Indian J Microbiol
https://doi.org/10.1007/s12088-023-01143-z

REVIEW ARTICLE

Next-Gen Dual Transcriptomics for Adult Extrapulmonary Tuberculosis Biomarkers and Host–Pathogen Interplay in Human Cells: A Strategic Review

Manohar Nesakumar[1] · Elizabeth Hanna Luke[1] · Umashankar Vetrivel[1]

Received: 27 April 2023 / Accepted: 9 November 2023

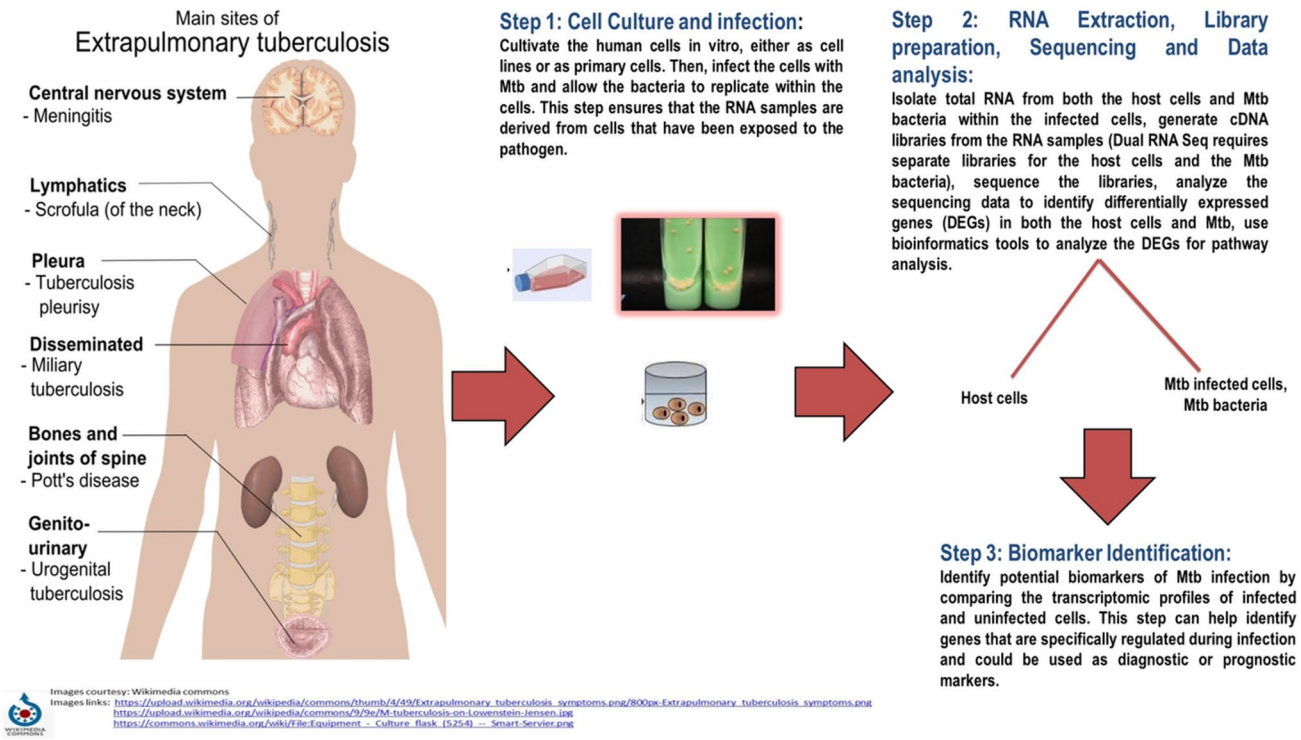
Abstract Tuberculosis (TB) is a major public health concern that results in significant morbidity and mortality, particularly in middle- to low-income countries. Extrapulmonary tuberculosis (EPTB) in adults is a form of TB that affects organs other than the lungs and is challenging to diagnose and treat due to a lack of accurate early diagnostic markers and inadequate knowledge of host immunity. Next-generation sequencing-based approaches have shown potential for identifying diagnostic biomarkers and host immune responses related to EPTB. This strategic review discusses on the significance using primary human cells and cell lines for in vitro transcriptomic studies on common forms of EPTB, such as lymph node TB, brain TB, bone TB, and endometrial TB to derive potential insights. While organoids have shown promise as a model system, primary cell lines still remain a valuable tool for studying host–pathogen interplay due to their conserved immune system, non-iPSC origin, and lack of heterogeneity in cell population. This review outlines a basic workflow for researchers interested in performing transcriptomics studies in EPTB, and also discusses the potential of cell-line based dual RNA-Seq technology for deciphering comprehensive transcriptomic signatures, host–pathogen interplay, and biomarkers from the host and *Mycobacterium tuberculosis*. Thus, emphasizing the implementation of this technique which can significantly contribute to the global anti-TB effort and advance our understanding of EPTB.

✉ Umashankar Vetrivel
umashankar.v@icmr.gov.in

[1] Department of Virology and Biotechnology, Bioinformatics Division, Indian Council for Medical Research-National Institute for Research in Tuberculosis (ICMR-NIRT), Chennai, India

Published online: 14 December 2023

Graphical Abstract



Keywords Transcriptomics · Biomarkers · Dual-RNA sequencing · Extra-pulmonary tuberculosis · Primary human cell cultures · Human cell lines · Tuberculosis

Introduction

Until the coronavirus (COVID-19) pandemic struck, tuberculosis (TB) was the leading cause of mortality worldwide arising due to single infecting pathogen [1]. The impact of COVID-19 is evident in the WHO Global TB Report 2022 [1]. The report reveals a concerning trend: the estimated number of TB-related deaths increased between 2019 and 2021, reversing the declining trend observed from 2005 to 2019. In 2021, approximately 1.4 million deaths occurred among HIV-negative individuals, and 187,000 deaths were recorded among HIV-positive individuals, resulting in a combined total of 1.6 million deaths. Furthermore, an estimated 10.6 million people contracted TB in 2021, marking a 4.5% increase from the 10.1 million cases reported in 2020. Notably, the TB incidence rate, measured as new cases per 100,000 population per year, rose by 3.6% between 2020 and 2021, reversing a two-decade decline of approximately 2% per year. Apart from infecting the lungs, TB bacilli can disseminate to virtually any tissue in the human body and affect other organs, causing extra-pulmonary tuberculosis (EPTB). Disease manifestations of EPTB includes TB meningitis (Brain-TB), cervical-lymphadenitis (neck-lymph node TB), ocular TB (TB of eye), oral TB (TB of the tongue or mouth), TB pleuritis (pleura covering the lungs), TB pericarditis (TB of the pericardium of heart), cutaneous TB (skin TB), musculoskeletal TB (bone/muscle TB), abdominal TB (TB of the gastrointestinal tract and peritoneum), genitourinary TB (TB of kidneys and genital tracts of both sexes) and miliary TB. EPTB contributed to 16% of the 7.5 million incident TB cases reported globally in 2019 [2]. In the recent years, there has not only been an increase in the number of EPTB cases, but also a raise in the number of drug-resistant (DR) EPTB cases. Recent studies have reported 16–20% drug resistance among EPTB cases in general, with maximum resistance reported to be isoniazid (INH) (8–14%) [3–6]. As a complex and multifaceted disease condition by itself, the management of drug resistance in EPTB warrants more research efforts for the development of novel anti-TB drugs as well as newer biomarkers and tests for improved EPTB diagnosis.

