Immunological Reviews   WILEY

Transcriptomics for child and adolescent tuberculosis*

Myrsini Kaforou[1] | **Claire Broderick[1]** | **Ortensia Vito[1]** | **Michael Levin[1]** | **Thomas J. Scriba[2]** | **James A. Seddon[1,3]**

[1]Department of Infectious Disease, Imperial College London, London, UK

[2]South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, Cape Town, South Africa

[3]Desmond Tutu TB Centre, Department of Paediatrics and Child Health, Stellenbosch University, Cape Town, South Africa

Correspondence
Myrsini Kaforou, Department of Infectious Disease, Imperial College London, Norfolk Place, London W2 1PG, UK.
Email: m.kaforou@imperial.ac.uk

Summary

Tuberculosis (TB) in humans is caused by *Mycobacterium tuberculosis* (*Mtb*). It is estimated that 70 million children (<15 years) are currently infected with *Mtb*, with 1.2 million each year progressing to disease. Of these, a quarter die. The risk of progression from *Mtb* infection to disease and from disease to death is dependent on multiple pathogen and host factors. Age is a central component in all these transitions. The natural history of TB in children and adolescents is different to adults, leading to unique challenges in the development of diagnostics, therapeutics, and vaccines. The quantification of RNA transcripts in specific cells or in the peripheral blood, using high-throughput methods, such as microarray analysis or RNA-Sequencing, can shed light into the host immune response to *Mtb* during infection and disease, as well as understanding treatment response, disease severity, and vaccination, in a global hypothesis-free manner. Additionally, gene expression profiling can be used for biomarker discovery, to diagnose disease, predict future disease progression and to monitor response to treatment. Here, we review the role of transcriptomics in children and adolescents, focused mainly on work done in blood, to understand disease biology, and to discriminate disease states to assist clinical decision-making. In recent years, studies with a specific pediatric and adolescent focus have identified blood gene expression markers with diagnostic or prognostic potential that meet or exceed the current sensitivity and specificity targets for diagnostic tools. Diagnostic and prognostic gene expression signatures identified through high-throughput methods are currently being translated into diagnostic tests.

KEYWORDS
children, diagnosis, differential expression, transcriptomics, tuberculosis

1 | INTRODUCTION

Tuberculosis (TB) is a disease caused by *Mycobacterium tuberculosis* (*Mtb*) and the organism has evolved with humans for >10 000 years.[1] Individuals with infectious TB generate droplet nuclei containing *Mtb* which, when coughed or breathed out, can remain in the air for several hours. Following inhalation of infected droplets, the bacilli need to overcome physical barriers and antimicrobial peptides, to reach the terminal alveoli, where they encounter elements of the innate immune system, most commonly macrophages and dendritic

---

This article is part of a series of reviews covering SARS-CoV-2 Immunity appearing in Volume 309 of *Immunological Reviews*.

cells.[2] If *Mtb* survives this encounter with the innate system and sensitizes the adaptive immune system, as measured by tuberculin skin testing (TST) or interferon gamma release assays (IGRAs), the individual is said to have *Mtb* infection (sometimes termed TB infection or latent TB infection). Commonly, the mycobacteria are contained by the immune system with only low numbers of organisms persisting. However, if the bacilli overcome these constraints and multiply, symptoms and signs of TB disease develop, accompanied by radiological changes in the lungs or other sites of disease. It may be possible to isolate mycobacteria from respiratory samples or samples taken from other sites of disease that can be cultured or identified using molecular tests. Overall, about 10% of individuals with *Mtb* infection will progress to TB disease. Drug therapy given to individuals with *Mtb* infection is effective at preventing this progression and is termed TB preventive therapy (TPT or latent TB treatment). If an individual develops TB disease, then TB disease treatment can be given, which is again successful in most patients.

The natural history of TB in children and adolescents is different to adults. The risk of *Mtb* infection rises with age in a relatively linear way dependent on the prevailing TB prevalence in that context, reflecting cumulative exposure. However, the risk of progressing from *Mtb* infection to TB disease varies with age.[3] Young children (<5 years of age) are at high risk of disease progression, with the risk falling to a nadir in primary school age children.[4] This risk rises as children enter puberty, increasing earlier in females but with males then following and eventually overtaking females in adulthood. Work done in the era prior to antibiotics suggests that there is a "timetable" for TB in children, with almost all children who progress to disease doing so within a year or two of infection. In addition, the type of disease seen in children is different to adults. Young children typically have paucibacillary disease, a term implying that few or no organisms are commonly found in respiratory samples that undergo microbiological evaluation. TB disease in this age group usually presents as either intrathoracic lymph node disease or disseminated disease, including TB meningitis or miliary TB.[5] As children enter adolescence, they begin to develop adult-type disease with extensive parenchymal involvement and cavities. Large numbers of organisms are commonly isolated from respiratory samples that undergo microbiological evaluation.

The World Health Organization (WHO) defines children as <15 years and adolescents as 10 to <20 years. It is estimated that currently 70 million children have *Mtb* infection and each year 1.2 million develop TB disease.[6] It is estimated that an additional 535 000 15 to <20-year-old develop disease each year.[7] Only about half of the estimated number of incident cases of child TB each year are diagnosed and treated, and WHO suggests that 230 000 children die of TB annually,[6] with 96% of these being undiagnosed.[8] For many years, global and most national guidance has been that children <5 years and those living with HIV who have been exposed to an infectious case of TB should receive TPT. Guidance has been to treat following exposure given that tests of *Mtb* infection are rarely available in high TB-burden, low-resource settings. However, of the 1.27 million children estimated to be eligible for TPT in 2017, only 23% received treatment.[9] The most recent WHO guidance expands TPT provision to also say that all household contacts of infectious TB patients can be given TPT following exclusion of TB disease.[10] Very few of these individuals receive TPT each year. The WHO End TB Strategy seeks to reduce TB deaths by 95% by 2035, compared with 2015 levels, as well as reduce incidence by 90% over the same period, with children and adolescent comprehensively included in the Strategy. However, it is recognized that using current approaches, global progress will fall far short of these targets.[11] The COVID-19 pandemic has severely disrupted health services and TB programs. This disruption is almost certainly a factor behind the observed 25–50% fall in the detection and treatment of new TB cases which was observed in just a 3-month period in 2020.[6] Though the impact of COVID-19 control measures on *Mtb* transmission has not yet been defined, it is predicted that reduced case-finding and treatment during the pandemic will lead to increased TB mortality.[12–15]
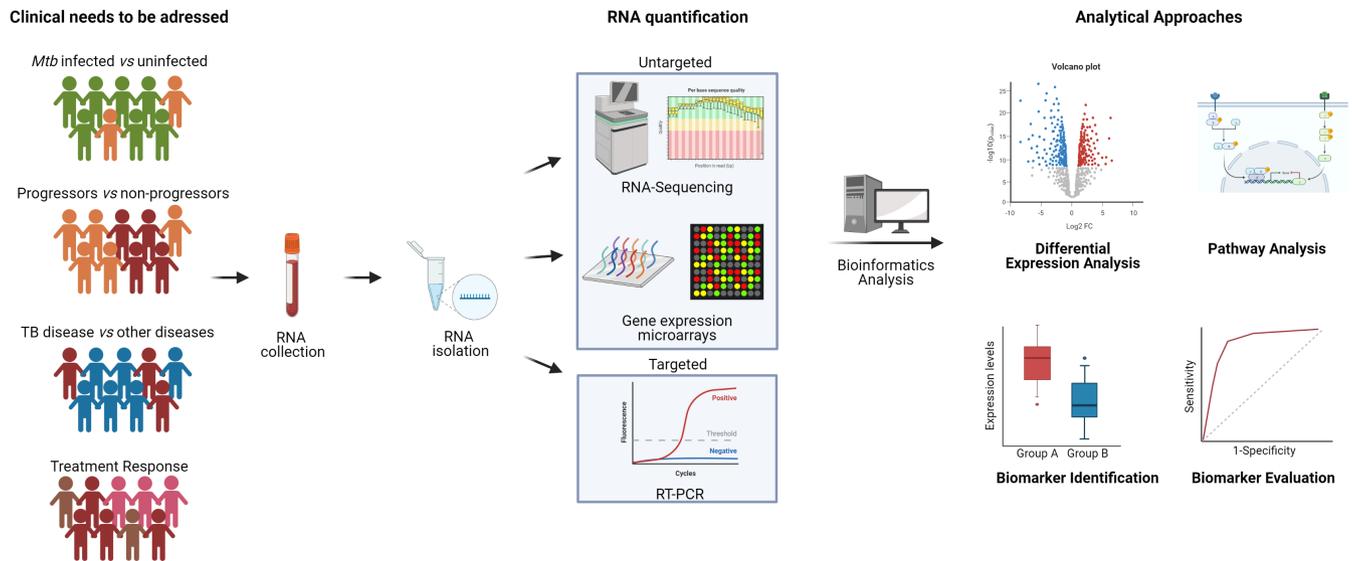
The End TB Strategy suggests that to achieve their ambitious targets new tools will be required. These include point-of-care diagnostic tests for TB disease, new tests to identify which individuals will progress to disease in the future, as well as new vaccines and new drugs. To develop these tools, it is increasingly recognized that a more complete understanding of the immune response to *Mtb* is required, as well as a better understanding of how to use host responses to discriminate between clinical groups. One area of host response biology that has evolved substantially over the last 10 years is transcriptomics or the study of RNA expression. In this article, we will discuss the role of transcriptomics in child and adolescent TB, with a particular focus on transcriptomics in blood. We review the literature on studies that have employed transcriptomics to both better understand child and adolescent TB as well as develop diagnostic tests.

2 | CHALLENGES IN CHILD AND ADOLESCENT TB

There are multiple challenges facing the field of childhood TB, many of which could benefit from the application of transcriptomic approaches (Figure 1). Some of these challenges relate to decision-making around the clinical management of children, some to the development of new therapies or vaccines, and some to the way that clinical research in child TB is conducted. Below we first outline these challenges and then we systematically describe studies that have been done in this field.

2.1 | Discriminating children with TB from children with other diseases

Children with TB disease generally have non-specific symptoms and signs that overlap with other common conditions seen in childhood. Radiological investigations, the most used being chest X-ray, are also frequently non-specific with abnormalities that could be consistent

Clinical needs to be adressed

RNA quantification

Analytical Approaches

Mtb infected *vs* uninfected

Progressors *vs* non-progressors

TB disease *vs* other diseases

Treatment Response

RNA collection

RNA isolation

Untargeted

RNA-Sequencing

Gene expression microarrays

Targeted

RT-PCR

Bioinformatics Analysis

Volcano plot

Differential Expression Analysis

Pathway Analysis

Biomarker Identification

Biomarker Evaluation

FIGURE 1 Overview of the role of transcriptomics in pediatric and adolescent TB, together with steps required for RNA quantification and bioinformatics analysis. Created with BioRender.com

with TB but also with other conditions. While TB disease in adults is usually diagnosed through the identification of *Mtb* in the sputum, this type of diagnostic confirmation is uncommon in children. There are two reasons for this. The first is that samples for evaluation can be challenging to obtain. As young children are unable to spontaneously expectorate sputum, more invasive sampling is required, including induced sputum (requiring nebulized hypertonic saline to stimulate coughing followed by aspiration of sputum from the pharynx through the nose), gastric aspiration (the suctioning of stomach contents using a nasogastric tube to collect sputum coughed and swallowed during sleeping), and the collection of stool samples (to identify swallowed *Mtb* that has passed through the digestive system). The second is that even if good quality respiratory samples are collected, microbiological evaluation using culture or molecular diagnostic testing only identifies *Mtb* in a relatively low proportion of children determined clinically to have TB disease (commonly about 20%).[16] It is assumed that this is due to the presence of few organisms in respiratory samples. Given these challenges, new diagnostic tests are required, ones that do not rely on the microbiological evaluation of respiratory samples.

Of the children who present to health facilities for evaluation of their clinical symptoms and signs, some will have TB, but many will not. Established symptoms required to classify a child as a presumptive TB case include any of the following[17,18]: (a) cough ≥2 weeks, (b) persistent, unexplained lethargy, (c) unexplained fever ≥1 week, (d) poor growth/weight loss over the preceding 3 months, or (e) cough <1 week with a known TB exposure in the previous 12 months. A positive TST or chest X-ray suggestive of TB are also criteria. A test that can discriminate presumptive TB cases who have TB from presumptive TB cases who have other causes for their symptoms would make a substantial impact on the vast under-diagnosis of child TB that is seen on a global scale. In turn, this could impact dramatically on child TB mortality.

In addition to children being brought to healthcare services for evaluation (termed passive case finding), active case finding strategies seek to identify new, undiagnosed cases amongst high-risk populations. Children with HIV, those with malnutrition and those with any degree of immunosuppression should be considered at high risk. It is recommended that these children are regularly screened for TB disease. A further high-risk group is children recently exposed to infectious cases of TB, given the high proportion with infection and the substantial risk of disease progression in young children with recent exposure. Global guidance and almost all national guidelines suggest that after a new infectious adult TB case is diagnosed, the house should be visited, and all household members evaluated for TB disease. Systematic reviews of the yield of household contact tracing suggest that between 5 and 10% of children screened at home visits have prevalent TB disease.[19,20] Deciding which children have TB disease, which have other diseases that are causing symptoms, and which are well, can be challenging as again, symptoms, signs, and radiology are non-specific. A test that could assist in identifying those children who need TB disease treatment would be very beneficial.

2.2 | Predicting disease progression in TB-exposed children and adolescents

Following exposure to an infectious case of TB, and following exclusion of TB disease, TPT is advised for young children and children living with HIV.[10] However, even though these children are at increased risk of developing TB compared with HIV-negative older children and adults, most of these "high-risk" children will not progress to TB disease. The only tests that are currently available that assist in decision-making are the tests of *Mtb* infection, namely the TST and IGRA. These tests signify immunological sensitization by

detecting *Mtb*-specific T-cell responses, and do not indicate if there are either viable bacilli in the child or if there is a high risk of future disease progression. TB-exposed children with a positive TST or IGRA are at higher risk of disease progression than children with negative tests but still only a small proportion of those with positive tests will progress to disease.[21] Children under 5 years with a positive IGRA/TST have a 2-year incidence of ~20% while in children over 5 years this risk is ~10%. Overall, for all TB-exposed children (irrespective of infection status), the risk is substantially below 10%. Increasingly there is recognition that older HIV-negative children and adults should also be given TPT following household exposure.[10] For these individuals, the risk of future disease progression is even lower.

This means that many well children, adolescents, and adults need to be given treatment to prevent each TB case. TPT involves several months of medication which can be challenging for children and families and puts additional strain on health systems. TPT is also seen as a low priority by health services and families and treatment completion rates for children are very low.[22] Although rare, adverse events do occur, there are non-specific effects on the microbiome, and any unnecessary pill burden is ideally avoided. A biomarker that could identify which TB-exposed children and adolescents are at high risk of disease progression would allow targeting of interventions to those most needing them.

Increasingly the concept of *Mtb* infection and TB disease being dichotomous disease states is being challenged and a dynamic continuum recognised.[23] Under the most commonly accepted model of this continuum,[24] those with incipient TB have detectable metabolic activity of *Mtb*, but without any clinical symptoms or signs, radiological abnormalities or positive microbiology for *Mtb*. Those with sub-clinical disease do not have clinical symptoms or signs but may have radiological changes or it may be possible to isolate *Mtb* from respiratory samples. Individuals with incipient or sub-clinical disease are more likely to progress to clinically apparent disease and so any biomarker that could identify these clinical states, allowing appropriate treatment, would be valuable.

2.3 | Identifying children and adolescents who are not responding to TB disease treatment

Treatment outcomes for children diagnosed and treated for TB are generally good with low rates of mortality.[25] However, there are several groups of children in whom the proportion with unfavorable outcome is higher. These include children with certain forms of TB, most notably TB meningitis,[26] and children with other conditions complicating their TB treatment, such as HIV co-infection or malnutrition. Also, if a child is treated for drug-susceptible TB, based on clinical criteria, while having drug-resistant TB, then treatment is unlikely to be effective. Children with TB who are not given their TB drugs, or who do not take their drugs, do not do well and if a child is diagnosed with TB based on clinical criteria and fails to respond, one potential reason is that they do not have TB but have another cause

for their symptoms, signs, and radiology. In contrast to younger children, treatment outcomes for adolescents are less good,[27] with high rates of treatment failure and death. In all these instances, it would be useful identify that a child or adolescent is not responding to treatment as early as possible to allow appropriate interventions. In clinical practice, it often takes several months to identify that a child or adolescent is not responding well to treatment. Failure to put on weight, persistence or worsening of symptoms, and worsening radiology all indicate treatment failure, but these take time to detect and are not specific. For adult TB, where most patients are microbiologically confirmed at baseline, sputum smear, or culture conversion at 2 months is used as a surrogate marker.[28] While this marker is associated with favorable outcome, it is not a sensitive or specific indicator and 2 months is late to be identifying a patient failing therapy. Although this 2-month microbiological conversion can be used for individuals who were microbiologically confirmed at baseline, this represents a small proportion of children treated for TB.

