB (44.4%) and TB/DM (55.5%), while the second was comprised by a mixture of HC (36.8%), TB (34.2%) and TB/DM (28.9%) (Figure 2A). Furthermore, the k-fold cross validation applied to the discovery set appointed an accuracy of 1, 95% C.I. [0.95, 1], a no-information rate of 0.42, sensitivity of 1, specificity of 1, positive and negative predictive values of 1. To check the accuracy of the identified DEncRNA signature in each different region dataset (India, Romania, and South Africa), we used Receiver Operating Characteristic (ROC) curves (Figures 2B and 2C and 2D). When classifying Indian samples, the lncRNA signature achieved an AUC of 0.86, 95% C.I. [0.78, 0.93], when classifying TB and HC samples and an AUC of 0.90, 95% C.I. [0.84, 0.97], when classifying TB/DM and HC samples (Figure 2B). A similar classifying performance was observed in the Romanian dataset, as the biomarker achieved AUCs of 0.96, 95% C.I. [0.90, 1.00], and 0.94, 95% C.I. [0.85, 1.00], when classifying TB and TB/DM from HC samples, respectively (Figure 2C). Lastly, the lncRNA signature was tested with South African samples, achieving AUCs of 0.90, 95% C.I. [0.79, 1.00], and 0.94, 95% C.I. [0.85, 1.00], when classifying TB and TB/DM from HC samples, respectively (Figure 2D). This finding has shown that the DEncRNA signature, herein identified using RF, could discriminate TB/DM with an accuracy higher than 90% in the clinical sites included in our study. Compared to the previously published TB signatures, this lncRNA signature had an overall similar performance, but could provide insights regarding potentially important lncRNAs in TB (Figure S3). Further information regarding the model is available at Table S1.

Impact of lncRNA in the overall gene expression

To assess how the signature affects the overall gene expression in each condition and region, we performed a spearman correlation analysis between the expressions of the five selected ncRNAs and all mRNA genes. The correlated genes and their respective rho and p values are available in supplemental material S2. Brazil was the site with the highest number of strong correlations, |rho| > 0.7, with 2292 in TB and 1214 in TB/DM. Additionally, most of them were positive correlations (85% and 89%, respectively) (Figure 3). The Indian region revealed 1484 interactions, 807 in TB and 677 in TB/DM, with 92% and 81% positive correlation, respectively (Figure 3).

As for the Romanian region, 317 and 425 correlations were identified in TB and TB/DM groups, respectively. Despite the decreased number of strong correlations, compared to Brazil and India, positive correlations were also predominant, with 74% in the TB group and 57% in the

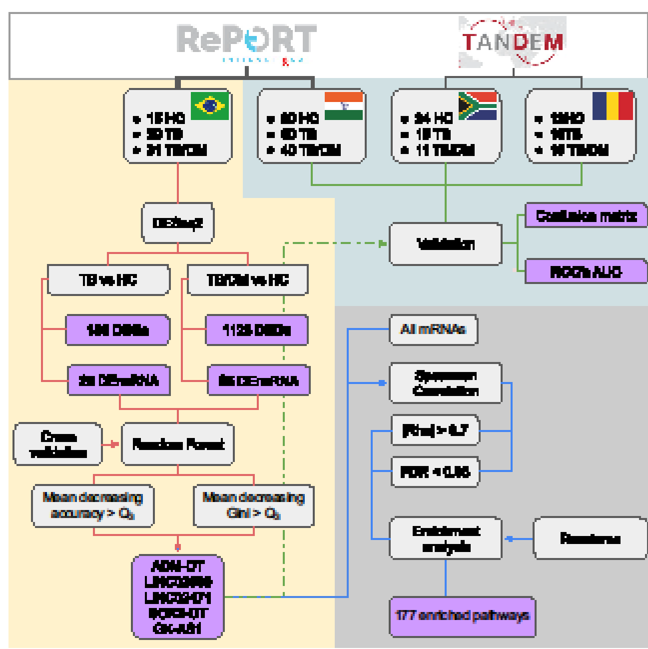


Figure 1. Flowchart of study's procedures from data acquisition through differentially expressed genes detection, feature selection and validation until enrichment analysis of genes correlated with the top five most informative DEncRNA

Red lines represent the analysis with the training dataset, Brazilian samples; Green lines represent the validation of ncRNA signature found in the random forest model; Blue lines represent the correlation and enrichment analysis. Purple boxes are illustrating the results.

TB/DM group. Lastly, the least number of correlations were observed in South Africa, being only four for the TB group (three positive and one negative) and 130 for TB/DM (78 positives and 52 negatives) (Figure 3). The result suggests a considerable impact of the ncRNA in the overall gene expression, marked mainly by positive correlations.

Impact of the correlated mRNA genes in the biological pathways

Next, to evaluate the impact of the above-mentioned correlated genes on biological pathways, we performed an enrichment analysis (Figure 4). Thus, all strongly correlated genes' Entrez IDs and their respective fold changes were used as input, grouped by the region in which the genes were correlated. A total of 177 pathways were enriched by the genes which were strongly correlated with our lncRNA TB signature. Further information about all enriched pathways is available at supplemental material S3. The top 10 gene ratio pathways for each region were identified and if the pathway was also enriched in the other regions, its respective gene ratios were retrieved. The genes which comprise these pathways had their respective correlated lncRNA assessed, to check each lncRNAs impact on the enrichment. Within the five most informative lncRNAs, three were correlated with the majority of the correlated genes comprising these pathways: *LINC02471*, *ADM-DT*, and *GK-AS1* (Figure S4). Both Neutrophil degranulation and Signaling by Interleukins were among the top 10 gene ratios in all TB infected groups (TB and TB/DM) in Brazil and India regions (Figure 4). The pathways which were among the top 10 gene ratios for at least one group, but also enriched by the other ones were Interleukin (IL)−4 and IL−13 signaling, Regulated Necrosis, Signaling by CSF3 (G−CSF), Inactivation of CSF3 (G−CSF) signaling, Diseases associated with the TLR signaling cascade, Diseases of Immune System, IRAK4 deficiency (TLR2/4) and MyD88 deficiency (TLR2/4) (Figure 4). The pathways commonly enriched by the correlated genes identified in Brazil TB, Brazil TB/DM and India TB were Programmed Cell Death, Toll−like Receptor Cascades, Interferon gamma signaling, Antigen processing−Cross presentation and ER−Phagosome pathway (Figure 4). The Interferon alpha/beta signaling was enriched only by the Brazil TB, Brazil TB/DM, and India TB/DM correlated genes. Interleukin−3, Interleukin−5, and GM−CSF signaling was enriched by the Brazil TB, India TB and India TB/DM correlated genes. The FCGR activation pathway was enriched by the Brazil TB/DM, India TB and India TB/DM correlated genes. Lastly, the Class I MHC mediated antigen processing & presentation pathway was exclusively enriched by the Brazil region (Figure 4).

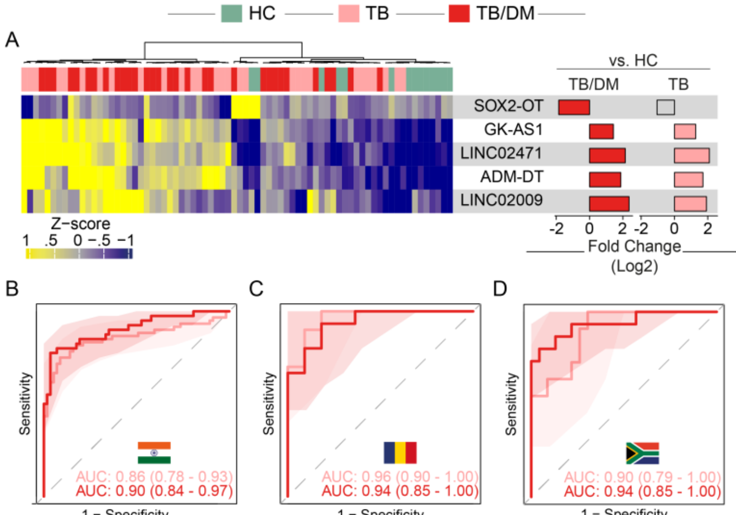


Figure 2. Random forest model and validation

(A) Heatmap displaying the Brazil region Z-scaled VST normalized expression data of the 5 classifying lncRNAs. The barplot alongside the heatmap is displaying the 5 lncRNAs log2 fold change when compared to the HC group. The bars are colored in red (TB/DM) and pink (TB) for statistically significant fold changes, while gray bars represent non significant fold changes.

(B–D) ROC curves displaying the lncRNA biomarker overall classifying performance when classifying samples from each validation dataset. The AUC with confidence interval values are depicted in each ROC curve. (B) India region dataset.

(C) Romania region dataset.

(D) South Africa region dataset.

DISCUSSION

This study aimed to evaluate the role of a TB signature composed of lncRNAs in TB and TB/DM, improving the molecular knowledge and providing further insights regarding the pathophysiology. Such insights could contribute, in the future, toward the development of new TB host directed therapy, regardless of DM comorbidity. We use the data from samples collected by the RePORT-Brazil in Salvador as the main discovery dataset, due to this data being paired end and presenting only one batch. Datasets from other regions were maintained as test datasets, as they were single-end RNA-seq data (Romania and South Africa) and were sequenced in more than one batch, demanding the application of a batch effect correction algorithm (India). This approach of using machine learning algorithms to identify biomarkers has been used before with HTLV-1,[29] mosquitoes with dengue, Zika, Chikungunya, and Yellow Fever[30,31] and to identify a predictive model in cardiovascular diseases.[32] The model composed of five lncRNAs (*ADM-DT, LINC02009, LINC02471, SOX2-OT,* and *GK-AS1*) achieved AUCs >0.85 when discriminating patients with TB from HC and >0.9 with TB/DM from HC, even in samples from other populations, exhibiting an outstandingly consistent accuracy. When compared with the previously proposed TB signatures, this model had similar accuracy, but provides unique insights regarding important lncRNAs in TB.

The identification of a concise transcriptomic signature to characterize TB/DM interaction has been the focus of several groups.[33,34] Recently, by applying a similar methodology, we identified a signature for TB and TB/DM composed of four mRNA genes using samples from the same cohort. Despite the signature's accuracy, it was noted that the expression of these four genes had a high degree of variability across the study regions, suggesting a strong influence of population-specific expression pattern.[24] Here we identified a more consistent pattern of expression, as three lncRNAs (ADM-DT, GK-AS1, and LINC02471) had similar fold changes for TB and TB/DM in all regions and LINC2009 fold changes were similar in Brazil, India, and South Africa. The persistent pattern of expression observed in our signature across all regions corroborates its consistency. Future studies are required to validate our findings. Despite that, the role of lncRNA emerges as a possible component promoting changes in the immunopathogenesis associated with the increased risk of persons with DM to develop active TB.[35] Once they are infected with *Mycobacterium tuberculosis*, higher is the transmission of TB among DM person[36] and more severe is the presentation of TB/DM, followed by an increased risk of unfavorable TB outcomes.[37]

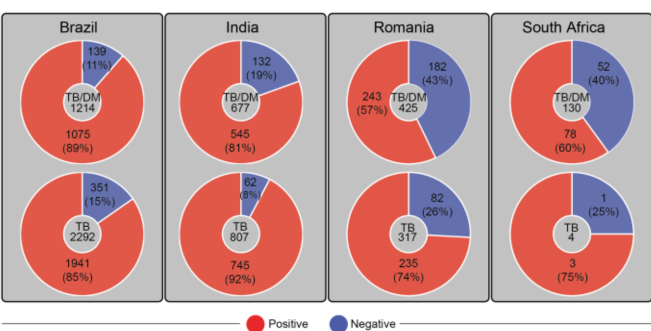


Figure 3. Correlations per region

Number of strong correlations between the lncRNA signature and overall mRNA gene expression in each region, featuring TB and TB/DM groups.

To gain insights about the role of lncRNA in the biological pathways, we used an enrichment analysis of the strongly correlated genes in Brazil and India. Most of the genes associated with the pathways were correlated to three lncRNAs (*LINC02471*, *ADM-DT* and *GK-AS1*), while the other two presented minor influence on the enrichment. Our results revealed pathways that are related to the host's immune response against Mtb. Neutrophil degranulation is one of the main neutrophil activities when facing Mtb, as they release proteins which are antimicrobial, proteolytic or even structural. These proteins are incorporated into the neutrophil's membrane to change the cellular response toward the environment.[38] The granule released molecules can inhibit the bacterial replication within the contacted macrophages,[39] but can also harm the host, as they damage both bacterial and host cells.[40] Neutrophils are the first immune cells to enter the lungs during Mtb infection, in the immunopathological side, and they are critical cells for granuloma cavitation in the active TB.[41] On the other hand, they are indispensable to control the Mtb and to induce the anti-Mtb adaptive immune response.[42] Both Signaling by and Inactivation of CSF3 (G−CSF) pathways regulate the hematopoietic proliferation of neutrophils, by the cytokine Granulocyte colony-stimulating factor (G-CSF).[43] During infections, G-CSF is induced by inflammatory cytokines, such as IL-1, TNFα and lipopolysaccharide (LPS).[44] This pathway causes its own inactivation to prevent an overpopulation of neutrophils, explaining both pathways representing its activation and inactivation being enriched simultaneously.[45]

The role of interleukins (IL) in modulating the inflammatory response toward Mtb have been largely explored.[46,47] It is known that IL-12 and IFN-γ play a crucial role in protecting the host against Mtb infection, as both molecules and their induced Th1 immune response have been extensively explored in TB.[48] There are also interleukins that can be induced by Mtb to impair the host Th1 response, such as IL-10, a potent immune regulatory interleukin. This interleukin reduces the antigen presentation and IL-12 production, enhancing intracellular bacteria survivability by inhibiting macrophages phagosomal maturation and cellular apoptosis.[49] Regarding the second enriched pathway, both IL-4 and IL-13 are associated with either the Th2 arm and have been associated with lung damage in TB.[50] IL-4 enhances macrophage endocytosis by mannose receptor, a major route of Mtb infection[51] and its suppression enhances the host resistance against Mtb in mice animal models.[52] Moreover, IL-13 upregulation in TB enhances Mtb replication and necrotizing granulomas in TB mice experimental model.[53]

The Interferon signaling is crucial for anti-TB immune response,[54] as known by exploring the IL-12/IFN-γ-mediated Th1 immune response,[55] IFN-γ macrophage and $CD8^+$ T cells activation to kill intracellular Mtb and to lyse host infected cells, respectively. Nevertheless, the excessive Th1 response activation through IFN-γ can be detrimental to the host, often leading to tecidual damage and necrosis.[56] Thus, the Interferon signaling is related to the commonly enriched pathway of Regulated Necrosis, which is mainly induced by the Th1 response byproduct and Neutrophil-produced reactive oxygen species.[57]

Regarding the last four commonly enriched pathways, MyD88 deficiency (TLR2/4) and IRAK4 deficiency (TLR2/4) are part of the Diseases associated with the TLR signaling cascade, which is a participant of the Diseases of Immune System pathway (stable ID R-HSA-5260271 in the Reactome database). Thus, all four pathways are related to Diseases of Immune System, perhaps this major pathway has been enriched due to the alterations in the host's immune system caused by the Mtb, as this pathogen impairs the host's adaptive immunological response.[58] Despite the lack of information about the biological functions regarding the five selected lncRNAs, the enrichment of important pathways to the immune response toward TB by the correlated genes is a great indicator of biological consistency in our findings.

In the present study, we have analyzed data from samples collected in Brazil, India, Romania, and South Africa and identified a set of lncRNAs with consistent accuracy at all four countries' study populations. Despite the lack of information regarding its biological functions in the literature, the five most informative lncRNAs were strongly correlated with genes associated with pathways related to immune response regulation against TB. This suggests that these lncRNAs may play an important role in the regulation of genes related to TB response, but

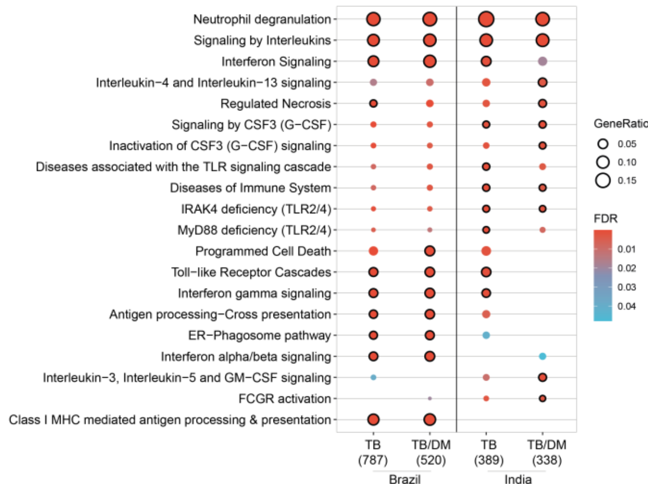


Figure 4. Correlated genes' enrichment

Dot plot displaying the top 10 gene ratio enriched pathways in each region, using the REACTOME database. Dot colors represent the statistical significance (FDR), while dot size represents the gene ratio of each enriched pathway. Dots with black circles around represent the top 10 gene ratio pathway for its respective region. The pathway names are displayed at the Y axis, while the region and group is displayed at the X axis.

further studies are still required to enlighten their biological functions and regulation mechanisms. We propose this highly consistent set of lncRNAs as biomarkers for TB, regardless of DM status.

Limitations of study

This work has some limitations, starting with the methodology used in the RNA-sequencing, as the data from South Africa and Romania were sequenced in single-end platforms, while the data from the India region has been sequenced in a paired-end platform. The employed negative binomial model for differential gene expression analysis has limitations related to multiple variable adjustments. Furthermore, as this present work employs the same samples as our previous work, all limitations regarding the metadata related to samples have been inherited as well, such as the observed differences in BMI, age, sex, smoking and alcohol use between the Brazilian and Indian populations.[24] Moreover, some patients were under treatment with metformin and statin, which could affect the inflammatory responses.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Participants enrollment and data acquisition
- METHOD DETAILS
  - Data preprocessing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Differential expression analysis and ncRNAs filtering
  - Machine learning - Random forest application and validation on independent datasets
  - Correlations lncRNAs - mRNAs and enrichment analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108662.

ACKNOWLEDGMENTS

We thank Ms. Daphne Martin and Ms. Samyra Cox for outstanding administrative support, Mr. Paul Simon and Mr. Art Garfunkel for the amazing inspiration they give to us for reporting the poetry of silent RNAs. This work was funded by: OISE-17-63459-1 from the National Institutes of Health, administered by CRDF Global; DAA3-18-64718-1, formerly USB1-31149-XX-13 from the Indo-US Vaccine Action Initiative on TB Research, administered by CRDF Global. The Brazilian site was supported by the National Institutes of Health (NIH U01AI069923 and R01AI120790), CCASAnet, RePORT-Brazil Tennessee Center for AIDS Research (TN-CFAR). The study was also supported by the Intramural Research Program of the Fundação José Silveira, and the Intramural Research Program of the Oswaldo Cruz Foundation, Brazil. BBA and ATLQ are senior investigators from the Brazilian Council for Science and Technology (CNPq). ERF and MBA received research fellowship from the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB).

AUTHOR CONTRIBUTIONS

C.L.V., E.R.F. T.F.M, M.A.P, and M.B.A. performed the data curation, analysis, interpretation, and draft of the first version of the article. A.N.G., N.G., V.K, S.G., R.K., V.M., M. P., and A.G. performed the data interpretation, revising article critically for important intellectual content, final approval of the article. N.P.K. performed sample preparation and curation for the Chennai cohort. T.R.S. supervised the Brazilian study and helped with data interpretation. S.B. and V.V. coordinated the clinical study in Chennai. A.T.Q.L. and B.B.A. performed the study conceptualization, data analysis and interpretation, and draft of the article. H.K. performed data curation, analysis and interpretation, revising article critically for important intellectual content and coordinated all sites studies.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 5, 2023
Revised: November 27, 2023
Accepted: December 5, 2023
Published: December 8, 2023

REFERENCES

1. Global Tuberculosis Report 2021 (2021 (World Health Organization).
2. Houben, R.M.G.J., and Dodd, P.J. (2016). The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. PLoS Med. 13, e1002152.
3. Narasimhan, P., Wood, J., MacIntyre, C.R., and Mathai, D. (2013). Risk factors for tuberculosis. Pulm. Med. 2013, 828939.
4. Zimmet, P., Alberti, K.G., and Shaw, J. (2001). Global and societal implications of the diabetes epidemic. Nature 414, 782–787.
5. American Diabetes Association (2013). Diagnosis and classification of diabetes mellitus. Diabetes Care 36, S67–S74.
6. Ruslami, R., Aarnoutse, R.E., Alisjahbana, B., van der Ven, A.J.A.M., and van Crevel, R. (2010). Implications of the global increase of diabetes for tuberculosis control and patient care. Trop. Med. Int. Health 15, 1289–1299.
7. Mantovani, A., and Garlanda, C. (2023). Humoral Innate Immunity and Acute-Phase Proteins. N. Engl. J. Med. 388, 439–452.
8. Jiménez-Corona, M.E., Cruz-Hervert, L.P., García-García, L., Ferreyra-Reyes, L., Delgado-Sánchez, G., Bobadilla-Del-Valle, M., Canizales-Quintero, S., Ferreira-Guerrero, E., Báez-Saldaña, R., Téllez-Vázquez, N., et al. (2013). Association of diabetes and tuberculosis: impact on treatment and post-treatment outcomes. Thorax 68, 214–220.
9. Liu, T., Wang, Y., Gui, J., Fu, Y., Ye, C., Hong, X., Chen, L., Li, Y., Zhang, X., and Hong, W. (2022). Transcriptome analysis of the impact of diabetes as a comorbidity on tuberculosis. Medicine 101, e31652.
10. Anastasiadou, E., Jacob, L.S., and Slack, F.J. (2018). Non-coding RNA networks in cancer. Nat. Rev. Cancer 18, 5–18.
11. Kazimierczyk, M., Kasprowicz, M.K., Kasprzyk, M.E., and Wrzesinski, J. (2020). Human Long Noncoding RNA Interactome: Detection, Characterization and Function. Int. J. Mol. Sci. 21, 1027.
12. Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. 81, 145–166.
13. Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P.P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 146, 353–358.
14. Ulitsky, I. (2018). Interactions between short and long noncoding RNAs. FEBS Lett. 592, 2874–2883.
15. Romero-Barrios, N., Legascue, M.F., Benhamed, M., Ariel, F., and Crespi, M. (2018). Splicing regulation by long noncoding RNAs. Nucleic Acids Res. 46, 2169–2184.
16. Gott, J.M., and Emeson, R.B. (2000). Functions and mechanisms of RNA editing. Annu. Rev. Genet. 34, 499–531.
17. Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., et al. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol. Cell 39, 925–938.
18. Fathizadeh, H., Hayat, S.M.G., Dao, S., Ganbarov, K., Tanomand, A., Asgharzadeh, M., and Kafil, H.S. (2020). Long non-coding RNA molecules in tuberculosis. Int. J. Biol. Macromol. 156, 340–346.
19. Wang, Y., Zhong, H., Xie, X., Chen, C.Y., Huang, D., Shen, L., Zhang, H., Chen, Z.W., and Zeng, G. (2015). Long noncoding RNA derived from CD244 signaling epigenetically controls CD8+ T-cell immune responses in tuberculosis infection. Proc. Natl. Acad. Sci. USA 112, E3883–E3892.
20. Zhang, Q., Chao, T.-C., Patil, V.S., Qin, Y., Tiwari, S.K., Chiou, J., Dobin, A., Tsai, C.-M., Li, Z., Dang, J., et al. (2019). The long noncoding RNA regulates inflammatory gene expression. EMBO J. 38, e100041.
21. Ismail, N., Abdullah, N., Abdul Murad, N.A., Jamal, R., and Sulaiman, S.A. (2021). Long Non-Coding RNAs (lncRNAs) in Cardiovascular Disease Complication of Type 2 Diabetes. Diagnostics 11, 145.
22. Yang, X., Yang, J., Wang, J., Wen, Q., Wang, H., He, J., Hu, S., He, W., Du, X., Liu, S., and Ma, L. (2016). Microarray analysis of long noncoding RNA and mRNA expression profiles in human macrophages infected with. Sci. Rep. 6, 38963.