The major limitation for such research is the challenge associated with sampling from the different EPTB sites that involves invasive, costly, and laborious procedures. The aim of this narrative strategic review is to hint the readers on the complexity associated with diagnosis of adult EPTB, and also to briefly summarize the role of human cell lines and primary cells in generating holistic *Mycobacterium tuberculosis* (Mtb) -host transcriptomic data that could be exploited to identify new biomarkers and identification of human immune signatures at multiple infecting sites. By choosing an appropriate cell line or primary cells and carefully designing the experimental study, researchers can delineate valuable insights on host–pathogen molecular interplay occurring at various EPTB infection sites, thereby shall open up potential avenues to ponder towards novel and effective therapeutic modalities.

The Challenge of Transcriptomics Studies in Extra-Pulmonary Tuberculosis

EPTB represents Mtb infection that primarily spreads either via the lymphatic system or through blood from the lung to other parts of the body. Some of the major challenges encountered in the diagnosis of EPTB are listed below:

(a) Since EPTB is relatively rare most clinicians lack experience in the diagnosis of EPTB.
(b) EPTB is usually encountered with a variety of symptoms that mimic symptoms of not only respiratory but also other organ pathologies based on the site of EPTB.
(c) Obtaining an appropriate sample is difficult in EPTB cases.

Diagnosis of EPTB depends much on smear studies, bacterial culture and nucleic acid amplification tests (NAAT). The major drawback of the currently available microbiological tests for the diagnosis of EPTB, as these methods feature poor sensitivity and specificity [7]. Given this scenario, the research fraternity is now poised with a need to identify novel diagnostic markers, as well as surrogate markers for predicting the treatment outcome and also to assess the efficacy of new TB therapeutics and vaccines. This brings exceptional value for transcriptomics studies in EPTB. However, the biggest challenge for conducting transcriptomics studies for EPTB is sample collection from extrapulmonary sites, as it involves invasive and laborious procedures. For example, specimen collection for diagnosis of extra-thoracic lymph node TB involves direct puncture of the lymph node; similarly, for intra-thoracic lymph node TB, the sample collection procedure involves endobronchial ultrasound or surgical mediastinoscopy. Again for diagnosis of pleural TB, surgical or medical thoracoscopy and or needle biopsy are required. Hence, alternate non-invasive procedures for collecting diagnostic specimens becomes safer options.

In recent years, several studies have investigated potential biomarkers that could aid in differential diagnosis and treatment monitoring of tubercular lymphadenitis (TBL). Mustafa et al. [8] reported a five blood biomarker biosignature (MIG, IP-10, MIF, CCL22 and CCL23) that shall be used as an indicator of treatment success in TBL and TB pleuritis. A study by Babu et al. [9] inferred that systemic levels of matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinase (TIMPs): MMP-1, 9 and TIMP-2, 4 could clearly discriminate PTB from EPTB, LTB and HCs, wherein, MMP-13 and TIMP-2 could clearly discriminate EPTB from LTB and HCs. Hence, MMPs and TIMPs are

Table 1 Comparison of primary human cells and immortalized cell lines (This comparative table describes various factors that researchers should have in mind while planning in vitro transcriptomics studies for EPTB, so as to decide on the choice of primary human cells or immortalized cell lines for the proposed research)

Parameter	Primary human cells	Immortalized cell lines
Source	Isolated directly from human tissues	Derived from primary cells or tissues
Lifespan	Limited lifespan	Unlimited proliferation potential
Budget requirements	Higher initial buying and culturing costs; higher budget required for study	Lower and affordable
Expertise required	Requires extensive training, expertise in cell culturing, and careful standardization	Easy to work with
Genetic stability	Genetically stable	Prone to genetic alterations
Culturing conditions	Requires specialized media, growth factors, low serum or serum-free media; conditions must be standardized for each cell type	Standard media and conditions provided by the manufacturer or literature
Result relevance	High relevance to in vivo microenvironment; outcomes closely reflect in vivo host outcomes	Low relevance
Reproducibility of results	Subject to donor-to-donor variation	More standardized and reproducible
Availability	Limited	Unlimited
Ethical consideration	Obtained with ethical considerations	Independent of ethical concerns

considered as potential candidates for non-sputum-based biomarkers for differentiating PTB and EPTB from LTB and HC individuals. In other studies, the authors have also shown that the blood levels of adipocytokines (except resistin) to have significantly elevated in TBL condition than pulmonary-TB and to get reversed on post treatment [10]. Moreover, it also decreased anti-microbial peptides (AMPs) levels on ATT [11]. In addition, Abhimanyu et al. [12] proposed a three analyte serum biosignature comprising of IL-8, IL-10 and TNF-β as blood biomarkers in tubercular lymphadenitis. However, many of these inferences need to be explored and validated larger prospective studies. Earlier studies have shown that strains of Mtb differ in their affinity for different infection sites [13], and thereby identification of novel biomarkers for different EPTB site becomes critical. In this line, transcriptomics data generated from different human cell lines and primary cell culture models infected with Mtb becomes a prominent method for identification of host/pathogen bio-signatures, that could be effectively pondered for early diagnosis and treatment of different types of EPTB. However, transcriptomics data generated from human cell lines and primary cell culture models infected with Mtb for host/pathogen biosignature identification in EPTB also have limitations. Firstly, these models may not fully replicate in vivo complexity, leading to divergent gene expression patterns and potential misrepresentation of EPTB responses. Secondly, artificial culture conditions in vitro can introduce non-physiological influences on gene expression, yielding biased results. Moreover, the absence of a functional immune microenvironment hampers accurate representation of intricate pathogen-immune interactions in EPTB. Batch variability among cell lines or primary cells can further confound results, challenging reproducibility. Additionally, the dynamic nature of EPTB over time is often inadequately captured by static in vitro models. This limitation extends to the translation of in vitro gene expression changes to direct correlation with clinical outcomes or EPTB biomarkers in patients. In addition to this, identification of reliable biomarkers from in vitro models necessitates rigorous validation in animal models and patient cohorts. Despite these limitations, in vitro models can provide valuable insights into the molecular mechanisms underlying EPTB. Integrating transcriptomics data from different model systems with other omics data and clinical information is essential for a more comprehensive understanding of EPTB pathogenesis and the identification of meaningful biomarkers.

Study Design (Infection-Model) Considerations

A variety of host model systems are available for TB studies. Of these, cultured cell lines constitute the most commonly used models because of their simplicity, affordability, scalability and stability. However, the major drawback of cell lines is that they do not fully recapitulate the host–pathogen interactions occurring in primary human cells. The difference between primary cells and cell lines have been described elsewhere in literature [14–16], but a few important points are listed in Table 1 below. The decision to use human primary cells or cell lines for a particular study has to be made based on the nature of research and study requirements, although primary human cells (PHCs) provide an improved and valuable tool for analyzing host–pathogen interactions, and identifying valuable biomarkers for diagnosis and prognosis of disease.. The comparative Table 1 provides a condensed representation of the essential factors to weigh when selecting an in vitro model system, as well as the strengths and limitations inherent in each of these model systems.