If it were possible early in therapy to identify children and adolescents who were not responding to treatment, then that individual could be evaluated thoroughly. The patient could be counselled intensively and supported to take their treatment if poor adherence was found to be a problem. Management of co-morbidities could be enhanced. Samples could be taken to evaluate for drug resistance and further investigations carried out to look for other diagnoses. Ultimately, it may be possible to use a change in biomarker status after several days to support the diagnosis of TB in those clinically diagnosed. It might be possible to conclude that those without any change in TB-specific biomarkers might not have had TB at baseline or have drug-resistant disease.

2.4 | Tailoring therapy to disease severity and treatment response

In many areas of medicine, personalized precision therapy is becoming more common with treatment targeted to host genotype, disease type, disease site, disease severity, and response to treatment. Yet for programmatic reasons, almost all TB cases are given the same combination of drugs, at the same dosages and for the same duration.[18] Most children with TB do not need 6 months of therapy and some might be successfully treated with substantially shorter durations. Early TB trials in adults demonstrated that although most patients with sputum smear-negative TB were cured after even 2 months, an unacceptably high proportion relapsed.[29] As it was not possible in those early studies to predict which patients might relapse, it was felt preferable to treat all patients with the minimum duration to achieve relapse free cure in >95%. This ultimately means overtreating most individuals and almost all children.

Increasingly there is recognition that different patients with different forms of TB may be appropriately treated with different drug combinations, dosages, or durations. A recently completed phase 3 clinical trial called SHINE (Shorter Treatment for Minimal Tuberculosis in Children) recruited children with minimal

TB and randomized them to either the conventional 6 months of treatment or a new 4-month treatment duration using the same drugs.[30] In children, minimal or paucibacillary disease accounts for two thirds of all childhood TB, and so, many children would be spared the additional and unnecessary 2 months of treatment. The trial found that 4 months of treatment were not inferior to the longer treatment duration. This exciting development has led to a revision to WHO guidance,[31] but a key challenge is to reliably define non-severe disease. For the trial, non-severe disease was classified as extra-thoracic lymph node TB or pulmonary TB which was sputum smear-negative and non-severe on chest X-ray. These are not easy to determine and are subject to substantial inter-investigator variability.

A biomarker that could discriminate severe disease from non-severe disease would pave the way for decision-making at baseline that could be stratified, or ultimately personalized. In addition to stratifying children at baseline into different phenotypes that may benefit from different therapeutic approaches, it may also be possible to tailor treatment duration to therapeutic response. A biomarker that modelled the trajectory of response to treatment would make it possible to decide when a child has returned to a "normal" state and at that point it might be possible to stop treatment.

2.5 | Identifying which children and adolescents will develop disease-related morbidity

There is increasing recognition that many TB survivors suffer substantial morbidity. A fifth of children with TB meningitis die, but of survivors over half have permanent long-term neurological impairment.[26] Although data in children and adolescents are limited, over half of adults who have survived pulmonary TB have substantial respiratory morbidity, and those surviving TB have increased risk of death.[32,33] Although severity of disease at baseline is a strong indicator of long-term morbidity, the reasons why some individuals develop post-TB morbidity while others do not is poorly understood and is likely to be due to the host inflammatory response causing host tissue damage and scarring.

If it were possible to determine either at baseline or during treatment, which individuals were likely to develop morbidity, it may be possible to intervene. This might include host-directed therapies (HDTs) at baseline to prevent future morbidity,[34] or the early identification of those with impairment and provision of supportive therapy.

2.6 | Using biomarkers in clinical research for the evaluation of new drugs or vaccines

Demonstrating the efficacy of new anti-TB drugs or TB vaccines requires trial entry and exit points. Entry points are inclusion or exclusion criteria while exit points are trial outcomes. For TB disease

treatment trials, the entry point is TB disease, and the exit points include cure, treatment completion, treatment failure, death, or TB relapse. For TPT trials and most vaccine trials that aim to prevent TB, the entry point is the exclusion of TB disease, with the outcome of interest being TB disease or death. If a biomarker were able to distinguish children with TB disease from those without, the ascertainment of these entry and exit points would be made much easier. In addition, if a biomarker was available that indicated children and adolescents with *Mtb* infection who were at higher risk of disease progression, then TPT trials might opt to focus only on those individuals, making the sample size required for a trial much smaller. In addition, if surrogate biomarkers were identified that served as a correlate of disease or protection, ones which indicated future disease progression or treatment failure but at a much earlier timepoint than clinical outcomes, then the duration, sample size, and cost of clinical trials of new TB drugs (for both TPT and TB disease treatment) as well as for new vaccines, could be reduced.

2.7 | Biological insight

The development of new vaccines and new HDTs requires an insight into the biological interaction between the host and *Mtb*. The aim of most vaccines is to either prime or modulate adaptive immune responses, so that when *Mtb* is encountered, the response is more effective and can either contain or eradicate the organisms before they proliferate and causes disease. The aim of most TB HDTs is to promote helpful inflammatory processes that assist the immune response in containing or eradicating *Mtb*, while inhibiting the damaging, destructive components of the host response that either assist bacterial proliferation or cause substantial host tissue damage that leads to mortality or long-term morbidity beyond the impact of *Mtb*.

By comparing the transcriptomic response of children who have been exposed to *Mtb* and do not progress to disease, with TB-exposed children who do progress, it may be possible to better understand the immune response that is effective in *Mtb* containment. Vaccines that seek to prevent TB, should aim to promote those immune responses. When evaluating the impact of HDTs on TB pathogenesis, it would be informative to compare children with TB and extensive host damage with child with minimal damage. In this way, destructive inflammatory pathways might be identified that could be treated with targeted HDTs.

2.8 | Impact of coinfections on TB susceptibility and disease progression

Children and adolescents are infected with multiple pathogens during the first two decades. Consequently, TB-coinfections are common in high TB-prevalence regions,[35] and there is growing recognition that coinfections may influence TB susceptibility, natural history, and the performance of diagnostics.[36,37]

Infection with a co-pathogen may provoke an immune response which may disrupt anti-mycobacterial immunological pathways important for controlling and containing *Mtb* infection.

Growing evidence suggests that viral coinfections such as influenza may increase susceptibility to *Mtb* infection or increase the risk of progression to disease.[37–42] Understanding the biology and immunological consequences of coinfection on *Mtb* infection and disease is important as it could impact the performance of host immune-based TB diagnostics including transcriptomic signatures, and might also have implications for clinical decision-making and development of vaccines and immunotherapies.[37]

TB-HIV coinfection has been extensively studied, and it is estimated that HIV increases the risk of incident TB disease in children eight-fold, with the risk higher in those more immunosuppressed.[43] Understanding how HIV increases TB susceptibility in children is incompletely understood and requires pediatric-specific studies. Comparing transcriptomic responses with TB in children living with HIV and in children without HIV may shed light on the mechanisms underlying their differing susceptibilities. In addition, the impact of HIV coinfection on the performance of transcriptomic-based diagnostic biomarkers must be understood for them to be clinically useful in coinfected populations, and it is possible that different transcriptomic TB signatures are required for HIV-positive and -negative children.

Since the COVID-19 pandemic began in early 2020, over 540 million confirmed cases of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection have been reported worldwide,[44] with the true number of infections likely to be substantially higher. Countries with high TB incidence such as India and South Africa have reported tens of millions of SARS-CoV-2 infections to date.[44] The pandemic has disrupted health services and TB control programs,[12] with significant drops in case finding and treatment[6] and predictions that this will lead to increased disease burden and mortality.[12–15] With their overlapping epidemiology, risk factors, and clinical presentations, studying the effects of *Mtb*-SARS-CoV-2 coinfection should be a research priority.

Studies of *Mtb*-SARS-CoV-2 coinfections are limited and focus on adults.[13,45–47] Their overlapping clinical presentations mean that distinguishing between SARS-CoV-2, *Mtb*, and *Mtb*-SARS-CoV-2 coinfection can be diagnostically challenging. As with other non-*Mtb* pathogens, when a TB diagnosis is under consideration, the detection of SARS-CoV-2 may exacerbate diagnostic uncertainty as current microbiological tests cannot distinguish between a colonizer, pathogen, and co-pathogen. A biomarker that measures the host immune response could provide insight as to whether the detected pathogen is contributing to the clinical presentation or just a bystander. As with HIV coinfection, SARS-CoV-2 will need to be considered when developing and evaluating host-based diagnostic biomarkers for TB.

In adults, TB disease is a risk factor for severe COVID-19 disease and associated mortality,[48–51] but very few studies have considered the immunological effects of *Mtb*-SARS-CoV-2 coinfection. A South African study has suggested an interaction via alterations in T-cell function including reduced *Mtb*-specific CD4+ cells in adults with COVID-19 disease.[52] Whole blood interferon-γ responses to SARS-CoV-2-peptide stimulation have been observed to be lower in adults with TB and COVID-19 disease compared with those with COVID-19 disease only.[53] A recently published study of a *Mtb*-SARS-CoV-2 coinfection mouse model reported the surprising finding that chronic pulmonary *Mtb* infection protected mice from the effects of a SARS-CoV-2 challenge, with an associated expansion of pulmonary T and B cell subsets observed. SARS-CoV-2 infection did not, however, affect *Mtb*.[54] Further immunological studies of *Mtb*-SARS-CoV-2 interactions are needed to understand the potential impact of the COVID-19 pandemic on TB susceptibility and should include pediatric and adolescent populations. Comparisons of transcriptomic responses to *Mtb* in SARS-CoV-2 infected and uninfected patients will highlight immunological pathways for further functional studies.

3 | THE ROLE OF TRANSCRIPTOMICS

Induction of tissue and immune responses against *Mtb* and the intracellular signaling between immune cells trigger a biochemical chain of events, which leads to the production of molecules needed for defence against *Mtb*. Untargeted host molecular profiling (transcriptomic, proteomic, and metabolomic) in tissues from the site of disease using high-throughput methodologies can be employed to shed light on biological processes and identify key biological molecules. These methods can help identify biomarkers and provide an understanding of the dynamics of infection and inflammation, for example, through highlighting the key pathways involved in the biological process. Studying the host transcriptome in the lungs[55] or other disease sites[56] can elucidate local pathogenic characteristics of the disease and the host immune response at the tissue or cellular level. Host transcriptomic analysis of bronchoalveolar lavage (BAL) or sputum samples has revealed strong type I/II interferon-mediated cytokine responses and T-cell activation in adults with TB.[57]

Due to accessibility, minimal invasiveness in collection, and the key role peripheral blood and its compartments play in host defense and immunity, peripheral blood has become the focus tissue for host transcriptomic studies in child and adolescent TB. This follows work done in adult studies which first established that immune changes associated with pulmonary disease can be identified and quantified in RNA derived from peripheral blood.[58] In this review, we focus on transcriptomic profiling of whole blood cells and peripheral blood mononuclear cells (PBMC) in the context of *Mtb* infection and disease.

As RNA is sensitive to degradation, which can hamper the quantification of results, specific RNA stabilizing systems are used for sample collection, shipping and storage that allow preservation of RNA. Different RNA stabilizing blood-sampling systems may introduce differences in the downstream quantification results, with thousands of genes being reported as significantly differentially expressed (SDE) between RNA stabilizing reagents. This needs to be taken into consideration in study design and meta-analyses.[59]

Subsequently, fine-tuned protocols for RNA purification can ensure the quality, integrity and yield of isolated RNA, along with minimization of potential DNA contamination.[60] Recent studies have shown in vitro whole blood stimulation with *Mtb* antigen peptides, which has been used in proteomic biomarker discovery studies, can unmask transcriptomic signals that are not detectable in unstimulated samples[61] or enhance the diagnostic potential of single gene markers in high burden settings.[62]

In terms of sample volume, although blood–volume dependent reduction in gene levels has been reported, transcriptomic profiling can be achieved with small volumes of blood, which is particularly important in young children. High quality RNA-Sequencing results have been reported in neonates using volumes as low as 0.5 mL of peripheral venous or arterial blood.[63]

4 | TRANSCRIPTOMICS METHODOLOGY

4.1 | Methods for RNA quantification

The "candidate gene" approach focuses on measuring expression for small numbers of genes and can be used when the genes of interest are already known. Reverse transcription quantitative polymerase chain reaction (RT-qPCR) is a sensitive, accurate, highly reproducible method able to detect very small amounts of RNA. It is considered a benchmark technology and forms the basis of various nucleic acid identification platforms with bedside use. Reverse transcription as a methodological step is shared between most protocols for RNA quantification. The fluorescence emitted at the end of each cycle of PCR is used for the estimation of the quantity of the starting material.[64] RT-qPCR has enabled multiple scientific breakthroughs and can now be used to simultaneously detect and quantify multiple nucleic acids through multiplexing.[65,66] The relatively new NanoString nCounter gene expression system first reported in 2008 (Nanostring Technologies, WA, USA) has introduced a new method for targeted RNA quantification based on color-coded probe pairs' ability to hybridize with complementary mRNA and fluorescence.[67] It can be used for quantification of custom genes or inventoried gene panels according to application area or biological process of interest.[68]

Transcriptome-wide profiling methods allow for systematic analysis of thousands of RNA molecules simultaneously in a high-throughput manner and provide a global quantitative profile of the cell or tissue of interest, without the need of a pre-existing research hypothesis. Microarray technology exploits the principles of specific hybridization between two DNA strands and the emission and detection of fluorescence proportional to the amount of nucleic acid that is bound. For gene expression quantification, it involves reverse transcription of the isolated RNA to cDNA, followed by amplification and labelling with a fluorescent dye. Subsequently, the dyed cDNA is hybridized to the array under specific conditions, which permit its binding to complementary sequences already printed on the array, followed by washing to eliminate non-specific binding events, and insertion into a scanner that both excites the fluorescent dyes and records the emitted intensity. The fluorescence signal emitted by the probes reflects the abundance of the corresponding RNA transcript in the sample. Prior knowledge of the position of reference transcripts allows for relative quantification of the transcripts bound to specific probes.[69,70] Gene expression microarrays have been used as the discovery tool in many TB transcriptomic studies,[71,72] and although superseded by RNA sequencing (RNA-Seq), they are still being extensively used due to lower cost, years of standardization, and less intensive bioinformatic analysis needed.

Over the past decade, RNA-Seq has become an indispensable and popular tool for transcriptome-wide analysis with many applications in studying TB host response as it is independent from pre-existing sequence information. However, of the total RNA in a cell, 80% is ribosomal RNA (rRNA), 15% is transfer RNA (tRNA), leaving only 5% as mRNA and all the other RNA forms.[73] To focus on the RNA molecules of interest, RNA-Seq libraries are prepared using either polyA+ selection (mRNA enrichment) or rRNA depletion. rRNA depletion allows the detection of more transcripts including non-coding RNA (ncRNA), small nucleolar RNA (snoRNA), and small nuclear RNA (snRNA). A comparison between the two methods has shown that although rRNA depletion captured more unique transcriptome features, for blood-derived RNAs, 220% more reads would have to be sequenced to achieve the same level of exonic coverage in the rRNA depletion method compared with the polyA+ selection approach.[74] Globin transcript depletion is another critical step to obtaining data suited for blood transcriptome analysis. Total RNA from whole blood contains a large portion of globin transcripts, which originate from red blood cells and account for 80–90% of total transcripts.[75] These affect the quality and accuracy of gene expression profiling and mask the quantification of genes with low expression levels. Next Generation Sequencing (NGS) technology is based on detecting and recording light that is emitted when a complimentary nucleotide is added to a particular fragment of cDNA. The light detected will determine the identity of the nucleotide ("base calling") and subsequently the sequence of the whole "read" in single base-resolution. The reads are either mapped bioinformatically to a reference genome or assembled de novo to produce the transcriptome, a base-resolution expression profile. Read mapping allows for the quantification of RNA and providing abundance estimates.[76] RNA-Seq has revolutionized the field of transcriptomics, also allowing discovery of novel transcripts, alternative splicing, the detection of gene fusion events and allele-specific expression. RNA-Seq also permits simultaneous sequencing of pools of transcripts that may come from different organisms that coexist in the same environment, termed metatranscriptomics.[77,78]

4.2 | Analytical approaches for transcriptome-wide profiling

The quality control, pre-processing and analysis steps for microarrays are mostly standardized, while RNA-Seq data analysis pipelines, which are more complex and computationally demanding, can

consist of a greater variety of steps and tools. The abundance per feature per sample is the input for differential expression analysis for both quantification methods. After quality assessment and exclusion of poor-quality samples, normalized microarray expression values are used for downstream analysis. For RNA-Seq, sequence reads need to be adapter- and quality-trimmed, and then aligned either to the human genome or transcriptome.[79] Features then are quantified, with low abundance features filtered and followed by normalization processes to account for biases, noise, and sequencing depth variation. Subsequently analytical approaches follow according to the biological and clinical questions that are being addressed (i.e., differential gene expression analysis and alternative splicing analysis). Both microarray and RNA-Seq differential expression analysis workflows are followed by multiple testing corrections to control for false positive errors.