23. Hu, X., Liao, S., Bai, H., Gupta, S., Zhou, Y., Zhou, J., Jiao, L., Wu, L., Wang, M., Chen, X., et al. (2020). Long Noncoding RNA and Predictive Model To Improve Diagnosis of Clinically Diagnosed Pulmonary Tuberculosis. J. Clin. Microbiol. 58, e01973-19.
24. Queiroz, A.T.L., Vinhaes, C.L., Fukutani, E.R., Gupte, A.N., Kumar, N.P., Fukutani, K.F., Arriaga, M.B., Sterling, T.R., Babu, S., Gaikwad, S., et al. (2023). A multi-center, prospective cohort study of whole blood gene expression in the tuberculosis-diabetes interaction. Sci. Rep. 13, 7769.
25. van der Heijden, Y.F., Abdullah, F., Andrade, B.B., Andrews, J.R., Christopher, D.J., Croda, J., Ewing, H., Haas, D.W., Hatherill, M., Horsburgh, C.R., Jr., et al. (2018). Building capacity for advances in tuberculosis research; proceedings of the third RePORT international meeting. Tuberculosis 113, 153–162.
26. Kornfeld, H., West, K., Kane, K., Kumpatla, S., Zacharias, R.R., Martinez-Balzano, C., Li, W., and Viswanathan, V. (2016). High Prevalence and Heterogeneity of Diabetes in Patients With TB in South India: A Report from the Effects of Diabetes on Tuberculosis Severity (EDOTS) Study. Chest 149, 1501–1508.
27. Gupte, A., Padmapriyadarsini, C., Mave, V., Kadam, D., Suryavanshi, N., Shivakumar, S.V.B.Y., Kohli, R., Gupte, N., Thiruvengadam, K., Kagal, A., et al. (2016). Cohort for Tuberculosis Research by the Indo-US Medical Partnership (CTRIUMPH): protocol for a multicentric prospective observational study. BMJ Open 6, e010542.
28. Eckold, C., Kumar, V., Weiner, J., Alisjahbana, B., Riza, A.-L., Ronacher, K., Coronel, J., Kerry-Barnard, S., Malherbe, S.T., Kleynhans, L., et al. (2021). Impact of Intermediate Hyperglycemia and Diabetes on Immune Dysfunction in Tuberculosis. Clin. Infect. Dis. 72, 69–78.
29. Fukutani, E.R., Ramos, P.I.P., Kasprzykowski, J.I., Azevedo, L.G., Rodrigues, M.M.d.S., Lima, J.V.d.O.P., de Araujo Junior, H.F.S., Fukutani, K.F., and de Queiroz, A.T.L. (2019). Meta-Analysis of HTLV-1-Infected Patients Identifies CD40LG and GBP2 as Markers of ATLL and HAM/TSP Clinical Status: Two Genes Beat as One. Front. Genet. 10, 1056.
30. Fukutani, K.F., Kasprzykowski, J.I., Paschoal, A.R., Gomes, M.d.S., Barral, A., de Oliveira, C.I., Ramos, P.I.P., and de Queiroz, A.T.L. (2017). Meta-Analysis of Expression Datasets: Comparing Virus Infection and Blood-Fed Transcriptomes to Identify Markers of Virus Presence. Front. Bioeng. Biotechnol. 5, 84.
31. Fukutani, E., Rodrigues, M., Kasprzykowski, J.I., Araujo, C.F.d., Paschoal, A.R., Ramos, P.I.P., Fukutani, K.F., and Queiroz, A.T.L.d. (2018). Follow up of a robust meta-signature to identify Zika virus infection in Aedes aegypti: another brick in the wall. Mem. Inst. Oswaldo Cruz 113, e180053.
32. Yang, L., Wu, H., Jin, X., Zheng, P., Hu, S., Xu, X., Yu, W., and Yan, J. (2020). Study of cardiovascular disease prediction model based on random forest in eastern China. Sci. Rep. 10, 5245.
33. Prada-Medina, C.A., Fukutani, K.F., Pavan Kumar, N., Gil-Santana, L., Babu, S., Lichtenstein, F., West, K., Sivakumar, S., Menon, P.A., Viswanathan, V., et al. (2017). Systems Immunology of Diabetes-Tuberculosis Comorbidity Reveals Signatures of Disease Complications. Sci. Rep. 7, 1999.
34. van Doorn, C.L.R., Eckold, C., Ronacher, K., Ruslami, R., van Veen, S., Lee, J.-S., Kumar, V., Kerry-Barnard, S., Malherbe, S.T., Kleynhans, L., et al. (2022). Transcriptional profiles predict treatment outcome in patients with tuberculosis and diabetes at diagnosis and at two weeks after initiation of anti-tuberculosis treatment. EBioMedicine 82, 104173.
35. Restrepo, B.I. (2016). Diabetes and Tuberculosis. Microbiol. Spectr. 4.
36. Arriaga, M.B., Araujo-Pereira, M., Barreto-Duarte, B., Nogueira, B., Freire, M.V.C.N.S., Queiroz, A.T.L., Rodrigues, M.M.S., Rocha, M.S., Souza, A.B., Spener-Gomes, R., et al. (2022). The Effect of Diabetes and Prediabetes on Antituberculosis Treatment Outcomes: A Multicenter Prospective Cohort Study. J. Infect. Dis. 225, 617–626.
37. Calderon, R.I., Arriaga, M.B., Aliaga, J.G., Barreda, N.N., Sanabria, O.M., Barreto-Duarte, B., Franco, J.P.D., Lecca, L., Andrade, B.B., Carvalho, A.C.C., and Kritski, A.L. (2022). Persistent dysglycemia is associated with unfavorable treatment outcomes in patients with pulmonary tuberculosis from Peru. Int. J. Infect. Dis. 116, 293–301.
38. Borregaard, N., Sørensen, O.E., and Theilgaard-Mönch, K. (2007). Neutrophil granules: a library of innate immunity proteins. Trends Immunol. 28, 340–345.
39. Tan, B.H., Meinken, C., Bastian, M., Bruns, H., Legaspi, A., Ochoa, M.T., Krutzik, S.R., Bloom, B.R., Ganz, T., Modlin, R.L., and Stenger, S. (2006). Macrophages acquire neutrophil granules for antimicrobial activity against intracellular pathogens. J. Immunol. 177, 1864–1871.
40. Dallenga, T., and Schaible, U.E. (2016). Neutrophils in tuberculosis–first line of defence or booster of disease and targets for host-directed therapy? Pathog. Dis. 74, ftw012.
41. Ong, C.W.M., Elkington, P.T., Brilha, S., Ugarte-Gil, C., Tome-Esteban, M.T., Tezera, L.B., Pabisiak, P.J., Moores, R.C., Sathyamoorthy, T., Patel, V., et al. (2015). Neutrophil-Derived MMP-8 Drives AMPK-Dependent Matrix Destruction in Human Pulmonary Tuberculosis. PLoS Pathog. 11, e1004917.
42. Blomgran, R., and Ernst, J.D. (2011). Lung neutrophils facilitate activation of naive antigen-specific CD4+ T cells during Mycobacterium tuberculosis infection. J. Immunol. 186, 7110–7119.
43. Roberts, A.W. (2005). G-CSF: a key regulator of neutrophil production, but that's not all. Growth Factors 23, 33–41.
44. Demetri, G.D., and Griffin, J.D. (1991). Granulocyte colony-stimulating factor and its receptor. Blood 78, 2791–2808.
45. Beekman, R., and Touw, I.P. (2010). G-CSF and its receptor in myeloid malignancy. Blood 115, 5131–5136.
46. Kalsum, S. (2019). Characterizing Phenotypes of Mycobacterium tuberculosis and Exploring Anti-mycobacterial Compounds through High Content Screening (Linköping University Electronic Press).
47. He, X.-Y., Xiao, L., Chen, H.-B., Hao, J., Li, J., Wang, Y.-J., He, K., Gao, Y., and Shi, B.-Y. (2010). T regulatory cells and Th1/Th2 cytokines in peripheral blood from tuberculosis patients. Eur. J. Clin. Microbiol. Infect. Dis. 29, 643–650.
48. Cooper, A.M., Kipnis, A., Turner, J., Magram, J., Ferrante, J., and Orme, I.M. (2002). Mice lacking bioactive IL-12 can generate protective, antigen-specific cellular responses to mycobacterial infection only if the IL-12 p40 subunit is present. J. Immunol. 168, 1322–1327.
49. Abdalla, A.E., Lambert, N., Duan, X., and Xie, J. (2016). Interleukin-10 Family and Tuberculosis: An Old Story Renewed. Int. J. Biol. Sci. 12, 710–717.
50. van Crevel, R., Karyadi, E., Preyers, F., Leenders, M., Kullberg, B.J., Nelwan, R.H., and van der Meer, J.W. (2000). Increased production of interleukin 4 by CD4+ and CD8+ T cells from patients with tuberculosis is related to the presence of pulmonary cavities. J. Infect. Dis. 181, 1194–1197.
51. Ernst, J.D. (1998). Macrophage receptors for Mycobacterium tuberculosis. Infect. Immun. 66, 1277–1281.
52. Buccheri, S., Reljic, R., Caccamo, N., Ivanyi, J., Singh, M., Salerno, A., and Dieli, F. (2007). IL-4 depletion enhances host resistance and passive IgA protection against tuberculosis infection in BALB/c mice. Eur. J. Immunol. 37, 729–737.
53. Heitmann, L., Abad Dar, M., Schreiber, T., Erdmann, H., Behrends, J., Mckenzie, A.N.J., Brombacher, F., Ehlers, S., and Hölscher, C. (2014). The IL-13/IL-4Rα axis is involved in tuberculosis-associated pathology. J. Pathol. 234, 338–350.
54. Chin, K.L., Anis, F.Z., Sarmiento, M.E., Norazmi, M.N., and Acosta, A. (2017). Role of Interferons in the Development of Diagnostics, Vaccines, and Therapy for Tuberculosis. J. Immunol. Res. 2017, 5212910.
55. O'Garra, A., Redford, P.S., McNab, F.W., Bloom, C.I., Wilkinson, R.J., and Berry, M.P.R. (2013). The immune response in tuberculosis. Annu. Rev. Immunol. 31, 475–527.
56. Sakai, S., Kauffman, K.D., Sallin, M.A., Sharpe, A.H., Young, H.A., Ganusov, V.V., and Barber, D.L. (2016). CD4 T Cell-Derived IFN-γ Plays a Minimal Role in Control of Pulmonary Mycobacterium tuberculosis Infection and Must Be Actively Repressed by PD-1 to Prevent Lethal Disease. PLoS Pathog. 12, e1005667.
57. Dallenga, T., Repnik, U., Corleis, B., Eich, J., Reimer, R., Griffiths, G.W., and Schaible, U.E. (2017). M. tuberculosis-Induced Necrosis of Infected Neutrophils Promotes Bacterial Growth Following Phagocytosis by Macrophages. Cell Host Microbe 22, 519–530.e3.
58. Chandra, P., Grigsby, S.J., and Philips, J.A. (2022). Immune evasion and provocation by Mycobacterium tuberculosis. Nat. Rev. Microbiol. 20, 750–766.
59. Hamilton, C.D., Swaminathan, S., Christopher, D.J., Ellner, J., Gupta, A., Sterling, T.R., Rolla, V., Srinivasan, S., Karyana, M., Siddiqui, S., et al. (2015). RePORT International: Advancing Tuberculosis Biomarker Research Through Global Collaboration. Clin. Infect. Dis. 61Suppl 3, S155–S159.
60. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

61. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.
62. Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 4, 1521. F1000Res.
63. Jeffrey, T.L., Evan, W.J., Hilary, S.P., Elana, J.F., Andrew, E.J., John, D.S., Yuqing, Z., and Leonardo, C.T. (2017). sva: Surrogate Variable Analysis. Bioconductor R package.
64. Melissa, L., Pedro, R., and Helder, N. (2018). mdp: Molecular Degree of Perturbation calculates scores for transcriptome data samples based on their perturbation from controls. Bioconductor R package.
65. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550.
66. Rainer, J., Gatto, L., and Weichenberger, C.X. (2019). ensembldb: an R package to create and use Ensembl-based annotation resources. Bioinformatics 35, 3151–3153.
67. Liaw, A., and Wiener, M. (2002). Classification and Regression by randomForest.
68. Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. J. Stat. Softw. 28.
69. Johnson, W.E., Odom, A., Cintron, C., Muthaiah, M., Knudsen, S., Joseph, N., Babu, S., Lakshminarayanan, S., Jenkins, D.F., Zhao, Y., et al. (2021). Comparing tuberculosis gene signatures in malnourished individuals using the TBSignatureProfiler. BMC Infect. Dis. 21, 106.
70. (2023). Hmisc. https://hbiostat.org/R/Hmisc/.
71. Mukaka, M.M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med. J. 24, 69–71.
72. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation 2, 100141.
73. Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., Griss, J., Sevilla, C., Matthews, L., Gong, C., et al. (2022). The reactome pathway knowledgebase 2022. Nucleic Acids Res. 50, D687–D692.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw RNAseq data from TANDEM	Eckold et al.[28]	Bioproject ID: PRJNA470512
Raw RNAseq from RePORT	Kornfeld et al.,[26] Gupte et al.[27], Hamilton et al.[59]	GEOncbi ID: GSE181143
Software and algorithms		
R version 4.2.2.	R Core Team	https://cran.r-project.org/
sra-tools version 3.0.6	Available at https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software	https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view = software
Trimmomatic version 0.32	https://doi.org/10.1093/bioinformatics/btu170	http://www.usadellab.org/cms/index.php?page = trimmomatic
STAR version 2.7.10	https://doi.org/10.1093/bioinformatics/bts635	https://code.google.com/archive/p/rna-star/
tximportv version 1.28.0	https://doi.org/10.18129/B9.bioc.tximport	https://bioconductor.org/packages/release/bioc/html/tximport.html
Sva version 3.48.0	https://doi.org/10.18129/B9.bioc.sva	https://bioconductor.org/packages/release/bioc/html/sva.html
mdp version 1.20.0	https://doi.org/10.18129/B9.bioc.mdp	https://bioconductor.org/packages/release/bioc/html/mdp.html
DESeq2 version 1.40.2	https://doi.org/10.1186/s13059-014-0550-8	https://www.bioconductor.org/packages/release/bioc/html/DESeq2.html
ensembldb version 2.24.0	https://doi.org/10.1093/bioinformatics/btz031	https://bioconductor.org/packages/release/bioc/html/ensembldb.html
randomForest version 4.7–1.1	Liaw et al.[67]	https://cran.r-project.org/web/packages/randomForest/index.html
caret version 6.0–94	Kuhn et al.[68]	https://cran.r-project.org/web/packages/caret/index.html
clusterProfiler version 4.8.2	https://doi.org/10.18129/B9.bioc.clusterProfiler	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests regarding the packages employed for the analysis performed in this study should be directed to and will be fulfilled, if possible, by the lead contact, Artur Trancoso Lopo de Queiroz (arturlopo@gmail.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

(1) The data from TANDEM have been previously deposited at the SRA database and are publicly available as of the date of publication (BioProject: PRJNA470512). The MSTDI gene expression data have been deposited at the GEO ncbi database and is publicly available as of the date of publication (GEOncbi: GSE181143, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE181143).
(2) All employed packages' references are available at the key resources table and methodology.
(3) This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Participants enrollment and data acquisition

The current study used an already published data (GEO NCBI accession number GSE181143), Regional Prospective Observational Research in Tuberculosis (RePORT), from India and Brazil consortia. Protocols have been approved by the Ethics Committees of the Prof. M. Viswanathan Diabetes Research Center and the Institutional Review Boards of Byramjee Jeejeebhoy Government Medical College, Pune and National Institute for Research in Tuberculosis and Johns Hopkins University. Participants enrolled from the RePORT Brazil had their protocols approved by Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, as well as Vanderbilt University Medical Center institutional review boards. Written informed consent was obtained from all participants. The enrollment of participants was prospective at two sites of the RePORT-India Consortium and one site of RePORT-Brazil, with organizational support provided by RePORT-International.[25] The Indian sites were located in Chennai (EDOTS study)[26] and Pune (CTRIUMPh study),[27] while the Brazilian site was in Salvador (RePORT International Common Protocol).[59] Furthermore, data from samples enrolled at the TANDEM study[28] were also retrieved, featuring samples from Indonesia, Peru, Romania and South Africa. However, only the Romania and South Africa regions had site-specific control patient data, being the only regions included in this study. Participant groups included active pulmonary TB disease, with or without DM (TB and TB/DM groups, respectively) and one control group, composed of healthy controls (HC). Inclusion criteria were age 18-65 and new diagnosis of active pulmonary TB (or absence of pulmonary TB for the control group participants). Drug-resistant TB, retreatment, treatment of incident TB for >7 days prior to enrollment, pregnancy, immunosuppressive medications and HIV infection were the exclusion criteria.

The combined 322 cohort comprised 160 participants from India, being 90 participants from Chennai and 70 from Pune (60 HC, 60 TB and 40 TB/DM), 75 participants from Brazil (15 HC, 29 TB and 31 TB/DM), 37 participants from Romania (12 HC, 10 TB and 15 TB/DM) and 50 participants from South Africa (24 HC, 11 TB and 15 TB/DM).

METHOD DETAILS

Data preprocessing

Raw RNA-seq data from the MSTDI cohort were retrieved from *Illumina HiSeq 2500* platform.[24] Sequence data from the TANDEM cohort was retrieved from the SRA database using BioProject PRJNA470512. Both MSTDI and TANDEM raw data were retrieved using the SRA tools and fastq files were processed identically to MSTDI data. Sequence data from MSTDI and TANDEM were prepared by removing low-quality bases and trimming adapters using *Trimmomatic V0.32*.[60] After the quality check, sequences were aligned against the human transcriptome (GRCh38 version) comprising both mRNA and ncRNA with the STAR algorithm v2.7.10.[61] After mapping, the outputs were converted to count tables using tximport package.[62] The India dataset was obtained by merging the Chennai and Pune individual datasets, followed by batch effect correction using the *sva* package.[63] Outliers detection among the samples in each individual dataset was performed by the *mdp* package,[64] after performing data normalization in each dataset using the variance stabilizing transformation, from the *DESeq2* package.[65] All data processing, post mapping, and downstream analysis have been performed in R environment v4.2.2.

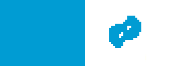
QUANTIFICATION AND STATISTICAL ANALYSIS

Differential expression analysis and ncRNAs filtering

The data analysis and biomarker discovery were performed in the Brazilian dataset, consisting of 89 non-outlier samples (15 DM, 14 HC, 29 TB and 31 TB/DM), whereas the other datasets were used as validation dataset. Differential expression analysis has been employed to compare the gene expressions between the groups: HC vs. TB and HC vs. TB/DM, identifying the differentially expressed genes (DEGs). This analysis has been performed using the *DESeq2* package using the raw transcriptomic count tables. The log2 fold change and significance values of all genes were calculated by applying generalized linear models to the data, considering the mean and dispersion values of each gene. For a gene to be considered a DEG, we used the threshold of ± 1.4 log2 fold change and false discovery rate (FDR) < 0.05. Afterward, the non-coding RNAs (miRNAs and lncRNAs) were identified within the DEGs, using the *ensembldb* package query with the *homo sapiens ensdb* version AH109336 [66]. The miRNAs and lncRNAs were retrieved using the gene biotype variable and the assessed database was the Ensembl 108 EnsDb for Homo sapiens.[66]

Machine learning - Random forest application and validation on independent datasets

Afterward, the transcriptomic data containing miRNAs and lncRNAs was normalized using the varianceStabilizingTransformation function from the DESeq2 package. Following, we applied the random forest algorithm[67] using the DEncRNAs normalized expression data, alongside the disease categories TB and TB/DM, plus the healthy (HC) as factors for performing the classification. This algorithm aims to identify the best variables to distinguish the sample groups and has been employed due to its ability to handle multicollinearity better than linear models, such as Lasso in example. Collinearity is frequently observed in gene expression data, since different genes can be associated with the same pathway. A total of 10000 decision trees were performed by the RF, mtry parameter was set to 50. The best variables were selected using the Mean decreasing accuracy and Mean decrease gini > third quartile as criteria, which are directly related to the variable importance when classifying samples. Selected variables were retrieved from the dataset and their accuracy was evaluated using the area under the curve (AUC) value, using receiver operating characteristic (ROC) curves. The k-fold cross validation was also performed to evaluate the RF model, using the caret package's confusion matrix to assess the model's overall performance, with 100-folds and 25 repetitions.[68] The inputs for the confusion matrix were the real classes for each sample, alongside with predictions performed by the model trained with the "rf" method. The



trControl parameter was set using the trainControl function with method "repeatedcv", fold of 100 and repeating 25 times. Furthermore, the sample overall dispersion among the groups was assessed in a heatmap using the biomarkers Z-scaled expression values with Manhattan distance calculation and Ward test.

To validate the random forest model's accuracy and consistency in other populations, independent datasets with samples from India, Romania and South Africa were employed. Thus, the biomarker genes expression values were used to classificate the samples in each dataset. In order to assess the classification overall performance, ROC curves were employed for each independent dataset. The previously proposed TB signatures were retrieved from the TBSignatureProfiles package,[69] and the same method was used in order to evaluate the perfomance of our identified lncRNA signature in comparison to the previous TB signatures.

Correlations lncRNAs - mRNAs and enrichment analysis

To assess how the lncRNAs selected by RF could impact the overall gene expression in each region and group, we perfomed a correlation analysis using the Spearman rho rank coefficient between these selected lncRNAs and all mRNA genes. The correlations have been made for each TB infected group (TB and TB/DM) in each region datasets (Brazil, India, Romania and South Africa). Only highly positive/negative correlations ($|rho| > 0.7$) and False Discovery Rate (FDR) < 0.05 were considered.[70,71] Afterward, the fold changes and entrez IDs of each correlated transcript were used as input to perform the enrichment analysis, using the clusterProfiler package.[72] The enrichment analysis was performed with the REACTOME database,[73] other parameters were: Minimum gene set size = 10, Maximum gene set size = 500, q value cutoff = 0.2, p value cutoff = 0.05, with FDR as p value adjustment method.

Supplementary Material

The sound of silent RNA: The role of lncRNAs on TB infection in four different populations

Fukutani ER, Vinhaes CL, Araújo-Pereira M, Mota TF, Gupte AN, Kumar NP, Arriaga MB, Sterling TR, Babu S, Gaikwad S, Karyakarte R, Mave V, Kulkarni V, Gupte N, Paradkar M, Viswanathan V, Kornfeld H, Gupta A, Andrade BB, Queiroz ATL, RePORT Brazil, RePORT India Consortia.

Supplementary Tables

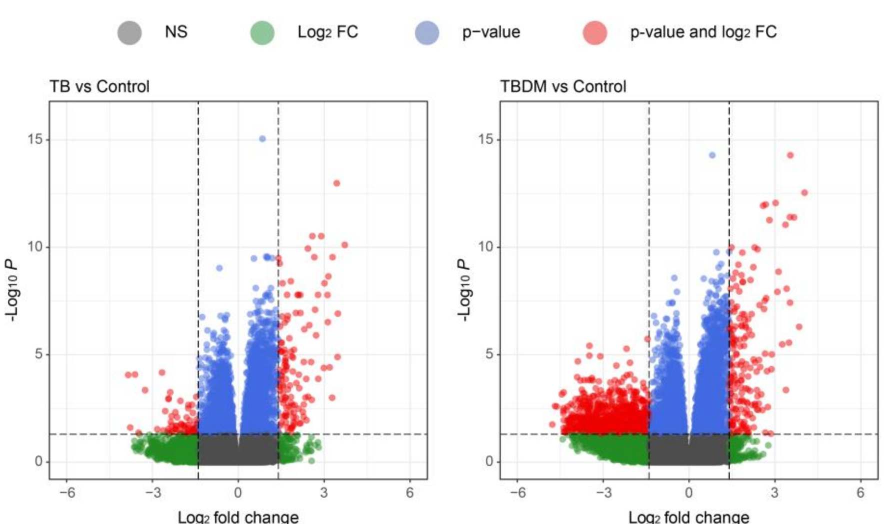
Supplementary table 1. Top 20 most informative variables identified by the random forest to discriminate between TB infected (TB and TB/DM) and healthy samples.