Primary Human Cell Cultures and Immortal Cell Lines as Important Tools to Study Pathogenic Mechanisms in EPTB

Quite a number of transcriptomics studies have been conducted in different forms of EPTB, including TB of the lymph node [21–23], brain [24] and ocular [34] sites, but these studies have a few limitations. Firstly, transcriptomics studies on lymph node and brain TB only included patients with confirmed TB but does not comprise appropriate controls. Secondly, studies on ocular TB does not well utilize "big omics data" generated using current NGS approaches. However, the observed potential of these studies prompts the need for further discovery and validation studies on host-transcriptional signatures for various forms of EPTB, so as to identify robust candidates that can be incorporated into point-of-care diagnostic tests for EPTB.

Based on the literature review, we collated the most commonly used host–pathogen models for transcriptomics studies namely, primary human cell cultures and cell lines, in order to apprise the researchers who intend to perform studies on transcriptome profiling in different forms of EPTB (Table 2). Majority of the invitro models have been cited in various studies and therefore can be useful for any prospective researcher to use as experimental platforms with minimal or no standardisation.

*As far as HGC-27 cell line referenced [26] in Table-2, there aren't many well-established human cell lines derived from lymph node, and no single cell line can serve as complete invitro model for biomarker purposes because of its presence at multiple sites in the system, we have referenced the HGC-27 cell line derived from metastatic lymph node of gastric cancer for the study of TB Lymphadenitis.

Table 2 Summary of few human primary cells and cell lines for EPTB research (with a focus on TB lymphadenitis, TB meningitis, Mucoskeletal, genital and ocular TB)

Types of EPTB	Infection site	Human primary cells			Immortalized human cell lines		
		Name	Media	Reference	Name	Media	References
TB lymphadenitis	Lymph node	Primary human lymphatic endothelial cells (LECs)	Endothelial cell medium (ECM)	Lerner et al. [25]	HGC-27*	NA	Akagi et al. [26]
TB Meningitis	Brain	Primary human brain microvascular endothelial cells (HBM ECs)	RPMI 1640 media	Jain et al. [27]	Human neuroblastoma SK-N-SH cells	DMEM	Randal et al. [28]
Muculoskeletal TB	Bone	Primary human osteoblast culture	DMEM supplemented with 2% Gluta MAX-I, 1 mM sodium pyruvate, and 1% FCS at 37 °C and 5% CO2	Duewelhenke et al., 2007 [29]	SaOS-2 human osteoblast cells (ATCC HTB-85)	McCoy's 5a medium with 15% FBS	Sarkar et al. [13]
					MC3T3 osteoblast cell line (ATCC CRL-2593)	α-MEM media containing 10% FBS at 37 °C in the presence of 5% CO2	Khushpreet Kaur et al. [30]
Genital TB	Endometrium	Primary human endometrial cells	DMEM:F12 (1:1) medium containing 10% (v/v) FCS, gentamicin (10 µg/ml), penicillin (100 IU/ml), streptomycin (100 µg/ml) and fungizone (2.5 µg/ml) in a 5% CO2 atmosphere at 37 °C on rat-tail collagen type I	Srivastava et al. (2013) [31]	Ishikawa cell lines	Minimal Essential medium supplemented with 2 mm L glutamine, 1X nonessential amino acids, 10 mM HEPES, 10% fetal bovine serum, and 1% penicillin streptomycin solution at 37 °C in 5% CO2	Meenu et al. [32]
Ocular TB	Eye	Primary human RPE cells (OZR1)	Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10% fetal bovine serum	Nora (2018) [33]	Adult retinal pigment epithelium-19 cell line (ARPE-19) (ATCC® CRL-2302™)	Dulbecco's Modified Eagle Medium Nutrient Mixture F-12 (DMEM/F-12) media (Thermo Scientific #12,400,024) supplemented with sodium bicarbonate (1.2 g/L), penicillin (100 U/ml), streptomycin (100 µg/ml), 10% FBS, and 5% CO2 at 37 °C	Abhishek et al. [34]

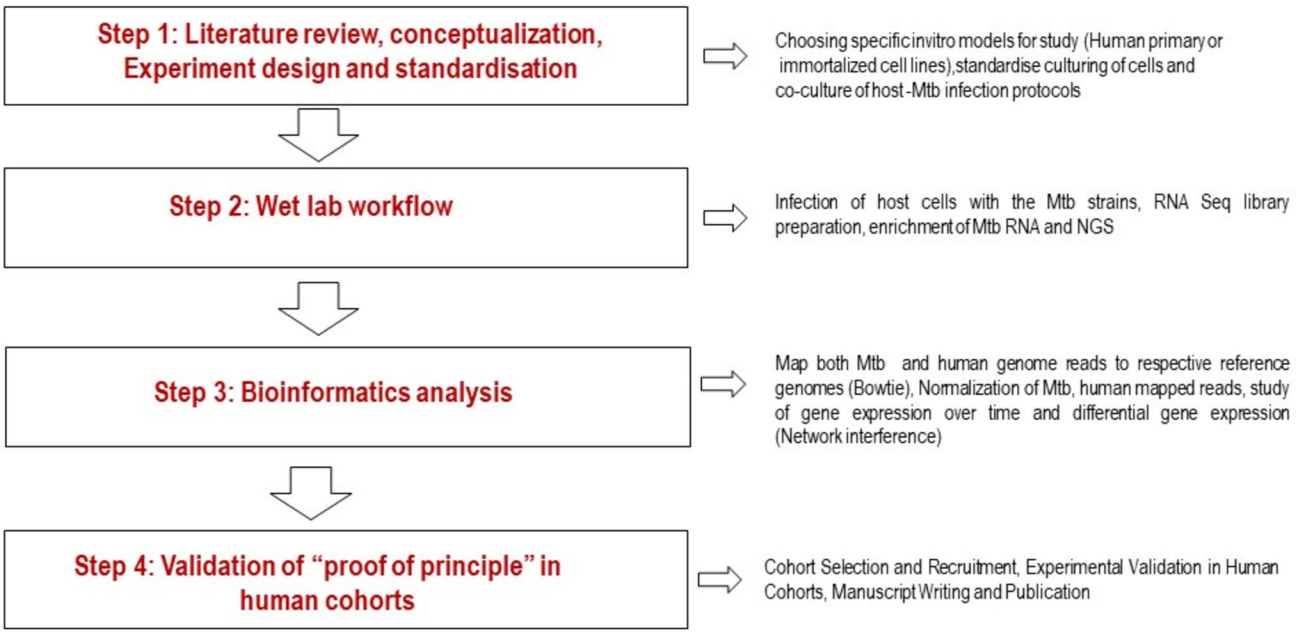


Fig. 1 Schematic flow chart of the proposed experimental approach to carry out host–pathogen interaction studies/biomarker/transcriptomics studies using dual RNA-seq approach in EPTB

Workflow for Transcriptomics Approaches

Transcriptomics studies in EPTB can be a powerful approach for discovery of both host and pathogen biomarkers. Designing transcriptomics studies in EPTB involves four key steps:

1. Selection of suitable cell lines/cell culture models relevant to the extra pulmonary site of infection
2. Propagation of cell lines/cell cultures and the use of appropriate infection protocols
3. Standardization of laboratory experimental conditions
4. Transcriptome profiling and biomarker identification

The basic workflow of experimental approaches for the identification of novel biomarkers and transcriptomic signatures for various forms of EPTB is outlined in Fig. 1.