The data analysis workflow is quite different for studies intending to discover diagnostic vs. mechanistic transcriptional signatures of disease. The one shared component is a set of initial algorithms to identify gene sets associated with different disease states, termed differentially expressed genes. For biomarker discovery, feature selection methods are employed to identify the marker or the combination of markers that minimize the classification error or maximize the accuracy of classification for patient subgroups, while eliminating noise and redundant genes. Feature selection methods are divided into filter, wrapper, and embedded methods.[80] Filter methods select a feature subset from the original dataset by evaluating the relation between each input variable and the target variable (e.g., statistical methods or feature importance methods). They are usually used as a pre-processing step, followed by a machine learning algorithm. The wrapper methods search for a well-performing subset of features by training a model on it. These iterative methods add (e.g., forward selection) or remove (e.g., backward elimination) features based on the performance of the trained model. The embedded methods combine the advantageous aspects of both filter and wrapper methods since they perform feature selection as part of the classifier construction. Many of these methods have an inbuilt penalization function to shrink the coefficients of the least important features toward zero (e.g., Least Absolute Shrinkage and Selection Operation regression) and keep in the model the most relevant features. Appropriate partitioning of datasets, representing samples, into training and test sets, and the use of iterative resampling methods (e.g., cross-validation) can minimize overfitting, while independent validation datasets can ensure the robustness and generalizability of the findings. To evaluate the performance of biomarker models, different point estimates and interval metrics are typically used for diagnostic tests.[81–83] Results are benchmarked against a "gold standard" diagnosis, which poses a particular challenge for pediatric TB. However, for binary output case–control studies the minimal set of point estimates reported are as follows: (a) the area under the operating receiver characteristic curve (AUC), (b) sensitivity, which reflects the probability of test being positive with disease present and (c) specificity, which reflects the probability of being test negative with disease absent. In prospective recruitment studies, the positive predictive value

(PPV), which reflects the probability of a patient having the disease when the test is positive, the negative predictive value (NPV), which reflects the probability of a patient not having the disease when the test is negative and likelihood ratios are reported in addition.[84] Confidence intervals are calculated to measure the reliability of the estimates. For case–control studies, the ratio of gold standard positive and negative individuals does not reflect the real prevalence of the disease in a community or hospital setting, as in observational studies. Given the dependency of NPV/PPV on the prevalence of the disease in the population, it is important to provide estimates of these values specific to scenarios in which such a diagnostic test would be applied. In this case, prevalence can be interpreted as "the probability before the test is carried out that the subject has the disease."[82]

4.3 | Understanding TB biology

Apart from biomarker discovery, interpreting differential expression results in terms of higher order biological processes or molecular pathways is a key outcome in transcriptomic analysis. One of the most commonly used resources is gene ontology (GO) databases, which annotate genes according to a dictionary of annotation terms, to identify the terms that are over-represented or enriched.[85] Another commonly used annotation database is the Kyoto Encyclopedia of Genes and Genomes (KEGG), a curated database of molecular pathways and disease signatures.[86] Ingenuity Pathway Analysis (IPA- QIAGEN) provides a series of different functionalities, allowing for the identification of significantly enriched canonical pathways, network analysis, and upstream regulating molecules.[87] In comparison to methods using overlap statistics such as the cumulative hypergeometric distribution to identify whether a group of differentially expressed genes is enriched for a pathway or ontology term, a different method can be used termed Gene Set Enrichment Analysis (GSEA) which considers all of the genes in an experiment, rather than only those above specific cut-offs.[88] There are different tools for performing GSEA analysis including MSigDB,[89] g:Profiler,[90] and DAVID.[91]

Understanding the cellular composition of bulk tissues is critical to investigate the underlying mechanisms of many biological processes. Molecular profiling using bulk RNA-Seq in heterogeneous tissues, such as blood, is confounded by the relative proportions of different cell types in the tissue. Single cell RNA-Seq data is quickly becoming the "gold standard" technique for cell specific expression profiles but is an expensive and data-analysis intensive technique. Flow cytometry is also widely used to estimate cell-type composition in bulk tissue experimentally and data on cell fractions obtained can inform differential gene expression analysis and allow model adjustment for cell fractions. However, flow cytometry requires isolation of cells, laboratory equipment, and additional sample sacrifice.

When flow cytometry or single cell RNA-Seq data are not available, cell-type specific gene expression profiles and sample-specific cell type proportions can be estimated from bulk gene

expression data using computational deconvolution methods also termed *in silico* deconvolution methods. These methods are usually based on matrix factorization, which is employed to deconvolve a matrix of gene expression profiles from bulk gene expression data into two matrices, one for the cell-type proportions for each sample and another containing the gene expression profiles for each cell type. Some of the developed algorithms performing partial deconvolution need either cell-specific gene expression profiles or cell-type proportions as input, and use marker genes or regression techniques to estimate the matrix of interest.[92-94] Other methods are able to estimate both matrices using only bulk gene expression data.[95]

5 | BIOMARKERS IN CHILD AND ADOLESCENT TB

Parsimonious gene expression signatures that have the potential to be measured at a clinical setting, and particularly in a low resource clinical setting in TB endemic countries, can assist in successfully addressing the multiple challenges faced in the field of child and adolescent TB, and in particular disease diagnosis, prognosis, and treatment decision-making. A number of research studies have been published presenting results from gene expression analysis for the identification of diagnostic, prognostic, and treatment response biomarkers (Table 1).

5.1 | Transcriptomics as a diagnostic tool

The field of infectious disease diagnostics has embraced molecular tools that profile the host response and can enhance disease diagnostic pipelines, particularly when the detection of the pathogen of interest is challenging, as in pediatric TB.[96] In clinical practice, a gene signature measured in blood that can distinguish pediatric TB from other diseases with similar presentation to TB would be of great value in evaluating symptomatic patients presenting to medical services with symptoms of TB.

Two studies to date have discovered diagnostic gene signatures specific for pediatric TB in a hypothesis-free transcriptome-wide manner. Verhagen and colleagues in 2013 published the first microarray profiling study for pediatric TB biomarker identification in Warao Amerindian children.[97] A signature of 116 genes identified by the random forest algorithm separated 9 TB cases from 9 with *Mtb* infection and 9 healthy controls in the training set, which was then subsequently validated in publicly available adult datasets.[72,98] Following random forest bootstrapping, the list was reduced to 10 genes that was validated using RT-qPCR in the discovery cohort, and in 20 children with *Mtb* infection, 16 healthy children, and 18 children with non-TB pneumonia. Further decision tree analysis indicated that five genes classified 78% of the TB cases correctly with a 0% false positives for the *Mtb* infection group, 4% for the healthy controls, and 11% for the non-TB pneumonia cases.

In 2014, Anderson and colleagues described the discovery of transcriptional signatures for distinguishing culture-confirmed TB from diseases other than TB in a multicohort pediatric population (<15 years of age), comprising of HIV-positive and HIV-negative children with symptoms and clinical findings that were suggestive of TB, using microarray analysis of host blood.[99] Patient data from South Africa and Malawi were combined into a discovery set (80% training and 20% test). SDE transcripts were subjected to feature selection using elastic net that resulted in a 51-transcript signature with abundance combined into a Disease Risk Score. The 51-transcript signature had a sensitivity of 82.9% ($CI_{95\%}$ 68.6–94.3), and a specificity of 83.6% ($CI_{95\%}$ 74.6–92.7) in an independent validation cohort from Kenya (Figure 2). The sensitivity exceeded that of the Xpert MTB/RIF assay which was 54.3% ($CI_{95\%}$ 37.1–68.6). Additional analysis provided estimates for sensitivity in the culture-negative groups of highly probable, probable, and possible TB. Using a similar discovery pipeline, a 42-transcript signature discriminated confirmed TB from *Mtb* infection with sensitivity of 94.3% and specificity of 100.0% in the independent validation cohort from Kenya.

In 2017, Gjoen and colleagues employed a dual color reverse transcription multiplex ligation dependent probe amplification assay (dcRT-MLPA) to assess the performance of 198 genes in a training set, comprising 47 pediatric TB cases (19 definite and 28 probable) and 36 asymptomatic household controls, and identified a 7- and a 10-transcript signature by using a combination of logistic regression and LASSO for feature selection.[100] The 10-transcript signatures had an AUC of 0.94 ($CI_{95\%}$ 0.88–1.00), correctly classifying 22 of 24 TB cases and 23 of 26 symptomatic non-TB cases, corresponding to a sensitivity of 91.7% ($CI_{95\%}$ 71.5–98.5) and a specificity of 88.5% ($CI_{95\%}$ 68.7–96.9). The 7-transcript signature also provided an AUC of 0.94 ($CI_{95\%}$ 0.88–1.00) but classified correctly only 21 of 26 symptomatic non-TB cases.

A multicohort meta-analysis by Sweeney and colleagues in 2016 included 14 datasets containing 2572 samples from 10 countries from both adult and pediatric patients and identified a 3-gene signature reporting an AUC for discriminating TB from other diseases of 0.84 ($CI_{95\%}$ 0.80–0.95) and from *Mtb* infection 0.88 ($CI_{95\%}$ 0.84–0.92). For the comparison of TB vs. other diseases in the previously reported pediatric datasets from South Africa/Malawi the Sweeney 3-gene signature had sensitivity of 68.5% and specificity of 74.0% and a sensitivity of 77.1% and specificity of 87.5% in the Kenyan dataset.[99] The metrics for the pediatric dataset from South Africa/Malawi are lower in comparison to the global metrics reported for TB vs. other diseases including the adult datasets (sensitivity 81%, specificity 74%), highlighting the need for pediatric-specific signatures or the inclusion of genes that boost the performance in the pediatric datasets to account for the different nature of the disease, the different host response and the fact that pediatric TB needs to be discriminated from a different set of other diseases than adult TB in clinical practice.

A larger number of studies have been published that focus on discriminating TB from healthy uninfected controls and children with *Mtb* infection (Table 1). Although important as a proof-of-concept,

TABLE 1 Studies that have used transcriptomic approaches in child and adolescent TB, presenting original patient recruitment data and analysis

First author	Year published	Country -Population	Description of study	Number of children/adolescents analyzed (original data)	Main findings of study
Verhagen[97]	2013	Venezuela	Discovery (microarray) and validation (RT-qPCR) of gene expression signature to distinguish TB from children with *Mtb* infection, healthy controls and non-TB pneumonia (validation only)	9 TB patients, 29 with *Mtb* infection, 25 healthy controls and 18 non-TB pneumonia	A 116-gene signature for TB vs. *Mtb* infection and healthy controls with an average prediction error of 11%, and a 9-gene signature for TB vs *Mtb* infection. An optimized set of 5 genes had 4% false positive rates for healthy controls and 11% for non-TB pneumonia cases.
Dhanasekaran[101]	2013	India	Whole-blood mRNA from 210 children was examined by dcRT-MLPA for the expression of 45 genes	13 children with TB disease, 90 with *Mtb* infection, 107 uninfected controls	A single gene discriminated between TB and *Mtb* infection (AUC 0.78). A 5-gene signature discriminated TB disease from controls (AUC 0.92). An 11-gene signature distinguished *Mtb* infection from controls (AUC 0.72).
Anderson[99]	2014	South Africa, Malawi, Kenya	Discovery and validation of gene expression biomarkers from microarray data to distinguish TB from *Mtb* infection and TB from Other Diseases	149 children with culture-confirmed TB, 44 with unconfirmed TB, 71 with *Mtb* infection and 139 with Other Diseases	A 51-trasncript signature for TB vs Other Diseases and a 42-transcript signature for TB vs *Mtb* infection with sensitivity of 82.9% and specificity of 83.6% in the independent validation cohort.
Li[62]	2015	China	Quantification by RT-qPCR of 7 genes in PBMC after ESAT-6 stimulation in children with PTB, EPTB and healthy controls	39 children with TB (3 smear culture positive and 26 culture negative) and 25 healthy controls	The expression of *IL-9* separated children with TB vs healthy controls with an AUC of 0.92 after ESAT-6 stimulation.
Wang[102]	2015	China	RT-qPCR was used to quantify miR-31 expression in PBMCs from children with TB and healthy controls	65 children with TB and 60 healthy controls	The expression of miRNA-31 distinguished children with TB from healthy controls with sensitivity of 98.5% and a specificity of 86.7%.

TABLE 1 (Continued)

First author	Year published	Country -Population	Description of study	Number of children/adolescents analyzed (original data)	Main findings of study
Zak[110]	2016	South Africa	Discovery using blood RNA-Sequencing, and validation (RNA-Sequencing and RT-qPCR) of a signature to predict progression of *Mtb* infection to disease in adolescents (12–18 years)	46 progressors and 107 matched controls	A 16-gene signature for TB progression which had sensitivity of 53.7% and specificity of 82.8% in the 12 months preceding TB in independent South African and Gambian cohorts.
Fletcher[137]	2016	South Africa	Infants were vaccinated with BCG at birth and followed for 2 years. Blood was collected at 10 weeks. Host responses from the 10-week samples were compared between those who developed TB disease within 2 years and controls who remained healthy.	5726 infants were recruited. 29 cases of confirmed TB were compared to 110 controls (55 household controls and 55 community controls).	Gene expression analysis did not show a difference between cases and controls.
Jenum[104]	2016	India	Targeted analysis of transcriptional immune biomarkers in *Mtb*-antigen stimulated whole blood using dcRT-MLPA	88 children with intra-thoracic TB (6 months - 15 years); 40 culture-confirmed, 48 unconfirmed and 39 asymptomatic	An 8-gene biomarker signature separated children with TB from asymptomatic siblings (AUC 0.88) in stimulated blood. 12 genes were found associated with clinical groups toward culture-positive TB or toward a decreased likelihood of TB disease on the TB disease spectrum.
Zhou[103]	2016	China	Identification of circulating miRNAs that can differentiate between TB and healthy controls	14 culture-positive TB cases, 14 culture-negative TB cases and 25 children with TB and 21 healthy controls for validation	An 8-miRNA signature provided 95.8% sensitivity and 100% specificity for the discrimination of children with TB vs uninfected healthy controls
Gjoen[100]	2017	India	Selection and optimization of 2 signatures for TB vs asymptomatic household controls, and other symptomatic non-TB cases, from a set of 198 genes using dcRT-MLPA.	71 TB cases (36 definite/35 probable) and 36 asymptomatic household controls, and 26 symptomatic non-TB cases.	A 7-and a 10-transcript signature with AUC of 0.94 in separating TB-cases from symptomatic non-TB cases regardless of culture status, and 100% sensitivity for definite TB.

TABLE 1 (Continued)

First author	Year published	Country -Population	Description of study	Number of children/adolescents analyzed (original data)	Main findings of study
Hemingway[121]	2017	South Africa	Longitudinal microarray blood gene expression analysis in children with TBM and comparison with children with PTB	9 children with TBM, (4 timepoints) and 9 healthy controls; 13 children with TBM and 28 with PTB.	Reduced abundance of 68% of SDE genes in TBM vs healthy controls. The difference in abundance was less in PTB than in TBM.
Rohlwink[122]	2019	South Africa	RNA-Sequencing on whole blood as well as on ventricular and lumbar cerebrospinal fluid of pediatric patients treated for TBM	20 TBM cases 20, 7 Non *Mtb* infection controls, and 24 healthy controls	2230 genes were SDE in TBM cases vs healthy controls in blood, and 312 genes were SDE in ventricular CSF in TBM vs infection controls. TB disease processes differ between the periphery and the central nervous system, and within brain compartments.
Penn-Nicholson[113]	2020	South Africa, The Gambia, Ethiopia, Peru, Brazil	Identification of a parsimonious signature from the RNA-Sequencing Zak et al. data using RT-qPCR data, and subsequent validation as a signature for diagnosis, progression and treatment response.	46 progressors and 107 matched controls (Adolescent cohort study)	A 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response
Tornheim[116]	2020	India	Longitudinal RNA-Sequencing from whole blood in cases during treatment and controls for the identification of differentially expressed genes.	16 TB cases and 32 TB-exposed controls	A 71 gene signature for TB diagnosis and a 25 gene signature for treatment response
Johnson[138]	2021	India	Performance of TB gene signatures in malnourished individuals (including children) with TB and *Mtb* infection	23 severely malnourished individuals with TB and 15 severely malnourished TST positive household contacts	4913 significant differentially expressed protein coding genes in TB vs *Mtb* infection in malnourished individuals; 56.9% of the genes overlap with the 45 TB gene signatures included on the paper.