Variable	Mean Decrease Accuracy	Mean Decrease Gini
ADM-DT	79.0320117	8.37621125
LINC02009	41.1253564	3.89386917
LINC02471	32.8127581	2.73420516
SOX2-OT	24.3243278	1.44404636
GK-AS1	22.8644554	1.78390002
GBP1P1	22.7993965	1.76137836
MIR634	19.900712	1.15368108
FAM66E	17.7154344	0.82191718
RNASE2CP	17.1820544	1.58252723
PRKAR1B-AS2	14.3998826	1.61164586
CARMN	14.335319	0.63802151
CLRN1-AS1	13.4143081	0.95281672
CATIP-AS1	12.8489688	1.09357793
LINC00882	11.4597303	0.60009585
LINC00824	11.1180142	1.34911805
KIF23-AS1	11.0589127	0.65188955
NR2F1-AS1	7.38277416	0.34185621
BASP1-AS1	7.16056428	0.55554485
LINC01088	6.86485928	0.80397838
MED28-DT	5.98847452	1.27584097

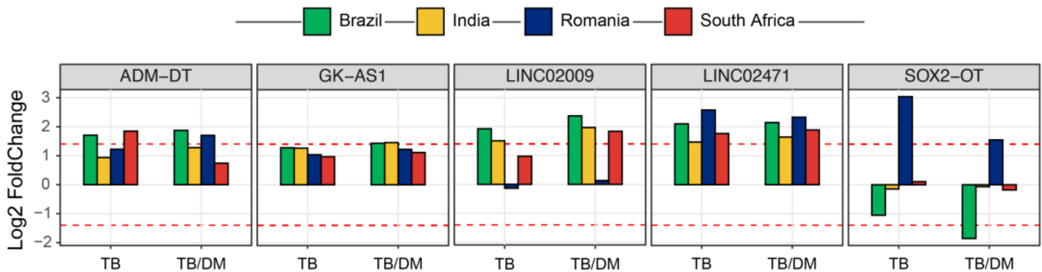
Supplementary table 2. Packages employed in this study

Package	Objective	Version	Platform	Reference
sra-tools	Download SRA dataset	3.0.6	Ubuntu Shell	Available at https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software
Trimmomatic	RNA-seq read quality control	0.32	Ubuntu Shell	DOI: 10.1093/bioinformatics/btu170
STAR	Align sequences	2.7.10	Ubuntu Shell	DOI: 10.1093/bioinformatics/bts635
tximport	Create count table	1.28.0	R 4.3.1	DOI: 10.18129/B9.bioc.tximport
sva	Correct batch effect	3.48.0	R 4.3.1	DOI: 10.18129/B9.bioc.sva
mdp	Detect potential outliers	1.20.0	R 4.3.1	DOI: 10.18129/B9.bioc.mdp
DESeq2	Identify DEGs and VST normalization	1.40.2	R 4.3.1	DOI: 10.1186/s13059-014-0550-8
ensembldb	Retrieve the list of miRNAs and lncRNAs	2.24.0	R 4.3.1	DOI: 10.1093/bioinformatics/btz031
randomForest	Select the most informative features	4.7-1.1	R 4.3.1	Liaw, A. & Wiener, M. Classification and Regression by randomForest. Preprint at https://CRAN.R-project.org/doc/Rnews/ (2002).
caret	K fold cross validation	6.0-94	R 4.3.1	Kuhn, M. Building Predictive Models in R Using the caret Package. J. Stat. Softw. 28, (2008).
ComplexHeatmap	Heatmap	2.16.0	R 4.3.1	DOI: 10.18129/B9.bioc.ComplexHeatmap
ggplot2	graphics	3.4.2	R 4.3.1	Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.
pROC	ROC curve	1.18.4	R 4.3.1	Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J, Müller M (2011). "pROC: an open-source package for R and S+ to analyze and compare ROC curves." *BMC Bioinformatics*, 12, 77.
clusterProfiler	Enrichment analysis	4.8.2	R 4.3.1	DOI: 10.18129/B9.bioc.sva

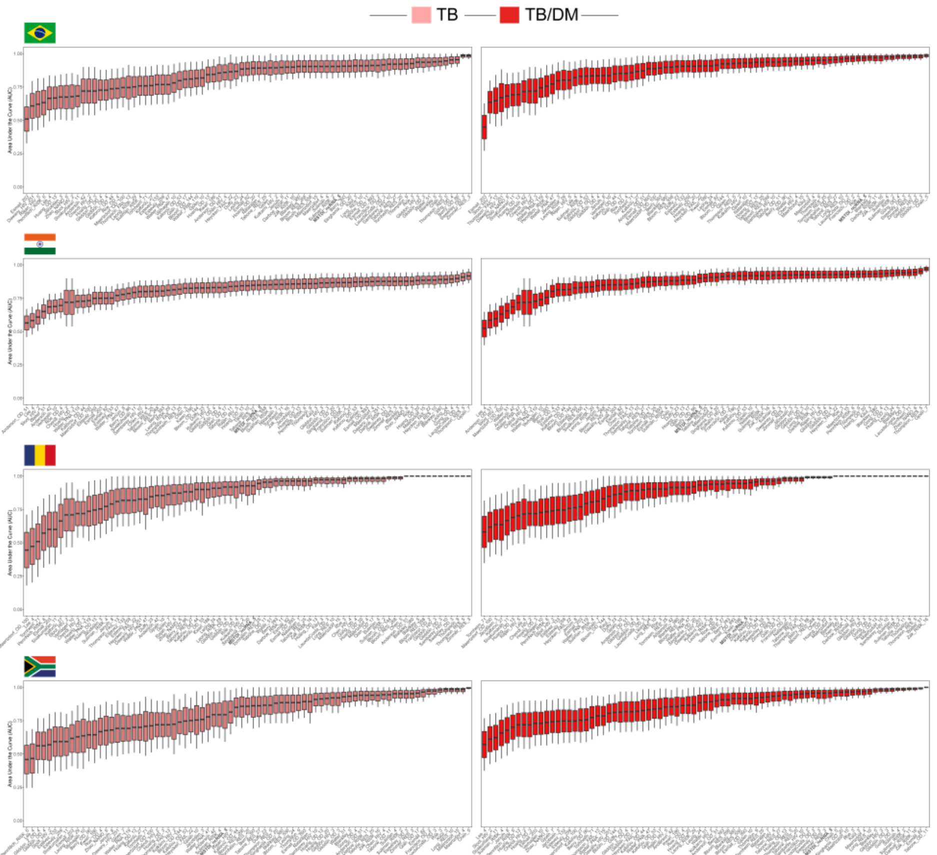
Supplementary figures



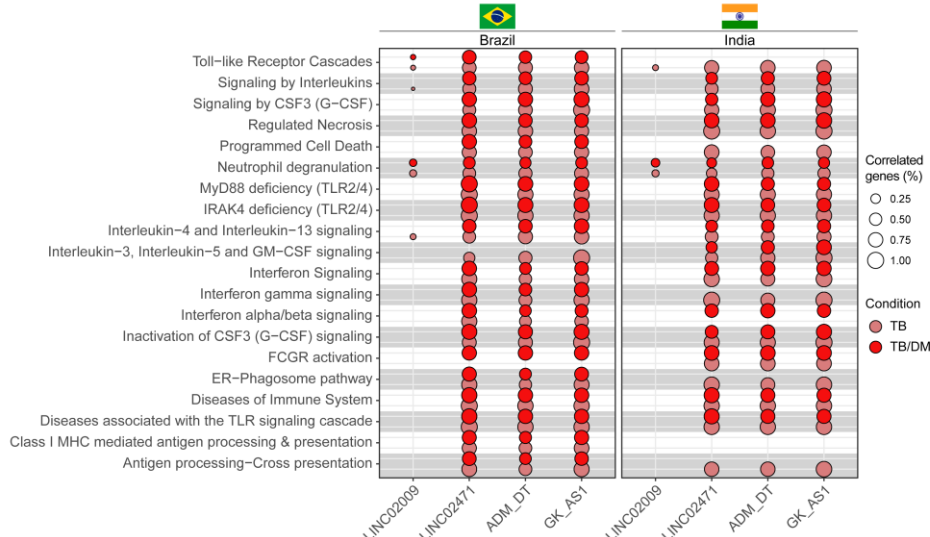
Supplementary figure 1. Volcano plot displaying the differentially expressed genes' log2 fold change and -Log10 p-value. Comparisons were made between the TB vs control and TB/DM vs control in the Brazil region dataset. Gray dots represent genes with no significance as DEG, green dots represent genes with sufficient log2 fold change but non statistical significance, in blue there's genes with significant p-value but no log2 fold change, red dots represent the DEGs.



Supplementary figure 2: Most informative lncRNAs fold changes across regions. Barplots displaying the 5 selected lncRNAs fold change per region (Brazil in green, India in yellow, Romania in blue and South Africa in red) and group (DM, TB and TB/DM). Dashed red lines represent the threshold of 1.4 and -1.4 log2 fold change.



Supplementary figure 3: Previously published TB signatures and most informative lncRNAs' overall performance when classifying the samples comprising this study. The AUCs for each previously proposed TB signatures were summarized in the boxplot when classifying TB and TB/DM samples.



Supplementary figure 4: Genes comprising each enriched pathway and their respective correlated lncRNA. Dot plot displaying the percentage of genes correlated with each lncRNA. Dot size is related to the percentage of genes comprising the pathway which are correlated with lncRNA. The X axis displays the lncRNAs, while the pathway names are displayed on the Y axis.

Artigo IV

Intersecting Epidemics: Deciphering the Complexities of Tuberculosis-Diabetes Comorbidity

Mariana Araujo-Pereira[1,2,3#], Caian L. Vinhaes[1,2,3,4,5#], Beatriz Barreto-Duarte[1,2,3,6], Klauss Villalva-Serra[1,2,7], Luís A. B. Cruz[1,2] and Bruno B. Andrade[1,2,3,4,6,7*]

[1]Laboratório de Pesquisa Clínica e Translacional (LPCT), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Bahia, Brazil.

[2]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Bahia, Brazil.

[3]Instituto de Pesquisa Clínica e Translacional (IPCT), Faculdade Zarns, Clariens Educação, Salvador, Bahia, Brazil.

[4]Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador, Bahia, Brazil.

[5]Departamento de Infectologia, Hospital Português da Bahia, Salvador, Bahia Brazil.

[6]Programa de Pós-Graduação em Clínica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

[7]Curso de Medicina, Universidade Salvador (UNIFACS), Salvador, Bahia, Brazil.

*Correspondent author: bruno.andrade@fiocruz.br

#These authors equally contributed to the work and share the first authorship.

Abstract

Amid the global health landscape, tuberculosis (TB) presents an ongoing challenge, demanding innovative strategies for its control. This review spotlights the intersection of TB with diabetes mellitus (DM), recognized by the World Health Organization as a key risk factor in the TB epidemic. Particularly prevalent in low and middle-income nations, the TB-DM comorbidity drives up TB rates through a nexus of chronic inflammation. By delving into the epidemiological, clinical, and inflammatory dimensions, we elucidate the impact of TB-DM on patient prognosis and the multifaceted complications it introduces to disease transmission, diagnosis, and treatment protocols. Our synthesis aims to offer a fresh lens on TB-DM, fostering a nuanced understanding that could inform future healthcare policies and interventions.

Introduction

In the contemporary global health landscape, tuberculosis (TB), a persistent challenge, intersects intricately with another widespread condition, diabetes mellitus (DM)[1,2]. Recognized by the World Health Organization (WHO) as a crucial risk factor in the TB epidemic, TB-DM comorbidity emerges as a significant concern, especially in low and middle-income countries[1]. The coexistence of TB-DM presents a unique challenge in global health, demanding a nuanced understanding of their interplay. This complex synergy significantly affects disease progression and individual outcomes. An integrated approach, incorporating both traditional epidemiological methods and advanced molecular techniques, is essential to fully comprehend and effectively address the TB-DM comorbidity.

TB remains one of the major causes of death by a single pathogen worldwide, leading to a global health concern, with the disease affecting around 10 million people each year[1]. The interplay between the *Mycobacterium tuberculosis* (Mtb) and the host, mediated by the inflammatory responses, determines a wide spectrum of clinical presentations. The challenge of reducing the disease burden ascends from different factors, including the influence of different comorbidities, mainly those that affect the quality of inflammatory responses, such as HIV[3], malnutrition[4] and DM.

DM is a chronic metabolic disorder that represents a dramatically high burden to healthcare systems worldwide. It has been reported that around half a billion persons live with this disease[2], mainly in low- and middle-income countries. Additionally, DM is a significant contributor to global morbidity and mortality, leading to an array of severe complications, including kidney failure, stroke, and heart disease, and is directly responsible for a substantial number of deaths annually[2]. DM affects the metabolism through multiple mechanisms, many of them related to the activation of poorly controlled pro-inflammatory pathways [5,6], that influence disease progression and susceptibility to infectious diseases, such as SARS-CoV-2[7] and Mtb[8].

In this context, the TB-DM comorbidity has garnered significant attention in the last years. Interestingly, some studies revealed regional disparities in the immune profile resulting from this interaction and emphasized the influence of socio-demographic and clinical factors on both diseases[9,10]. This article delves into the multifaceted relationship between TB and DM, exploring how this interplay exacerbates TB incidence. Our review compiles recent findings on the epidemiological, clinical, and inflammatory aspects of this nexus, highlighting its profound implications on patient outcomes and the broader challenges it poses in disease management. By scrutinizing the TB-DM interconnection, we aim to provide a comprehensive multiplatform perspective that not only sheds light on the complexities of this comorbidity but also suggests pathways for innovative healthcare strategies and policy formulations.

Epidemiological Trends of DM Among TB Patients

The intersection of DM and TB poses a global health challenge, as DM is recognized by the WHO as a primary risk factor for new TB cases. Interestingly,

many of the top ten countries with the highest DM rates are also significant contributors to the global TB burden, creating a dual challenge that has serious public health implications[11]. Grasping the epidemiological and clinical patterns of DM in TB patients is vital for delivering patient-centered care and managing these diseases to reduce the global TB impact[1].

The global prevalence of DM in TB cases exceeds 15%, surpassing the prevalence in the general adult population by over 50% in 2021[12,13]. A recent metanalysis, including 2.3 million TB patients, identified that this trend in notably higher in North America (19.7%), Western Pacific (19.4%), Southeast Asia (19.0%) and Middle East and North Africa (17.5%)[12]. While these variations in DM prevalence among TB patients appear to correlate with the general DM prevalence in each region, disparities within the same regions suggest more intricate underlying factors. India and Sri-Lanka, for example, possess a significantly higher burden of TB-DM than the other countries in the South Asia region [12,14]. This points to the potential influence of regional epidemiological patterns, as well consumption habits, environmental factors, and genetic predispositions, now being explored through multiplatform studies, indicating a complex interplay of epidemiological elements in TB-DM comorbidity.

The epidemiological factors associated with the TB-DM dynamicity has been the focus of several studies. Among the known risk factors, advanced age was consistently identified as a prominent risk factor for TB-DM comorbidity [15,16]. Lifestyle habits, such as illicit drug use [17,18] and sedentarism[19,20] behavior, along with socio-demographic factors like urban living and high-income status[21,22] have also been recognized as elements that increases the TB-DM prevalence. Clinical variables such as family history of DM[20,23,24] and hypertension[25], further contribute to increase the risk of TB-DM. However, the relationship between Body Mass Index (BMI) and TB-DM is unclear, with both low and high BMI values, as well as obesity and malnutrition, having been separately associated with TB-DM[20,26]. Furthermore, genetic factors at the population level may significantly influence the occurrence of TB-DM comorbidity. A study on Indian TB-DM individuals and their household contacts (HHC) observed that HHCs who developed TB, had a specific genetic pattern called the 'GG genotype' of the interleukin(IL)-6 -174G>C gene, indicating a genetic component in TB susceptibility[27].

The insights from epidemiological and molecular studies on TB-DM comorbidity have significant implications for future healthcare policies and interventions. The relationship between clinical and lifestyle factors, along with the genetic factors like the IL-6 GG genotype influencing TBDM susceptibility, pave the way for more personalized medicine approaches. These findings encourage the design of molecular epidemiology projects to identify specific SNPs associated with increased disease risk. Such targeted research can explain varied disease burdens across populations and inform tailored public health strategies, potentially reducing the incidence and improving the management of TB-DM comorbidity.

Dissecting the Clinical Interplay and Implications of TB-DM Comorbidity

The relationship between TB-DM represents a complex, bidirectional nexus significantly impacting clinical presentation, and disease dynamics and

outcomes. DM is not only associated with the prevalence of TB but also exacerbates its progression[28,29]. By 2050, projections suggest that one-third of TB incidence and mortality within the Asia-Pacific region and similar environments, will be attributable to DM[30]. This alarming trend underscores the need for integrated health strategies that address both TB and DM, particularly in regions with high prevalence rates.

Clinically, DM complicates TB management, manifesting in higher mycobacterial loads and more pronounced lung lesions in patients. In recent research spearheaded by our team, we uncovered a positive association between DM and increased mycobacterial loads, as well heightened Acid-Fast Bacilli (AFB) positivity, at diagnosis in TB-DM participants if compared with those without DM[29,31]. Furthermore, DM has been implicated in the presence of distinct lung lesions in chest radiographs of TB-DM patients in our Brazilian cohort[32]. The presence of DM also influences the clinical presentation of TB, with studies indicating alterations in symptoms and increased challenges in diagnosis and treatment initiation. DM increases the occurrence of TB symptoms such as hemoptysis, night sweats, and weight loss, as well as elevates the scores on clinical severity indices if compared with those without DM[18,31-34]. In contrast, TB-DM patients exhibited lower fever, reduced cough, and less sputum production, leading to delays in diagnosis and initiating appropriate treatment[35]. This results in more severe disease progression and heightened risks of adverse outcomes, including death, treatment failure and recurrence of TB[34]. The changes in clinical presentations and the consequences of a late diagnosis and start of TB treatment emerge as a clinical challenge in TB management and disease burden control.

It is important to emphasize that DM also impacts on anti-TB treatment, being associated with unfavorable outcomes, such as death, treatment failure and relapse cases[14,36]. Specifically, mortality and anti-TB treatment failure have been consistently linked with DM and higher glycated hemoglobin (HbA1c) levels[14,37,38,39,40]. Of note, treatment failure is more pronounced in TB-DM comorbidity within low- and middle-income countries[41], where DM conferred a 3.9 times increased risk of treatment failure in contrast with TB-only patients[42,43]. A pooled metanalysis also had shown that TB relapse risk increases in TB-DM individuals[14,44].

As a consequence of late diagnosis and challenges in treatment, DM also affects another important nexus to TB control, the transmission cascade. TB-DM patients shown an increased risk of Mtb transmission to their close contacts, evidenced by an increased likelihood of TB infection among contacts of pulmonary TB patients with dysglycemia if compared with the contacts of those TB patients normoglycemic[45]. The challenges of TB-DM regarding clinical manifestations and, consequently, outcomes play pivotal role in the future direction of TB management. First, is necessary to expand the glycemic tests among TB patients, identifying and treating DM with better glucose control. On the other hand, among those DM patients, an active search for TB is fundamental to early diagnosis and treatment. In addition, it is important to focus on screening close contacts of TB-DM individuals, given the higher risk of TB infection. Finally, improve the understand of the molecular mechanisms associated with the worse

clinical presentation and unfavorable outcomes could help to tailoring effective and patient-centered interventions, improving treatment, and contributing with the diseases burden control.

Cellular and Molecular Mechanisms Underlying TB-DM Comorbidity

The varying global incidences of TB-DM and the notable impact of DM on the clinical presentation, outcomes, and transmission of TB underscore a complex and intricate synergy between these conditions. In this landscape, multi-platform approaches are essential for dissecting the intricate cellular and molecular interactions in TB-DM comorbidity, and provide a comprehensive view of the inflammatory processes involved. Such advancements have the potential to significantly enhance prognosis, follow-up, and contribute to a reduction in the burden of TB. In this context, several multimolecular biomarkers have been explored with the goal of enhancing diagnosis and clinical management of TB-DM patients.

Cellular immunology aspects

Focusing on the cellular immunology aspects of TB-DM comorbidity, it becomes evident that the interplay between TB and DM significantly alters immune cell function and response. In TB infection, the effector functions of alveolar macrophages are crucial to containing the infection within the lungs[46]. However, in DM patients, the functionality of these cells is decreased due to metabolic alterations associated with the hyperglycemia[47,48]. DM impairs the functional activity of neutrophils[49] and reduces macrophage migration to sites of infection[46]. Animal studies with diabetic mice have shown a delay in innate immune response initiation, which include a compromise of nitric oxide production and phagocytic cell functionality, notwithstanding cytokine stimulation[46]. In this same model, alveolar macrophages exhibited increased expression of CCR2, which potentially hampers the migration of monocytes to the lungs. Therefore, this may result in a compromised capability to kill intracellular Mtb, thus further contributing to both infection susceptibility and increased bacterial load[46,50,51].

Similarly, the T cell response is notably affected, with DM patients often displaying dysregulated T cell responses. This dysregulation is hallmarked by imbalance in T helper (Th) cell subsets, with decreased Th1 responses and increased Th2 and Th17 responses. Such an imbalance can significantly alter the host ability to mount an effective response against TB. In a study comparing euglycemic and diabetic mice, it was observed that at the onset of infection diabetic mice exhibited a delayed activation of the adaptive immune system. This delay was indicated by decreased production of IFN-γ and fewer Mtb antigen (ESAT-6) presence compared to euglycemic mice[52,53].These cellular alterations contribute to a weakened immune defense against TB in diabetic individuals, underlining the importance of targeted interventions that address these specific cellular immune challenges in TB-DM comorbidity.

Genomics

Multi-omics research has delved deeper into the layers of complexity in TB-DM comorbidity. Genomics, proteomics, and transcriptomics, each provide unique insights into the pathophysiological mechanisms in TB-DM. Genomic studies, for instance, have identified genetic variants that predispose individuals to TB-DM, revealing potential targets for personalized medicine approaches. Polymorphisms on IL-6 and IL-18 genes were associated with TB-DM comorbidity and the occurrence of TB in close contacts[27]. Another study analyzed the interferon-gamma gene variants and found that the TACCCAGA haplotype was negatively associated with TB-DM. The frequency of this haplotype was high in the healthy controls compared to TB-DM patients, what may denote the importance of genetic variation in TB-DM predisposition, as well as facilitate to identify individuals at risk.

Transcriptomics

In the realm of transcriptomics, investigations in samples from TB-DM patients have illuminated the molecular pathways that may are dysregulated in this comorbidity. The superposition of TB-DM is marked by chronic inflammation, alongside qualitative and quantitative changes in immune activation characterized by distinct gene expression patterns. In a recent multi-center cohort study involving TB-DM individuals[54], our group identified a concise transcriptomic signature specific to TB regardless DM status[54]. In a more detailed analysis, pathway enrichment analysis had shown a notable trend towards heightened neutrophil and innate immune pathway activation in TB-DM participants, as well as higher expression of BATF2, VAMP5 and ANKRD22 genes in TB-DM even after anti-TB treatment commencement[54]. This finding might reflect persistent inflammation in TB-DM.

The findings of this same study unveiled that the genetic and immune responses may vary across different geographical regions[54], as discussed in the sections above. Another important aspect of this study was the positive correlation founded between HbA1c levels and pathways associated with insulin resistance, metabolic dysfunction, diabetic complications, and chromosomal instability[54]. These correlations may play a pivotal role in the pathophysiology of TB-DM, contributing to a more severe clinical presentation and unfavorable outcomes.

Another multicentric study conducted by an international consortium has found that DM amplifies the expression levels of genes related to the innate inflammatory response and reducing genes related with the adaptative immune response in TB individuals[55]. A decreased type I interferon (IFN) response was identified in TB-DM participants if compared to TB-only patients, suggesting an uncoupling in IFN pathways[55]. Additionally, a Chinese study revealed 952 differentially expressed genes (DEGs) in TB-DM, enriched in pathways associated with the cell cycle, homeostasis, and immunological processes, shedding light on the intricate interplay between TB and DM[56].