The choice of cell line for transcriptomics studies should take into account the following: (i) the cell type should be present in abundance at the site of extra-pulmonary infection, and (ii) the cell should be a proven replicative niche for Mtb. Another key consideration for designing a successful transcriptomics study is the use of cell lines and primary cell cultures obtained from good commercial sources, as well as the use of well-standardized/published infection protocols and ideal experimental conditions. Besides, the study design should include adequate number of biological and technical replicates, and take into account optimization of workflow, biological variation between models of study, sampling and handling variation, RNA quality, transcriptome changes, appropriate quality controls for each type of experiment, sequencing depth as per the need of the study, use of validated computational pipelines for data analysis, etc.

Another key consideration in the design of transcriptomics studies is to determine whether the transcriptome data of the TB bacilli or the human cell or both need to be captured. For instance, studies aimed at understanding the mechanisms used by the TB bacillus for successfully infecting human cells and escaping the immunological surveillance mechanisms of the host, requires transcriptomics data for both the TB bacilli and the human cells. Thus, depending up on the research question, many types of data may be generated from transcriptomics studies. The main advantage of transcriptomics studies lie in its ability to capture novel isoforms occurring due to alternative splicing, single nucleotide polymorphisms and low abundant gene transcripts in the cell population. However, the major disadvantage is the inability to understand post-transcriptional modifications in the protein. Yet, due to the cost and complexity of proteome analysis, measuring gene expression is often considered a viable preliminary proxy in the scientific community for proteomic data, and is contextual based on the research question. In order to get a precise picture of the pathology of the disease at the cellular level and the host response to the pathogen, along with transcriptomics and proteomics analyses, additional immunological studies such as immune cell profiling, multi-cytokine profiling, etc. should be carefully planned. This would provide a complete picture of the "immunome"

of the host along with the "transcriptome and proteome" at the EPTB infection sites.

The traditional approach to unravel the host–pathogen transcriptomics involved on the RNA Seq datasets obtained on both host and the bacterium separately post-infection, followed by their individual analysis through sequencing- or probe-based technologies. In contrast, Dual RNA-seq allows the capturing and analysis of the transcriptomes of infecting pathogen, as well as the infected host simultaneously without the need of the separation of the both organisms. As discussed elsewhere [36–38], the typical workflow of a dual RNA sequencing experiment slightly differs from the original RNA sequencing experiment, in that the infected host cells and infecting pathogens are lysed together, which results in the release of total RNA, followed by ribosomal RNA (rRNA) depletion, library preparation and sequencing. Raw sequencing reads are stored in the form of FASTQ files and needs to be subjected to various quality checks, adapter trimming and filtering processes; the resulting datasets derived from host and the bacterial reads are aligned to their respective reference genomes and the transcripts are finally quantified.

A balanced and a holistic discerning of the transcriptomes of both host and pathogen can render valuable insights on identifying new pathogenic virulence factors, or new pathways in the host cells triggered due to the exposure to precise pathogens or pathogen-associated molecular patterns (PAMPs). This approach has made a major influence on better understanding of the transcriptomic interplay between host and the pathogen. Such studies have provided very valuable insights on the pathogenic mechanisms of Salmonella species, *S. pneumoniae*, uropathogenic *E. coli* and *Haemophilus influenza*. Dual rna-seq is also utilized to probe the host–pathogen interplay in intracellular pathogens like *Chlamydia trachomatis* and Mtb [38]. In case of Mtb, Dual rna-seq technology is applied in many recent studies: Pisu et al. [39] explored the *invivo* molecular dynamics of Mtb infection on Mtb infected murine-macrophage model; Rachel et al. [40] attempted to get a comprehensive picture of the entire ecosystem by generating RNA sequence data from sputum and analysing the transcripts of the pathogen, host and the microbiome community; Hong et al. [41] investigated the host–pathogen interactions at the transcriptomic level in an in vitro macrophage infection model; Victor et al. [42] showed that two closely related clinical strains spanning the Latin American and Mediterranean family of Mtb to arbitrate a discrete host response in human splenic macrophages despite their genetic similarity. Despite all these studies, there is dearth of knowledge on whole transcriptome (including small RNA) interplay snapshot on the direct infection of Mtb on different human cells and it becomes essential to know the transcriptomic landscape for deep understanding of pathogenic mechanisms, especially in EPTB. Hence, through this review, we explicitly provide the insights for those researchers who intend to unravel and study the host–pathogen interactions in conditions, where there is limited access to infected tissues like EPTB. In the study of Mtb-host interactions, Dual RNA-Seq emerges as a valuable technique by simultaneously profiling host and Mtb transcriptomes. This approach, however, presents certain limitations when compared with methods like bulk/single-cell RNA-Seq. Challenges include the intricate transcriptome complexity and associated computational demands, potential sensitivity issues towards low-abundance Mtb transcripts in infected tissues, and the dominance of host RNA, often masking pathogen-specific signals. Furthermore, accurately quantifying Mtb gene expression within the dynamic range of host–pathogen interactions can be intricate. While dual RNA-Seq may lack single-cell resolution, its capability to offer unique insights into Mtb-host interactions remains evident. In spite of these limitations, Dual RNA Seq continued to serve as a valuable tool for unraveling new insights into the dynamics of host–pathogen interactions during infection, identify critical pathogen adaptations to the intracellular host environment and corresponding host responses modulated by invading pathogen. We propose properly designed studies that harness the state of the art dual RNA sequencing coupled with cell culture technologies shall lead to the discovery of specific biomarkers and potential drug targets.

"Omics" Data Could be a Game-Changer in the Anti-TB Efforts Against EPTB

Eradication of TB and achieving the END-TB goal depends on identification and treatment of active TB cases. The strategic framework to END TB in India [43]. consists of four pillars: Pillar 1: build, strengthen and sustain an enabling environment for TB Elimination, Pillar 2: prevent the emergence of TB in susceptible populations using a combination of biomedical, behavioural, social and structural interventions, Pillar 3: detect all, that is, Early identification of presumptive TB, at the first point of contact be it private or public sectors, and prompt diagnosis using high sensitivity diagnostic tests to provide universal access to quality TB diagnosis including drug resistant TB in the country and Pillar 4: treat all. Delivering high-quality, patient-centered diagnostics and care at the initial healthcare facility visit is essential for augmenting the detection of TB cases and mitigating patient attrition during the early stages of the care cascade that is vital in the pursuit of End-TB target.

The management of TB treatment is further complicated by the potential for unfavorable outcomes such as treatment failure or disease recurrence. Hence, it is of paramount importance to closely monitor treatment response and ensure

its success. In this context, the significance of host/pathogen biomarkers becomes evident. These biomarkers hold promise for their potential to be translated into rapid point-of-care tests. Such tests can be employed for triaging and diagnosing TB, predicting the progression from infection to active disease, and monitoring treatment response. These capabilities constitute essential attributes for advancing TB elimination efforts. Nonetheless, it remains challenging to achieve early diagnosis of EPTB with existing diagnostic approaches, thereby presenting a major obstacle in realizing the objectives outlined in Pillar 3 of the END-TB/TB elimination strategy. In order to achieve the END-TB goal, intensified research towards the identification of new diagnostic biomarkers for various forms of EPTB is acutely needed. Data derived from dual RNA sequencing experiments which simultaneously profile the host and bacterial transcripts at the site of infection would unveil new sights into the dynamics of host–pathogen interactions during the disease, identify pathogen adaptations to the host environment [44, 45], determine host responses to infection, and identify novel biomarkers for early diagnosis of the disease [46, 47].