Abbreviations: AUC, area under the curve; CSF, cerebrospinal fluid; dcRT-MLPA, dual colour reverse transcription multiplex ligation dependent probe amplification assay; EPTB, extrapulmonary TB; PBMC, peripheral blood mononuclear cell; PTB, pulmonary TB; RT-qPCR, reverse transcription quantitative polymerase chain reaction; SDE, significantly differentially expressed; TB, tuberculosis; TBM, TB meningitis; TST, tuberculin skin test.

this discrimination has less clinical utility. As this is a disease vs. healthy comparison, the host response is more perturbed, and the groups can be discriminated more easily and with fewer genes. In 2013, Dhanasekaran and colleagues identified a 5-gene signature that can discriminate children with TB from healthy controls and an 11-transcript signature that can discriminate children with TB from children with *Mtb* infection, using dcRT-MLPA in a pediatric population in India.[101] Wang and colleagues in 2015 reported that a single micro RNA (miRNA) can distinguish children with TB from healthy controls with sensitivity of 98.5% and a specificity of 86.7%,[102] while

in 2016 Zhou and colleagues identified a 8-miRNA signature that had 95.8% sensitivity and 100% specificity for the discrimination of children with TB vs uninfected healthy controls.[103] Antigen specific stimulation has also been employed prior to gene expression quantification. Li and colleagues reported that a single gene (IL-9 mRNA) can separate children with TB vs. healthy controls with an AUC of 0.92 after ESAT-6 stimulation,[62] while Jenum and colleagues reported that an 8-gene biomarker signature separated children with TB from asymptomatic siblings in stimulated blood with an AUC of 0.88.[104]

It is important to put these novel transcriptomic signatures in the context of the tools that are currently used to diagnose TB in children and compare performance. In a study of children with presumptive TB who were living with HIV, chest X-ray features that were consistent with TB provided a sensitivity of 71.4% and specificity of 50% when comparing confirmed TB against unlikely TB.[105] A systematic review and meta-analysis of the accuracy of IGRAs to discriminate children with confirmed TB disease from children with other diseases demonstrated a sensitivity of 75% and a specificity of 66%.[106] Xpert MTB/RIF in sputum demonstrated a sensitivity of 64.6% and specificity of 99% in a systematic review and meta-analysis of performance against a microbiological (culture) reference standard. However, against a composite reference standard (including children diagnosed both clinically and microbiologically), the sensitivity was 19.7% with a specificity of 100%.[107] Finally, in 2022, the WHO revised their child and adolescent TB guideline and, in the operational handbook accompanying the guideline, suggested a clinical decision-making algorithm. The sensitivity of this algorithm (against a composite reference standard) was 85% with a specificity of 37%.[108]

5.2 | Transcriptomics as a prognostic tool

A test that identifies children and adolescents who are at greater risk of disease progression would transform TB control by enabling targeted preventive therapy for the population at risk in high-burden settings. While several studies have identified signatures in adults,[109] few have included children or adolescents. In 2016, Zak and colleagues reported discovery of a 16-gene signature by blood RNA-Seq profiling in a cohort of 153 South African adolescents with *Mtb* infection, using samples collected at different timepoints prior to TB diagnosis, which was able to discriminate between TB progressors from non-progressors (Figure 3).[110] The prognostic performance of the signature was dependent on the time interval between sampling and disease diagnosis. Signature performance was better when measured in samples collected more proximal to disease diagnosis than in samples collected at distal time points. At 6 months prior to diagnosis, the AUC was 0.79 (CI$_{95\%}$ 0.76–0.82). Subsequent validation using RT-qPCR, in the adolescent cohort, and in independent Gambian, Ethiopian, and South African adult cohorts replicated the findings. As the genes in the signature overlap with signatures of TB disease, it is evident that progressors have a host response similar to

that of TB disease, albeit at lower magnitude, long before diagnosis is established. It is therefore most likely that such signatures identify incipient or sub-clinical disease, as recently confirmed in the CORTIS study.[111] A subsequent study developed a more parsimonious signature in the adolescent cohort based on 6 genes (RISK6),[112,113] which was further evaluated in an additional multi-country adult study for TB diagnosis.[114] As for diagnostic tools, it is important to compare these transcriptomic prognostic signatures with other tools used for clinical decision-making. Although children <5 years with positive IGRA or TST tests are at ~8 times greater risk of TB disease progression than those with negative IGRA tests, only about 20% of children of this age with a positive IGRA/TST will progress to disease.[21] A positive test therefore may have relatively high sensitivity, but low specificity.

5.3 | Treatment response

Currently available tests have very low accuracy for monitoring TB treatment response and predicting failure or relapse in pulmonary TB even in adults.[115] To improve disease outcomes, we need better biomarkers to identify appropriate responses to treatment, that will allow us to identify treatment failure early and enable shortening treatment. It has been shown that the RISK6 signature tracks treatment response in adults,[113] and has been evaluated in monitoring treatment response in a multi-country cohort (Figure 4),[114] but studies in children are limited.

In 2020, Tornheim and colleagues identified a 25-gene list of differentially i have spotted expressed genes for treatment response in children, in 16 pediatric Indian TB cases and 32 age- and sex-matched TB-exposed controls. The 25-gene treatment response list was evaluated against adult datasets with AUCs of 0.90 (CI$_{95\%}$ 0.74–1.00), 0.72 (CI$_{95\%}$ 0.85–0.94), and 0.50 (CI$_{95\%}$ 0.36-0.63).[116]

5.4 | Mtb-SARS-CoV-2 coinfections and how they impact transcriptomic biomarkers

COVID-19 and TB disease are both primarily respiratory pathogens and cause overlapping clinical syndromes in children and adolescents. A clinically useful transcriptomic diagnostic biomarker will need to discriminate between these two infections as well as other common respiratory pathogens. A recently published large scale systematic review and meta-analysis of COVID-19 transcriptomic signatures investigated their commonality with the transcriptome in TB disease. The authors identified 35 eligible COVID-19 signatures associated with disease severity, derived from whole blood, PBMCs, and BAL fluid. Expression of these signatures was profiled on whole blood RNA-Seq samples from 3 independent adult TB cohorts, comprising of patients with TB disease, *Mtb* infection, including those progressing ("progressors") and not progressing ("non-progressors") to TB disease over the next 2 years, and uninfected controls. Twenty COVID-19 gene
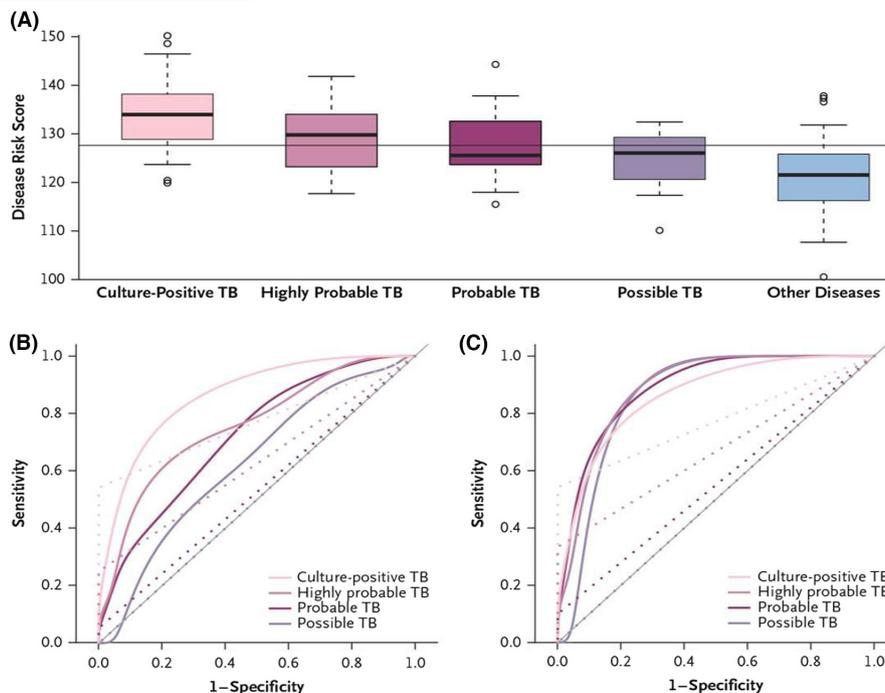
FIGURE 2 Risk Scores and Sensitivity and Specificity in the Kenyan Validation Cohort, According to Diagnostic Group. Panel A shows the risk scores for tuberculosis according to study group, calculated with the use of a 51-transcript signature applied to the independent Kenyan validation cohort, in which culture-positive tuberculosis was reported in 35 patients, diseases other than tuberculosis were reported in 55 patients, and culture-negative tuberculosis was reported as highly probable in 5 patients, probable in 19 patients, and possible in 17 patients. The bar within each box indicates the median score, the bottom and top of the box indicate the interquartile range, the bars below and above the box are at a distance of 0.8 times the interquartile range from the upper and lower edges of the box, and the circles indicate outliers; the horizontal line across the graph indicates the mean score. Panel B shows smoothed receiver-operating-characteristic (ROC) curves for the sensitivity and specificity of the risk score (solid lines) and the Xpert MTB/RIF assay (dotted lines). Panel C shows ROC curves based on an adjusted analysis in which the actual prevalence of disease was assumed to be 80% among patients in whom the disease was highly probable, 50% among those in whom it was probable, and 40% among those in whom it was possible. From Anderson and colleagues. New Eng J Med 2014; 370: 1712–23

signatures were significantly associated with higher "COVID-19 risk scores" in TB disease patients and progressors, compared with non-progressors ($P < 0.005$). By comparison, an influenza signature score was zero across the cohort, regardless of TB status. From profiling 20 single-cell (sc)RNA-Seq immune cell population signatures, the authors found that neutrophil and monocyte COVID-19 signatures generated the highest COVID-19 risk scores in TB disease and progressors, in contrast to adaptive immune cell signatures which were higher in non-progressors. The authors also performed a meta-pathway enrichment analysis using whole blood transcriptomic data from one of the adult TB cohorts, a COVID-19 whole blood scRNA-Seq dataset plus an influenza viral control cohort. In contrast to the influenza group, the COVID-19, TB disease and TB progressor groups were all strongly enriched for interferon-γ and tumor necrosis factor response pathways. Of the top 100 enriched pathways, just one (Hallmark of mTORC signaling) was unique to COVID-19 and absent in TB and influenza.[117] Not only does this study highlight the great potential for transcriptomic studies to elucidate underlying biological mechanisms of infectious diseases, but this work also suggests that pre-existing TB diagnostic signatures may perform

less well in the context of COVID-19 due to shared immune responses. Thus, when discovering and developing future diagnostic transcriptome-based biomarkers, it will be important to include COVID-19 disease as a comparator group. In addition, the impact of SARS-CoV-2 coinfection on the performance of pre-existing and future pediatric biomarkers should be determined, to enable them to be interpreted and deployed in real-world settings where *Mtb*-SARS-CoV-2 coinfection is likely to be common. Pediatric immune responses to SARS-CoV-2 and *Mtb* infection differ from adults, and therefore, it is imperative that pediatric-specific studies are undertaken.

6 | USING TRANSCRIPTOMICS TO UNDERSTAND CHILD AND ADOLESCENT TB BIOLOGY

Recent blood gene expression profiling studies have highlighted differences between adult and childhood TB in terms of pathways involved in host response during infection and disease and progression of infection to disease.

6.1 | Children with TB vs children with other diseases vs well children

Parsimonious gene expression signatures are the product of feature selection algorithms, which try to maximize accuracy in segregation and simultaneously minimize the number of genes in the selected set based on specific criteria. Although they include key molecules that are differentially expressed between groups, they do not capture the full picture of perturbation in blood gene expression in different TB states. Here, we present the results of pathway analysis using Ingenuity Pathway Analysis (IPA) on SDE genes in TB vs. Other Diseases from the datasets presented in the Anderson and colleagues' study, as well as an overview of concordance/discordance of gene expression in TB and Other Diseases in relation to *Mtb* infection as the baseline. The analytical steps for quality control and pre-processing were followed as reported in the original study, including background correction, variance stabilization and normalization, while the data from the HIV-negative children from South Africa, Malawi and Kenya were combined.[118,119] Using a linear model corrected for age, gender and site, we identified 1021 transcripts that were SDE (adj. *P*-value <0.0001, |log2FoldChange|>0.25) between TB and other diseases (Figures 5 and 6). After performing ID mapping in IPA, 742 molecules (324 under-expressed in TB and 418 over-expressed in TB) were included in the pathway analysis. Figure 7 shows the most significant canonical pathways, with Benjamini-Hochberg corrected *P*-values, *z*-scores indicating activation or inhibition. The significance indicates the probability of association of

molecules from the list of SDE genes in the dataset with the canonical pathway by random chance alone. While Figure 8 demonstrates the interferon signaling pathway.

Differential expression analysis compares two patient groups in a direct way, identifying genes which are significantly different between the two in terms of expression. However, elements of the immune response as captured by the transcriptome may exhibit similarity or dissimilarity in comparison to a baseline (control) group. Here, we present the concordant and discordant genes in TB and other diseases in relation to *Mtb* infection as a baseline.[120]

The concordance-discordance between the statistically differentially expressed genes (adjusted *P*-value <0.01) between culture confirmed TB vs *Mtb* infection (11 533 SDE) and other diseases vs *Mtb* infection (5982 SDE) is shown in Figure 9, where a discordance and concordance score for homologous gene pairs was calculated based on their log2FoldChange and adjusted *P*-value. The plot highlights the high degree of concordance of over- or under-expression of SDE genes between TB disease and Other Diseases in relation to baseline. Remarkably, only 5 genes were found to be discordant (over-expressed in TB vs. *Mtb* infection and under-expressed in OD vs. *Mtb* infection), while 4 out these 5 genes have been previously proposed as TB-specific transcriptomic biomarkers.[99,110] Normalized expression of these 5 genes is shown for TB disease, *Mtb* infection, and Other Diseases (Figure 10), while network analysis using Ingenuity Pathway Analysis indicated the 5 discordant genes form a network downstream of IFN-a and IFN-g (Figure 11).
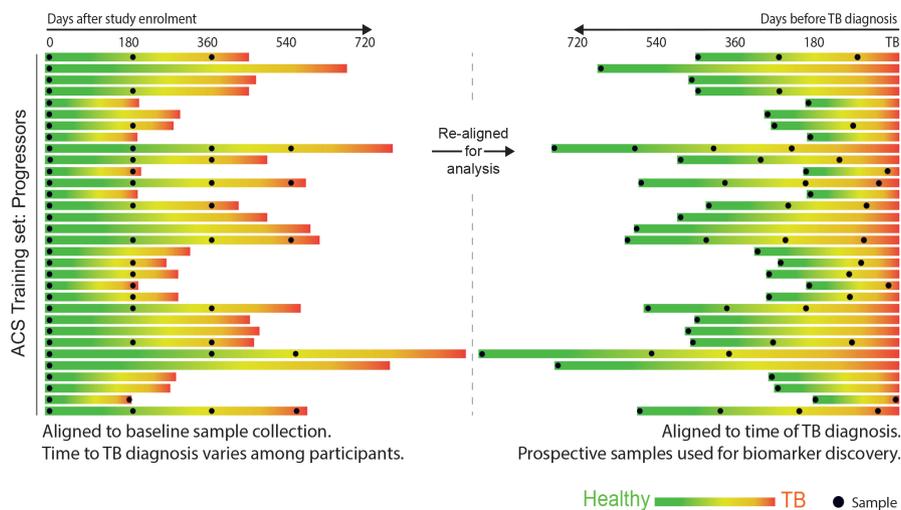
FIGURE 3 Strategy for discovery and validation of the tuberculosis risk signature. Synchronization of the adolescent cohort study training set in terms of the clinical outcome. To ensure optimal extraction of a tuberculosis risk signature from the adolescent cohort study training set, the timescale of the RNA-Sequencing dataset was realigned according to tuberculosis diagnosis instead of study enrolment, allowing gene expression differences to be measured before disease diagnosis. Each progressor within the adolescent cohort study training set is represented by a horizontal bar. The length of the bar represents the number of days between study enrolment and diagnosis with active tuberculosis. During follow-up, each progressor transitioned from an asymptomatic healthy state (green) to pulmonary disease (red). The left graph shows alignment of PAXgene sample collection (black points) with respect to study enrolment. The right graph shows alignment of PAXgene sample collection with respect to diagnosis with active tuberculosis, for use in analysis. From Zak and colleagues. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet 2016; 387: 2312–22
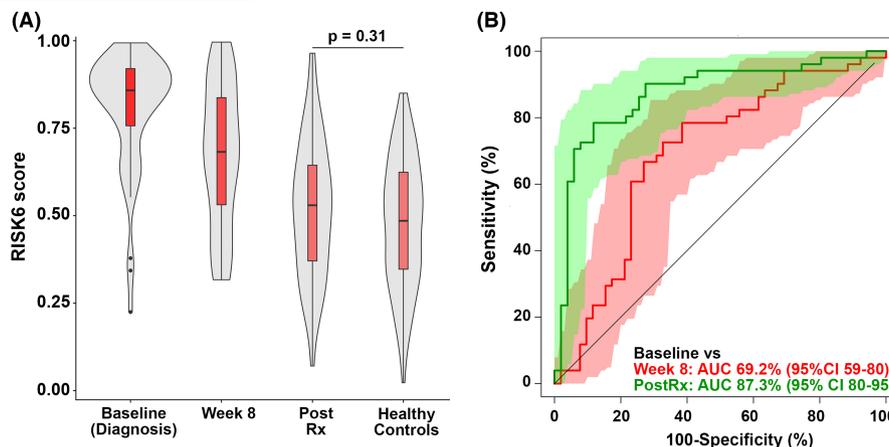
FIGURE 4 Diagnostic performance and treatment monitoring in South American cohorts. (A) Comparison of RISK6 signature scores in TB cases at baseline, Week 8 after treatment initiation and after treatment completion (Post Rx). Also shown are the RISK6 signature scores in healthy controls from Brazil. Horizontal lines depict medians, the boxes the IQR, and the whiskers the range. Violin plots depict the density of data points. The *P*-value, computed by Mann–Whitney U test, compares RISK6 signature scores after treatment completion with those in controls. (B) ROC curves depicting performance of RISK6 for discriminating between baseline samples from TB cases and samples collected 8 weeks after treatment initiation, or upon completion of TB treatment (Post Rx). From Penn-Nicholson and colleagues. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis, and treatment response. Sci Rep 2020; 10: 8629

6.2 | Disseminated disease

Hemingway and colleagues employed blood gene expression profiling in children with TB meningitis (TBM) to identify temporal patterns of RNA during the course of the disease in comparison to healthy controls, and a second cohort of TBM and pulmonary TB (PTB) to explore differences in the host response in these two disease manifestations in children.[121] The authors identified a decreased abundance of multiple genes in TBM patients in comparison with healthy controls, 68% of SDE genes were under-expressed. The magnitude of differential abundance was less in PTB than in TBM. Most genes were involved in activation of leucocytes ($P = 2.67E^{-11}$) and T-cell receptor signaling ($P = 6.56E^{-07}$). Multiple genes involved in T-cell activation showed decreased abundance in children with TB, suggesting that childhood TB is associated with an acquired immune defect. Although in silico *deconvolution* revealed a reduction in CD4+ T lymphocytes in TBM, when compared with healthy controls, the differences in gene expression were not explained by the differing cell proportions. The suppression of genes in the T-cell response pathway was associated with reduced T-cell proliferative responses in vitro, indicating a functional impairment in responses is associated with the RNA transcriptional suppression.