Expanding the scope of transcriptomic studies in TB-DM, our group evaluated the role of the long non-coding RNA (lncRNA) in this synergic condition[57]. Using transcriptomic data from a Brazilian cohort, the lncRNA expression profiles were compared across TB, TB-DM and healthy control (without TB and DM) groups[57]. This comparison led to the identification of a distinct lncRNA signature, which effectively distinguishes TB-DM from TB-only cases with an accuracy of 90-94%. Notably, the lncRNAs included in the signature (LINCO2009, LINCO2471, ADM-DT, and GK-AS1) hold a critical role in the pathways related to inflammatory activation against Mtb[57].

These studies demonstrate that transcriptomics has shed light on the field of TB-DM, revealing that there are consistently altered pathways in TB-DM patients compared to those with TB only. These pathways are often linked to the regulation of the inflammatory response, potentially reflecting increased and persistent inflammation in patients with both diseases. This indicates that further research in the transcriptomics of TB-DM is necessary, but it also suggests the need to integrate data from this platform with that of other platforms for a more comprehensive understanding.

Proteomics

Proteomic analyses, through the quantification of cytokines, chemokines, and other immune-related proteins, have significantly advanced in the understanding of immune responses in TB-DM. This approach has been essential in revealing how hyperglycemia-induced metabolic alterations in DM patients impair the functionality of the innate and adaptive immune responses in TB-DM. Proteomic data in TB-DM also have the potential to reveal key alterations in protein expression and pointing toward potential novel biomarkers. These insights are crucial in delineating the complex dynamics of immune dysfunction in TB susceptibility, transmission, and treatment outcomes[46,52,53,58-61].

In TB-DM patients, proteomic data indicate an increase in complement component 2 (C2), as well APOB and APOC2 levels if compared to TB-only group. These proteins are involved in the complement and coagulation cascade, as well as in the cholesterol metabolism. This elevation suggests a potential link between lipid metabolism dysregulation and the heightened inflammatory state observed in TB-DM comorbidity. In another study, 18 differentially expressed protein spots were identified in TB-DM patients. These alterations were associated with potential metabolic complications specific to TB-DM and shifts in proteins governing cell cycle and growth regulation hint at disrupted processes like cell proliferation and apoptosis. These findings not only provide a deeper understanding of TB-DM pathophysiology but also open avenues for new diagnostic, monitoring, and treatment strategies.

DM patients also present an altered cytokine milieu that favors immune dysregulation. In several studies, with TB and TB-DM individuals, it was observed that DM participants experienced higher levels of inflammatory activation than those without DM[61,62]. Patients with TB-DM have higher levels of pro-inflammatory cytokines such as IFN-γ, IL-1β, and IL-17, as well as lower levels of anti-inflammatory cytokines such as IL-10, compared to patients without DM. Additionally, throughout the anti-TB treatment, these markers remained elevated

for a longer period in TB-DM patients compared to non-DM individuals, which characterized persistent hyperinflammation in this group of individuals[62]. In summary, the comorbidity of TB-DM is hallmarked by chronic and unbalanced inflammation, reflected in abnormal levels of proteins, and the superposition of these disorders lead to qualitative and quantitative changes in immune activation[54].

Lipidomic

In a previous study of our group, we utilized lipidomic approaches to identify persistent hyperinflammation by evaluating urinary lipid mediator profiles of participants with TB and TB-dysglycemia. In this study, levels of a urinary metabolite of prostaglandin 2 (PGE-M) and leukotriene 4 (LTE4) were consistently higher during anti-TB treatment in the DM group compared to the normoglycemic group. These lipid mediators play a crucial role in modulating the immune response[63]. Interestingly, in an adjusted multivariable model TB-DM was independently associated with increased concentrations of PGD-M, PGI-M, and LTE4 at baseline[63]. This profile of exacerbated inflammation in TB-DM patients help explain why these individuals present severe symptoms and more enduring lung damage more often[62], which can be associated with unfavorable outcomes and Mtb dissemination.

Metabolomics

Metabolomics, by analyzing the small molecule metabolites present in TB-DM patients, offers insights into the metabolic disruptions caused by the interplay of TB and DM. In studies exploring metabolomics from an omics perspective, specific metabolic changes induced by TB-DM were delineated[64]. Plasma amine and acylcarnitine levels were measured in TB and TB-DM patients, with partial least squares discrimination analysis showing robust group discrimination[64]. Notably, TB-DM exhibited lower levels of choline, glycine, serine, threonine, and homoserine compared to TB-only patients. Of note, the levels of these metabolites did not normalize during treatment[64]. In a recent Korean study, plasma metabolic profiles of TB and TB-DM were investigated using metabolomics and lipidomics[65]. TB-DM participants presented higher concentrations of bile acids and molecules related to carbohydrate metabolism, as well as the depletion of glutamine, retinol, lysophosphatidylcholine, and phosphatidylcholine[65]. Arachidonic acid metabolism, crucial for eicosanoid production, emerged as a key factor in TB-DM pathophysiology[65]. Eicosanoids, extensively studied in TB and TB-DM[63,66], were found to mirror disease severity and extent[60], with potential as markers for disease extension.

Integrative analysis

Multi-omics investigations have significantly advanced our comprehension of the complex interplay between these two diseases. By integrating data from various omics layers such as genomics, transcriptomics, proteomics, metabolomics, and epigenomics, it is possible to achieve a holistic understanding of the biological processes involved in TB-DM interaction and consequently prognosis. Funding multi-omic studies is fundamental to better understanding the pathophysiology of TB-DM and its impact on anti-TB treatment outcomes. Additionally, the findings

on multi-omic studies can help not only in new prognosis strategies, but also in the identification of new targets to host directed therapies.

In a groundbreaking study aimed at deciphering the mechanisms of TB susceptibility in DM patients and assessing the impact of TB on DM complications, a sophisticated integrative analysis was employed. Using whole blood gene expression and plasma analytes[61], our group identified that DM in comorbidity with TB intensifies the neutrophilic inflammatory response, possibly indicative of a higher bacterial load or a distinct disruption in immune function. This heightened response was marked by increased plasma levels of cytokines and growth factors, as well differentially expressed genes, that differentiating individuals with TB-DM from the majority of those with only TB or DM[61]. Intriguingly, the expression patterns of TB-DM exclusive genes were linked to critical biological processes and therapeutic targets. They were associated with endoplasmic reticulum stress, a vital cell stress response, and showed connections to the mechanisms of action of the antibiotic doxycycline and anti-cancer drugs such as 5-fluorouracil and semaxanib[61]. This insightful research not only sheds light on the complex interplay between TB and DM at the molecular level but also opens new avenues for potential therapeutic strategies and better understanding of these comorbidities.

In a recent article, our group employed a multi-platform approach, integrating clinical, transcriptomic, lipidomic, and proteomic data from a Brazilian TB-DM cohort, to unravel the molecular interactions in this synergistic scenario. Utilizing machine learning to analyze combined data from cytokines, gene expression, and eicosanoids, we identified several multimolecular baseline markers — MMP-28, LTE-4, 11-dTxB2, PGDM, FBXO6, SECTM1, and LINCO2009 — that effectively differentiate between TB-DM, TB-only, DM-only, and healthy control groups. After anti-TB treatment onset, a notable decrease in these markers was observed, correlating with microbiological cure. Significantly, markers such as 11-dTxB2, SECTM1, and LINCO2009 not only emerged as indicators for new host-directed TB treatments but also as potential predictors of treatment outcomes. Furthermore, this integrated molecular signature demonstrated high accuracy in distinguishing TB-DM cases in Brazil and was validated in three external cohorts, outperforming signatures derived solely from transcriptomic data[54]. Crucially, these findings highlight that multimolecular signatures can be more predictive and impactful for precision medicine compared to single-omic approaches, underscoring the enhanced potential of multi-omic platforms in advancing our understanding of inflammatory and infectious diseases, as well in finding markers that can be implemented in the clinical practice.

Paving the Path for Future Breakthroughs

The studies included in this review provide substantial evidence of the interplay between TB and DM and highlight the need for advanced research methodologies. Current evidence in epidemiology demonstrates a global prevalence of DM in TB cases, emphasizing age, lifestyle, socio-economic factors, family history, and hypertension as key risk factors. However, there exists a knowledge gap that needs addressing to understand the regional disparities in TB-DM comorbidity. Invest in molecular epidemiology studies is crucial for this understanding and is pivotal for developing targeted public health strategies. This

approach would not only elucidate regional differences but also aid in formulating more effective, region-specific interventions.

The clinical nexus of TB-DM presents a bidirectional impact, with DM complicating TB management and exacerbating disease progression. Research shows a positive association between DM and increased mycobacterial loads and distinct lung lesions, underscoring the need for integrated health strategies addressing both diseases. The next step in addressing TB-DM comorbidity in the clinical point of view would involve developing more targeted public health policies for individuals with both conditions. This could include enhanced TB screening in DM patients and the other way around, as well as expanding research into contacts of these patients to assess transmission dynamics. These strategies would improve individual patient care and contribute to broader public health efforts in managing and preventing the spread of TB-DM comorbidity.

It is also known that DM impacts immune cell function and response in TB, with specific genetic variations associated with TB susceptibility. This points to the potential of using advanced technologies like single-cell analysis to uncover new therapeutic targets and biomarkers.

This review provides several cellular and molecular insights associated with TB-DM comorbidity. We discussed the altered immune cell function in DM patients which are crucial in containing TB infection, as well as the influence of genetic factors, such as polymorphisms, and the role of multi-omics in understanding molecular pathways disrupted in TB-DM. Building on this, the use of multi-platforms, as well as the addition cutting-edge technologies such as single-cell analysis could be instrumental. This technology can allow for a more granular understanding of cellular responses in TB-DM comorbidity at an individual cell level, potentially uncovering new pathways and therapeutic targets. In addition, the development, validation and implementation of point-of-care testing for specific biomarkers already identified through these advanced methods could revolutionize early detection and monitoring of TB-DM comorbidity. This approach aligns with the development of predictive scores, integrating genetic, molecular, and clinical data to accurately assess disease progression and treatment outcomes.

Moreover, the creation and improvement of comprehensive risk scores, incorporating socio-demographic, lifestyle, and clinical variables, could greatly enhance the precision of public health interventions. These scores, derived from multi-omic and epidemiological data, could be tailored to specific populations, considering regional variations in TB-DM comorbidity. **Figure 1** encapsulates the current state of knowledge and future directions in TB-DM comorbidity research. In essence, leveraging these innovative technologies and approaches could bridge the gap between current knowledge and the untapped potential in managing TB-DM comorbidity, leading to more effective, personalized treatment and prevention strategies.

Concluding remarks

The intricate relationship between tuberculosis and diabetes is a worldwide health threat, impacting treatment outcomes and mortality rates. This intersecting

epidemic demands a multifaceted research approach. The synthesis of epidemiological, clinical, genomic, transcriptomic, proteomic, and lipidomic studies is vital for understanding the complexities of TB-DM comorbidity. A new approach is required to deepen the inflammatory profile in TB-DM interactions and understanding how the multimolecular markers interact and impacts in the epidemiological and clinical fields. The study of multi-omic platforms emerges as an opportunity to gain insights into disease pathogenesis, given that it simultaneously explores several components of immune responses through multiple assay platforms. The identification of precise biomarkers for diagnosis and individualized treatment, along with public health strategies informed by molecular and epidemiological findings, is crucial. This area of research holds the promise of significant advancements, offering enhanced management of TB-DM comorbidity and contributing to global public health outcomes.

References

1. Organization, W.H. (2022). Global Tuberculosis Report.
2. Organization, W.H. (2016). Global report on diabetes.
3. Bell, L.C.K., and Noursadeghi, M. (2018). Pathogenesis of HIV-1 and Mycobacterium tuberculosis co-infection. Nat Rev Microbiol *16*, 80-90. 10.1038/nrmicro.2017.128.
4. Gupta, K.B., Gupta, R., Atreja, A., Verma, M., and Vishvkarma, S. (2009). Tuberculosis and nutrition. Lung India *26*, 9-16. 10.4103/0970-2113.45198.
5. Tsalamandris, S., Antonopoulos, A.S., Oikonomou, E., Papamikroulis, G.A., Vogiatzi, G., Papaioannou, S., Deftereos, S., and Tousoulis, D. (2019). The Role of Inflammation in Diabetes: Current Concepts and Future Perspectives. Eur Cardiol *14*, 50-59. 10.15420/ecr.2018.33.1.
6. Berbudi, A., Rahmadika, N., Tjahjadi, A.I., and Ruslami, R. (2020). Type 2 Diabetes and its Impact on the Immune System. Curr Diabetes Rev *16*, 442-449. 10.2174/1573399815666191024085838.
7. Roy, B., and Runa, S.A. (2022). SARS-CoV-2 infection and diabetes: Pathophysiological mechanism of multi-system organ failure. World J Virol *11*, 252-274. 10.5501/wjv.v11.i5.252.
8. Restrepo, B.I. (2016). Diabetes and Tuberculosis. Microbiol Spectr *4*. 10.1128/microbiolspec.TNMI7-0023-2016.
9. Al-Rifai, R.H., Pearson, F., Critchley, J.A., and Abu-Raddad, L.J. (2017). Association between diabetes mellitus and active tuberculosis: A systematic review and meta-analysis. PLoS One *12*, e0187967. 10.1371/journal.pone.0187967.
10. Foe-Essomba, J.R., Kenmoe, S., Tchatchouang, S., Ebogo-Belobo, J.T., Mbaga, D.S., Kengne-Nde, C., Mahamat, G., Kame-Ngasse, G.I., Noura, E.A., Mbongue Mikangue, C.A., et al. (2021). Diabetes mellitus and tuberculosis, a systematic review and meta-analysis with sensitivity analysis for studies comparable for confounders. PLoS One *16*, e0261246. 10.1371/journal.pone.0261246.
11. Restrepo, B.I., Fisher-Hoch, S.P., Crespo, J.G., Whitney, E., Perez, A., Smith, B., McCormick, J.B., and Nuevo Santander Tuberculosis, T. (2007). Type 2 diabetes and tuberculosis in a dynamic bi-national border population. Epidemiol Infect *135*, 483-491. 10.1017/S0950268806006935.
12. Noubiap, J.J., Nansseu, J.R., Nyaga, U.F., Nkeck, J.R., Endomba, F.T., Kaze, A.D., Agbor, V.N., and Bigna, J.J. (2019). Global prevalence of diabetes in active tuberculosis: a systematic review and meta-analysis of data from 2.3 million patients with tuberculosis. Lancet Glob Health *7*, e448-e460. 10.1016/S2214-109X(18)30487-X.
13. Workneh, M.H., Bjune, G.A., and Yimer, S.A. (2017). Prevalence and associated factors of tuberculosis and diabetes mellitus comorbidity: A systematic review. PLoS One *12*, e0175925. 10.1371/journal.pone.0175925.
14. Gautam, S., Shrestha, N., Mahato, S., Nguyen, T.P.A., Mishra, S.R., and Berg-Beckhoff, G. (2021). Diabetes among tuberculosis patients and its impact on tuberculosis treatment in South Asia: a systematic review and meta-analysis. Sci Rep *11*, 2113. 10.1038/s41598-021-81057-2.
15. Mabula, P.L., Kazinyingia, K.I., Chavala, E.C., Mosha, V., Msuya, S.E., and Leyaro, B.J. (2021). Prevalence and risk factors for diabetes mellitus among tuberculosis patients in Moshi Municipal Council, Kilimanjaro Tanzania. East Afr Health Res J *5*, 69-74. 10.24248/eahrj.v5i1.653.
16. Alturki, S., Al Amad, M., Mahyoub, E., Al Hanash, N., and Alhammadi, A. (2023). Prevalence of Diabetes Mellitus among Patients with Tuberculosis and Its Associated Factors in Sana'a, Yemen, 2021. Epidemiologia (Basel) *4*, 202-211. 10.3390/epidemiologia4020021.
17. Alavi, S.M., and Khoshkhoy, M.M. (2012). Pulmonary tuberculosis and diabetes mellitus: Co-existence of both diseases in patients admitted in a teaching hospital in the southwest of Iran. Caspian J Intern Med *3*, 421-424.
18. Arriaga, M.B., Araujo-Pereira, M., Barreto-Duarte, B., Sales, C., Miguez-Pinto, J.P., Nogueira, E.B., Nogueira, B.M.F., Rocha, M.S., Souza, A.B., Benjamin, A., et al. (2021). Prevalence and Clinical Profiling of Dysglycemia and HIV Infection in Persons With Pulmonary Tuberculosis in Brazil. Front Med (Lausanne) *8*, 804173. 10.3389/fmed.2021.804173.
19. Balde, N.M., Camara, A., Camara, L.M., Diallo, M.M., Kake, A., and Bah-Sow, O.Y. (2006). Associated tuberculosis and diabetes in Conakry, Guinea: prevalence and clinical characteristics. Int J Tuberc Lung Dis *10*, 1036-1040.
20. Viswanathan, V., Kumpatla, S., Aravindalochanan, V., Rajan, R., Chinnasamy, C., Srinivasan, R., Selvam, J.M., and Kapur, A. (2012). Prevalence of diabetes and pre-diabetes and associated risk factors among tuberculosis patients in India. PLoS One *7*, e41367. 10.1371/journal.pone.0041367.
21. Wang, Q., Ma, A., Han, X., Zhao, S., Cai, J., Ma, Y., Zhao, J., Wang, Y., Dong, H., Zhao, Z., et al. (2013). Prevalence of type 2 diabetes among newly detected pulmonary tuberculosis patients in China: a community based cohort study. PLoS One *8*, e82660. 10.1371/journal.pone.0082660.

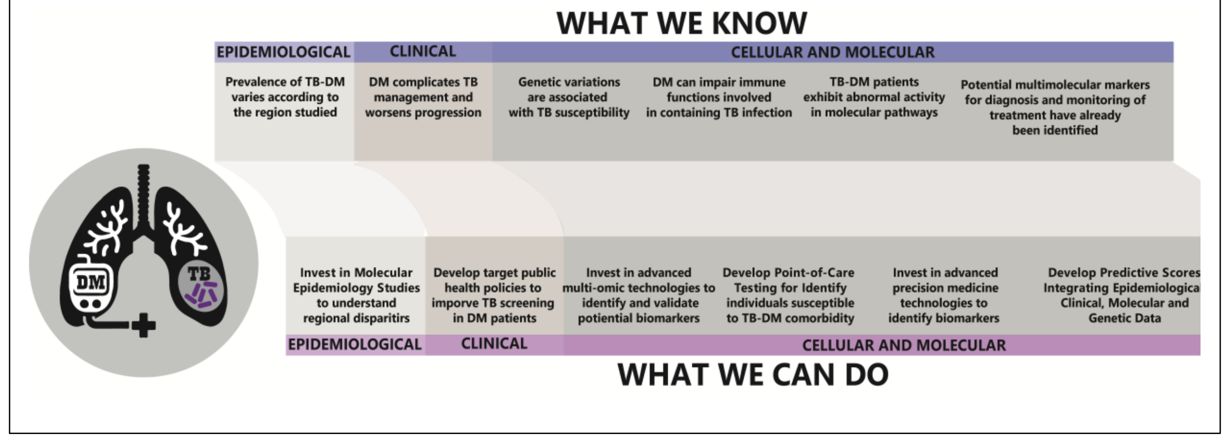
22. Amare, H., Gelaw, A., Anagaw, B., and Gelaw, B. (2013). Smear positive pulmonary tuberculosis among diabetic patients at the Dessie referral hospital, Northeast Ethiopia. Infect Dis Poverty *2*, 6. 10.1186/2049-9957-2-6.
23. Workneh, M.H., Bjune, G.A., and Yimer, S.A. (2016). Prevalence and Associated Factors of Diabetes Mellitus among Tuberculosis Patients in South-Eastern Amhara Region, Ethiopia: A Cross Sectional Study. PLoS One *11*, e0147621. 10.1371/journal.pone.0147621.
24. Shidam, U.G., Roy, G., Sahu, S.K., Kumar, S.V., and Ananthanarayanan, P.H. (2015). Screening for diabetes among presumptive tuberculosis patients at a tertiary care centre in Pondicherry, India. Int J Tuberc Lung Dis *19*, 1163-1168. 10.5588/ijtld.14.0971.
25. Adegbite, B.R., Edoa, J.R., Agbo Achimi Abdul, J., Epola, M., Mevyann, C., Dejon-Agobe, J.C., Zinsou, J.F., Honkpehedji, Y.J., Mpagama, S.G., Alabi, A.S., et al. (2022). Non-communicable disease co-morbidity and associated factors in tuberculosis patients: A cross-sectional study in Gabon. EClinicalMedicine *45*, 101316. 10.1016/j.eclinm.2022.101316.
26. Lin, Y.H., Chen, C.P., Chen, P.Y., Huang, J.C., Ho, C., Weng, H.H., Tsai, Y.H., and Peng, Y.S. (2015). Screening for pulmonary tuberculosis in type 2 diabetes elderly: a cross-sectional study in a community hospital. BMC Public Health *15*, 3. 10.1186/1471-2458-15-3.
27. Ponnana, M., Sivangala, R., Joshi, L., Valluri, V., and Gaddam, S. (2017). IL-6 and IL-18 cytokine gene variants of pulmonary tuberculosis patients with co-morbid diabetes mellitus and their household contacts in Hyderabad. Gene *627*, 298-306. 10.1016/j.gene.2017.06.046.
28. Huber, F.G., Kristensen, K.L., Holden, I.K., Andersen, P.H., Bakir, B., Jorgensen, A., Lorentsson, H.J.N., Bjorn-Mortensen, K., Johansen, I.S., and Ravn, P. (2022). The prevalence of diabetes among tuberculosis patients in Denmark. BMC Infect Dis *22*, 64. 10.1186/s12879-022-07048-4.
29. Almeida-Junior, J.L., Gil-Santana, L., Oliveira, C.A., Castro, S., Cafezeiro, A.S., Daltro, C., Netto, E.M., Kornfeld, H., and Andrade, B.B. (2016). Glucose Metabolism Disorder Is Associated with Pulmonary Tuberculosis in Individuals with Respiratory Symptoms from Brazil. PLoS One *11*, e0153590. 10.1371/journal.pone.0153590.
30. Awad Susanne F , H.P., Ayoub Houssein H, Pearson Fiona , Dargham Soha R, Critchley Julia A, Abu-Raddad Laith J (2019). Forecasting the impact of diabetes mellitus on tuberculosis disease incidence and mortality in India. Journal of Global Health 10.7189/jogh.09.020415.
31. Gil-Santana, L., Almeida-Junior, J.L., Oliveira, C.A., Hickson, L.S., Daltro, C., Castro, S., Kornfeld, H., Netto, E.M., and Andrade, B.B. (2016). Diabetes Is Associated with Worse Clinical Presentation in Tuberculosis Patients from Brazil: A Retrospective Cohort Study. PLoS One *11*, e0146876. 10.1371/journal.pone.0146876.
32. Chiang, C.Y., Lee, J.J., Chien, S.T., Enarson, D.A., Chang, Y.C., Chen, Y.T., Hu, T.Y., Lin, C.B., Suk, C.W., Tao, J.M., and Bai, K.J. (2014). Glycemic control and radiographic manifestations of tuberculosis in diabetic patients. PLoS One *9*, e93397. 10.1371/journal.pone.0093397.
33. Wu, Z., Guo, J., Huang, Y., Cai, E., Zhang, X., Pan, Q., Yuan, Z., and Shen, X. (2016). Diabetes mellitus in patients with pulmonary tuberculosis in an aging population in Shanghai, China: Prevalence, clinical characteristics and outcomes. J Diabetes Complications *30*, 237-241. 10.1016/j.jdiacomp.2015.11.014.
34. Chiang, C.Y., Bai, K.J., Lin, H.H., Chien, S.T., Lee, J.J., Enarson, D.A., Lee, T.I., and Yu, M.C. (2015). The influence of diabetes, glycemic control, and diabetes-related comorbidities on pulmonary tuberculosis. PLoS One *10*, e0121698. 10.1371/journal.pone.0121698.
35. Faurholt-Jepsen, D., Aabye, M.G., Jensen, A.V., Range, N., Praygod, G., Jeremiah, K., Changalucha, J., Faurholt-Jepsen, M., Jensen, L., Jensen, S.M., et al. (2014). Diabetes is associated with lower tuberculosis antigen-specific interferon gamma release in Tanzanian tuberculosis patients and non-tuberculosis controls. Scand J Infect Dis *46*, 384-391. 10.3109/00365548.2014.885657.
36. Baker, M.A., Harries, A.D., Jeon, C.Y., Hart, J.E., Kapur, A., Lonnroth, K., Ottmani, S.E., Goonesekera, S.D., and Murray, M.B. (2011). The impact of diabetes on tuberculosis treatment outcomes: a systematic review. BMC Med *9*, 81. 10.1186/1741-7015-9-81.
37. Bezerra, A.L., Moreira, A., Isidoro-Goncalves, L., Lara, C., Amorim, G., Silva, E.C., Kritski, A.L., and Carvalho, A.C.C. (2022). Clinical, laboratory, and radiographic aspects of patients with pulmonary tuberculosis and dysglycemia and tuberculosis treatment outcomes. J Bras Pneumol *48*, e20210505. 10.36416/1806-3756/e20210505.
38. Viswanathan, V., Vigneswari, A., Selvan, K., Satyavani, K., Rajeswari, R., and Kapur, A. (2014). Effect of diabetes on treatment outcome of smear-positive pulmonary tuberculosis--a report from South India. J Diabetes Complications *28*, 162-165. 10.1016/j.jdiacomp.2013.12.003.