Thus, a properly designed and executed in vitro "tissue or cell-transcriptomics" studies would help to advance our knowledge in EPTB research in the following ways:

(i) Characterizing tissue/cell-specific infection patterns leading to the identification of infection-specific metabolic alterations induced by the TB bacilli as in the case of pulmonary TB [48]
(ii) Predicting changes in gene expression levels in vivo by extrapolating the results obtained from the closest in vitro model system
(iii) Understanding the mechanism of latency
(iv) Characterizing how each mutant/clinical isolate infects various sites in the body
(v) Characterizing the difference in host-immune response at different infection sites
(vi) Understanding the function of host non-coding regulatory RNAs
(vii) Comprehending host-Mtb interactions at the cellular level
(viii) Identifying biomarkers for EPTB
(ix) Identifying biomarkers for treatment response
(x) Understanding the effect of BCG vaccination at various EPTB infection sites
(xi) Deciphering the mechanism of action of novel anti-TB compounds and evaluating their therapeutic efficacy.

Future Perspectives

Traditionally, two-dimensional (2D) cell cultures have been used as in vitro models for various cell response studies. However, the 2D systems cannot completely mimic the real-time in vivo cellular responses. To overcome this limitation, 3D cell culture models also referred to as spheroid or organoid cultures were developed [49, 50]. The goal of the 3D cell culture model is to recapitulate as closely as possible the microenvironment present in the in vivo system. By having adequate amount of biological/technical replicates, one can get statistically significant results using these models depending on the type of study. However, in macro-scale cultures, one may have to sacrifice either the accuracy of the model or the throughput. Microfluidic devices and micro patterned plates with extracellular matrix components, and cultures that are performed on spheroids embedded in ECM scaffolds, offer hybrid culture systems that unify the advantages of both platforms to form a complex microenvironment for 3D cell culture. These 3D cell culture technologies have the potential to improve model accuracy without compromising on the high-throughput capability. Although many reports suggest that 3D systems should be applied whenever possible, the choice of the technology or platform has to be decided by the researcher based on the specific application/objective of the experiment. There is no universal 3D platform for the study of EPTB; a researcher must choose a 3D cell culture model that would simulate a tissue-specific physiological or pathophysiological disease-specific microenvironment wherein the cells can proliferate, aggregate and differentiate. The model chosen by the researcher should essentially take into consideration cell-to-cell and cell-to-extracellular matrix (ECM) interactions, nutrient and metabolic waste gradients, and a combination of tissue-specific scaffolding cells [51]. The use of 3D cellular systems such as in vitro human granuloma model, bio electrospray 3D model [52], or "organ on a chip" model [53, 54] has a definitive edge over the 2D models. Overall, the 3D platforms are likely to provide an increasingly attractive alternative for 2D cell culture as the technology has the potential to capture biologically relevant factors and recapitulate the micro-environmental factors that resemble the in vivo condition when Mtb interrogates the host cell to cause disease pathology.

Discussion and Conclusion

Globally, EPTB accounts for a significant proportion of TB cases, and yet the diagnosis of EPTB suffers from poor performance of laboratory tests. This remains an obstacle to the END-TB goal even in low-incident countries. Various blood-based biomarkers hold promise for improving TB diagnosis, but many of these have been identified in case–control studies of pulmonary TB alone, for which reason the findings cannot be extrapolated to EPTB and culture-negative TB. While next-generation sequencing (NGS)-led "omic" approaches are now being consistently used in the field of

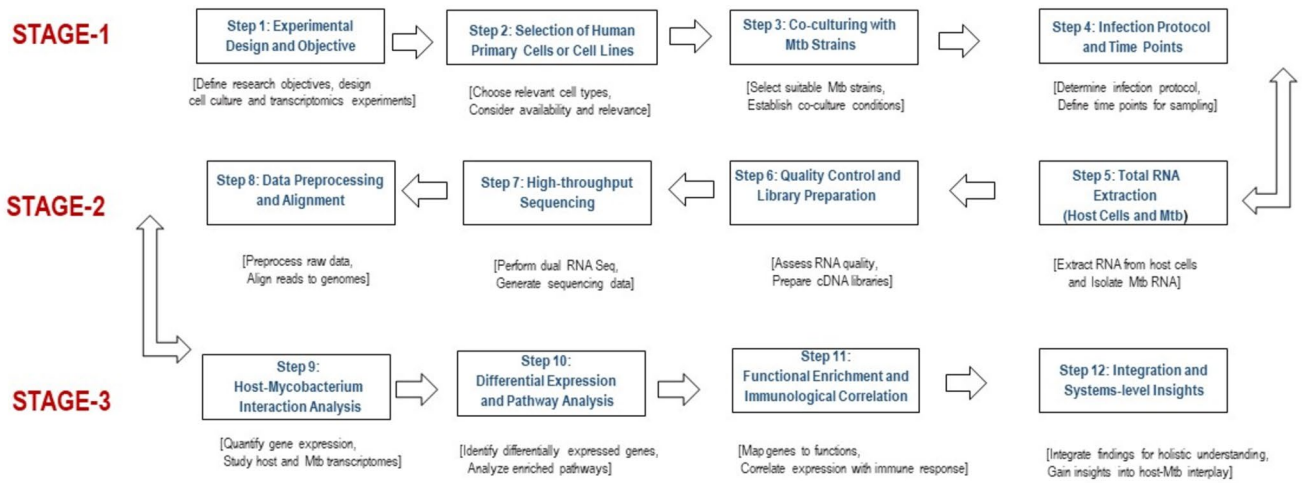


Fig. 2 Schematic roadmap of experimental approach, scientific and methodological considerations in transcriptomics workflow in EPTB research

biomedical research to address a variety of questions such as novel target discovery, protective biomarker identification, diagnostic biomarkers identification, understanding latency and reactivation, and so on, conducting such studies for EPTB remains a serious challenge due to sampling issues from extra-pulmonary infection sites in patients. This constitutes a major stumbling block in the control of EPTB disease. For a researcher interested in deciphering specific transcriptomes from different forms of EPTB for the purpose of understanding host–pathogen interactions or identification of novel diagnostic markers, the most critical step in the design of the study is to mine the literature for ideal cell culture models and infection protocols that can be used for the study of various types of EPTB. While organoids have shown promise as a model system, primary cell lines remain a valuable tool for studying host–pathogen interplay due to their conserved immune system, non-iPSC origin, and lack of heterogeneity in cell population. To our knowledge, this is the first snapshot review that provides a road-map for high throughput –omics studies aimed at biomarker discovery by offering the readers a catalogue of optimal host–pathogen infection models (primary human cell cultures and cell lines) and infection protocols relevant for the five most common forms of EPTB.