In 2019, Rohlwink and colleagues conducted RNA-Seq both on whole blood as well as brain and spinal CSF of children with TBM to identify patterns of differential expression against healthy controls and non-TB cerebral infection controls in CSF.[122] Differential gene expression analysis, functional canonical pathways and disease networks revealed that TBM is characterized by a significant increase in inflammasome activation, IL-1 signaling and decrease in T-cell activation in blood and neural injury in CSF. The findings corroborated the lack of an effective T-cell response in pediatric TBM.

6.3 | Transcriptomic profiles in children vs adolescents vs adults

The risk of progression from *Mtb* infection to disease is substantially lower in pre-adolescent children above 4 years of age than in post-pubescent adolescents and young adults.[123] Blood gene expression profile analysis revealed that pre-adolescent children had lower levels of myeloid-associated pro-inflammatory mediators than young adults. When compared with young adults, pre-adolescent children had higher levels of IFN-stimulated genes IFNAR2, MX2, OAS1, and STAT2, as well as B cells and M2 macrophages than adults.

To explore the difference between the immune response to *Mtb* between children and adults, Bah and colleagues analyzed 11 datasets comprising 1073 patients from Africa, Europe, and South America, including three datasets from children.[124] Pathway analyses indicated that adults showed more IFN-driven innate immune pathways and downregulated adaptive pathways. Childhood specific upregulated genes were associated with inflammasome IL1RN– IL1R2 suppression axis, glucose transport, CO_2 and O_2 release, and cell surface interaction pathways while downregulated genes were associated with mRNA translation, protein metabolism, and amino acid transport. In childhood TB, the immune inhibitory molecules, IL1RN and IL1R2 which inhibit functional IL1 signaling, and molecules involved in generation of an adaptive immune response (CD40LG, HAL-DOB, CD28) that were downregulated is consistent with emerging evidence linking the cross talk between IL1 and type 1 IFN to TB and which provide potential targets for host directed therapy. This could explain the observation that children have less pronounced adaptive immune responses to *Mtb* compared with adults, potentially resulting in disseminated forms of the diseases seen more frequently in childhood TB.
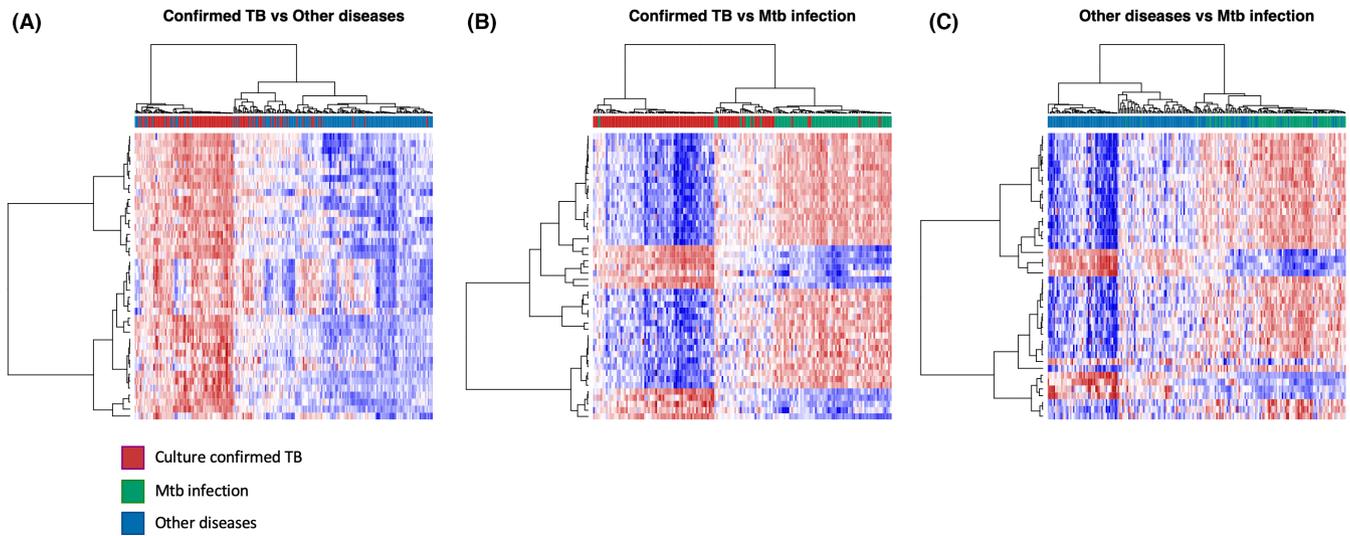
FIGURE 5 Heatmaps showing clustering (Unweighted Pair Group Method with Arithmetic Mean or UPGMA method) of the top 50 SDEs in (A) culture confirmed TB vs other diseases, (B) culture confirmed TB vs *Mtb* infection and (C) other diseases vs *Mtb* infection of the patients from South Africa, Malawi, and Kenya in Anderson et al datasets. Patients' clinical groups are highlighted at the bar on the top of each heatmap, with culture confirmed TB patients in red, patients with *Mtb* infection in green and patients with other diseases in blue. Only HIV uninfected patients have been included and shown on these heatmaps. Under-expression is depicted in blue and over expression in red
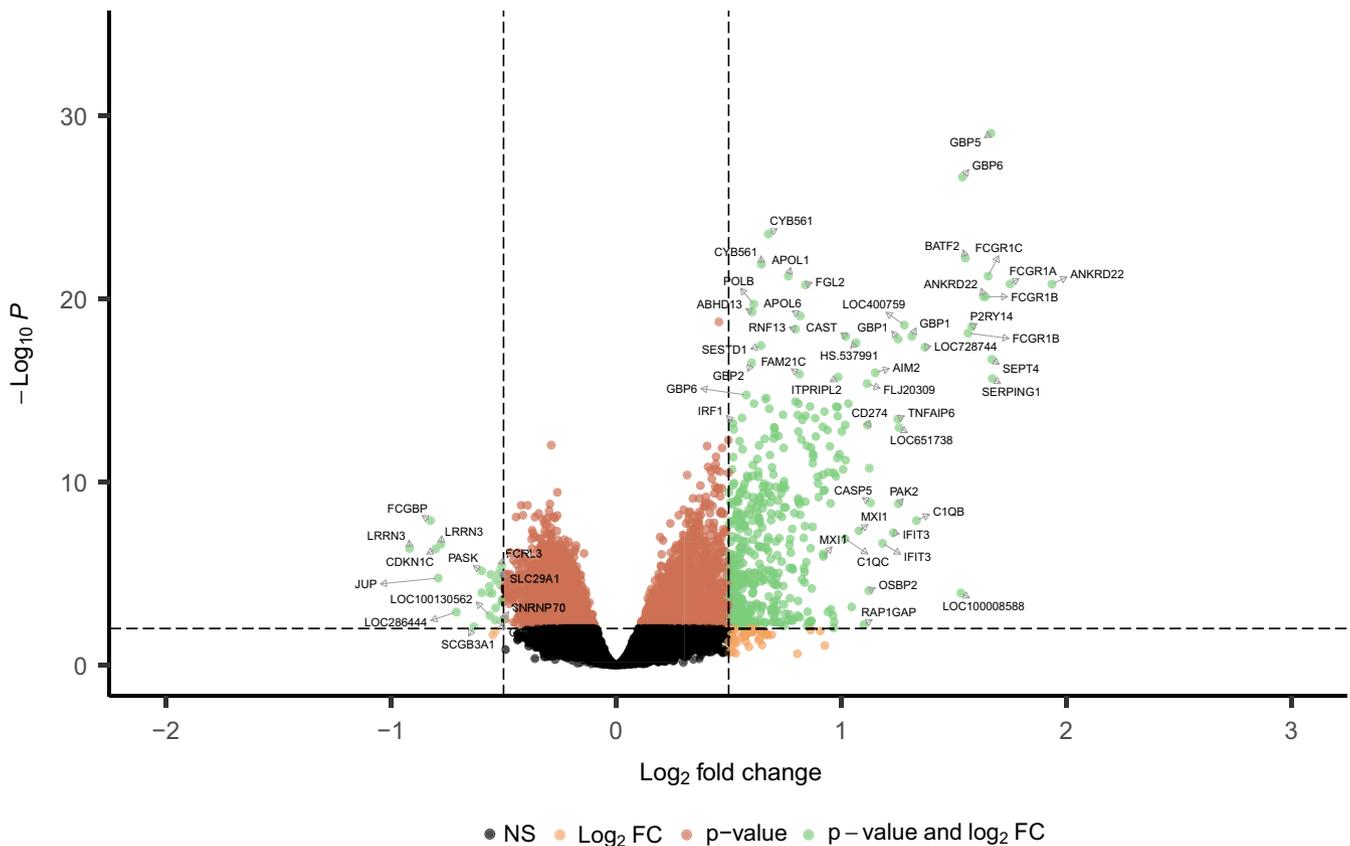
FIGURE 6 Volcano plot showing the significant genes identified comparing culture confirmed TB cases vs other diseases (OD) from Anderson et al South Africa, Malawi, and Kenya HIV uninfected patient datasets. Genes that passed the thresholds for absolute value of log2FoldChange >0.5 and adj-p-value <0.05, were colored in green

(A)

© 2000-2022 QIAGEN. All rights reserved.

(B)
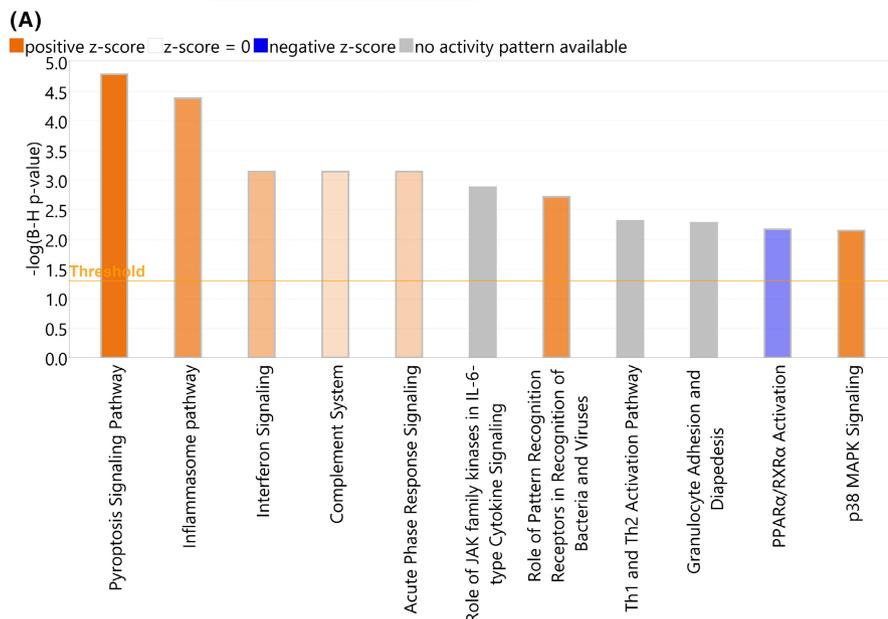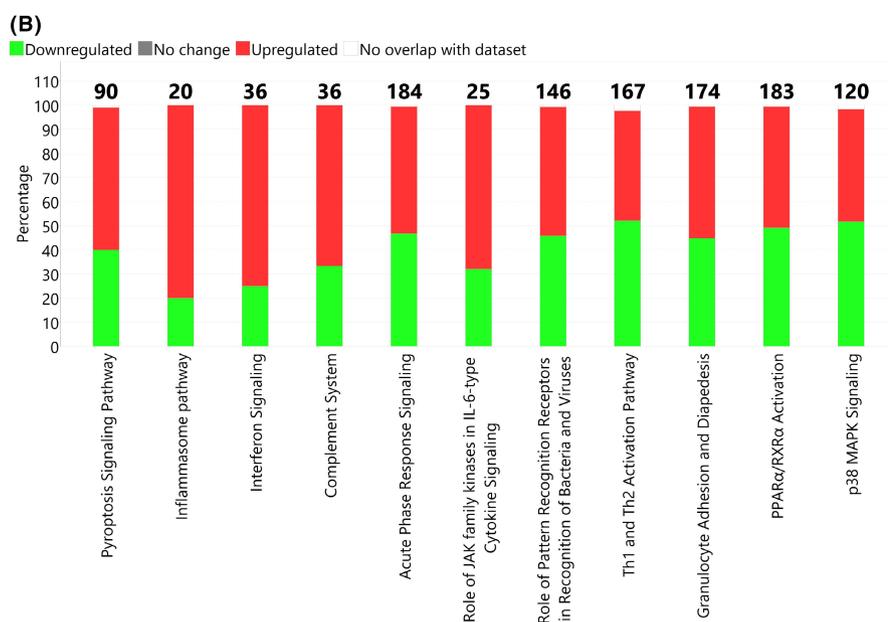
© 2000-2022 QIAGEN. All rights reserved.

FIGURE 7 Canonical pathway analysis results comparing confirmed TB to other diseases (OD) using Anderson et al datasets. (A) The orange and blue-colored bars in the bar chart indicate predicted pathway activation or predicted inhibition, respectively. Gray bars indicate pathways for which no prediction can be made due to insufficient evidence in the Knowledge Base for confident activity predictions across datasets. (B) Displays the number of molecules in the list of SDE genes, showing the up-regulated (red), down-regulated (green). The y-axis represents the percentage of molecules that are present in a specific Canonical Pathway. The total number of molecules in the pathway is shown

7 | MULTI-OMICS APPROACHES

Despite the numerous insights obtained already by the analysis of host transcriptomic datasets, a single-layer of -omics is unlikely to capture the complete biological complexity of disease.[125] Integrating multiple-omic levels, which represent multiple levels of biological organization, in a cross-layer manner, enables a more accurate reconstruction of the dynamic molecular networks underpinning healthy and diseased states.[126] The popularity of multi-omic integration has grown rapidly during recent years, resulting in a vast range of integration methods for supervised and unsupervised analysis, using a range of statistical and machine learning approaches. By combining weak yet consistent alterations across different data layers, new insights have been obtained in cancer[127] and COVID-19 disease pathogenesis,[128] while a recent study used three different -omic layers for biomarker identification for tuberculosis in the serum of adult patients with advanced HIV.[129] Multi-omic integrative studies could be key in furthering the understanding of pathogenesis and disease progression in childhood TB and can be used to provide combinations of host markers across different layers that could increase their discriminatory performance.