39. Wang, C.S., Yang, C.J., Chen, H.C., Chuang, S.H., Chong, I.W., Hwang, J.J., and Huang, M.S. (2009). Impact of type 2 diabetes on manifestations and treatment outcome of pulmonary tuberculosis. Epidemiol Infect *137*, 203-210. 10.1017/S0950268808000782.
40. Chiang, C.Y., Lee, J.J., Yu, M.C., Enarson, D.A., Lin, T.P., and Luh, K.T. (2009). Tuberculosis outcomes in Taipei: factors associated with treatment interruption for 2 months and death. Int J Tuberc Lung Dis *13*, 105-111.
41. Huangfu, P., Ugarte-Gil, C., Golub, J., Pearson, F., and Critchley, J. (2019). The effects of diabetes on tuberculosis treatment outcomes: an updated systematic review and meta-analysis. Int J Tuberc Lung Dis *23*, 783-796. 10.5588/ijtld.18.0433.
42. Singla, R., Khan, N., Al-Sharif, N., Ai-Sayegh, M.O., Shaikh, M.A., and Osman, M.M. (2006). Influence of diabetes on manifestations and treatment outcome of pulmonary TB patients. Int J Tuberc Lung Dis *10*, 74-79.
43. Singla, R., Raghu, B., Gupta, A., Caminero, J.A., Sethi, P., Tayal, D., Chakraborty, A., Jain, Y., and Migliori, G.B. (2021). Risk factors for early mortality in patients with pulmonary tuberculosis admitted to the emergency room. Pulmonology *27*, 35-42. 10.1016/j.pulmoe.2020.02.002.
44. Arriaga, M.B., Amorim, G., Queiroz, A.T.L., Rodrigues, M.M.S., Araujo-Pereira, M., Nogueira, B.M.F., Souza, A.B., Rocha, M.S., Benjamin, A., Moreira, A.S.R., et al. (2021). Novel stepwise approach to assess representativeness of a large multicenter observational cohort of tuberculosis patients: The example of RePORT Brazil. Int J Infect Dis *103*, 110-118. 10.1016/j.ijid.2020.11.140.
45. Arriaga, M.B., Rocha, M.S., Nogueira, B.M.F., Nascimento, V., Araujo-Pereira, M., Souza, A.B., Andrade, A.M.S., Costa, A.G., Gomes-Silva, A., Silva, E.C., et al. (2021). The Effect of Diabetes and Prediabetes on Mycobacterium tuberculosis Transmission to Close Contacts. J Infect Dis *224*, 2064-2072. 10.1093/infdis/jiab264.
46. Martinez, N., Ketheesan, N., West, K., Vallerskog, T., and Kornfeld, H. (2016). Impaired Recognition of Mycobacterium tuberculosis by Alveolar Macrophages From Diabetic Mice. J Infect Dis *214*, 1629-1637. 10.1093/infdis/jiw436.
47. Martinez, N., and Kornfeld, H. (2014). Diabetes and immunity to tuberculosis. Eur J Immunol *44*, 617-626. 10.1002/eji.201344301.
48. Ayelign, B., Negash, M., Genetu, M., Wondmagegn, T., and Shibabaw, T. (2019). Immunological Impacts of Diabetes on the Susceptibility of Mycobacterium tuberculosis. J Immunol Res *2019*, 6196532. 10.1155/2019/6196532.
49. Alba-Loureiro, T.C., Hirabara, S.M., Mendonca, J.R., Curi, R., and Pithon-Curi, T.C. (2006). Diabetes causes marked changes in function and metabolism of rat neutrophils. J Endocrinol *188*, 295-303. 10.1677/joe.1.06438.
50. Lecube, A., Pachon, G., Petriz, J., Hernandez, C., and Simo, R. (2011). Phagocytic activity is impaired in type 2 diabetes mellitus and increases after metabolic improvement. PLoS One *6*, e23366. 10.1371/journal.pone.0023366.
51. Pavlou, S., Lindsay, J., Ingram, R., Xu, H., and Chen, M. (2018). Sustained high glucose exposure sensitizes macrophage responses to cytokine stimuli but reduces their phagocytic activity. BMC Immunol *19*, 24. 10.1186/s12865-018-0261-0.
52. Martens, G.W., Arikan, M.C., Lee, J., Ren, F., Greiner, D., and Kornfeld, H. (2007). Tuberculosis susceptibility of diabetic mice. Am J Respir Cell Mol Biol *37*, 518-524. 10.1165/rcmb.2006-0478OC.
53. Vallerskog, T., Martens, G.W., and Kornfeld, H. (2010). Diabetic mice display a delayed adaptive immune response to Mycobacterium tuberculosis. J Immunol *184*, 6275-6282. 10.4049/jimmunol.1000304.
54. Queiroz, A.T.L., Vinhaes, C.L., Fukutani, E.R., Gupte, A.N., Kumar, N.P., Fukutani, K.F., Arriaga, M.B., Sterling, T.R., Babu, S., Gaikwad, S., et al. (2023). A multi-center, prospective cohort study of whole blood gene expression in the tuberculosis-diabetes interaction. Sci Rep *13*, 7769. 10.1038/s41598-023-34847-9.
55. Eckold, C., Kumar, V., Weiner, J., Alisjahbana, B., Riza, A.L., Ronacher, K., Coronel, J., Kerry-Barnard, S., Malherbe, S.T., Kleynhans, L., et al. (2021). Impact of Intermediate Hyperglycemia and Diabetes on Immune Dysfunction in Tuberculosis. Clin Infect Dis *72*, 69-78. 10.1093/cid/ciaa751.
56. Liu, T., Wang, Y., Gui, J., Fu, Y., Ye, C., Hong, X., Chen, L., Li, Y., Zhang, X., and Hong, W. (2022). Transcriptome analysis of the impact of diabetes as a comorbidity on tuberculosis. Medicine (Baltimore) *101*, e31652. 10.1097/MD.0000000000031652.

57. Rocha, E.F., Vinhaes, C.L., Araujo-Pereira, M., Mota, T.F., Gupte, A.N., Kumar, N.P., Arriaga, M.B., Sterling, T.R., Babu, S., Gaikwad, S., et al. (2024). The sound of silent RNA in tuberculosis and the lncRNA role on infection. iScience *27*, 108662. 10.1016/j.isci.2023.108662.
58. DeFronzo, R.A., Ferrannini, E., Groop, L., Henry, R.R., Herman, W.H., Holst, J.J., Hu, F.B., Kahn, C.R., Raz, I., Shulman, G.I., et al. (2015). Type 2 diabetes mellitus. Nat Rev Dis Primers *1*, 15019. 10.1038/nrdp.2015.19.
59. Arriaga, M.B., Araujo-Pereira, M., Barreto-Duarte, B., Nogueira, B., Freire, M., Queiroz, A.T.L., Rodrigues, M.M.S., Rocha, M.S., Souza, A.B., Spener-Gomes, R., et al. (2022). The Effect of Diabetes and Prediabetes on Antituberculosis Treatment Outcomes: A Multicenter Prospective Cohort Study. J Infect Dis *225*, 617-626. 10.1093/infdis/jiab427.
60. Pavan Kumar, N., Moideen, K., Nancy, A., Viswanathan, V., Shruthi, B.S., Shanmugam, S., Hissar, S., Kornfeld, H., and Babu, S. (2019). Plasma Eicosanoid Levels in Tuberculosis and Tuberculosis-Diabetes Co-morbidity Are Associated With Lung Pathology and Bacterial Burden. Front Cell Infect Microbiol *9*, 335. 10.3389/fcimb.2019.00335.
61. Prada-Medina, C.A., Fukutani, K.F., Pavan Kumar, N., Gil-Santana, L., Babu, S., Lichtenstein, F., West, K., Sivakumar, S., Menon, P.A., Viswanathan, V., et al. (2017). Systems Immunology of Diabetes-Tuberculosis Comorbidity Reveals Signatures of Disease Complications. Sci Rep *7*, 1999. 10.1038/s41598-017-01767-4.
62. Kumar, N.P., Fukutani, K.F., Shruthi, B.S., Alves, T., Silveira-Mattos, P.S., Rocha, M.S., West, K., Natarajan, M., Viswanathan, V., Babu, S., et al. (2019). Persistent inflammation during anti-tuberculosis treatment with diabetes comorbidity. Elife *8*. 10.7554/eLife.46477.
63. Arriaga, M.B., Karim, F., Queiroz, A.T.L., Araujo-Pereira, M., Barreto-Duarte, B., Sales, C., Moosa, M.S., Mazibuko, M., Milne, G.L., Maruri, F., et al. (2022). Effect of Dysglycemia on Urinary Lipid Mediator Profiles in Persons With Pulmonary Tuberculosis. Front Immunol *13*, 919802. 10.3389/fimmu.2022.919802.
64. Vrieling, F., Alisjahbana, B., Sahiratmadja, E., van Crevel, R., Harms, A.C., Hankemeier, T., Ottenhoff, T.H.M., and Joosten, S.A. (2019). Plasma metabolomics in tuberculosis patients with and without concurrent type 2 diabetes at diagnosis and during antibiotic treatment. Sci Rep *9*, 18669. 10.1038/s41598-019-54983-5.
65. Yen, N.T.H., Anh, N.K., Jayanti, R.P., Phat, N.K., Vu, D.H., Ghim, J.L., Ahn, S., Shin, J.G., Oh, J.Y., Long, N.P., and Kim, D.H. (2023). Multimodal plasma metabolomics and lipidomics in elucidating metabolic perturbations in tuberculosis patients with concurrent type 2 diabetes. Biochimie *211*, 153-163. 10.1016/j.biochi.2023.04.009.
66. Vinhaes, C.L., Oliveira-de-Souza, D., Silveira-Mattos, P.S., Nogueira, B., Shi, R., Wei, W., Yuan, X., Zhang, G., Cai, Y., Barry, C.E., 3rd, et al. (2019). Changes in inflammatory protein and lipid mediator profiles persist after antitubercular treatment of pulmonary and extrapulmonary tuberculosis: A prospective cohort study. Cytokine *123*, 154759. 10.1016/j.cyto.2019.154759.

Figure Legend

Figure 1: Overview of the interplay between Tuberculosis (TB) and Diabetes Mellitus (DM) and future directions. Upper: Current epidemiological, clinical, and cellular/molecular knowledge concerning TB-DM co-morbidity. Down: Prospective actions for advancing research and healthcare strategies.



7 DISCUSSÃO

A presente tese traz resultados de três artigos originais que avaliam o impacto da disglicemia em paciente com tuberculose. O estudo necessário no processo do doutoramento resultou ainda em um artigo de revisão. Em consonância com a era da Medicina de precisão, os nossos achados, em conjunto, trazem novas perspectivas moleculares que podem servir de base para o desenvolvimento de técnicas inovadoras e eficazes para o diagnóstico, prognóstico e seguimento terapêutico de pacientes com TB e TBDM. Adicionalmente, destacamos o papel e valor de estudos envolvendo múltiplas plataformas ômicas em doenças infecciosas e inflamatórias, emergindo com grande potencial de estabelecer insights para melhor entendimento dos processos fisiopatológicos e desenvolvimento de terapias guiadas.

O primeiro trabalho da tese foi parte de um coorte multicêntrico, com dados extraídos de dois sítios da Índia e um no Brasil, em Salvador, sendo utilizados dados transcriptômicos do MSTDI (Molecular Signature of Tuberculosis-Diabetes Interaction). A fim de melhor estudar a variabilidade entre populações, incluímos em nossas análises dados públicos do consórcio TANDEM, projeto que estudou possível relação causal entre TB e DM (37). Nossas análises objetivaram a avaliação da intensidade e qualidade na ativação inflamatória de pacientes com TBDM, além de avaliar o impacto dos níveis de HbA1c nas vias de expressão biológicas, aquelas que regulam o funcionamento homeostático do organismo.

O primeiro resultado demonstra variabilidade nos DEGs entre os países do estudo. Encontramos 6 DEGs compartilhados entre os 4 países em TB, 12 em TBDM e nenhum DEG foi encontrado compartilhado entre os países do estudo nos participantes com DM. Esses achados sugerem que a TB leva à alteração em alguns genes estudados no nosso trabalho, independente do sítio de estudo, mantendo, porém, algum grau de especificidade na resposta de acordo com o país de origem. Além disso, a ausência de DEGs compartilhados nos participantes com DM demonstram efeitos epigenéticos importantes associados à condição. Esses resultados não nos causam surpresas. Em estudo anterior, identificamos diferenças no perfil de ativação inflamatória em pacientes chineses e indianos com tuberculose (11). Utilizando o MDP, evidenciamos que pacientes do grupo controle eram indistinguíveis do ponto de vista de perturbação molecular, enquanto aqueles com

PTB demonstravam uma distinção na perturbação molecular capaz de distinguir por completo os dois países (11). O impacto da expressão genica, da epigenética e de fatores ambientais, é foco de vários trabalhos no campo da DM. Está bem descrito que o diabetes e suas complicações surgem da combinação de alterações genéticas e epigenéticas, resultando em mudanças no fenótipo celular (64).

Posteriormente, ainda no primeiro artigo, buscamos quantificar a expressão dos DEGs nos países entre os grupos de estudo e, nos grupos de estudo entre os países. Para isso, utilizamos o MDP, técnica amplamente utilizada pelo nosso grupo em estudos em doenças infecciosas e não infecciosas (11, 12, 38-41). Notamos que a superposição de processos TB e DM leva a um elevado grau de perturbação molecular, independente do país estudado, sugerindo impactos quantitativos na associação TBDM. Além disso, notamos que os pacientes da Romênia apresentam os maiores valores de perturbação molecular, exceto no grupo controle. Diferenças quantitativas na ativação inflamatória entre países foram também encontradas em publicação previa, onde um elevado grau de perturbação molecular foi encontrado em pacientes PTB na Índia quando comparados com pacientes da China (11).

Após identificadas alterações quantitativas na resposta, nosso próximo passo foi avaliar qualitativamente o impacto da disglicemia na ativação inflamatória de TBDM. Utilizamos então o enriquecimento de vias utilizando os DEGs previamente identificados. Apesar de não identificarmos vias unicamente expressas em TBDM nos sítios clínicos estudados, obtivemos resultados compatíveis com a literatura ao encontrar via de sinalização do interferon enriquecidas entre os pacientes com TB no Brasil, África do Sul e Romênia. Estudos prévios utilizando transcriptoma reportaram as vias de IFN identificadas em pacientes com TB ativa, mas não naqueles com TB latente (54). Complementarmente, vias associadas ao IFN-γ e ao IFN do tipo I (alfa e beta), foram identificadas em uma assinatura de TB, onde dominou o perfil de indução de IFN dirigido por neutrófilos (65). Por outro lado, identificamos nas vias de TBDM expressão de degranulação de neutrófilos, peptídeos antimicrobianos (AMPs) e organização de matriz extracelular nos participantes de Brasil e Índia. A associação das vias de AMPs e TBDM foi foco de estudo anterior, que revelou níveis aumentados de catelicidina e defensina humana beta-2 em TBDM (66). As vias de organização de matriz extracelular, e as metaloproteínases de matriz (MMP) presentes nesse resultado e em análises do nosso segundo manuscrito, vem sendo amplamente

estudadas na TB e na TBDM, sendo inclusive tema de um artigo de revisão do nosso grupo publicado em 2021 (67). As MMPs e os seus inibidores (TIMPs) foram foco de estudo preditivo de prognóstico de TB pulmonar (68) e, ao que parece, nossos resultados do segundo manuscrito dessa tese reforçam a relevância das MMPs também na interação TBDM.

Conflitando resultados obtidos no enriquecimento de vias do artigo I com os resultados identificados no enriquecimento de vias do artigo III, onde utilizamos RNA longos não codificadores (lncRNA), observamos maior heterogeneidade na ativação das vias biológicas em vias reguladas por lncRNA, destacando potencial papel dessas moléculas na fisiopatologia da TB e TBDM.

A fim de aprimorar a avaliação dos impactos da disglicemia na ativação inflamatória, estudamos a correlação dos níveis de HbA1c com a expressão das vias biológicas. Identificamos correlação positiva da HbA1c com a via de ativação das MMP. Portanto, aqui mostramos alterações quantitativas e identificamos alterações qualitativas que podem justificar as consequências da associação TBDM. A presente tese traz a participação das vias de organização de matriz celular, no primeiro manuscrito e, a MMP-28 no segundo manuscrito. No processo imunopatogênico da TB, o balanço entre MMP e TIMP está ligado ao remodelamento tecidual e a MMP-28 vem sendo descrita em condições homeostáticas (69, 70). Nossos achados demonstraram aumento desse marcador em pacientes do grupo controle, o que reforça o papel da MMP-28 na homeostase e argumenta a favor da quebra desse equilíbrio na cascata que envolve inflamação, remodelamento tecidual e estresse oxidativo (67), que culmina por fim na disseminação bacilar após destruição tissular. Esses achados contribuem para o entendimento do risco aumentado de conversão no QuantiFERON em contactantes de pessoas com TB e disglicemia (21).

O próximo passo dessa tese traz abordagem científica inovadora utilizando inteligência artificial baseada em aprendizado por máquina (machine-learning). A técnica escolhida se aplicou aos três trabalhos originais, com pequenas adaptações para o objetivo desejado.

No primeiro estudo, utilizamos os dados do consórcio TANDEM como set de descoberta, aplicando o Randon Forrest. O modelo selecionou 4 variáveis que foram posteriormente aplicadas nos dados do coorte MSTDI como validação. Os 4 genes

encontrados pelo nosso modelo, SMARCD3, VAMP5, ANKRD22 e BATF2 apresentaram boa acurácia na identificação de TBDM, com 97% de acurácia no Brasil e na África do Sul, 92% na Índia e 89% na Romênia. Esses achados podem embasar estudos futuros na identificação de uma assinatura transcriptômica que caracterize a associação TBDM.

Após avaliar possíveis assinaturas transcriptômicas relacionadas ao TBDM, estudando quantitativamente e qualitativamente a resposta, nos propomos a estudar mais de uma plataforma ômica, a fim de delinear o expressão multi-ômica da interação TBDM. No segundo manuscrito dessa tese, que ao nosso conhecimento é o primeiro trabalho que traz avaliação multi-ômica da interação TBDM, utilizamos a IA como ponto de partida para reduzir a dimensão do dado, haja vista integração de três plataformas: transcriptoma, dosagem de biomarcadores inflamatórios em sangue periférico por LUMINEX e dosagem de eicosanoides na urina, coletados dos mesmos pacientes e nos mesmos pontos temporais do estudo. Aqui nosso modelo selecionou 7 variáveis: MMP-28, previamente citada, o leucotrieno E4 (LTE4), o 11-dehidrotromboxano (dTx) B2, o metabólito D da prostaglandina (PGDM) e os genes FBXO6, SECTM1 e LINCO2009, RNA não codificador (ncRNA), com poucos estudos no contexto TB.

O papel dos eicosanoides na patogênese da TB foi previamente descrito pelo nosso grupo (10, 71). O LTE4 vem sendo associado à inflamação pulmonar (72), apesar de o seu papel na TB permanecer incerto. Nos resultados obtidos no segundo artigo dessa tese, demonstramos elevação dos níveis do analito nos grupos de TB, DM e TBDM em comparação com o grupo controle. Digno de nota, os níveis apresentaram queda após o início da ATT, principalmente nos pacientes com TB sem DM, sugerindo uma redução nos níveis do marcador com a atenuação da inflamação tissular após início da terapia. Estudos anteriores demonstram que pacientes com DM apresentam níveis elevados de LTE4 (73), o que pode justificar a redução no grupo de TB sem DM ter sido mais expressiva do que naqueles com TBDM. Outro eicosanoide encontrado no nosso segundo estudo foi o PGDM, com níveis elevados em nossos grupos de estudo quando comparados com os controles. Pouco se sabe a respeito desse marcador no contexto da TB e de patologia pulmonar, com um estudo demonstrando um papel de estímulo à mecanismos fungicidas em modelo de *Histoplasma capsulatum* (74). Futuros estudos mecanísticos são necessários para um melhor

entendimento do PGDM na TB e o seu papel como marcador de progressão de doença ou controle terapêutico.

Por outro lado, o 11dTxB2, um metabólito urinário da quebra do tromboxano A2 vem sendo utilizado como marcador indireto da atividade plaquetária em uma série de doenças (75-77). O status de hipercoagulabilidade vem sendo correlacionado com a gravidade da TB (78), justificando o uso de alguns antiplaquetários como potenciais agentes terapêuticos nas terapias alvo para PTB (79), conforme reportado em - recente artigo de revisão do nosso grupo. O nosso achado reforça o papel do marcador associado com dano pulmonar, uma vez que os níveis mostram queda após a ATT, independente do status glicêmico, destacando o 11dTxB2 como potencial marcador para estudos futuros em progressão de doença pulmonar e resposta terapêutica.

Ademais, encontramos dois genes após aplicação do nosso modelo de IA, o FBXO6 e o SECTM1, além de um RNA não codificador, o LINCO2009. O achado nos intrigou e motivou o terceiro trabalho dessa tese. Desses três genes, o mais estudado no campo das doenças infecciosas é o SECTM1, sendo reportado em estudos prévios como um gene associado a respostas precoces de IFN (80), marcador crucial para ativação macrofágica e proteção contra o Mtb (81, 82). Nossos achados revelaram uma correlação entre o SECTM1 e a graduação do escarro em pacientes TBDM, destacando uma possível utilidade do marcador na avaliação da extensão da doença pulmonar em indivíduos com ambas as condições.

Estendemos nossa análise a fim de avaliar a acurácia dos genes identificados na nossa abordagem multi-ômica na identificação das condições estudadas. De maneira curiosa, a assinatura genica identificada pelo modelo de aprendizado de máquina treinado com as multi-plataformas revelou acurácia oscilando de 91%-99% na identificação de TBDM e 89%-96% na identificação de TB, a depender do sitio testado, revelando melhor performance do que a assinatura genica identificada no primeiro trabalho dessa tese, onde apenas dados transcriptômicos foram fornecidos no aprendizado de máquina. Dessa forma, ressaltamos a relevância de trabalhos multi-ômicos nos estudos de doenças inflamatórias, sejam ou não infecciosas, para uma avaliação mais completa dos estratos biológicos envolvidos na fisiopatologia das doenças.

Destacamos o achado do LINCO2009, que apesar de não ter sua função esclarecida no campo da TB, é um ncRNA, com potencial de modular vias biológicas e, potencial analito para estudos futuros.

Esse intrigante achado foi o ponto de partida para o terceiro trabalho dessa tese, que visou avaliar o papel dos RNAs não codificadores no contexto da TB e TBDM, um objetivo inovador haja vista a escassez de trabalhos que avaliam tais moléculas no contexto das doenças infecciosas. ncRNA representam aproximadamente 60% dos produtos transcricionais, sendo o os ncRNA longos (lncRNA) os mais comuns. Apesar de pouco estudados e com funções pouco elucidadas, os lncRNA exercem papel crucial na regulação de uma gama de processos biológicos, como regulação da expressão gênica e das atividades de RNA mensageiros, afetando a tradução de proteínas e consequentemente processos biológicos e homeostáticos.