We used PubMed and Google Scholar for collecting articles related to this topic using relevant search terms reported in literature for identifying primary cells and cell lines that have been used as models for EPTB. In the PubMed database, we employed a variety of general search strings, such as ("human cell lines" OR "human cell cultures") AND "*Mycobacterium tuberculosis*" AND ("transcriptomics" OR "gene expression") AND ("Musculoskeletal TB" OR "Bone TB" OR "TB Meningitis" OR "Brain TB" OR "Ocular TB" OR "Genital TB" OR "TB lymphadenitis" OR "Lymph node TB"), along with specific search strings tailored to each EPTB subtype. These specific strings used for the search were as follows: ("human cell lines" OR "human cell cultures") AND "*Mycobacterium tuberculosis*" AND ("transcriptomics" OR "gene expression") AND "Bone TB"; ("human cell lines" OR "human cell cultures") AND "*Mycobacterium tuberculosis*" AND ("transcriptomics" OR "gene expression") AND "Brain TB"; ("human cell lines" OR "human cell cultures") AND "*Mycobacterium tuberculosis*" AND ("transcriptomics" OR "gene expression") AND "Ocular TB"; ("human cell lines" OR "human cell cultures") AND "*Mycobacterium tuberculosis*" AND ("transcriptomics" OR "gene expression") AND "Genital TB"; ("human cell lines" OR "human cell cultures") AND "*Mycobacterium tuberculosis*" AND ("transcriptomics" OR "gene expression") AND "Lymph node TB". These search strings were used to effectively combine different concepts using the "AND" operator and to encompass synonyms or related terms within each concept using the "OR" operator. For the Google Scholar database search, we employed the following search terms: "human cell lines" AND Mtb AND transcriptomics AND EPTB; Musculoskeletal TB cell lines AND Mtb AND transcriptomics; TB Meningitis cell lines AND Mtb AND transcriptomics; Ocular TB cell lines AND Mtb AND transcriptomics; Genital TB cell lines AND Mtb AND transcriptomics; TB lymphadenitis cell lines AND Mtb AND transcriptomics. The search and the selection process were not systematic. We meticulously reviewed the abstracts and methodologies of the identified key articles and subjected those to thorough scrutiny. The objective was to identify cell lines and cell culture models that could serve as valuable in vitro tools for researchers planning to undertake studies related to transcriptomic biomarker identification or host-Mtb interaction analysis within the context of the

five frequently occurring forms of EPTB. In our strategic-review, we have meticulously catalogued only those cell lines and culture models that have been previously explored and utilized in various EPTB studies. This compilation is intended to assist researchers during the initial conceptualization phase, aiding in the identification of suitable cell lines and co-culture conditions for transcriptomic biomarker identification or host-Mtb interaction studies on these five commonly occurring forms of EPTB. During data curation, we asked the following research questions:

- Q1: Why are human primary cells and immortalized cell lines important tools for transcriptomics studies in EPTB?
- Q2: Which model human primary cells or immortal cell lines could be the best option for transcriptomics studies on EPTB?
- Q3: What are the different human cell lines that have been used in literature to study various types of EPTB?
- Q4: What are the general technical considerations for conducting transcriptomics studies?
- Q5: What are the different applications of "tissue/cell-transcriptomics" studies in EPTB in advancing anti-TB efforts?

Through literature-mining, we identified answers to Q1–Q5 and have presented them in appropriate sections of this review, respectively. We have also provided the readers with references of studies that have utilized primary human cell cultures and immortalized human cell lines for EPTB studies (Table 1). These references can serve as a guiding resource for researchers planning to conduct transcriptomics studies in EPTB to choose the most appropriate cell /cell line, culture medium and infection protocol. The step-wise approach that may be adopted for generating transcriptomics data from EPTB sites for both the host and the TB bacilli is schematically represented in a flowchart (Fig. 1). The schematic roadmap of the experimental approach and and methodological considerations in the transcriptomics workflow are outlined in Fig. 2.

In conclusion, EPTB constitutes a significant proportion of global TB burden and curbing the pathogen in its "multifaceted dwelling sites" is vital to achieving the END-TB goal. EPTB still poses significant diagnostic challenges, and the development of new biomarkers using transcriptomic approaches is essential for reducing the mortality and morbidity associated with the disease. The use of human cell cultures and cell lines, and integrative exploitation of recent developments such as 3D cell culture technologies along with next generation sequencing (NGS) enabled approaches have great potential to support the study of genome-wide transcriptional changes in both Mtb and the infected host cell in EPTB sites. A properly designed and conducted biomarker study using an appropriate human cell line can generate valuable data on comprehensive transcriptomics signatures and biomarkers that would serve as proof-of principle for subsequent human studies.

Author Contributions MN: Literature search, writing original draft of the manuscript. LEH: Writing, Critical review and editing of the manuscript. UV: Conceptualization, Writing, Critical review and editing of the manuscript.

Funding We did not receive any funding for this study.

Declarations

Ethical approval No ethical approval was required as this is a review article with no original research data.

Conflict of Interest The authors declare no competing interests.

References:

1. WHO. Global Tuberculosis Report. (2022). https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022. (Accessed on Aug 23 2023).
2. WHO. Global Tuberculosis Report 2020. https://www.who.int/publications/i/item/9789240013131 (Accessed on MAY 07, 2022).
3. Raveendran R, Oberoi JK, Wattal C (2015) Multidrug-resistant pulmonary & extrapulmonary tuberculosis: a 13 years retrospective hospital-based analysis. Indian J Med Res 142(5):575–582. https://doi.org/10.4103/0971-5916.171285
4. Boonsarngsuk V, Mangkang K, Santanirand P (2018) Prevalence and risk factors of drug-resistant extrapulmonary tuberculosis. Clin Respir J 12(6):2101–2109. https://doi.org/10.1111/crj.12779
5. Mok J, Kang BH, Kim HJ, Lee SJ, Lee T, Lee HK, Cho YJ, Jeon D (2019) Drug resistance in extra-pulmonary tuberculosis in South Korea: comparison with pulmonary tuberculosis. Int J Tuberc Lung Dis Off J Int Union Against Tuberc Lung Dis 23(2):151–156. https://doi.org/10.5588/ijtld.18.0109
6. Diriba G, Kebede A, Tola HH, Yenew B, Moga S, Addise D, Alemu A, Mohammed Z, Getahun M, Fantahun M, Tadesse M, Dagne B, Amare M, Assefa G, Abera D, Desta K (2020) Molecular characterization and drug resistance patterns of *Mycobacterium tuberculosis* complex in extrapulmonary tuberculosis patients in Addis Ababa. Ethiopia PloS one 15(12):e0243493. https://doi.org/10.1371/journal.pone.0243493
7. Norbis L, Alagna R, Tortoli E, Codecasa LR, Migliori GB, Cirillo DM (2014) Challenges and perspectives in the diagnosis of extrapulmonary tuberculosis. Expert Rev Anti Infect Ther 12(5):633–647. https://doi.org/10.1586/14787210.2014.899900
8. Ambreen A, Khaliq A, Naqvi SZH, Tahir A, Mustafa M, Chaudhary SU, Mirza S, Mustafa T (2021) Host biomarkers for monitoring therapeutic response in extrapulmonary tuberculosis. Cytokine 142:155499. https://doi.org/10.1016/j.cyto.2021.155499
9. Kathamuthu GR, Kumar NP, Moideen K, Nair D, Banurekha VV, Sridhar R, Baskaran D, Babu S (2020) Matrix metalloproteinases and tissue inhibitors of metalloproteinases are potential biomarkers of pulmonary and extra-pulmonary tuberculosis. Front Immunol 11:419. https://doi.org/10.3389/fimmu.2020.00419