8 | CLINICAL EVALUATION OF BIOMARKERS IN CHILDREN AND ADOLESCENTS

8.1 | The challenge of unconfirmed TB

An important challenge for the development of biomarkers in childhood TB is the composition of children with unconfirmed TB. This is particularly relevant for the development of diagnostic tests for TB disease but has implications for prediction of future disease

Path Designer Interferon Signaling
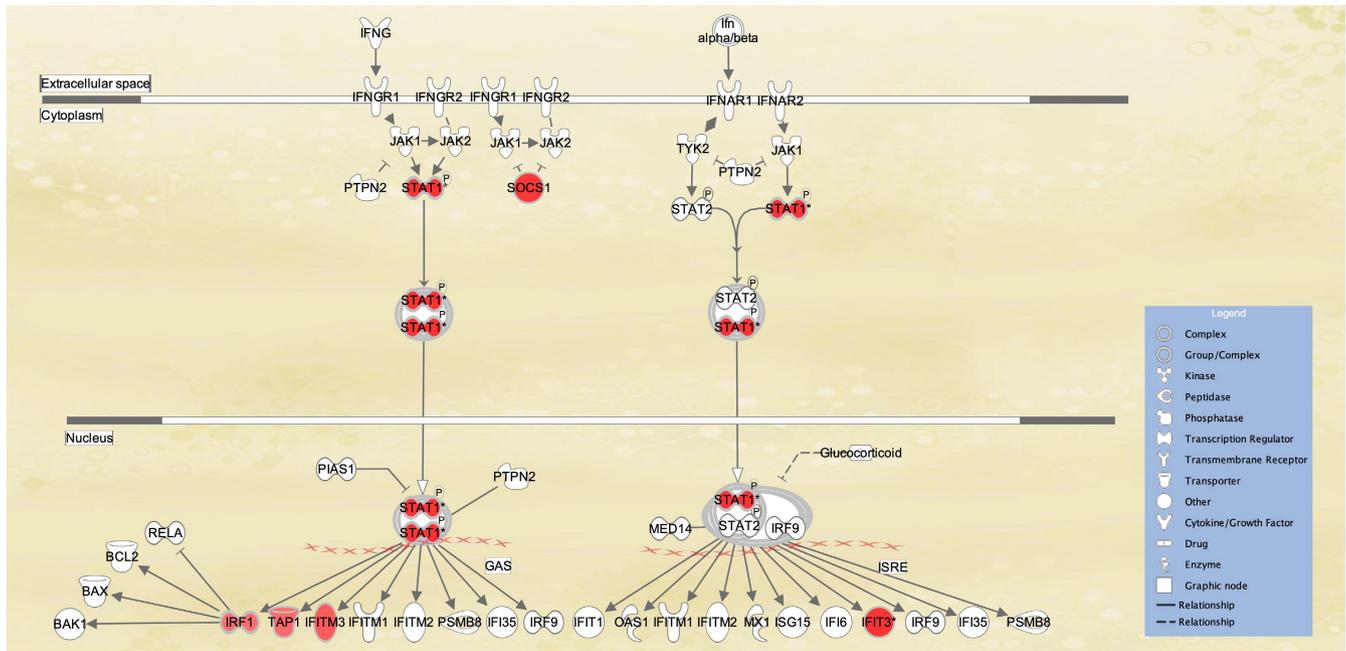
FIGURE 8 Canonical pathway example: Interferon Signaling Pathway. The molecules that are different shades of red color indicate up-regulation in the comparison of TB vs other diseases

FIGURE 9 Concordance and discordance of the log2FoldChange of statistically differentially expressed genes in TB vs *Mtb* infection against the log2FoldChange of corresponding genes in OD vs *Mtb* infection from Anderson et al datasets. Each dot is colored according to the disco score and represents a gene: the stronger the red color the more concordantly regulated is the gene pair; the stronger the blue color, the more discordantly regulated is the gene pair
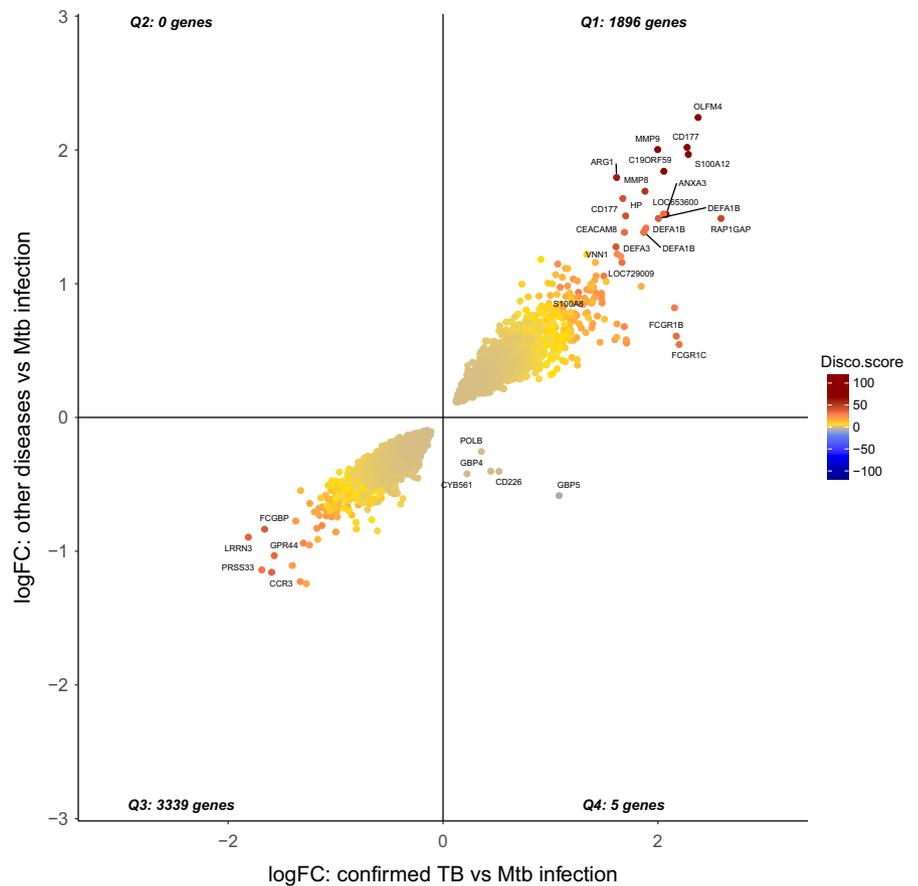
progression and treatment response on therapy for TB disease. Children with unconfirmed TB fulfil established consensus clinical and radiological criteria for disease but without microbiological confirmation.[17,130] Children with confirmed disease tend to have more severe disease than those with unconfirmed disease, the disease has usually been present for longer, symptoms are often more
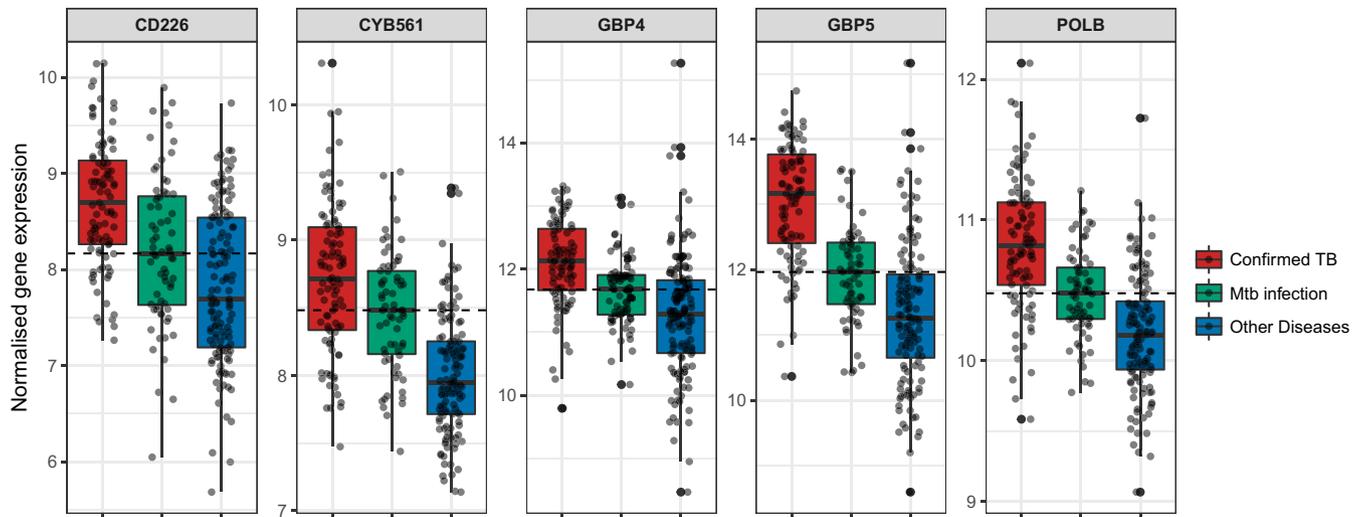
FIGURE 10 Boxplots of the 5 discordant genes between two comparisons TB vs *Mtb* infection (baseline) and OD vs *Mtb* infection (baseline) from Anderson et al datasets. The bar within each box indicates the median score, the bottom and top of the box indicate the interquartile range, the bars below and above the box are at a distance of 1.5 times the interquartile range from the upper and lower edges of the box, and the circles indicate outliers; the horizontal line across the graph indicates the median score of the *Mtb* infected group

pronounced and there are likely more bacilli in the child. Many biomarker studies have been developed comparing children with confirmed disease to children felt very unlikely to have TB disease. When tests are evaluated in children with a diagnosis of unconfirmed TB, performance is generally less good. However, it is in these children that tests are most needed. There are several explanations for why performance may be less good, and it is likely that all are to some extent evident. First, all children with unconfirmed disease have TB but have a milder form of TB that has a less pronounced transcriptomic signature than children with confirmed disease. Second, some of the children with unconfirmed disease have TB but the overall signature for children with unconfirmed disease is "diluted" by the children who do not actually have TB at all. Finally, the interaction between host and pathogen in unconfirmed TB is actually subtly different to the interaction seen in confirmed disease.

One approach to try to statistically deal with the analysis of diagnostic studies in which the "gold" standard is imperfect is using latent class analysis.[131] Schumacher and colleagues have previously employed Bayesian latent class analysis to estimate the diagnostic accuracy of mycobacterial culture, smear microscopy, Xpert MTB/RIF (Cepheid Inc.), TST, and chest radiography in childhood PTB, in a cohort of 748 hospitalized South African children. Transcriptomic signatures have not yet been evaluated using latent class analysis to define the reference standard. In the Anderson study,[99] the authors addressed the challenge of evaluating the 51-gene expression biomarker signature in children treated for TB, but without microbiological confirmation, by considering different scenarios where this population of children included varying combinations of "true" TB cases, and children incorrectly assigned as having TB by the clinical features. They used a range of estimates of the true prevalence of TB amongst those considered clinically as having highly probable TB, probable TB or possible TB, and used these

estimates to evaluate the sensitivity of the TB Disease Risk Score, as well as to identify the children likely to have TB (High TB Disease Risk Score) and those unlikely to have TB (Low Disease Risk Score). The gradient in the performance of the risk score in the culture-negative groups was consistent with the different degrees of diagnostic certainty in each group. Another approach would be to evaluate change in transcriptomic signature early in treatment to see if some children in the unconfirmed group have changes that cluster with the confirmed group and some have signatures that cluster with the unlikely TB group.

8.2 | Development of point-of-care tests

The use of these biomarkers in a clinical decision process either as standalone diagnostic tools or in conjunction with other tools needs to be studied further. Ultimately, prospective studies would be required in which the decisions about whether and when to initiate TB treatment are evaluated when using the new biomarkers. A concern in using transcriptional signatures as clinical diagnostic tools in resource poor settings is the complexity, cost, and time needed for the current methodologies for isolating and quantifying RNA from blood. The approaches described above have generally collected samples from children/adolescents and then divided the population into distinct clinical groups before identifying the minimal number of transcripts that can effectively discriminate these groups. A discovery/test approach is commonly used prior to validation in an external cohort. However, the next step necessary to make a signature like this useful clinically is to translate it into a true point-of-care test (POCT). For RT-qPCR based platforms, individual primers to each transcript need to be designed and these need to be then tested experimentally and validated externally. Platforms need to be robust, affordable,
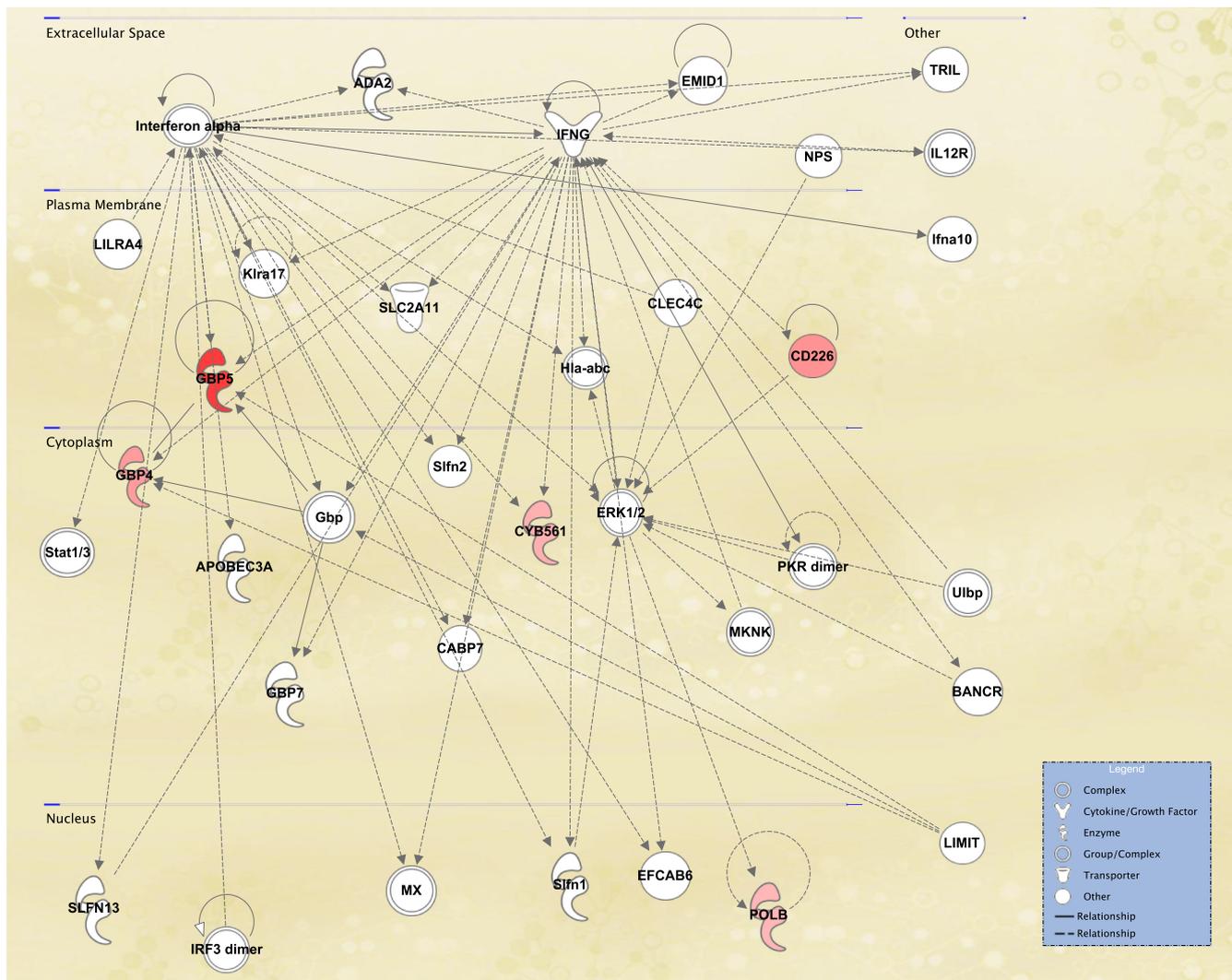
FIGURE 11 Molecular network, constructed from the discordant genes identified in the disco plot from Anderson et al dataset. The molecules that are different shades of red color indicate up-regulation in the comparison of TB vs other diseases. The default number of molecules per network has been used, $n = 35$

reliable, and able to be used in low resource settings with intermittent power supply and frequently at the extremes of temperature and humidity. Only a relatively small number of transcripts (typically fewer than ten) can be included in RT-qPCR POCTs for feasibility of engineering. POCTs have mainly, but not exclusively, been developed by commercial groups, with expertise in diagnostic test development and steps are currently being made toward cost-effective ways of measuring gene expression in the field efficiently by using recent advances in biotechnology and engineering.[132] An early-prototype cartridge-based assay, the "Xpert MTB Host Response" Cepheid (Sunnyvale, CA, USA) measuring the Sweeney 3-gene signature discriminated adults with TB disease from controls amongst people living with HIV (>18 years) with an AUC of 0.89 (CI$_{95\%}$ 0.83–0.94) and AUC of 0.84 (CI$_{95\%}$ 0.79–0.89) for active case finding in a high burden setting.[133,134] Similar studies on pediatric populations are needed to assess the potential of these signatures in pediatric TB.

8.3 | Evaluation in real-time as point-of-care tests

Once a POCT has been developed it must be evaluated clinically to determine discriminatory accuracy in real-world conditions, ease-of-use, resource utilization, time to clinical decision-making and ultimately clinical outcome. Direct translation into clinical studies in which treatment decisions have been led by biomarker-based POCTs can guide preventive therapy in healthy individuals. The CORTIS trial recruited 2923 HIV-negative adults. Individuals with a positive 11-transcript PT-qPCR test (the RISK11 signature) were randomized to TPT or no treatment while a sub-set of individuals with a negative RISK11 signature were followed with no treatment.[111] Those with a positive RISK11 signature were at increased risk of having prevalent TB and also of progressing to TB in the subsequent 15 months. However, when comparing rates of incident TB between the RISK11-positive adults treated with TPT and those not treated,

the risk of disease was not reduced. The research team are continuing to explore reasons for this unexpected finding.