O ponto de partida do nosso trabalho foi a identificação de DEGs em um set de descoberta, os dados transciptômicos do RePORT Brasil, previamente citados aqui. Foram identificados 189 DEGs na comparação de TB com os controles e 1128 DEGs entre TBDM e controles. Após isso, buscamos identificar entre esses genes diferencialmente expressos quantos eram ncRNA, resultando em 120 ncRNA. Então, aplicamos o modelo de aprendizado de máquina afim de identificar os 5 ncRNA que caracterizassem TB e TBDM. ADM-DT, LINC02009, LINC02471, SOX2-OT e GK-AS1 foram os top ncRNA. Posteriormente avaliamos a expressão desses marcadores, além de avaliar a acurácia dessa assinatura de ncRNA na identificação de TB e TBDM nos sítios Índia, Romenia e África do Sul, a fim de validar nosso achado. Os resultados mostraram robustez na identificação de TB e TBDM nos sítios avaliados.

Buscando insights sobre o papel dos ncRNA nas vias biológicas, performamos uma correlação entre os ncRNA e a expressão dos demais genes, selecionando aqueles fortemente correlacionados com os ncRNA no Brasil e Índia. A maior parte dessas correlações foram positivas, e encontramos expressão dos genes fortemente correlacionados com os lncRNA em diversas vias de relevância no contexto da TB, como a via de degranulação de neutrófilos, sinalização por interleucinas e interferons, vias de regulação de necrose dentre outras.

Nossos achados revelam a influência dos lncRNA na patogênese da TB e TBDM, destacando o impacto na expressão de genes associados a vias fundamentais na

resposta inflamatória à TB, evidenciando potencial papel dos lncRNA como biomarcadores de TB que podem futuramente ser estudados.

A presente tese possui algumas limitações. Primeiro, encontramos algumas diferenças em fatores epidemiológicos, como idade, sexo, tabagismo, uso de álcool e Índice de Massa Corporal (IMC) entre os grupos de estudo utilizados aqui, que podem afetar o processo de ativação inflamatória. No entanto, performamos análises suplementares a fim de avaliar o efeito dessas variáveis na distinção dos grupos. Outra limitação foi a ausência de alguns dados epidemiológicos da amostra pública do TANDEM, limitando o ajuste das variáveis apenas idade e sexo. Além disso, alguns pacientes estavam em uso de metformina e estatina, que também podem afetar a expressão inflamatória, podendo constituir-se como fator de confusão. Podemos ainda mencionar a diferença numérica das amostras do primeiro estudo, que podem influenciar tanto nas análises de expressão genica, com também na performance de possíveis assinaturas. Apesar dessas limitações, a presente tese traz conteúdo de grande relevância para o campo da TB e da interação TBDM, utilizando metodologia inovadora, identificando alguns aspectos do impacto da disglicemia na ativação inflamatória de pacientes com TBDM, que poderão futuramente ser utilizados como alvo para estratégias terapêuticas, além de trazer marcadores multi-ômicos e destacar o papel de lncRNA na patogênese da TB, que poderão embasar novos estudos mecanísticos na busca de marcadores para definição de prognóstico e evolução da TB em portadores de DM.

8 CONCLUSÕES

O conjunto de achados dos três manuscritos que compõe a tese levam as seguintes conclusões:

Artigo 1:

- As condições clínicas TB, DM e TBDM compartilham poucos genes diferencialmente expressos na comparação entre os pacientes do Brasil, Índia, África do Sul e Romênia;
- Pacientes com TBDM apresentam maior grau de perturbação molecular;
- A influência dos genes diferencialmente expressos nas vias biológicas é distinta, não apresentando um mesmo padrão entre os países, exceto as vias de degranulação de neutrófilos, peptídeos antimicrobianos e a via de organização da matriz extracelular;
- A assinatura gênica encontrada após aplicação do modelo de Randon Forest caracterizou TBDM com acurácia de 89%-97% nos diferentes sítios;
- Níveis de HbA1c estão correlacionados com importantes vias biológicas, como a das metaloproteinases de matriz.

Artigo 2:

- A disglicemia está associada a impactos em múltiplas plataformas ômicas;
- Existem potenciais marcadores multi-ômicos que podem ser úteis no desenvolvimento de ferramenta de predição de desfecho e prognósticos de pacientes com TB e TBDM;
- Estratégias multi-ômicas podem somar na avaliação de doenças inflamatórias.

Artigo 3:

- RNAs não codificadores, apesar de pouco estudados, podem influenciar diversas vias inflamatórias na TB e TBDM;
- O set de ncRNA selecionados pela IA apresentou acurácia robusta na identificação de TB, independente do status glicêmico;

Por fim, o conhecimento gerado por essa tese de doutorado é de grande contribuição ao meio científico, principalmente ao campo da interação TBDM, trazendo aspectos quantitativos e qualitativos do impacto da disglicemia na ativação inflamatória de pacientes TBDM. Além disso, utilizando nosso modelo de inteligência artificial identificamos sete marcadores multi-ômicos associados com a interação TBDM. Em conjunto, nossos achados poderão servir de embasamento para futuros estudos que visem a utilização da medicina de precisão para rastreio, seguimento de tratamento, definição de prognóstico e, também, no desenvolvimento de terapias alvo, contribuindo para sociedade científica na estratégia de erradicação da TB.

REFERÊNCIAS

1. Organization WH. Global Tuberculosis Report. 2022.

2. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. Nat Rev Dis Primers. 2016;2:16076.

3. Silveira-Mattos PS, Barreto-Duarte B, Vasconcelos B, Fukutani KF, Vinhaes CL, Oliveira-De-Souza D, et al. Differential Expression of Activation Markers by Mycobacterium tuberculosis-specific CD4+ T Cell Distinguishes Extrapulmonary From Pulmonary Tuberculosis and Latent Infection. Clin Infect Dis. 2020;71(8):1905-11.

4. Narendran G, Jyotheeswaran K, Senguttuvan T, Vinhaes CL, Santhanakrishnan RK, Manoharan T, et al. Characteristics of paradoxical tuberculosis-associated immune reconstitution inflammatory syndrome and its influence on tuberculosis treatment outcomes in persons living with HIV. Int J Infect Dis. 2020;98:261-7.

5. Vinhaes CL, Sheikh V, Oliveira-de-Souza D, Wang J, Rupert A, Roby G, et al. An Inflammatory Composite Score Predicts Mycobacterial Immune Reconstitution Inflammatory Syndrome in People with Advanced HIV: A Prospective International Cohort Study. J Infect Dis. 2021;223(7):1275-83.

6. Breglio KF, Vinhaes CL, Arriaga MB, Nason M, Roby G, Adelsberger J, et al. Clinical and Immunologic Predictors of Mycobacterium avium Complex Immune Reconstitution Inflammatory Syndrome in a Contemporary Cohort of Patients With Human Immunodeficiency Virus. J Infect Dis. 2021;223(12):2124-35.

7. Vinhaes CL, Araujo-Pereira M, Tiburcio R, Cubillos-Angulo JM, Demitto FO, Akrami KM, et al. Systemic Inflammation Associated with Immune Reconstitution Inflammatory Syndrome in Persons Living with HIV. Life (Basel). 2021;11(1).

8. Kauffman KD, Sakai S, Lora NE, Namasivayam S, Baker PJ, Kamenyeva O, et al. PD-1 blockade exacerbates Mycobacterium tuberculosis infection in rhesus macaques. Sci Immunol. 2021;6(55).

9. Silveira-Mattos PS, Narendran G, Akrami K, Fukutani KF, Anbalagan S, Nayak K, et al. Differential expression of CXCR3 and CCR6 on CD4(+) T-lymphocytes with distinct memory phenotypes characterizes tuberculosis-associated immune reconstitution inflammatory syndrome. Sci Rep. 2019;9(1):1502.

10. Vinhaes CL, Oliveira-de-Souza D, Silveira-Mattos PS, Nogueira B, Shi R, Wei W, et al. Changes in inflammatory protein and lipid mediator profiles persist after antitubercular treatment of pulmonary and extrapulmonary tuberculosis: A prospective cohort study. Cytokine. 2019;123:154759.

11. Oliveira-de-Souza D, Vinhaes CL, Arriaga MB, Kumar NP, Cubillos-Angulo JM, Shi R, et al. Molecular degree of perturbation of plasma inflammatory markers

associated with tuberculosis reveals distinct disease profiles between Indian and Chinese populations. Sci Rep. 2019;9(1):8002.

12. Oliveira-de-Souza D, Vinhaes CL, Arriaga MB, Kumar NP, Queiroz ATL, Fukutani KF, et al. Aging increases the systemic molecular degree of inflammatory perturbation in patients with tuberculosis. Sci Rep. 2020;10(1):11358.

13. Kathirvel M, Mahadevan S. The role of epigenetics in tuberculosis infection. Epigenomics. 2016;8(4):537-49.

14. Murray M, Oxlade O, Lin HH. Modeling social, environmental and biological determinants of tuberculosis. Int J Tuberc Lung Dis. 2011;15 Suppl 2:64-70.

15. Gupta KB, Gupta R, Atreja A, Verma M, Vishvkarma S. Tuberculosis and nutrition. Lung India. 2009;26(1):9-16.

16. Organization WH. Global report on diabetes. 2016.

17. Berbudi A, Rahmadika N, Tjahjadi AI, Ruslami R. Type 2 Diabetes and its Impact on the Immune System. Curr Diabetes Rev. 2020;16(5):442-9.

18. Lim S, Bae JH, Kwon HS, Nauck MA. COVID-19 and diabetes mellitus: from pathophysiology to clinical management. Nat Rev Endocrinol. 2021;17(1):11-30.

19. Jeon CY, Murray MB. Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. PLoS Med. 2008;5(7):e152.

20. Barreda NN, Arriaga MB, Aliaga JG, Lopez K, Sanabria OM, Carmo TA, et al. Severe pulmonary radiological manifestations are associated with a distinct biochemical profile in blood of tuberculosis patients with dysglycemia. BMC Infect Dis. 2020;20(1):139.

21. Arriaga MB, Rocha MS, Nogueira BMF, Nascimento V, Araujo-Pereira M, Souza AB, et al. The Effect of Diabetes and Prediabetes on Mycobacterium tuberculosis Transmission to Close Contacts. J Infect Dis. 2021;224(12):2064-72.

22. Arriaga MB, Karim F, Queiroz ATL, Araujo-Pereira M, Barreto-Duarte B, Sales C, et al. Effect of Dysglycemia on Urinary Lipid Mediator Profiles in Persons With Pulmonary Tuberculosis. Front Immunol. 2022;13:919802.

23. Calderon RI, Arriaga MB, Aliaga JG, Barreda NN, Sanabria OM, Barreto-Duarte B, et al. Persistent dysglycemia is associated with unfavorable treatment outcomes in patients with pulmonary tuberculosis from Peru. Int J Infect Dis. 2022;116:293-301.

24. Arriaga MB, Araujo-Pereira M, Barreto-Duarte B, Nogueira B, Freire M, Queiroz ATL, et al. The Effect of Diabetes and Prediabetes on Antituberculosis Treatment Outcomes: A Multicenter Prospective Cohort Study. J Infect Dis. 2022;225(4):617-26.

25. Johnson KB, Wei WQ, Weeraratne D, Frisse ME, Misulis K, Rhee K, et al. Precision Medicine, AI, and the Future of Personalized Health Care. Clin Transl Sci. 2021;14(1):86-93.

26. Orth M, Averina M, Chatzipanagiotou S, Faure G, Haushofer A, Kusec V, et al. Opinion: redefining the role of the physician in laboratory medicine in the context of emerging technologies, personalised medicine and patient autonomy ('4P medicine'). J Clin Pathol. 2019;72(3):191-7.

27. Nensa F, Demircioglu A, Rischpler C. Artificial Intelligence in Nuclear Medicine. J Nucl Med. 2019;60(Suppl 2):29S-37S.

28. Keskinbora K, Guven F. Artificial Intelligence and Ophthalmology. Turk J Ophthalmol. 2020;50(1):37-43.

29. Hashimoto DA, Witkowski E, Gao L, Meireles O, Rosman G. Artificial Intelligence in Anesthesiology: Current Techniques, Clinical Applications, and Limitations. Anesthesiology. 2020;132(2):379-94.

30. Hanko M, Grendar M, Snopko P, Opsenak R, Sutovsky J, Benco M, et al. Random Forest-Based Prediction of Outcome and Mortality in Patients with Traumatic Brain Injury Undergoing Primary Decompressive Craniectomy. World Neurosurg. 2021;148:e450-e8.

31. Yang L, Wu H, Jin X, Zheng P, Hu S, Xu X, et al. Study of cardiovascular disease prediction model based on random forest in eastern China. Sci Rep. 2020;10(1):5245.

32. Kulkarni V, Queiroz ATL, Sangle S, Kagal A, Salvi S, Gupta A, et al. A Two-Gene Signature for Tuberculosis Diagnosis in Persons With Advanced HIV. Front Immunol. 2021;12:631165.

33. Krishnan S, Queiroz ATL, Gupta A, Gupte N, Bisson GP, Kumwenda J, et al. Integrative Multi-Omics Reveals Serum Markers of Tuberculosis in Advanced HIV. Front Immunol. 2021;12:676980.

34. Hamilton CD, Swaminathan S, Christopher DJ, Ellner J, Gupta A, Sterling TR, et al. RePORT International: Advancing Tuberculosis Biomarker Research Through Global Collaboration. Clin Infect Dis. 2015;61Suppl 3(Suppl 3):S155-9.

35. Kornfeld H, West K, Kane K, Kumpatla S, Zacharias RR, Martinez-Balzano C, et al. High Prevalence and Heterogeneity of Diabetes in Patients With TB in South India: A Report from the Effects of Diabetes on Tuberculosis Severity (EDOTS) Study. Chest. 2016;149(6):1501-8.

36. Gupte A, Padmapriyadarsini C, Mave V, Kadam D, Suryavanshi N, Shivakumar SV, et al. Cohort for Tuberculosis Research by the Indo-US Medical Partnership (CTRIUMPH): protocol for a multicentric prospective observational study. BMJ Open. 2016;6(2):e010542.

37. van Crevel R, Dockrell HM, Consortium T. TANDEM: understanding diabetes and tuberculosis. Lancet Diabetes Endocrinol. 2014;2(4):270-2.

38. Vinhaes CL, Arriaga MB, de Almeida BL, Oliveira JV, Santos CS, Calcagno JI, et al. Newborns With Zika Virus-Associated Microcephaly Exhibit Marked Systemic Inflammatory Imbalance. J Infect Dis. 2020;222(4):670-80.

39. Vinhaes CL, Carmo TA, Queiroz ATL, Fukutani KF, Araujo-Pereira M, Arriaga MB, et al. Dissecting disease tolerance in Plasmodium vivax malaria using the systemic degree of inflammatory perturbation. PLoS Negl Trop Dis. 2021;15(11):e0009886.

40. Vinhaes CL, Cruz LAB, Menezes RC, Carmo TA, Arriaga MB, Queiroz ATL, et al. Chronic Hepatitis B Infection Is Associated with Increased Molecular Degree of Inflammatory Perturbation in Peripheral Blood. Viruses. 2020;12(8).

41. Vinhaes CL, Teixeira RS, Monteiro-Junior JAS, Tiburcio R, Cubillos-Angulo JM, Arriaga MB, et al. Hydroxyurea treatment is associated with reduced degree of oxidative perturbation in children and adolescents with sickle cell anemia. Sci Rep. 2020;10(1):18982.

42. Eckold C, Kumar V, Weiner J, Alisjahbana B, Riza AL, Ronacher K, et al. Impact of Intermediate Hyperglycemia and Diabetes on Immune Dysfunction in Tuberculosis. Clin Infect Dis. 2021;72(1):69-78.

43. Ayelign B, Negash M, Genetu M, Wondmagegn T, Shibabaw T. Immunological Impacts of Diabetes on the Susceptibility of Mycobacterium tuberculosis. J Immunol Res. 2019;2019:6196532.

44. Geerlings SE, Hoepelman AI. Immune dysfunction in patients with diabetes mellitus (DM). FEMS Immunol Med Microbiol. 1999;26(3-4):259-65.

45. Martinez N, Kornfeld H. Diabetes and immunity to tuberculosis. Eur J Immunol. 2014;44(3):617-26.

46. Yew WW, Leung CC, Zhang Y. Oxidative stress and TB outcomes in patients with diabetes mellitus? J Antimicrob Chemother. 2017;72(6):1552-5.

47. Kumar Nathella P, Babu S. Influence of diabetes mellitus on immunity to human tuberculosis. Immunology. 2017;152(1):13-24.

48. Kumar NP, Moideen K, Sivakumar S, Menon PA, Viswanathan V, Kornfeld H, et al. Modulation of dendritic cell and monocyte subsets in tuberculosis-diabetes co-morbidity upon standard tuberculosis treatment. Tuberculosis (Edinb). 2016;101:191-200.

49. Raposo-Garcia S, Guerra-Laso JM, Garcia-Garcia S, Juan-Garcia J, Lopez-Fidalgo E, Diez-Tascon C, et al. Immunological response to Mycobacterium tuberculosis infection in blood from type 2 diabetes patients. Immunol Lett. 2017;186:41-5.

50. Restrepo BI, Schlesinger LS. Host-pathogen interactions in tuberculosis patients with type 2 diabetes mellitus. Tuberculosis (Edinb). 2013;93 Suppl(0):S10-4.

51. Yamashiro S, Kawakami K, Uezu K, Kinjo T, Miyagi K, Nakamura K, et al. Lower expression of Th1-related cytokines and inducible nitric oxide synthase in mice with streptozotocin-induced diabetes mellitus infected with Mycobacterium tuberculosis. Clin Exp Immunol. 2005;139(1):57-64.

52. Kumar NP, Moideen K, George PJ, Dolla C, Kumaran P, Babu S. Coincident diabetes mellitus modulates Th1-, Th2-, and Th17-cell responses in latent tuberculosis in an IL-10- and TGF-beta-dependent manner. Eur J Immunol. 2016;46(2):390-9.

53. Hamada Y, Penn-Nicholson A, Krishnan S, Cirillo DM, Matteelli A, Wyss R, et al. Are mRNA based transcriptomic signatures ready for diagnosing tuberculosis in the clinic? - A review of evidence and the technological landscape. EBioMedicine. 2022;82:104174.

54. Singhania A, Wilkinson RJ, Rodrigue M, Haldar P, O'Garra A. The value of transcriptomics in advancing knowledge of the immune response and diagnosis in tuberculosis. Nat Immunol. 2018;19(11):1159-68.

55. Burel JG, Babor M, Pomaznoy M, Lindestam Arlehamn CS, Khan N, Sette A, et al. Host Transcriptomics as a Tool to Identify Diagnostic and Mechanistic Immune Signatures of Tuberculosis. Front Immunol. 2019;10:221.

56. Prada-Medina CA, Fukutani KF, Pavan Kumar N, Gil-Santana L, Babu S, Lichtenstein F, et al. Systems Immunology of Diabetes-Tuberculosis Comorbidity Reveals Signatures of Disease Complications. Sci Rep. 2017;7(1):1999.

57. Liu T, Wang Y, Gui J, Fu Y, Ye C, Hong X, et al. Transcriptome analysis of the impact of diabetes as a comorbidity on tuberculosis. Medicine (Baltimore). 2022;101(52):e31652.

58. van Doorn CLR, Eckold C, Ronacher K, Ruslami R, van Veen S, Lee JS, et al. Transcriptional profiles predict treatment outcome in patients with tuberculosis and diabetes at diagnosis and at two weeks after initiation of anti-tuberculosis treatment. EBioMedicine. 2022;82:104173.

59. Malta-Santos H, Fukutani KF, Sorgi CA, Queiroz ATL, Nardini V, Silva J, et al. Multi-omic Analyses of Plasma Cytokines, Lipidomics, and Transcriptomics Distinguish Treatment Outcomes in Cutaneous Leishmaniasis. iScience. 2020;23(12):101840.

60. Cuypers B, Meysman P, Erb I, Bittremieux W, Valkenborg D, Baggerman G, et al. Four layer multi-omics reveals molecular responses to aneuploidy in Leishmania. PLoS Pathog. 2022;18(9):e1010848.

61. Linh NN, Viney K, Gegia M, Falzon D, Glaziou P, Floyd K, et al. World Health Organization treatment outcome definitions for tuberculosis: 2021 update. Eur Respir J. 2021;58(2).

62. American Diabetes A. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2021. Diabetes Care. 2021;44(Suppl 1):S73-S84.

63. Breiman L. Random Forest. Machine Learning. 2001.

64. Rosen ED, Kaestner KH, Natarajan R, Patti ME, Sallari R, Sander M, et al. Epigenetics and Epigenomics: Implications for Diabetes and Obesity. Diabetes. 2018;67(10):1923-31.

65. Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature. 2010;466(7309):973-7.

66. Kumar NP, Moideen K, Viswanathan V, Sivakumar S, Menon PA, Kornfeld H, et al. Heightened circulating levels of antimicrobial peptides in tuberculosis-Diabetes co-morbidity and reversal upon treatment. PLoS One. 2017;12(9):e0184753.

67. Amaral EP, Vinhaes CL, Oliveira-de-Souza D, Nogueira B, Akrami KM, Andrade BB. The Interplay Between Systemic Inflammation, Oxidative Stress, and Tissue Remodeling in Tuberculosis. Antioxid Redox Signal. 2021;34(6):471-85.

68. Kumar NP, Moideen K, Nancy A, Viswanathan V, Thiruvengadam K, Sivakumar S, et al. Association of Plasma Matrix Metalloproteinase and Tissue Inhibitors of Matrix Metalloproteinase Levels With Adverse Treatment Outcomes Among Patients With Pulmonary Tuberculosis. JAMA Netw Open. 2020;3(12):e2027754.

69. Ra HJ, Parks WC. Control of matrix metalloproteinase catalytic activity. Matrix Biol. 2007;26(8):587-96.

70. Sabir N, Hussain T, Mangi MH, Zhao D, Zhou X. Matrix metalloproteinases: Expression, regulation and role in the immunopathology of tuberculosis. Cell Prolif. 2019;52(4):e12649.

71. Mayer-Barber KD, Andrade BB, Oland SD, Amaral EP, Barber DL, Gonzales J, et al. Host-directed therapy of tuberculosis based on interleukin-1 and type I interferon crosstalk. Nature. 2014;511(7507):99-103.

72. Paruchuri S, Tashimo H, Feng C, Maekawa A, Xing W, Jiang Y, et al. Leukotriene E4-induced pulmonary inflammation is mediated by the P2Y12 receptor. J Exp Med. 2009;206(11):2543-55.

73. Rafnsson A, Back M. Urinary leukotriene E4 is associated with renal function but not with endothelial function in type 2 diabetes. Dis Markers. 2013;35(5):475-80.

74. Pereira PAT, Assis PA, Prado MKB, Ramos SG, Aronoff DM, de Paula-Silva FWG, et al. Prostaglandins D(2) and E(2) have opposite effects on alveolar macrophages infected with Histoplasma capsulatum. J Lipid Res. 2018;59(2):195-206.

75. Lordkipanidze M, Pharand C, Schampaert E, Turgeon J, Palisaitis DA, Diodati JG. A comparison of six major platelet function tests to determine the prevalence of aspirin resistance in patients with stable coronary artery disease. Eur Heart J. 2007;28(14):1702-8.

76. Oosaki R, Mizushima Y, Mita H, Shida T, Akiyama K, Kobayashi M. Urinary leukotriene E4 and 11-dehydrothromboxane B2 in patients with aspirin-sensitive asthma. Allergy. 1997;52(4):470-3.

77. Catella F, Healy D, Lawson JA, FitzGerald GA. 11-Dehydrothromboxane B2: a quantitative index of thromboxane A2 formation in the human circulation. Proc Natl Acad Sci U S A. 1986;83(16):5861-5.

78. Kirwan DE, Chong DLW, Friedland JS. Platelet Activation and the Immune Response to Tuberculosis. Front Immunol. 2021;12:631696.