10. Kathamuthu GR, Moideen K, Sridhar R, Baskaran D, Babu S (2022) Plasma adipocytokines distinguish tuberculous lymphadenitis from pulmonary tuberculosis. Tuberculosis (Edinb) 132:102161. https://doi.org/10.1016/j.tube.2021.102161
11. Kathamuthu GR, Moideen K, Sridhar R, Baskaran D, Babu S (2020) Diminished systemic levels of antimicrobial peptides in tuberculous lymphadenitis and their reversal after anti-tuberculosis treatment. Tuberculosis (Edinb) 122:101934. https://doi.org/10.1016/j.tube.2020.101934
12. Abhimanyu B, M., Varma-Basil, M., Jain, A., Sethi, T., Tiwari, P. K., Agrawal, A., Banavaliker, J. N., & Bhowmick, K. T. (2016) Establishment of Elevated Serum Levels of IL-10, IL-8 and TNF-β as Potential Peripheral Blood Biomarkers in Tubercular Lymphadenitis: A Prospective Observational Cohort Study. PLoS ONE 11(1):e0145576. https://doi.org/10.1371/journal.pone.0145576
13. Sarkar S, Dlamini MG, Bhattacharya D, Ashiru OT, Sturm AW, Moodley P (2016) Strains of *Mycobacterium tuberculosis* differ in affinity for human osteoblasts and alveolar cells in vitro. Springerplus 5:163. https://doi.org/10.1186/s40064-016-1819-z
14. Mitra A, Mishra L, Li S (2013) Technologies for deriving primary tumor cells for use in personalized cancer therapy. Trends Biotechnol 31(6):347–354. https://doi.org/10.1016/j.tibtech.2013.03.006
15. Gillooly JF, Hayward A, Hou C, Burleigh JG (2012) Explaining differences in the lifespan and replicative capacity of cells: a general model and comparative analysis of vertebrates. Proc Biol Sci 279(1744):3976–3980. https://doi.org/10.1098/rspb.2012.1129
16. Miserocchi G, Mercatali L, Liverani C, De Vita A, Spadazzi C, Pieri F, Bongiovanni A, Recine F, Amadori D, Ibrahim T (2017) Management and potentialities of primary cancer cultures in preclinical and translational studies. J Transl Med 15(1):229. https://doi.org/10.1186/s12967-017-1328-z
17. Ali R, Samman N, Al Zahrani H, Nehdi A, Rahman S, Khan AL, Al Balwi M, Alriyees LA, Alzaid M, Al Askar A, Boudjelal M (2017) Isolation and characterization of a new naturally immortalized human breast carcinoma cell line, KAIMRC1. BMC Cancer 17(1):803. https://doi.org/10.1186/s12885-017-3812-5
18. Verma A, Verma M, Singh A (2020) Animal tissue culture principles and applications. Anim Biotechnol. https://doi.org/10.1016/B978-0-12-811710-1.00012-4
19. Weigand A, Boos AM, Tasbihi K, Beier JP, Dalton PD, Schrauder M, Horch RE, Beckmann MW, Strissel PL, Strick R (2016) Selective isolation and characterization of primary cells from normal breast and tumors reveal plasticity of adipose derived stem cells. Breast Cancer Res BCR 18(1):32. https://doi.org/10.1186/s13058-016-0688-2
20. Petrini C (2012) Ethical and legal considerations regarding the ownership and commercial use of human biological materials and their derivatives. J Blood Med 3:87–96. https://doi.org/10.2147/JBM.S36134
21. Phelan JJ, O'Leary S, Keane J (2021) Tuberculosis lymph node granulomas: using transcriptomics to discover immunopathology paradigms and guide host-directed therapy. J Clin Investig 131(15):e151810. https://doi.org/10.1172/JCI151810
22. Reichmann MT, Tezera LB, Vallejo AF, Vukmirovic M, Xiao R, Reynolds J, Jogai S, Wilson S, Marshall B, Jones MG, Leslie A, D'Armiento JM, Kaminski N, Polak ME, Elkington P (2021) Integrated transcriptomic analysis of human tuberculosis granulomas and a biomimetic model identifies therapeutic targets. J Clin Investig 131(15):e148136. https://doi.org/10.1172/JCI148136
23. Maji A, Misra R, Kumar Mondal A, Kumar D, Bajaj D, Singhal A, Arora G, Bhaduri A, Sajid A, Bhatia S, Singh S, Singh H, Rao V, Dash D, Baby Shalini E, Sarojini Michael J, Chaudhary A, Gokhale RS, Singh Y (2015) Expression profiling of lymph nodes in tuberculosis patients reveal inflammatory milieu at site of infection. Sci Rep 5:15214. https://doi.org/10.1038/srep15214
24. Rohlwink UK, Figaji A, Wilkinson KA, Horswell S, Sesay AK, Deffur A, Enslin N, Solomons R, Van Toorn R, Eley B, Levin M, Wilkinson RJ, Lai RPJ (2019) Tuberculous meningitis in children is characterized by compartmentalized immune responses and neural excitotoxicity. Nat Commun 10(1):3767. https://doi.org/10.1038/s41467-019-11783-9
25. Lerner TR, de Souza Carvalho-Wodarz C, Repnik U, Russell MR, Borel S, Diedrich CR, Rohde M, Wainwright H, Collinson LM, Wilkinson RJ, Griffiths G, Gutierrez MG (2016) Lymphatic endothelial cells are a replicative niche for *Mycobacterium tuberculosis*. J Clin Investig 126(3):1093–1108. https://doi.org/10.1172/JCI83379
26. Akagi T, Kimoto T (1976) Human cell line (HGC-27) derived from the metastatic lymph node of gastric cancer. Acta Med Okayama 30(3):215–219
27. Jain SK, Paul-Satyaseela M, Lamichhane G, Kim KS, Bishai WR (2006) *Mycobacterium tuberculosis* invasion and traversal across an in vitro human blood-brain barrier as a pathogenic mechanism for central nervous system tuberculosis. J Infect Dis 193(9):1287–1295. https://doi.org/10.1086/502631
28. Randall PJ, Hsu NJ, Lang D, Cooper S, Sebesho B, Allie N, Keeton R, Francisco NM, Salie S, Labuschagné A, Quesniaux V, Ryffel B, Kellaway L, Jacobs M (2014) Neurons are host cells for *Mycobacterium tuberculosis*. Infect Immun 82(5):1880–1890. https://doi.org/10.1128/IAI.00474-13
29. Duewelhenke N, Krut O, Eysel P (2007) Influence on mitochondria and cytotoxicity of different antibiotics administered in high concentrations on primary human osteoblasts and cell lines. Antimicrob Agents Chemother 51(1):54–63. https://doi.org/10.1128/AAC.00729-05
30. Kaur K, Sharma S, Abhishek S, Kaur P, Saini UC, Dhillon MS, Karakousis PC, Verma I (2023) Metabolic switching and cell wall remodelling of *Mycobacterium tuberculosis* during bone tuberculosis. J Infect 86(2):134–146. https://doi.org/10.1016/j.jinf.2022.12.014
31. Srivastava A, Sengupta J, Kriplani A, Roy KK, Ghosh D (2013) Profiles of cytokines secreted by isolated human endometrial cells under the influence of chorionic gonadotropin during the window of embryo implantation. Reproduct Biol Endocrinol RB&E 11:116. https://doi.org/10.1186/1477-7827-11-116
32. Meenu S, Thiagarajan S, Ramalingam S, Michael A, Ramalingam S (2016) Modulation of host ubiquitin system genes in human endometrial cell line infected with *Mycobacterium tuberculosis*. Med Microbiol Immunol 205(2):163–171. https://doi.org/10.1007/s00430-015-0432-z
33. La Distia Nora R, Walburg KV, van Hagen PM, Swagemakers SMA, van der Spek PJ, Quinten E, van Velthoven M, Ottenhoff THM, Dik WA, Haks MC (2018) Retinal pigment epithelial cells control early *mycobacterium tuberculosis* infection via interferon signaling. Invest Ophthalmol Vis Sci 59(3):1384–1395. https://doi.org/10.1167/iovs.17-23246
34. Abhishek S, Saikia UN, Gupta A, Bansal R, Gupta V, Singh N, Laal S, Verma I (2018) Transcriptional profile of *Mycobacterium tuberculosis* in an in vitro model of intraocular tuberculosis. Front Cell Infect Microbiol 8:330. https://doi.org/10.3389/fcimb.2018.00330
35. Hopewell PC (1994) Overview of clinical tuberculosis. Tuberculosis, 25–46.
36. Barrios-Payán J, Saqui-Salces M, Jeyanathan M, Alcántara-Vazquez A, Castañon-Arreola M, Rook G, Hernandez-Pando R (2012) Extrapulmonary locations of *Mycobacterium tuberculosis* DNA during latent infection. J Infect Dis 206(8):1194–1205. https://doi.org/10.1093/infdis/jis381