Unlike in adults, however, it can be challenging to conduct a study in children where clinical decisions, particularly around which symptomatic child should be treated for TB disease, are based on a transcriptomic POCT. In addition, any test that is to be used to guide clinical decision-making would need to be approved by medical regulators in the country used. A potential interim step would be to carry out a Simulated Clinical Impact study in which the biomarker POCT is conducted in real-time in parallel to the routine standard of care decision-making. In that way, real patient information (from children recruited to a prospective cohort) is later presented to independent clinicians, and their clinical decisions, use of investigations, and therapeutic choices are compared when POCT results are made available at certain timepoints post presentation of child. Ultimately, biomarker tests need to be evaluated in comparison to other diagnostic approaches with the hard clinical endpoints being the outcomes of interest. Additional data on time to treatment initiation, preferences of children, parents, and healthcare worker, and costs need also to considered.

8.4 | Integration into treatment decision algorithms

In most clinical decision-making, the pre-test probability of a disease is combined with a test result to arrive at a post-test probability that then informs whether treatment should be started. For many children with presumptive TB, information from the clinical history and examination alone is sufficient to either reassure the healthcare worker that the child does not have TB or that the child is highly likely to have TB and should be started on treatment even if further tests were negative. In these instances, it is not necessary or helpful to perform testing. If the healthcare worker is not sufficiently reassured or convinced of the need to start treatment, then biomarker testing can occur, but the interpretation of that test result should include data from the clinical history and presentation as pre-test probability of disease. In addition, if further tests are performed after the first biomarker (which might include other biomarkers or radiological investigations) the interpretation of these results should include clinical information as well as results from any tests already performed. So, in addition to evaluating biomarkers for their ability to distinguish clinical states in children, it is important going forward to evaluate how biomarkers can be incorporated into treatment-decision algorithms. Data driven approaches have been developed to combine clinical, radiological, and microbiological data for the evaluation of child TB.[135,136] To date, these have not included transcriptomic biomarkers. An illustrative example is shown in Figure 12.

9 | CONCLUSIONS

Over the last ten years, transcriptomic approaches have led to the generation of multiple biomarkers in adults that can predict future TB disease progression, discriminate TB from other disease and monitor treatment response. Transcriptomics has also allowed for novel insights into the pathogenesis of *Mtb* infection and disease as well as better understanding of the host response to this pathogen. Work in children and adolescents has
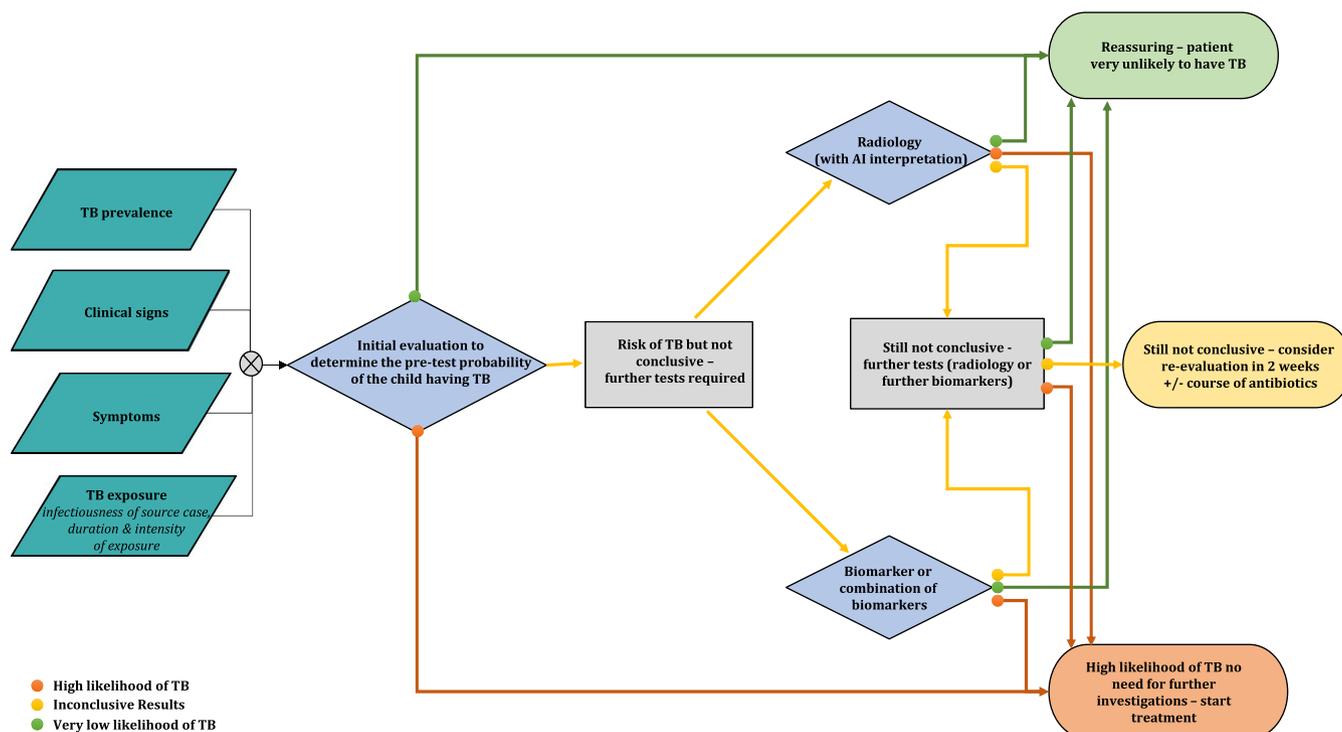

FIGURE 12 Illustration of an integrated TB treatment decision algorithm including biomarkers

lagged but several seminal studies have demonstrated that the host response to *Mtb* varies with age and the discovery of child-specific biomarkers requires child-specific studies. As these signatures are developed, they will need to be translated first into POCTs and then rigorously evaluated in the relevant clinical contexts, alone, and as part of integrated algorithms. In addition to the discovery of pediatric and adolescent biomarkers, transcriptomic studies in children are beginning to help us understand the biology of *Mtb* infection and disease in this age group, which will be vital to develop better vaccines and therapeutics. Transcriptomics has the potential to substantially contribute to meeting global End TB targets.

CONFLICTS OF INTEREST

MK, ML, and TS hold patents on several gene expression signatures. CB, OV, and JAS have no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study. Data presented in Figures 5-9 are publically available in Anderson et al [99].

ORCID

Myrsini Kaforou https://orcid.org/0000-0001-9878-4007
Claire Broderick https://orcid.org/0000-0003-2582-4527
Ortensia Vito https://orcid.org/0000-0001-5509-6351
Michael Levin https://orcid.org/0000-0003-2767-6919
Thomas J. Scriba https://orcid.org/0000-0002-0641-1359
James A. Seddon https://orcid.org/0000-0002-2296-2302

REFERENCES

1. Daniel TM. The history of tuberculosis. *Respir Med.* 2006;100(11):1862-1870.
2. Jones C, Whittaker E, Bamford A, Kampmann B. Immunology and pathogenesis of childhood TB. *Paediatr Respir Rev.* 2011;12(1):3-8.
3. Basu Roy R, Whittaker E, Seddon JA, Kampmann B. Tuberculosis susceptibility and protection in children. *Lancet Infect Dis.* 2019;19(3):e96-e108.
4. Seddon JA, Chiang SS, Esmail H, Coussens AK. The wonder years: what can primary school children teach us about immunity to mycobacterium tuberculosis? *Front Immunol.* 2018;9:2946.
5. Marais BJ, Gie RP, Schaaf HS, Hesseling AC, Enarson DA, Beyers N. The spectrum of disease in children treated for tuberculosis in a highly endemic area. *Int J Tuberc Lung Dis.* 2006;10(7):732-738.
6. WHO. *Global Tuberculosis Report 2020.* World Health Organization; 2020. https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf. Accessed February 3, 2021.
7. Snow KJ, Sismanidis C, Denholm J, Sawyer SM, Graham SM. The incidence of tuberculosis among adolescents and young adults: a global estimate. *Eur Respir J.* 2018;51(2):1702352. doi:10.1183/13993003.02352-2017
8. Dodd PJ, Yuen CM, Sismanidis C, Seddon JA, Jenkins HE. The global burden of tuberculosis mortality in children: a mathematical modelling study. *Lancet Glob Health.* 2017;5(9):e898-e906.
9. Hamada Y, Glaziou P, Sismanidis C, Getahun H. Prevention of tuberculosis in household members: estimates of children eligible for treatment. *Bull World Health Organ.* 2019;97(8):534-547D.
10. WHO. *WHO Consolidated Guidelines on Tuberculosis: Tuberculosis Preventive Treatment.* World Health Organization; 2020. https://www.ncbi.nlm.nih.gov/books/NBK554956/pdf/Bookshelf_NBK554956.pdf. Accessed September 8, 2020.
11. WHO. *The End TB Strategy.* World Health Organization; 2015. https://www.who.int/tb/End_TB_brochure.pdf?ua=1. Accessed February 23, 2020.
12. McQuaid CF, Vassall A, Cohen T, Fiekert K, White RG. The impact of COVID-19 on TB: a review of the data. *Int J Tuberc Lung Dis.* 2021;25(6):436-446.
13. Togun T, Kampmann B, Stoker NG, Lipman M. Anticipating the impact of the COVID-19 pandemic on TB patients and TB control programmes. *Ann Clin Microbiol Antimicrob.* 2020;19(1):21.
14. Zumla A, Marais BJ, McHugh TD, et al. COVID-19 and tuberculosis–threats and opportunities. *Int J Tuberc Lung Dis.* 2020;24(8):757-760.
15. *The Potential Impact Of The Covid-19 Response On Tuberculosis In High-Burden Countries: A Modelling Analysis.* Stop TB Partnership; 2020.
16. Zar HJ, Hanslo D, Apolles P, Swingler G, Hussey G. Induced sputum versus gastric lavage for microbiological confirmation of pulmonary tuberculosis in infants and young children: a prospective study. *Lancet.* 2005;365(9454):130-134.
17. Graham SM, Cuevas LE, Jean-Philippe P, et al. Clinical case definitions for classification of intrathoracic tuberculosis in children: an update. *Clin Infect Dis.* 2015;61(Suppl 3):S179-S187.
18. WHO. *Guidance for National Tuberculosis Programme on the Management of Tuberculosis in Children,* 2nd edition. World Health Organization; 2014. http://apps.who.int/iris/bitstream/10665/112360/1/9789241548748_eng.pdf?ua=1. Accessed June 11, 2018.
19. Fox GJ, Barry SE, Britton WJ, Marks GB. Contact investigation for tuberculosis: a systematic review and meta-analysis. *Eur Respir J.* 2013;41(1):140-156.
20. Morrison J, Pai M, Hopewell PC. Tuberculosis and latent tuberculosis infection in close contacts of people with pulmonary tuberculosis in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Infect Dis.* 2008;8(6):359-368.
21. Martinez L, Cords O, Horsburgh CR, Andrews JR, Pediatric TBCSC. The risk of tuberculosis in children after close exposure: a systematic review and individual-participant meta-analysis. *Lancet.* 2020;395(10228):973-984.
22. Marais BJ, van Zyl S, Schaaf HS, van Aardt M, Gie RP, Beyers N. Adherence to isoniazid preventive chemotherapy: a prospective community based study. *Arch Dis Child.* 2006;91(9):762-765.
23. Seddon JA, Whittaker E, Kampmann B, et al. The evolving research agenda for paediatric tuberculosis infection. *Lancet Infect Dis.* 2019;19(9):e322-e329.
24. Drain PK, Bajema KL, Dowdy D, et al. Incipient and subclinical tuberculosis: a clinical review of early stages and progression of infection. *Clin Microbiol Rev.* 2018;31(4):e00021-18. doi:10.1128/CMR.00021-18

25. Jenkins HE, Yuen CM, Rodriguez CA, et al. Mortality in children diagnosed with tuberculosis: a systematic review and meta-analysis. *Lancet Infect Dis*. 2017;17(3):285-295.

26. Chiang SS, Khan FA, Milstein MB, et al. Treatment outcomes of childhood tuberculous meningitis: a systematic review and meta-analysis. *Lancet Infect Dis*. 2014;14(10):947-957.

27. Snow K, Hesseling AC, Naidoo P, Graham SM, Denholm J, du Preez K. Tuberculosis in adolescents and young adults: epidemiology and treatment outcomes in the Western cape. *Int J Tuberc Lung Dis*. 2017;21(6):651-657.

28. Chaves Torres NM, Quijano Rodríguez JJ, Porras Andrade PS, Arriaga MB, Netto EM. Factors predictive of the success of tuberculosis treatment: a systematic review with meta-analysis. *PLoS ONE*. 2019;14(12):e0226507.

29. A controlled trial of a 2-month, 3-month, and 12-month regimens of chemotherapy for sputum smear-negative pulmonary tuberculosis: the results at 30 months. Hong Kong chest service/tuberculosis research Centre, Madras/British Medical Research Council. *Am Rev Respir Dis*. 1981;124(2):138-142.

30. Turkova A, Wills GH, Wobudeya E, et al. Shorter treatment for nonsevere tuberculosis in African and Indian children. *N Engl J Med*. 2022;386(10):911-922.

31. WHO. *Rapid Communication on Updated Guidance on the Management of Tuberculosis in Children and Adolescents*. World Health Organization; 2021. https://www.who.int/publications/i/item/9789240033450. Accessed November 2, 2021.

32. Meghji J, Simpson H, Squire SB, Mortimer K. A systematic review of the prevalence and pattern of imaging defined post-TB lung disease. *PLoS ONE*. 2016;11(8):e0161176.

33. Romanowski K, Baumann B, Basham CA, Ahmad Khan F, Fox GJ, Johnston JC. Long-term all-cause mortality in people treated for tuberculosis: a systematic review and meta-analysis. *Lancet Infect Dis*. 2019;19(10):1129-1137.

34. Wallis RS, Ginindza S, Beattie T, et al. Adjunctive host-directed therapies for pulmonary tuberculosis: a prospective, open-label, phase 2, randomised controlled trial. *Lancet Respir Med*. 2021;9(8):897-908.

35. van der Zalm MM, Walters E, Claassen M, et al. High burden of viral respiratory co-infections in a cohort of children with suspected pulmonary tuberculosis. *BMC Infect Dis*. 2020;20(1):924.

36. McArdle AJ, Turkova A, Cunnington AJ. When do co-infections matter? *Curr Opin Infect Dis*. 2018;31(3):209-215.

37. Whittaker E, Lopez-Varela E, Broderick C, Seddon JA. Examining the complex relationship between tuberculosis and other infectious diseases in children. *Front Pediatr*. 2019;7:233.

38. Dangor Z, Izu A, Moore DP, et al. Temporal association in hospitalizations for tuberculosis, invasive pneumococcal disease and influenza virus illness in south African children. *PLoS ONE*. 2014;9(3):e91464.

39. Yen YF, Pan SW, Su VY, Chuang PH, Feng JY, Su WJ. Influenza vaccination and incident tuberculosis among elderly persons, Taiwan(1). *Emerg Infect Dis*. 2018;24(3):498-505.

40. Redford PS, Mayer-Barber KD, McNab FW, et al. Influenza a virus impairs control of mycobacterium tuberculosis coinfection through a type I interferon receptor-dependent pathway. *J Infect Dis*. 2014;209(2):270-274.

41. Muller J, Tanner R, Matsumiya M, et al. Cytomegalovirus infection is a risk factor for tuberculosis disease in infants. *JCI. Insight*. 2019;4(23):e130090. doi:10.1172/jci.insight.130090

42. Martinez L, Nicol MP, Wedderburn CJ, et al. Cytomegalovirus acquisition in infancy and the risk of tuberculosis disease in childhood: a longitudinal birth cohort study in Cape Town, South Africa. *the Lancet Global Health*. 2021;9(12):e1740-e1749.

43. Dodd PJ, Prendergast AJ, Beecroft C, Kampmann B, Seddon JA. The impact of HIV and antiretroviral therapy on TB risk in children: a systematic review and meta-analysis. *Thorax*. 2017;72(6):559-575.

44. WHO. *WHO COVID-19 Dashboard*. World Health Organisation; 2020. https://covid19.who.int/. Accessed July 5, 2022.

45. Tadolini M, Codecasa LR, García-García JM, et al. Active tuberculosis, sequelae and COVID-19 co-infection: first cohort of 49 cases. *Eur Respir J*. 2020;56(1):2001398. doi:10.1183/13993003.01398-2020

46. Visca D, Ong CWM, Tiberi S, et al. Tuberculosis and COVID-19 interaction: a review of biological, clinical and public health effects. *Pulmonology*. 2021;27(2):151-165.

47. Bostanghadiri N, Jazi FM, Razavi S, Fattorini L, Darban-Sarokhalil D. Mycobacterium tuberculosis and SARS-CoV-2 coinfections: a review. *Front Microbiol*. 2022;12:747827. doi:10.3389/fmicb.2021.747827

48. Western cape Department of Health in collaboration with the National Institute for communicable diseases SA. Risk factors for coronavirus disease 2019 (COVID-19) death in a population cohort study from the Western Cape Province, South Africa. *Clin Infect Dis*. 2020;73(7):e2005-e2015.