79. Cubillos-Angulo JM, Nogueira BMF, Arriaga MB, Barreto-Duarte B, Araujo-Pereira M, Fernandes CD, et al. Host-directed therapies in pulmonary tuberculosis: Updates on anti-inflammatory drugs. Front Med (Lausanne). 2022;9:970408.

80. Huyton T, Gottmann W, Bade-Doding C, Paine A, Blasczyk R. The T/NK cell co-stimulatory molecule SECTM1 is an IFN "early response gene" that is negatively regulated by LPS in human monocytic cells. Biochim Biophys Acta. 2011;1810(12):1294-301.

81. Khan TA, Mazhar H, Saleha S, Tipu HN, Muhammad N, Abbas MN. Interferon-Gamma Improves Macrophages Function against M. tuberculosis in Multidrug-Resistant Tuberculosis Patients. Chemother Res Pract. 2016;2016:7295390.

82. Flynn JL, Chan J, Triebold KJ, Dalton DK, Stewart TA, Bloom BR. An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection. J Exp Med. 1993;178(6):2249-54.

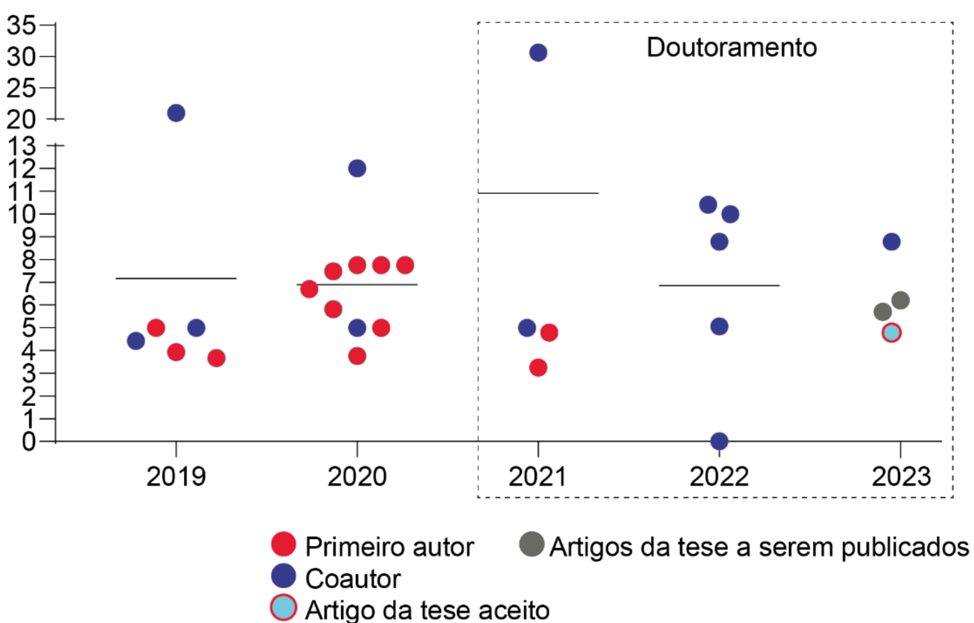
PERSPECTIVA HISTÓRICA DA TESE

Ingresso no programa de Iniciação Científica em 2016, ainda na faculdade de medicina, iniciei junto ao meu grupo estudos focados no entendimento de marcadores moleculares associados a ativação inflamatória na Tuberculose. Inicialmente caracterizamos o perfil de expressão inflamatória de pacientes com PTB e EPTB, utilizando 27 citocinas inflamatórias, receptores solúveis e mediadores lipídicos. Nos anos que seguiram, agora como bolsista do CNPq, adaptamos a ferramenta estatística MDP, utilizada em estudos de genoma, para o estudo de biomarcadores solúveis, em uma linha que se mantem até os dias atuais, de estudos que buscam caracterizar perfis de expressão inflamatória em várias doenças, como Malária, Hepatite B, Zika vírus, TB, HIV e Síndrome de Reconstituição Imune. Com o passar dos anos, novas linhas de pesquisa foram surgindo buscando ferramentas inovadoras na análise de dados que possam predizer desfechos dessas doenças. Ao final do período de iniciação cientifica, foram publicados 17 artigos em periódicos internacionais, incluindo revistas de grande relevância no meio das doenças infectocontagiosas, como o Clinical Infectious Diseases, o The Journal of Infectious Diseases, o Internacional Journal of Infectious Diseases, Science Immunology e Cell Reposrts. Com uma linha forte de pesquisa dentro do grupo que estudava a interação de Tuberculose e Diabetes escolhemos dar continuidade aos projetos agregando fatores que estiveram presentes durante a iniciação cientifica, agora aplicados ao entendimento do impacto das disglicemias em pacientes portadores de diabetes.

ANEXOS

Anexo I: Produção científica no Doutorado

O estudante iniciou os trabalhos no Laboratório de Inflamação e Biomarcadores em junho de 2016, como estudante voluntário de iniciação cientifica. Até 01 de março de 2023, o estudante foi autor de 25 trabalhos científicos em periódicos, incluindo 11 como primeiro autor e 14 colaborações. Abaixo está ilustrada a produção durante todo período, com uma lista dos manuscritos que foram produzidos no período do doutorado. O primeiro artigo dessa tese está publicado na revista Scientific Reports, do grupo Nature, com fator de impacto 4.99. O segundo e terceiro trabalhos estão em fase de revisão na revista iScience, com fator de impacto 5.8.



2018-2020: Iniciação científica (bolsista CNPq) – 02 anos. Total de Publicações: 16.

2021-2023: Doutorado – 02 anos. Total de publicações 11.

Média de publicações/ano em março/2023: 5.4.

Total de publicações listadas no PubMed: 27

Total de citações (Web of Science): 231

Média de citação por artigo: 9,24

Média do fator de impacto das publicações: 7,8

Mediana do fator de impacto das publicações: 5,4

Artigos aceitos para publicação 2021-2023

www.nature.com/scientificreports

scientific reports

OPEN

A multi-center, prospective cohort study of whole blood gene expression in the tuberculosis-diabetes interaction

Artur T. L. Queiroz[1,2,3,25], Caian L. Vinhaes[2,3,4,25], Eduardo R. Fukutani[1], Akshay N. Gupte[5], Nathella Pavan Kumar[6], Kiyoshi F. Fukutani[2], María B. Arriaga[2,3], Timothy R. Sterling[7], Subash Babu[6], Sanjay Gaikwad[8], Rajesh Karyakarte[9], Vidya Mave[10,11], Mandar Paradhkar[10,11], Vijay Viswanathan[12], Amita Gupta[5], Bruno B. Andrade[2,3,4,13,26✉], Hardy Kornfeld[14,15,26✉], the RePORT Brazil* & RePORT India Consortia*

Diabetes mellitus (DM) increases tuberculosis (TB) severity. We compared blood gene expression in adults with pulmonary TB, with or without diabetes mellitus (DM) from sites in Brazil and India. RNA sequencing (RNAseq) performed at baseline and during TB treatment. Publicly available baseline RNAseq data from South Africa and Romania reported by the TANDEM Consortium were also analyzed. Across the sites, differentially expressed genes varied for each condition (DM, TB, and TBDM) and no pattern classified any one group across all sites. A concise signature of TB disease was identified but this was expressed equally in TB and TBDM. Pathway enrichment analysis failed to distinguish TB from TBDM, although there was a trend for greater neutrophil and innate immune pathway activation in TBDM participants. Pathways associated with insulin resistance, metabolic dysfunction, diabetic complications, and chromosomal instability were positively correlated with glycohemoglobin. The immune response to pulmonary TB as reflected by whole blood gene expression is substantially similar with or without comorbid DM. Gene expression pathways associated with the microvascular and macrovascular complications of DM are upregulated during TB, supporting a syndemic interaction between these coprevalent diseases.

Diabetes mellitus (DM) has been associated with increased risk for tuberculosis (TB) progression and adverse TB treatment outcomes in most clinical studies[1]. Mirroring the human data, animal models combining chronic

[1]Centro de Integração de Dados e Conhecimentos para Saúde, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil. [2]Laboratório de Inflamação e Biomarcadores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil. [3]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador 41810-710, Brazil. [4]Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador 40290-150, Brazil. [5]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [6]National Institutes of Health- NIRT - International Center for Excellence in Research, Chennai, India. [7]Division of Infectious Diseases, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [8]Department of Pulmonary Medicine, Byramjee-Jeejeebhoy Government Medical College and Sassoon General Hospitals, Pune, India. [9]Department of Microbiology, Byramjee-Jeejeebhoy Government Medical College and Sassoon General Hospitals, Pune, India. [10]Byramjee-Jeejeebhoy Government Medical College-Johns Hopkins University Clinical Research Site, Pune, India. [11]Johns Hopkins Center for Infectious Diseases in India, Pune, India. [12]Prof. M. Viswanathan Diabetes Research Centre, Chennai, India. [13]Faculdade de Tecnologia e Ciências, Instituto de Pesquisa Clínica e Translacional, Salvador 41741-590, Brazil. [14]Department of Medicine, University of Massachusetts Medical School, Worcester, MA, USA. [15]UMass Chan Medical School, Worcester, MA, USA. [25]These authors contributed equally: Artur T. L. Queiroz and Caian L. Vinhaes. [26]These authors jointly supervised this work: Bruno B. Andrade and Hardy Kornfeld. *List of authors and their affiliations appears at the end of the paper. ✉email: bruno.andrade@fiocruz.br; hardy.kornfeld@umassmed.edu

TYPE Original Research
PUBLISHED 18 April 2023
DOI 10.3389/fimmu.2023.1177432

OPEN ACCESS

EDITED BY
Li Xing,
Shanxi University, China

REVIEWED BY
Maria Salomé Gomes,
Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
Stephen Cose,
University of London, United Kingdom

*CORRESPONDENCE
Bruno B. Andrade
Bruno.andrade@fiocruz.br
Mariana Araújo-Pereira
araujopereira.mariana@gmail.com

[†]These authors share first authorship

[‡]These authors share senior authorship

SPECIALTY SECTION
This article was submitted to Inflammation, a section of the journal Frontiers in Immunology

RECEIVED 01 March 2023
ACCEPTED 04 April 2023
PUBLISHED 18 April 2023

CITATION
Araújo-Pereira M, Schutz C, Barreto-Duarte B, Barr D, Villalva-Serra K, Vinhaes CL, Ward A, Meintjes G and Andrade BB (2023) Interplay between systemic inflammation, anemia, and mycobacterial dissemination and its impact on mortality in TB-associated HIV: a prospective cohort study. *Front. Immunol.* 14:1177432. doi: 10.3389/fimmu.2023.1177432

COPYRIGHT
© 2023 Araújo-Pereira, Schutz, Barreto-Duarte, Barr, Villalva-Serra, Vinhaes, Ward, Meintjes and Andrade. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Interplay between systemic inflammation, anemia, and mycobacterial dissemination and its impact on mortality in TB-associated HIV: a prospective cohort study

Mariana Araújo-Pereira[1,2,3,4*†], Charlotte Schutz[5,6†], Beatriz Barreto-Duarte[1,2,7,8], David Barr[6,9], Klauss Villalva-Serra[1,2,8], Caian L. Vinhaes[1,2,3,10], Amy Ward[5], Graeme Meintjes[5,6‡] and Bruno B. Andrade[1,2,3,4,7,8,10*‡]

[1]Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil, [2]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Brazil, [3]Programa de Pós-Graduação em Patologia Humana e Experimental, Universidade Federal da Bahia, Salvador, Bahia, Brazil, [4]Curso de Medicina, UNIFTC, Salvador, Bahia, Brazil, [5]Department of Medicine, University of Cape Town, Cape Town, South Africa, [6]Wellcome Centre for Infectious Diseases Research in Africa (CIDRI-Africa), Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa, [7]Programa de Pós-Graduação em Clínica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, [8]Curso de Medicina, Universidade Salvador (UNIFACS), Salvador, Bahia, Brazil, [9]Department of Infectious Diseases, NHS Greater Glasgow & Clyde, Glasgow, United Kingdom, [10]Bahiana School of Medicine and Public Health, Bahia Foundation for the Development of Sciences, Salvador, Brazil

Introduction: Anemia frequently affects people living with HIV (PLHIV). Nevertheless, the impact of anemia on treatment outcomes of patients with HIV-associated tuberculosis (TB) and the underlying molecular profiles are not fully characterized. The aim of this study was to investigate the interplay between anemia, the systemic inflammatory profile, dissemination of TB and death in HIV-TB patients in an ad hoc analysis of results from a prospective cohort study.

Methods: 496 hospitalized PLHIV ≥18 years old, with CD4 count <350 cells/μL and high clinical suspicion of new TB infection were enrolled in Cape Town between 2014-2016. Patients were classified according to anemia severity in non-anemic, mild, moderate, or severe anemia. Clinical, microbiologic, and immunologic data were collected at baseline. Hierarchical cluster analysis, degree of inflammatory perturbation, survival curves and C-statistics analyses were performed.

Results: Through the analysis of several clinical and laboratory parameters, we observed that those with severe anemia exhibited greater systemic inflammation, characterized by high concentrations of IL-8, IL-1RA and IL-6. Furthermore, severe anemia was associated with a higher Mtb dissemination score and a higher risk of death, particularly within 7 days of admission. Most of the patients who died had severe anemia and had a more pronounced systemic inflammatory profile.

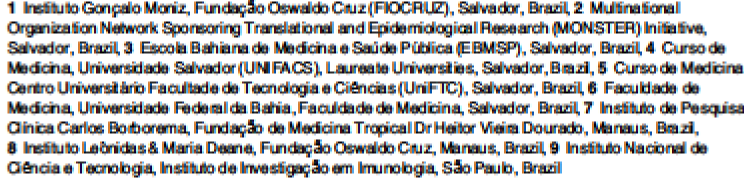
Discussion: Therefore, the results presented here reveal that severe anemia is associated with greater TB dissemination and increased risk of death in PLHIV.

PLOS NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Dissecting disease tolerance in *Plasmodium vivax* malaria using the systemic degree of inflammatory perturbation

Caian L. Vinhaes[1,2,3‡], Thomas A. Carmo[1,2,4‡], Artur T. L. Queiroz[1,2], Kiyoshi F. Fukutani[1,2,5], Mariana Araújo-Pereira[1,2,6], María B. Arriaga[1,2,6], Marcus V. G. Lacerda[7,8], Manoel Barral-Netto[1,6,9], Bruno B. Andrade[1,2,3,4,5,6]*



1 Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (FIOCRUZ), Salvador, Brazil, 2 Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Brazil, 3 Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador, Brazil, 4 Curso de Medicina, Universidade Salvador (UNIFACS), Laureate Universities, Salvador, Brazil, 5 Curso de Medicina, Centro Universitário Faculdade de Tecnologia e Ciências (UniFTC), Salvador, Brazil, 6 Faculdade de Medicina, Universidade Federal da Bahia, Faculdade de Medicina, Salvador, Brazil, 7 Instituto de Pesquisa Clínica Carlos Borborema, Fundação de Medicina Tropical Dr Heitor Vieira Dourado, Manaus, Brazil, 8 Instituto Leônidas & Maria Deane, Fundação Oswaldo Cruz, Manaus, Brazil, 9 Instituto Nacional de Ciência e Tecnologia, Instituto de Investigação em Imunologia, São Paulo, Brazil

‡ These authors share first authorship on this work.
* bruno.andrade@fiocruz.br

OPEN ACCESS

Citation: Vinhaes CL, Carmo TA, Queiroz ATL, Fukutani KF, Araújo-Pereira M, Arriaga MB, et al. (2021) Dissecting disease tolerance in *Plasmodium vivax* malaria using the systemic degree of inflammatory perturbation. PLoS Negl Trop Dis 15(11): e0009886. https://doi.org/10.1371/journal.pntd.0009886

Editor: Photini Sinnis, Johns Hopkins Bloomberg School of Public Health, UNITED STATES

Received: March 5, 2021

Accepted: October 8, 2021

Published: November 2, 2021

Copyright: © 2021 Vinhaes et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This work was supported by Financiadora de Estudos e Projetos (FINEP) (grant ward number: 0104009605) / Fundo Nacional de Desenvolvimento Científico e Tecnológico (FNDCTCT- Amazônia), Brazil (M.B.-N.). This study was also financed in part by Coordenacão de Aperfeiçoamento de Pessoal de Nível Superior

Abstract

Homeostatic perturbation caused by infection fosters two major defense strategies, resistance and tolerance, which promote the host's survival. Resistance relates to the ability of the host to restrict the pathogen load. Tolerance minimizes collateral tissue damage without directly affecting pathogen fitness. These concepts have been explored mechanistically in murine models of malaria but only superficially in human disease. Indeed, individuals infected with *Plasmodium vivax* may present with asymptomatic malaria, only mild symptoms, or be severely ill. We and others have reported a diverse repertoire of immunopathological events that potentially underly susceptibility to disease severity in *vivax* malaria. Nevertheless, the combined epidemiologic, clinical, parasitological, and immunologic features associated with defining the disease outcomes are still not fully understood. In the present study, we perform an extensive outlining of cytokines and inflammatory proteins in plasma samples from a cohort of individuals from the Brazilian Amazon infected with *P. vivax* and presenting with asymptomatic (n = 108) or symptomatic (n = 134) disease (106 with mild presentation and 28 with severe malaria), as well as from uninfected endemic controls (n = 128) to elucidate these gaps further. We employ highly multidimensional Systems Immunology analyses using the molecular degree of perturbation to reveal nuances of a unique profile of systemic inflammation and imbalanced immune activation directly linked to disease severity as well as with other clinical and epidemiologic characteristics. Additionally, our findings reveal that the main factor associated with severe cases of *P. vivax* infection was the number of symptoms, despite of a lower global inflammatory perturbation and parasitemia. In these participants, the number of symptoms directly correlated with perturbation of markers of inflammation and tissue damage. On the other hand, the main factor

Review

Systemic Inflammation Associated with Immune Reconstitution Inflammatory Syndrome in Persons Living with HIV

Caian L. Vinhaes [1,2,3,†], Mariana Araujo-Pereira [1,2,4,†], Rafael Tibúrcio [1,2,4], Juan M. Cubillos-Angulo [1,2,4], Fernanda O. Demitto [2], Kevan M. Akrami [1,2,4,5] and Bruno B. Andrade [1,2,3,4,6,*]

1 Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador 40296-710, Brazil; caianleal@gmail.com (C.L.V.); araujopereira.mariana@gmail.com (M.A.-P.); rafael.santos@aluno.bahia.fiocruz.br (R.T.); j.cubillosangulo@gmail.com (J.M.C.-A.); Kevan.akrami@gmail.com (K.M.A.)
2 Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador 40210-320, Brazil; fernandademitto@gmail.com
3 Bahiana School of Medicine and Public Health, Bahia Foundation for the Development of Sciences, Salvador 40290-000, Brazil
4 Faculdade de Medicina, Universidade Federal da Bahia, Salvador 40110-100, Brazil
5 Divisions of Infectious Diseases and Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, University of California, San Diego, La Jolla, CA 92093, USA
6 Curso de Medicina, Centro Universitário Faculdade de Tecnologia e Ciências (UniFTC), Salvador 41741-590, Brazil
* Correspondence: bruno.andrade@fiocruz.br; Tel.: +55-71-3176-2264
† These authors contributed equally to this work.

Citation: Vinhaes, C.L.; Araujo-Pereira, M.; Tibúrcio, R.; Cubillos-Angulo, J.M.; Demitto, F.O.; Akrami, K.M.; Andrade, B.B. Systemic Inflammation Associated with Immune Reconstitution Inflammatory Syndrome in Persons Living with HIV. *Life* 2021, *11*, 65. https://doi.org/10.3390/life11010065

Received: 16 November 2020
Accepted: 14 January 2021
Published: 18 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Abstract: Antiretroviral therapy (ART) has represented a major advancement in the care of people living with HIV (PLWHH), resulting in significant reductions in morbidity and mortality through immune reconstitution and attenuation of homeostatic disruption. Importantly, restoration of immune function in PLWH with opportunistic infections occasionally leads to an intense and uncontrolled cytokine storm following ART initiation known as immune reconstitution inflammatory syndrome (IRIS). IRIS occurrence is associated with the severe and rapid clinical deterioration that results in significant morbidity and mortality. Here, we detail the determinants underlying IRIS development in PLWH, compiling the available knowledge in the field to highlight details of the inflammatory responses in IRIS associated with the most commonly reported opportunistic pathogens. This review also highlights gaps in the understanding of IRIS pathogenesis and summarizes therapeutic strategies that have been used for IRIS.

Keywords: systemic inflammation; mycobacteria; HIV; immune reconstitution inflammatory syndrome (IRIS)

1. Introduction

Globally, nearly 38 million people are living with HIV (PLWH) [1]. The most critical advancement in this epidemic was the development and increased access to antiretroviral therapy (ART), which led to significant reductions in morbimortality through immune reconstitution and attenuation of homeostatic disruption [2]. This has reduced the incidence and severity of opportunistic infections (OI) such as *Mycobacterium tuberculosis* (Mtb) and *Avium complex* (MAC), *Cytomegalovirus* (CMV), Kaposi sarcoma-associated herpesvirus (KSHV), hepatitis C (HCV) and B (HBV) virus, *Cryptococcus neoformans*, *Pneumocystis jirovecci* and *Toxoplasma gondii*. However, paradoxically in a subset of PLWH, ART initiation may trigger clinical worsening with pathologic immune activation against these

IMMUNOTHERAPY

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

PD-1 blockade exacerbates *Mycobacterium tuberculosis* infection in rhesus macaques

Keith D. Kauffman[1], Shunsuke Sakai[1], Nickiana E. Lora[1], Sivaranjani Namasivayam[2], Paul J. Baker[3], Olena Kamenyeva[4], Taylor W. Foreman[1], Christine E. Nelson[1], Deivide Oliveira-de-Souza[5], Caian L. Vinhaes[5], Ziv Yaniv[6], Cecilia S. Lindestam Arleham[7], Alessandro Sette[7,8], Gordon J. Freeman[9], Rashida Moore[10], NIAID/DIR Tuberculosis Imaging Program*, Alan Sher[2], Katrin D. Mayer-Barber[3], Bruno B. Andrade[5], Juraj Kabat[4], Laura E. Via[11]*, Daniel L. Barber[1]†

Boosting immune cell function by targeting the coinhibitory receptor PD-1 may have applications in the treatment of chronic infections. Here, we examine the role of PD-1 during *Mycobacterium tuberculosis* (Mtb) infection of rhesus macaques. Animals treated with anti–PD-1 monoclonal antibody developed worse disease and higher granuloma bacterial loads compared with isotype control–treated monkeys. PD-1 blockade increased the number and functionality of granuloma Mtb-specific CD8 T cells. In contrast, Mtb-specific CD4 T cells in anti–PD-1–treated macaques were not increased in number or function in granulomas, expressed increased levels of CTLA-4, and exhibited reduced intralesional trafficking in live imaging studies. In granulomas of anti–PD-1–treated animals, multiple proinflammatory cytokines were elevated, and more cytokines correlated with bacterial loads, leading to the identification of a role for caspase 1 in the exacerbation of tuberculosis after PD-1 blockade. Last, increased Mtb bacterial loads after PD-1 blockade were found to associate with the composition of the intestinal microbiota before infection in individual macaques. Therefore, PD-1–mediated coinhibition is required for control of Mtb infection in macaques, perhaps because of its role in dampening detrimental inflammation and allowing for normal CD4 T cell responses.

Downloaded from http://immunology.sciencemag.org/ at National Institutes of Health on January 15, 2021

INTRODUCTION

Mycobacterium tuberculosis (Mtb) infection is the leading cause of death due to a single infectious agent worldwide, despite the availability of antibiotics that can effectively treat most Mtb infections (*1*). Drugs that target the host rather than the bacteria, i.e., host-directed therapies (HDTs), may be useful in shortening the standard 6-month-long course of antibiotic treatment, as well as providing sorely needed new options for the treatment of drug resistant infections (*2*–*4*). In particular, there is interest in developing strategies to boost host-protective immune responses or on the other hand limiting the detrimental inflammation that causes tissue destruction and promotes bacterial growth during tuberculosis (TB). However, the mechanisms of host resistance and tissue pathology during Mtb infection are incompletely understood, impeding the development of HDTs.