37. Westermann AJ, Gorski SA, Vogel J (2012) Dual RNA-seq of pathogen and host. Nat Rev Microbiol 10(9):618–630. https://doi.org/10.1038/nrmicro2852
38. Westermann AJ, Vogel J (2018) Host-pathogen transcriptomics by dual RNA-seq. Methods Mol Biol 1737:59–75. https://doi.org/10.1007/978-1-4939-7634-8_4
39. Pisu D, Huang L, Rin Lee BN, Grenier JK, Russell DG (2020) Dual RNA-sequencing of *Mycobacterium tuberculosis*-infected cells from a murine infection model. STAR protocols 1(3):100123. https://doi.org/10.1016/j.xpro.2020.100123
40. Lai RPJ, Cortes T, Marais S, Rockwood N, Burke ML, Garza-Garcia A, Horswell S, Sesay AK, O'Garra A, Young DB, Wilkinson RJ (2021) Transcriptomic characterization of tuberculous sputum reveals a host warburg effect and microbial cholesterol catabolism. mBio 12(6):e0176621. https://doi.org/10.1128/mBio.01766-21
41. Park HT, Lee SM, Ko S, Kim S, Park HE, Shin MK, Kim D, Yoo HS (2022) Delineating transcriptional crosstalk between *Mycobacterium avium* subsp. paratuberculosis and human THP-1 cells at the early stage of infection via dual RNA-seq analysis. Vet Res 53(1):71. https://doi.org/10.1186/s13567-022-01089-y
42. López-Agudelo VA, Baena A, Barrera V, Cabarcas F, Alzate JF, Beste DJV, Ríos-Estepa R, Barrera LF (2022) Dual RNA sequencing of *Mycobacterium tuberculosis*-infected human splenic macrophages reveals a strain-dependent host-pathogen response to infection. Int J Mol Sci 23(3):1803. https://doi.org/10.3390/ijms23031803
43. National Strategic Plan to end Tuberculosis in India (2020–25). Available online at: https://tbcindia.gov.in/index1.php?lang=1&level=2&sublinkid=5507&lid=3528
44. Rienksma RA, Suarez-Diez M, Mollenkopf HJ, Dolganov GM, Dorhoi A, Schoolnik GK, Martins Dos Santos VA, Kaufmann SH, Schaap PJ, Gengenbacher M (2015) Comprehensive insights into transcriptional adaptation of intracellular mycobacteria by microbe-enriched dual RNA sequencing. BMC Genomics 16(1):34. https://doi.org/10.1186/s12864-014-1197-2
45. Damron FH, Oglesby-Sherrouse AG, Wilks A, Barbier M (2016) Dual-seq transcriptomics reveals the battle for iron during *Pseudomonas aeruginosa* acute murine pneumonia. Sci Rep 6:39172. https://doi.org/10.1038/srep39172
46. Saliba AE, Li L, Westermann AJ, Appenzeller S, Stapels DA, Schulte LN, Helaine S, Vogel J (2016) Single-cell RNA-seq ties macrophage polarization to growth rate of intracellular Salmonella. Nat Microbiol 2:16206. https://doi.org/10.1038/nmicrobiol.2016.206
47. Avraham R, Haseley N, Brown D, Penaranda C, Jijon HB, Trombetta JJ, Satija R, Shalek AK, Xavier RJ, Regev A, Hung DT (2015) Pathogen cell-to-cell variability drives heterogeneity in host immune responses. Cell 162(6):1309–1321. https://doi.org/10.1016/j.cell.2015.08.027
48. Zimmermann M, Kogadeeva M, Gengenbacher M, McEwen G, Mollenkopf HJ, Zamboni N, Kaufmann SHE, Sauer U (2017) Integration of metabolomics and transcriptomics reveals a complex diet of *Mycobacterium tuberculosis* during early macrophage infection. mSystems 2(4):e00057-17. https://doi.org/10.1128/mSystems.00057-17
49. Scadden DT (2006) The stem-cell niche as an entity of action. Nature 441(7097):1075–1079. https://doi.org/10.1038/nature04957
50. Whitesides GM (2006) The origins and the future of microfluidics. Nature 442(7101):368–373. https://doi.org/10.1038/nature05058
51. Griffith LG, Swartz MA (2006) Capturing complex 3D tissue physiology in vitro. Nat Rev Mol Cell Biol 7(3):211–224. https://doi.org/10.1038/nrm1858
52. Elkington P, Lerm M, Kapoor N, Mahon R, Pienaar E, Huh D, Kaushal D, Schlesinger LS (2019) In vitro granuloma models of tuberculosis: potential and challenges. J Infect Dis 219(12):1858–1866. https://doi.org/10.1093/infdis/jiz020
53. Thacker VV, Dhar N, Sharma K, Barrile R, Karalis K, McKinney JD (2020) A lung-on-chip model of early *Mycobacterium tuberculosis* infection reveals an essential role for alveolar epithelial cells in controlling bacterial growth. eLife 9:e59961. https://doi.org/10.7554/eLife.59961
54. Ingber DE (2022) Human organs-on-chips for disease modelling, drug development and personalized medicine. Nat Rev Genet 23(8):467–491. https://doi.org/10.1038/s41576-022-00466-9