49. Mousquer GT, Peres A, Fiegenbaum M. Pathology of TB/COVID-19 co-infection: the phantom menace. *Tuberculosis (Edinb)*. 2021;126:102020.

50. Tamuzi JL, Ayele BT, Shumba CS, et al. Implications of COVID-19 in high burden countries for HIV/TB: a systematic review of evidence. *BMC Infect Dis*. 2020;20(1):744.

51. Gao Y, Liu M, Chen Y, Shi S, Geng J, Tian J. Association between tuberculosis and COVID-19 severity and mortality: a rapid systematic review and meta-analysis. *J Med Virol*. 2021;93(1):194-196.

52. Riou C, du Bruyn E, Stek C, et al. Relationship of SARS-CoV-2-specific CD4 response to COVID-19 severity and impact of HIV-1 and tuberculosis coinfection. *J Clin Invest*. 2021;131(12):e149125. doi:10.1172/JCI149125

53. Petrone L, Petruccioli E, Vanini V, et al. Coinfection of tuberculosis and COVID-19 limits the ability to in vitro respond to SARS-CoV-2. *Int J Infect Dis*. 2021;113(Suppl 1):S82-s87.

54. Rosas Mejia O, Gloag ES, Li J, et al. Mice infected with mycobacterium tuberculosis are resistant to acute disease caused by secondary infection with SARS-CoV-2. *PLoS Pathog*. 2022;18(3):e1010093.

55. Dheda K, Lenders L, Srivastava S, et al. Spatial network mapping of pulmonary multidrug-resistant tuberculosis cavities using RNA sequencing. *Am J Respir Crit Care Med*. 2019;200(3):370-380.

56. Maji A, Misra R, Kumar Mondal A, et al. Expression profiling of lymph nodes in tuberculosis patients reveal inflammatory milieu at site of infection. *Sci Rep*. 2015;5:15214.

57. Lai RP, Cortes T, Marais S, et al. Transcriptomic characterization of tuberculous sputum reveals a host Warburg effect and microbial cholesterol catabolism. *bioRxiv*. 2020:2020.2003.2009.983163.

58. Deffur A, Wilkinson RJ, Coussens AK. Tricks to translating TB transcriptomics. *Ann Transl Med*. 2015;3(Suppl 1):S43.

59. Skogholt AH, Ryeng E, Erlandsen SE, Skorpen F, Schonberg SA, Saetrom P. Gene expression differences between PAXgene and tempus blood RNA tubes are highly reproducible between independent samples and biobanks. *BMC Res Notes*. 2017;10(1):136.

60. Carrillo-Avila JA, de la Puente R, Catalina P, et al. Evaluation of RNA purification methods by using different blood stabilization tubes: identification of key features for epidemiological studies. *BMC Res Notes*. 2020;13(1):77.

61. Broderick C, Cliff JM, Lee JS, Kaforou M, Moore DA. Host transcriptional response to TB preventive therapy differentiates two sub-groups of IGRA-positive individuals. *Tuberculosis (Edinb)*. 2021;127:102033.

62. Li Q, Yu L, Xin T, et al. Increased IL-9 mRNA expression as a biomarker to diagnose childhood tuberculosis in a high burden settings. *J Infect*. 2015;71(2):273-276.

63. Montaldo P, Kaforou M, Pollara G, et al. Whole blood Gene expression reveals specific transcriptome changes in neonatal encephalopathy. _Neonatology_. 2019;115(1):68-76.

64. Bustin SA. _AZ of Quantitative PCR_. International University Line La Jolla; 2004.

65. Moniri A, Miglietta L, Holmes A, Georgiou P, Rodriguez-Manzano J. High-level multiplexing in digital PCR with intercalating dyes by coupling real-time kinetics and melting curve analysis. _Anal Chem_. 2020;92(20):14181-14188.

66. Moniri A, Miglietta L, Malpartida-Cardenas K, et al. Amplification curve analysis: data-driven multiplexing using real-time digital PCR. _Anal Chem_. 2020;92(19):13134-13143.

67. Geiss GK, Bumgarner RE, Birditt B, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. _Nat Biotechnol_. 2008;26(3):317-325.

68. Petrilli JD, Araujo LE, da Silva LS, et al. Whole blood mRNA expression-based targets to discriminate active tuberculosis from latent infection and other pulmonary diseases. _Sci Rep_. 2020;10(1):22072.

69. Duggan DJ, Bittner M, Chen YD, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. _Nat Genet_. 1999;21:10-14.

70. Trevino V, Falciani F, Barrera-Saldana HA. DNA microarrays: a powerful genomic tool for biomedical and clinical research. _Mol Med_. 2007;13(9–10):527-541.

71. Kaforou M, Wright VJ, Oni T, et al. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. _PLoS Med_. 2013;10(10):e1001538.

72. Berry MP, Graham CM, McNab FW, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. _Nature_. 2010;466(7309):973-977.

73. Lodish HF. _Molecular Cell Biology_. 7th ed. W.H. Freeman and Co.; 2013.

74. Zhao S, Zhang Y, Gamini R, Zhang B, von Schack D. Evaluation of two main RNA-seq approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus rRNA depletion. _Sci Rep_. 2018;8(1):4781.

75. Debey S, Zander T, Brors B, Popov A, Eils R, Schultze JL. A highly standardized, robust, and cost-effective method for genome-wide transcriptome analysis of peripheral blood applicable to large-scale clinical trials. _Genomics_. 2006;87(5):653-664.

76. Grabherr MG, Haas BJ, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. _Nat Biotechnol_. 2011;29(7):644-652.

77. Gilbert JA, Hughes M. Gene expression profiling: metatranscriptomics. _Meth Mol Biol (Clifton, N J)_. 2011;733:195-205.

78. Gosalbes MJ, Durban A, Pignatelli M, et al. Metatranscriptomic approach to analyze the functional human gut microbiota. _PloS One_. 2011;6(3):e17447.

79. Conesa A, Madrigal P, Tarazona S, et al. A survey of best practices for RNA-seq data analysis. _Genome Biol_. 2016;17:13.

80. Saeys Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. _Bioinformatics_. 2007;23(19):2507-2517.

81. Altman DG, Bland JM. Diagnostic tests. 1: sensitivity and specificity. _BMJ_. 1994;308(6943):1552.

82. Altman DG, Bland JM. Diagnostic tests 2: predictive values. _BMJ_. 1994;309(6947):102.

83. Altman DG, Bland JM. Diagnostic tests 3: receiver operating characteristic plots. _BMJ_. 1994;309(6948):188.

84. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. _BMJ_. 2004;329(7458):168-169.

85. Gene OC. The Gene ontology project in 2008. _Nucleic Acids Res_. 2008;36(Database issue):D440-D444.

86. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. _Nucleic Acids Res_. 2000;28(1):27-30.

87. Kramer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. _Bioinformatics_. 2014;30(4):523-530.

88. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. _Proc Natl Acad Sci U S A_. 2005;102(43):15545-15550.

89. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. _Bioinformatics_. 2011;27(12):1739-1740.

90. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. _Nucleic Acids Res_. 2007;35(Web Server issue):W193-W200.

91. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. _Nucleic Acids Res_. 2009;37(1):1-13.

92. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. _Nat Methods_. 2015;12(5):453-457.

93. Yu H, Gu D, Yue C, Xu J, Yan F, He X. An Immune Cell-Based Signature Associating With EMT Phenotype Predicts Postoperative Overall Survival of ESCC. _Front Oncol_. 2021;11:636479. doi:10.3389/fonc.2021.636479

94. Chikina M, Zaslavsky E, Sealfon SC. CellCODE: a robust latent variable approach to differential expression analysis for heterogeneous cell populations. _Bioinformatics_. 2015;31(10):1584-1591.

95. Kang K, Meng Q, Shats I, et al. CDSeq: a novel complete deconvolution method for dissecting heterogeneous samples using gene expression data. _PLoS Comput Biol_. 2019;15(12):e1007510.

96. Gliddon HD, Herberg JA, Levin M, Kaforou M. Genome-wide host RNA signatures of infectious diseases: discovery and clinical translation. _Immunology_. 2018;153(2):171-178.

97. Verhagen LM, Zomer A, Maes M, et al. A predictive signature gene set for discriminating active from latent tuberculosis in Warao Amerindian children. _BMC Genomics_. 2013;14:74.

98. Maertzdorf J, Repsilber D, Parida SK, et al. Human gene expression profiles of susceptibility and resistance in tuberculosis. _Genes Immun_. 2011;12(1):15-22.

99. Anderson ST, Kaforou M, Brent AJ, et al. Diagnosis of childhood tuberculosis and host RNA expression in Africa. _N Engl J Med_. 2014;370(18):1712-1723.

100. Gjoen JE, Jenum S, Sivakumaran D, et al. Novel transcriptional signatures for sputum-independent diagnostics of tuberculosis in children. _Sci Rep_. 2017;7(1):5839.

101. Dhanasekaran S, Jenum S, Stavrum R, et al. Identification of biomarkers for mycobacterium tuberculosis infection and disease in BCG-vaccinated young children in southern India. _Genes Immun_. 2013;14(6):356-364.

102. Wang JX, Xu J, Han YF, Zhu YB, Zhang WJ. Diagnostic values of microRNA-31 in peripheral blood mononuclear cells for pediatric pulmonary tuberculosis in Chinese patients. _Genet Mol Res_. 2015;14(4):17235-17243.

103. Zhou M, Yu G, Yang X, Zhu C, Zhang Z, Zhan X. Circulating microRNAs as biomarkers for the early diagnosis of childhood tuberculosis infection. _Mol Med Rep_. 2016;13(6):4620-4626.

104. Jenum S, Dhanasekaran S, Lodha R, et al. Approaching a diagnostic point-of-care test for pediatric tuberculosis through evaluation of immune biomarkers across the clinical disease spectrum. _Sci Rep_. 2016;6:18520.

105. Berteloot L, Marcy O, Nguyen B, et al. Value of chest X-ray in TB diagnosis in HIV-infected children living in resource-limited countries: the ANRS 12229-PAANTHER 01 study. _Int J Tuberc Lung Dis_. 2018;22(8):844-850.

106. Machingaidze S, Wiysonge CS, Gonzalez-Angulo Y, et al. The utility of an interferon gamma release assay for diagnosis of latent tuberculosis infection and disease in children: a systematic review and meta-analysis. *Pediatr Infect Dis J*. 2011;30(8):694-700. doi:10.1097/INF.0b013e318214b915

107. Kay AW, González Fernández L, Takwoingi Y, et al. Xpert MTB/RIF and Xpert MTB/RIF ultra assays for active tuberculosis and rifampicin resistance in children. *Cochrane Database Syst Rev*. 2020;8(8):CD013359. doi:10.1002/14651858.CD013359.pub2

108. WHO. *WHO Operational Handbook on Tuberculosis. Module 5: Management of Tuberculosis in Children and Adolescents*. World Health Organization; 2022. https://www.who.int/publications/i/item/9789240046832. Accessed June 15, 2022.

109. Mulenga H, Zauchenberger C-Z, Bunyasi EW, et al. Performance of diagnostic and predictive host blood transcriptomic signatures for tuberculosis disease: a systematic review and meta-analysis. *PLoS ONE*. 2020;15(8):e0237574.

110. Zak DE, Penn-Nicholson A, Scriba TJ, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet*. 2016;387(10035):2312-2322.

111. Scriba TJ, Fiore-Gartland A, Penn-Nicholson A, et al. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled trial. *Lancet Infect Dis*. 2021;21(3):354-365.

112. Darboe F, Mbandi SK, Thompson EG, et al. Diagnostic performance of an optimized transcriptomic signature of risk of tuberculosis in cryopreserved peripheral blood mononuclear cells. *Tuberculosis (Edinb)*. 2018;108:124-126.

113. Penn-Nicholson A, Mbandi SK, Thompson E, et al. RISK6, a 6-gene transcriptomic signature of TB disease RISK, diagnosis and treatment response. *Sci Rep*. 2020;10(1):8629.

114. Bayaa R, Ndiaye MDB, Chedid C, et al. Multi-country evaluation of RISK6, a 6-gene blood transcriptomic signature, for tuberculosis diagnosis and treatment monitoring. *Sci Rep*. 2021;11(1):13646.

115. Horne DJ, Royce SE, Gooze L, et al. Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. *Lancet Infect Dis*. 2010;10(6):387-394.

116. Tornheim JA, Madugundu AK, Paradkar M, et al. Transcriptomic profiles of confirmed Pediatric tuberculosis patients and household contacts identifies active tuberculosis, infection, and treatment response among Indian children. *J Infect Dis*. 2020;221(10):1647-1658.

117. Sheerin D, Abhimanyu, Peton N, et al. Immunopathogenic overlap between COVID-19 and tuberculosis identified from transcriptomic meta-analysis and human macrophage infection. *iScience*. 2022;25(6):104464. doi:10.1016/j.isci.2022.104464

118. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.

119. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics*. 2008;24(13):1547-1548.

120. Domaszewska T, Scheuermann L, Hahnke K, et al. Concordant and discordant gene expression patterns in mouse strains identify best-fit animal model for human tuberculosis. *Sci Rep*. 2017;7(1):12094.

121. Hemingway C, Berk M, Anderson ST, et al. Childhood tuberculosis is associated with decreased abundance of T cell gene transcripts and impaired T cell function. *PLoS ONE*. 2017;12(11):e0185973.

122. Rohlwink UK, Figaji A, Wilkinson KA, et al. Tuberculous meningitis in children is characterized by compartmentalized immune responses and neural excitotoxicity. *Nat Commun*. 2019;10(1):3767.

123. Baguma R, Mbandi SK, Rodo MJ, et al. Inflammatory determinants of differential tuberculosis risk in pre-adolescent children and young adults. *Front Immunol*. 2021;12:639965. doi:10.3389/fimmu.2021.639965

124. Bah SY, Forster T, Dickinson P, Kampmann B, Ghazal P. Meta-analysis identification of highly robust and differential immune-metabolic signatures of systemic host response to acute and latent tuberculosis in children and adults. *Front Genet*. 2018;9:457. doi:10.3389/fgene.2018.00457

125. Karczewski KJ, Snyder MP. Integrative omics for health and disease. *Nat Rev Genet*. 2018;19(5):299-310.

126. Singh A, Shannon CP, Gautier B, et al. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics*. 2019;35(17):3055-3062.

127. Mezlini AM, Wang B, Deshwar A, Morris Q, Goldenberg A. Identifying cancer specific functionally relevant miRNAs from gene expression and miRNA-to-gene networks using regularized regression. *PLoS One*. 2013;8(10):e73168.

128. Stephenson E, Reynolds G, Botting RA, et al. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat Med*. 2021;27(5):904-916.

129. Krishnan S, Queiroz ATL, Gupta A, et al. Integrative multi-omics reveals serum markers of tuberculosis in advanced HIV. *Front Immunol*. 2021;12:676980.

130. Graham SM, Ahmed T, Amanullah F, et al. Evaluation of tuberculosis diagnostics in children: 1. Proposed clinical case definitions for classification of intrathoracic tuberculosis disease. Consensus from an expert panel. *J Infect Dis*. 2012;205(Suppl 2):S199-S208.

131. Schumacher SG, van Smeden M, Dendukuri N, et al. Diagnostic test accuracy in childhood pulmonary tuberculosis: a Bayesian latent class analysis. *Am J Epidemiol*. 2016;184(9):690-700.

132. Pennisi I, Rodriguez-Manzano J, Moniri A, et al. Translation of a host blood RNA signature distinguishing bacterial from viral infection into a platform suitable for development as a point-of-care test. *JAMA Pediatr*. 2021;175(4):417-419.

133. Södersten E, Ongarello S, Mantsoki A, et al. Diagnostic accuracy study of a novel blood-based assay for identification of tuberculosis in people living with HIV. *J Clin Microbiol*. 2021;59(3):e01643-20. doi:10.1128/JCM.01643-20

134. Moreira FMF, Verma R, Pereira Dos Santos PC, et al. Blood-based host biomarker diagnostics in active case finding for pulmonary tuberculosis: a diagnostic case-control study. *EClinicalMedicine*. 2021;33:100776.

135. Gunasekera KS, Walters E, van der Zalm MM, et al. Development of a treatment-decision algorithm for HIV-uninfected children evaluated for pulmonary tuberculosis. *Clin Infect Dis*. 2021;73(4):e904-e912. doi:10.1093/cid/ciab018

136. Marcy O, Borand L, Ung V, et al. A treatment-decision score for HIV-infected children with suspected tuberculosis. *Pediatrics*. 2019;144(3):e20182065. doi:10.1542/peds.2018-2065

137. Fletcher HA, Filali-Mouhim A, Nemes E, et al. Human newborn bacille Calmette–Guérin vaccination and risk of tuberculosis disease: a case-control study. *BMC Med*. 2016;14(1):76.

138. Johnson WE, Odom A, Cintron C, et al. Comparing tuberculosis gene signatures in malnourished individuals using the TBSignatureProfiler. *BMC Infect Dis*. 2021;21(1):106.