PD-1 (programmed death-1) is a coinhibitory receptor primarily expressed on activated CD4 and CD8 T cells that has been shown to limit the function of pathogen-specific T cells during chronic infection and tumor-specific T cells during cancer (*5*, *6*). Blockade of PD-1 or its ligands with monoclonal antibodies (mAbs) enhances the number and function of antitumor cytotoxic T cells resulting in enhanced tumor control, and there are multiple PD-1–targeting drugs approved for use against various malignancies (*6*). The major success of immune checkpoint blockade–targeting drugs in cancer treatment has highlighted how potent such approaches can be in the treatment of human disease. Boosting T cell function by blocking PD-1 has been suggested as a therapy for TB (*7*).

Human Mtb-specific T cells in circulation can express low levels of PD-1 during disease, and in vitro blockade of PD-1 can enhance T cell responses, although the effects are modest (*8*). The first in vivo data on the role of PD-1 in Mtb infection came from knockout (KO) mouse studies where it was found that PD-$1^{-/-}$ mice die very rapidly after Mtb infection compared with wild-type (WT) mice (*9*, *10*). In the absence of PD-1, CD4 T cells and, to a lesser extent, CD8 T cells drive this early mortality (*10*). Although the T cell mechanisms that cause pathology in Mtb-infected PD-$1^{-/-}$ mice are not completely understood, we have shown that the overproduction of interferon-γ (IFNγ) by CD4 T cells is at least partly responsible (*11*), and in a human in vitro three-dimensional (3D) granuloma model, it was found that PD-1 blockade drives higher bacterial loads in a tumor necrosis factor (TNF)–dependent manner (*12*). Consistent with these data showing a host-protective role for PD-1 in Mtb infection, clinical case reports of checkpoint blockade–associated TB in patients treated with anti–PD-1 (αPD-1) are accumulating in the literature (*12*–*18*).

[1]T Lymphocyte Biology Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [2]Immunobiology Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [3]Inflammation and Innate Immunity Unit, Laboratory of Clinical Immunology and Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [4]Biological Imaging Section, Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [5]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Intituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil. [6]Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [7]Division of Vaccine Discovery, La Jolla Institute for Immunology, La Jolla, CA, USA. [8]Department of Medicine, University of California, San Diego, La Jolla, CA, USA. [9]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. [10]Comparative Medicine Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [11]Tuberculosis Research Section, Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA.
*The members of the NIAID/DIR Tuberculosis Imaging Program can be found at the end of the Acknowledgments.
†Corresponding author. Email: barberd@niaid.nih.gov

www.nature.com/scientificreports

scientific reports

OPEN

Zika-exposed microcephalic neonates exhibit higher degree of inflammatory imbalance in cerebrospinal fluid

Gustavo C. Nascimento-Carvalho[1✉], Eduardo C. Nascimento-Carvalho[1], Clara L. Ramos[1], Ana-Luisa Vilas-Boas[1], Otávio A. Moreno-Carvalho[2], Caian L. Vinhaes[1,3,4], Beatriz Barreto-Duarte[3,4,5], Artur T. L. Queiroz[3,4], Bruno B. Andrade[1,3,4,5,6,7,8,10] & Cristiana M. Nascimento-Carvalho[9,10]

Not every neonate with congenital Zika virus (ZIKV) infection (CZI) is born with microcephaly. We compared inflammation mediators in CSF (cerebrospinal fluid obtained from lumbar puncture) between ZIKV-exposed neonates with/without microcephaly (cases) and controls. In Brazil, in the same laboratory, we identified 14 ZIKV-exposed neonates during the ZIKV epidemic (2015–2016), 7(50%) with and 7(50%) without microcephaly, without any other congenital infection, and 14 neonates (2017–2018) eligible to be controls and to match cases. 29 inflammation mediators were measured using Luminex immunoassay and multidimensional analyses were employed. Neonates with ZIKV-associated microcephaly presented substantially higher degree of inflammatory perturbation, associated with uncoupled inflammatory response and decreased correlations between concentrations of inflammatory biomarkers. The groups of microcephalic and non-microcephalic ZIKV-exposed neonates were distinguished from the control group (area under curve [AUC] = 1; P < 0.0001). Between controls and those non-microcephalic exposed to ZIKV, IL-1β, IL-3, IL-4, IL-7 and EOTAXIN were the top CSF markers. By comparing the microcephalic cases with controls, the top discriminant scores were for IL-1β, IL-3, EOTAXIN and IL-12p70. The degree of inflammatory imbalance may be associated with microcephaly in CZI and it may aid additional investigations in experimental pre-clinical models testing immune modulators in preventing extensive damage of the Central Nervous System.

A widespread epidemic of Zika virus (ZIKV) infection was reported in South America in 2015[1]. In 2016, a causal relationship between ZIKV infection during pregnancy and microcephaly was strongly suspected[2], when a serious epidemic sparked a major concern given the hundreds of microcephalic neonates born in Brazil[3]. Subsequently, several studies have shown that ZIKV has high Central Nervous System (CNS) tropism and is harmful when the CNS is immature[4].

In 2017, microcephaly was revealed on postmortem examination of 7 neonates with congenital ZIKV infection (CZI)[5]. Leptomeningeal and cerebral parenchymal inflammation was described with varying intensity and distribution[5]. In 2018, an immunohistochemical analysis of neural parenchyma tissues from 8 deceased neonate/stillbirth babies, 4 with and 4 without ZIKV infection, showed significantly higher expression of several cytokines

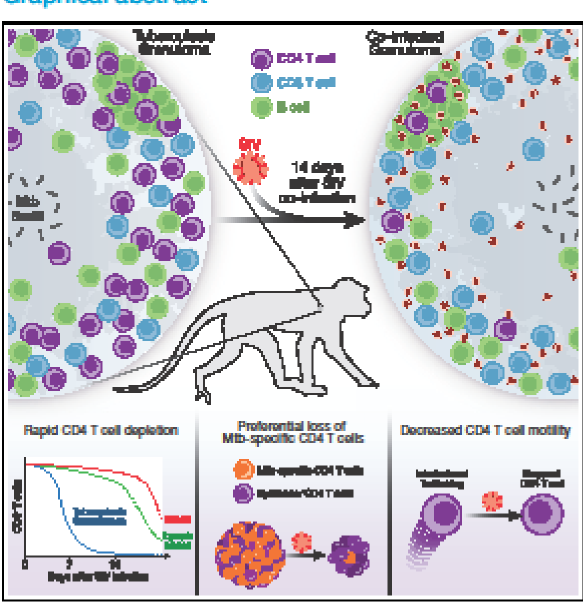
[1]Bahiana Foundation for Science Development, Bahiana School of Medicine, Salvador, Bahia 40290-000, Brazil. [2]Cerebrospinal Fluid Laboratory, José Silveira Foundation, Salvador, Bahia 40170-100, Brazil. [3]Gonçalo Moniz Institute, Oswaldo Cruz Foundation, Salvador, Bahia 40296-710, Brazil. [4]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Bahia 40296-710, Brazil. [5]University Salvador (UNIFACS), Laureate Universities, Salvador, Bahia 41820-021, Brazil. [6]School of Medicine, Faculdade de Tecnologia e Ciências (Uni-FTC), Salvador, Bahia 41741-590, Brazil. [7]Wellcome Centre for Infectious Disease Research in Africa, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town 7700, South Africa. [8]Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN 37232, USA. [9]Department of Pediatrics, Federal University of Bahia School of Medicine, Salvador, Bahia 40210-630, Brazil. [10]These authors contributed equally: Bruno B. Andrade and Cristiana M. Nascimento-Carvalho. ✉email: gcn.carvalho@hotmail.com

Cell Reports

Article

CD4 T cells are rapidly depleted from tuberculosis granulomas following acute SIV co-infection

Graphical abstract



Highlights

- SIV rapidly replicates in Mtb granulomas of co-infected macaques
- Granuloma CD4 T cells are depleted before those in blood, BAL, LNs, or spleen
- $CCR5^+Eomes^-$ Mtb-specific Th1 and Th1* cells in granulomas are preferentially depleted
- SIV co-infection reduces motility of CD4 T cells in granulomas

Authors

Taylor W. Foreman, Christine E. Nelson, Keith D. Kauffman, ..., Juraj Kabat, Laura E. Via, Daniel L. Barber

Correspondence

barberd@niaid.nih.gov

In brief

HIV-mediated destruction of CD4 T cells enhances susceptibility to *Mycobacterium tuberculosis*. Using macaques, Foreman et al. show that CD4 T cells in granulomas are depleted very rapidly after SIV co-infection, indicating that loss of immunity at the site of bacterial replication occurs long before signs of peripheral T cell depletion.

Foreman et al., 2022, Cell Reports 39, 110896
May 31, 2022
https://doi.org/10.1016/j.celrep.2022.110896

frontiers | Frontiers in Immunology

ORIGINAL RESEARCH
published: 23 June 2022
doi: 10.3389/fimmu.2022.916216

Relationship Between Anemia and Systemic Inflammation in People Living With HIV and Tuberculosis: A Sub-Analysis of the CADIRIS Clinical Trial

Mariana Araújo-Pereira[1,2,3], Beatriz Barreto-Duarte[1,2,4,5], María B. Arriaga[1,2,6], Laura W. Musselwhite[7], Caian L. Vinhaes[1,2,8], Pablo F. Belaunzaran-Zamudio[9], Adam Rupert[10], Luis J. Montaner[11], Michael M. Lederman[12], Irini Sereti[10], Juan G. Sierra Madero[9]* and Bruno B. Andrade[1,2,3,4,5,8]*

[1] Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil, [2] Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Brazil, [3] Programa de Pós-Graduação em Patologia Humana e Experimental, Universidade Federal da Bahia, Salvador, Brazil, [4] Programa de Pós-Graduação em Clínica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, [5] Curso de Medicina, Universidade Salvador (UNIFACS), Salvador, Brazil, [6] Instituto de Medicina Tropical Alexander Von Humboldt, Universidad Peruana Cayetano Heredia, Lima, Peru, [7] Department of Solid Tumor Oncology, Levine Cancer Institute, Charlotte, NC, United States, [8] Bahiana School of Medicine and Public Health, Bahia Foundation for the Development of Sciences, Salvador, Brazil, [9] Infectious Diseases Department, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico, [10] National Institutes of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, United States, [11] The Wistar Institute, Philadelphia, PA, United States, [12] Division of Infectious Diseases and HIV Medicine, Department of Medicine, Case Western Reserve University, Cleveland, OH, United States

OPEN ACCESS

Edited by:
Dawit Wolday,
Mekelle University, Ethiopia

Reviewed by:
Wondwossen Amogne Degu,
Addis Ababa University, Ethiopia
Theresa Marie Rossouw,
University of Pretoria, South Africa

***Correspondence:**
Bruno B. Andrade
bruno.andrade@fiocruz.br
Juan G. Sierra-Madero
jsmadero@yahoo.com

Specialty section:
This article was submitted to Inflammation,
a section of the journal
Frontiers in Immunology

Received: 08 April 2022
Accepted: 30 May 2022
Published: 23 June 2022

Citation:
Araújo-Pereira M, Barreto-Duarte B, Arriaga MB, Musselwhite LW, Vinhaes CL, Belaunzaran-Zamudio PF, Rupert A, Montaner LJ, Lederman MM, Sereti I, Madero JGS and Andrade BB (2022) Relationship Between Anemia and Systemic Inflammation in People Living With HIV and Tuberculosis: A Sub-Analysis of the CADIRIS Clinical Trial.
Front. Immunol. 13:916216.
doi: 10.3389/fimmu.2022.916216

People with HIV (PWH) are at increased risk of developing active tuberculosis (TB), and anemia is a common complication in both conditions. Anemia in TB patients has been linked to immune activation, levels of inflammatory biomarkers in blood, and risk for HIV disease progression and death. In this study we show that anemia was associated with a more pronounced inflammatory profile in HIV-TB coinfected persons in a cohort of 159 individuals with advanced HIV disease (CD4 count < 100 cells/μL) recruited as part of a randomized clinical trial (NCT00988780). A panel of plasma biomarkers was assessed on plasma obtained prior to combination antiretroviral therapy (cART) initiation. We performed a series of multidimensional analyses including clinical variables and concentrations of inflammatory biomarkers to profile systemic inflammation of PWH with and without anemia. We observed that TB participants presented with moderately lower levels of hemoglobin than non-TB participants. These participants also presented a higher Degree of Inflammatory Perturbation (DIP) score, related to increased levels of IFN-γ and TNF. The DIP was associated with TB coinfection and anemia before cART initiation. Future mechanistic studies are warranted to assess the determinants of such associations and the implications on treatment outcomes.

Keywords: HIV, Tuberculosis, inflammation, degree of inflammatory perturbation, anemia

Articles

Association between severe anaemia and inflammation, risk of IRIS and death in persons with HIV: A multinational cohort study

Mariana Araújo-Pereira,[a,b,c,m] Virginia Sheikh,[d,m] Irini Sereti,[d,m] Beatriz Barreto-Duarte,[a,b,f,g] María B. Arriaga,[a,b,c] Rafael Tibúrcio,[a,b,c] Caian L. Vinhaes,[a,b,h] Manuella Pinto-de-Almeida,[b,g] Jing Wang,[e] Adam Rupert,[d] Gregg Roby,[d] Douglas Shaffer,[i,j] Jintanat Ananworanich,[k] Nittaya Phanuphak,[l,n] Fred Sawe,[i,n] and Bruno B. Andrade[a,b,c,f,g,m,*]

[a]Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil
[b]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Brazil
[c]Programa de Pós-Graduação em Patologia Humana e Experimental, Universidade Federal da Bahia, Salvador, Bahia, Brazil
[d]National Institutes of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA
[e]Clinical Monitoring Research Program Directorate, Frederick National Laboratory for Cancer Research, Frederick, MD, USA
[f]Programa de Pós-Graduação em Clínica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[g]Curso de Medicina, Universidade Salvador (UNIFACS), Salvador, Bahia, Brazil
[h]Bahiana School of Medicine and Public Health, Bahia Foundation for the Development of Sciences, Salvador, Brazil
[i]Kenya Medical Research Institute, Henry Jackson Foundation Medical Research International, Bethesda, MD, USA
[j]Walter Reed Army Institute of Research/US Army Medical Research Directorate-Africa, Nairobi, Kenya
[k]South East Asia Research Collaboration with Hawaii, Henry M. Jackson Foundation for the Advancement of Military Medicine, United States Military HIV Research Program, Bethesda, MD, USA
[l]SEARCH, Institute of HIV Research and Innovation, Bangkok, Thailand

eBioMedicine
2022;85: 104309
Available online xxx
https://doi.org/10.1016/j.ebiom.2022.104309

Summary

Background After initiating antiretroviral therapy (ART), approximately 25% of people with HIV (PWH) may develop Immune Reconstitution Inflammatory Syndrome (IRIS), which is associated with increased morbidity and mortality. Several reports have demonstrated that low haemoglobin (Hb) levels are a risk factor for IRIS. To what extent the severity of anaemia contributes to the risk of IRIS and/or death is still insufficiently explored.

Methods We investigated both the presence and severity of anaemia in PWH in a multinational cohort of ART-naïve patients. A large panel of plasma biomarkers was measured pre-ART and patients were followed up for 6 months. IRIS or deaths during this period were considered as outcomes. We performed multidimensional analyses, logistic regression, and survival curves to delineate associations.

Findings Patients with severe anaemia (SA) presented a distinct systemic inflammatory profile, characterized by higher TNF, IL-6, and IL-27 levels. SA was independently associated with IRIS, with a higher risk of both early IRIS onset and death. Among IRIS patients, those with SA had a higher risk of mycobacterial IRIS.

Interpretation PWH with SA display a more pronounced inflammatory profile, with an elevated risk of developing IRIS earlier and a statistically significant higher risk of death.

Funding Intramural Research Program of National Institute of Allergy and Infectious Diseases/National Institutes of Health (NIAID/NIH). Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Finance code: 001) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil.

Copyright © 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Tuberculosis; Systemic inflammation; IRIS; Death; HIV

*Corresponding author. Laboratório de Inflamação e Biomarcadores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Rua Waldemar Falcão, 121, Candeal, Salvador, Bahia 40296-710, Brazil.
E-mail address: bruno.andrade@fiocruz.br (B.B. Andrade).
[m]M.A-P., V.S., and I.S. contributed equally.
[n]N.P., F.S., and B.B.A. contributed equally.

TYPE Specialty Grand Challenge
PUBLISHED 21 October 2022
DOI 10.3389/fitd.2022.1037913

OPEN ACCESS

EDITED BY
Emanuele Nicastri,
National Institute for Infectious Diseases Lazzaro Spallanzani (IRCCS), Italy

REVIEWED BY
Maria Teresa García-Romero,
National Institute of Pediatrics, Mexico

*CORRESPONDENCE
Bruno B. Andrade
bruno.andrade@fiocruz.br

SPECIALTY SECTION
This article was submitted to Major Tropical Diseases, a section of the journal Frontiers in Tropical Diseases

RECEIVED 06 September 2022
ACCEPTED 03 October 2022
PUBLISHED 21 October 2022

CITATION
Barreto-Duarte B, Araujo-Pereira M, Miguez-Pinto JP, Ferreira IBB, Menezes RC, Rosier GL, Vinhaes CL, Maggitti-Bezerril M, Villalva-Serra K and Andrade BB (2022) Grand challenges in major tropical diseases. *Front. Trop. Dis.* 3:1037913. doi: 10.3389/fitd.2022.1037913

COPYRIGHT
© 2022 Barreto-Duarte, Araújo-Pereira, Miguez-Pinto, Ferreira, Menezes, Rosier, Vinhaes, Maggitti-Bezerril, Villalva-Serra and Andrade. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Grand challenges in major tropical diseases

Beatriz Barreto-Duarte[1,2,3,4,5], Mariana Araújo-Pereira[3,4,5,6], João P. Miguez-Pinto[1,3], Isabella B. B. Ferreira[3,7], Rodrigo C. Menezes[3,6], Gabriela L. Rosier[3,4], Caian L. Vinhaes[3,4,5,7], Mateus Maggitti-Bezerril[3], Klauss Villalva-Serra[1,3] and Bruno B. Andrade[1,2,3,4,5,6,7]*

[1]Curso de Medicina, Universidade Salvador (UNIFACS), Salvador, Brazil, [2]Programa Pós-graduação de Clínica Médica, Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, [3]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Brazil, [4]Instituto de Pesquisa Clínica e Translacional (IPCT), Centro Universitário Faculdade de Tecnologia e Ciências (UniFTC), Salvador, Brazil, [5]Laboratório de Inflamação e Biomarcadores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil, [6]Faculdade de Medicina, Universidade Federal da Bahia, Salvador, Brazil, [7]Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador, Brazil

KEYWORDS
tropical disease, tuberculosis, malaria, HIV, grand challenges

Overview

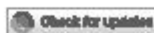
Approximately 15 million people die each year due to tropical diseases, which are caused by a variety of infectious agents, such as bacteria, viruses, parasites, or protozoa. Such diseases usually are a result of an intricate relationship between poverty, poor living conditions, malnutrition, and poor healthcare system infrastructure, affecting a large proportion of developing and underdeveloped countries (1–3). Therefore, these diseases are not restricted to infections that are uniquely reported in tropical regions, but also include illnesses that exhibit a very high burden in such zones of the globe.

Despite the great progress in governmental and private initiatives to improve measures to prevent and treat these infections, they remain the world's leading cause of premature death, highlighting the magnitude of this public health problem (4–6). Herein, we summarize top priorities that pose grand challenges to dampen the burden of major tropical diseases worldwide. We will especially focus on the big three pathogens, which are the human immunodeficiency virus (HIV), *Mycobacterium tuberculosis* (*Mtb*), and *Plasmodium* sp. although other important conditions are highlighted. Aspects of disease burden, diagnosis, treatment, and prophylaxis are discussed to define the scope of interest of the Section *Major Tropical Diseases* of the *Frontiers in Tropical Diseases* journal.

HIV

HIV infection is a major global public health problem, as it has been associated with a total of approximately 40.1 million deaths. Moreover, it is estimated that there are 38.4 million people living with HIV (PLWH), and at the end of 2021, 650 thousand people had died from HIV-related causes (7).

TYPE Review
PUBLISHED 23 September 2022
DOI 10.3389/fmed.2022.970408



OPEN ACCESS

EDITED BY
Mohammad Asif Sherwani,
University of Alabama at Birmingham,
United States

REVIEWED BY
Shashank Gupta,
Division of Intramural Research,
National Heart, Lung, and Blood
Institute (NIH), United States
Faraz Ahmad,
University of Missouri, United States

*CORRESPONDENCE
Bruno B. Andrade
bruno.andrade@fiocruz.br

†These authors have contributed
equally to this work

SPECIALTY SECTION
This article was submitted to
Infectious Diseases – Surveillance,
Prevention, and Treatment,
a section of the journal
Frontiers in Medicine

RECEIVED 15 June 2022
ACCEPTED 22 August 2022
PUBLISHED 23 September 2022

CITATION
Cubillos-Angulo JM, Nogueira BMF,
Arriaga MB, Barreto-Duarte B,
Araújo-Pereira M, Fernandes CD,
Vinhaes CL, Villalva-Serra K, Nunes VM,
Miguez-Pinto JP, Amaral EP and
Andrade BB (2022) Host-directed
therapies in pulmonary tuberculosis:
Updates on anti-inflammatory drugs.
Front. Med. 9:970408.
doi: 10.3389/fmed.2022.970408

COPYRIGHT
© 2022 Cubillos-Angulo, Nogueira,
Arriaga, Barreto-Duarte,
Araújo-Pereira, Fernandes, Vinhaes,
Villalva-Serra, Nunes, Miguez-Pinto,
Amaral and Andrade. This is an
open-access article distributed under
the terms of the Creative Commons
Attribution License (CC BY). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Host-directed therapies in pulmonary tuberculosis: Updates on anti-inflammatory drugs

Juan M. Cubillos-Angulo[1,2,3†], Betânia M. F. Nogueira[2,3†],
María B. Arriaga[1,2,3], Beatriz Barreto-Duarte[1,3,4,5],
Mariana Araújo-Pereira[1,2,3], Catarina D. Fernandes[3],
Caian L. Vinhaes[1,3,6], Klauss Villalva-Serra[1,2,3,4],
Vanessa M. Nunes[4], João P. Miguez-Pinto[4],
Eduardo P. Amaral[7] and Bruno B. Andrade[1,2,3,5,6*]

[1]Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil, [2]Faculdade de Medicina, Universidade Federal da Bahia, Salvador, BA, Brazil, [3]Multinational Organization Network Sponsoring Translational and Epidemiological Research Initiative, Salvador, Brazil, [4]Curso de Medicina, Universidade Salvador, Salvador, Brazil, [5]Programa de Pós-Graduação em Clínica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, [6]Bahiana School of Medicine and Public Health, Bahia Foundation for the Development of Sciences, Salvador, Brazil, [7]Immunobiology Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, United States

Tuberculosis (TB) is a lethal disease and remains one of the top ten causes of mortality by an infectious disease worldwide. It can also result in significant morbidity related to persistent inflammation and tissue damage. Pulmonary TB treatment depends on the prolonged use of multiple drugs ranging from 6 months for drug-susceptible TB to 6–20 months in cases of multi-drug resistant disease, with limited patient tolerance resulting from side effects. Treatment success rates remain low and thus represent a barrier to TB control. Adjunct host-directed therapy (HDT) is an emerging strategy in TB treatment that aims to target the host immune response to *Mycobacterium tuberculosis* in addition to antimycobacterial drugs. Combined multi-drug treatment with HDT could potentially result in more effective therapies by shortening treatment duration, improving cure success rates and reducing residual tissue damage. This review explores the rationale and challenges to the development and implementation of HDTs through a succinct report of the medications that have completed or are currently being evaluated in ongoing clinical trials.

KEYWORDS
host-directed therapy, *Mycobacterium*, tuberculosis, adjunct therapy, immunotherapies

Anexo II: Rede de colaboração do estudante

Durante o período de Iniciação Científica e Doutoramento, através do seu grupo de pesquisa, o MONSTER, o estudante pode participar de inúmeras colaborações com pesquisadores de diversos centros ao redor do mundo, como o National Institute oh Health (NIH), University of Cape Town, Organização Mundial de Saúde, University of California, Vanderbilt University, University of San Diego entre outras.