Correlates of tuberculosis risk: predictive biomarkers for progression to active tuberculosis

Elisa Petruccioli¹, Thomas J. Scriba²,³, Linda Petrone¹, Mark Hatherill²,³, Daniela M. Cirillo⁴, Simone A. Joosten⁵, Tom H. Ottenhoff⁵, Claudia M. Denkinger⁶ and Delia Goletti¹

Affiliations: ¹Dept of Epidemiology and Preclinical Research, Translational Research Unit, National Institute for Infectious Diseases "L. Spallanzani", Rome, Italy. ²South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine, Cape Town, South Africa. ³Division of Immunology, Dept of Pathology, University of Cape Town, Cape Town, South Africa. ⁴Emerging Bacterial Pathogens Unit, Division of Immunology and Infectious Diseases, San Raffaele Scientific Institute, HSR, Milan, Italy. ⁵Leiden University Medical Center, Leiden, The Netherlands. ⁶Tuberculosis and Hepatitis Programme, FIND, Geneva, Switzerland.

Correspondence: Delia Goletti, Translational Research Unit, National Institute for Infectious Diseases "L. Spallanzani", Dept of Epidemiology and Preclinical Research, Via Portuense 292, 00149, Rome, Italy. E-mail: delia.goletti@inmi.it

ABSTRACT  New approaches to control the spread of tuberculosis (TB) are needed, including tools to predict development of active TB from latent TB infection (LTBI). Recent studies have described potential correlates of risk, in order to inform the development of prognostic tests for TB disease progression. These efforts have included unbiased approaches employing "omics" technologies, as well as more directed, hypothesis-driven approaches assessing a small set or even individual selected markers as candidate correlates of TB risk. Unbiased high-throughput screening of blood RNAseq profiles identified signatures of active TB risk in individuals with LTBI, ≥1 year before diagnosis. A recent infant vaccination study identified enhanced expression of T-cell activation markers as a correlate of risk prior to developing TB; conversely, high levels of Ag85A antibodies and high frequencies of interferon (IFN)-γ specific T-cells were associated with reduced risk of disease. Others have described CD27⁻IFN-γ⁺CD4⁺ T-cells as possibly predictive markers of TB disease. T-cell responses to TB latency antigens, including heparin-binding haemagglutinin and DosR-regulon-encoded antigens have also been correlated with protection.

Further studies are needed to determine whether correlates of risk can be used to prevent active TB through targeted prophylactic treatment, or to allow targeted enrolment into efficacy trials of new TB vaccines and therapeutic drugs.

@ERSpublications

Promising biomarkers may allow accurate prediction of progression from infection to active TB disease http://ow.ly/OzCL304ezfk

Received: May 20 2016 | Accepted after revision: Sept 08 2016
Support statement: Funding for this article has been deposited with the Open Funder Registry.
Conflict of interest: Disclosures can be found alongside the online version of this article at erj.ersjournals.com
Copyright ©ERS 2016. This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0.
Introduction

Tuberculosis (TB) is the most significant infectious cause of mortality and morbidity worldwide. As reported by the World Health Organization (WHO), there are 10.4 million new cases of TB each year [1]. TB is a communicable disease, transmitted almost exclusively by cough aerosols carrying pathogens of the *Mycobacterium tuberculosis* complex. Pathogenesis is characterised by a period of asymptomatic subclinical infection, defined broadly as latent tuberculosis infection (LTBI), which might last for weeks or decades. From an operational point of view, LTBI may best be defined as a state of persistent immune response to *M. tuberculosis* antigens detected either by the tuberculin skin test (TST) or by interferon (IFN)-γ release assay (IGRA) without evidence of clinically manifest TB. Based on this definition, individuals with LTBI carry an increased risk of progression to TB. However, an unknown but large proportion of those with LTBI will not develop TB, either because their immune system persistently controls mycobacterial replication or because they are no longer infected with live bacteria. The risk of reactivation and subsequent disease and mortality is significantly increased in *M. tuberculosis*-infected individuals with immune suppression, due to HIV co-infection [2] or therapy with tumour necrosis factor (TNF)-α inhibitors [3–6], or other immune regulators used for inflammatory diseases and transplantation [7] or compromised immunity due to noncommunicable diseases, such as type 2 diabetes [8, 9]. Approaches to decrease TB morbidity and mortality, along with *M. tuberculosis* transmission, rely on correct diagnosis, effective treatment and prevention of infection and disease. Until recently, LTBI was thought to represent a uniform state [10]. However, it has become clear that LTBI has to be considered a broad spectrum of infection states that differ by the degree of the pathogen replication, host resistance and inflammation [10–13]. Similarly, TB disease presentation is similarly diverse and heterogeneous. This heterogeneity includes the type of pathology developed in the lungs (pulmonary TB) or outside the lungs (extrapulmonary TB), the spread of the lesions, the characteristics of immune activation and inflammation, *M. tuberculosis* replication and bacterial load. An increasing body of evidence supports the hypothesis that the genetic background of the host influences TB disease manifestations and the success of the bacterium in infecting large numbers of susceptible individuals [14–17]. An important hurdle in the field of TB research is that the bacteria cannot be directly detected *in vivo* during latent, asymptomatic *M. tuberculosis* infection. A test that could quantify bacterial numbers or levels of replication during LTBI would advance our ability to define the stages of infection, and allow more detailed studies of pathogenesis and immunity during asymptomatic *M. tuberculosis* infection.

Outcome of *M. tuberculosis* infection is therefore not a simple two-state distribution represented by either LTBI or active TB, but rather represents a continuous spectrum of states that differ by pathogen and host “activity”, which require different diagnostic and treatment strategies (figures 1 and 2).

From both a clinical and a research point of view, it would be important to correctly identify those *M. tuberculosis*-infected individuals who are most likely to progress to active disease in order to target them precisely with preventive treatment [18, 19]. Prevention of TB disease arising from latent infection is key to achieving WHO elimination targets [18, 20, 21], yet mass preventive therapy based on IGRA/TST screening in TB-endemic countries would need to treat 50–80% of the population, most of them unnecessarily, since 85–95% of latently infected persons will never develop disease in their lifetime [22–24]. The positive predictive value of IGRA/TST testing is thus too low, and better prognostic tests are required.

The TST and IGRA tests are based on immunological sensitisation to mycobacterial antigens. In the case of the TST this is quantified by the transverse diameter of skin induration resulting from intradermal injection of purified protein derivative, a crude mixture of antigens, many of which are shared by *M. tuberculosis*, *M. bovis*, bacille Calmette–Guérin (BCG) and several species of environmental mycobacteria. Blood-based IGRA, including QuantiFERON TB Gold in-tube (QFT-GIT; Qiagen, Hilden, Germany) and T-SPOT.TB (Oxford Immunotec, Abingdon, UK), measure *in vitro* IFN-γ production upon antigen stimulation of whole blood using ELISA, or peripheral blood mononuclear cells (PBMCs) using enzyme-linked immunospot (ELISPOT) assay, respectively [25]. In these assays, specificity for *M. tuberculosis* is derived from stimulation with peptides spanning the *M. tuberculosis* antigens ESAT-6, CFP-10 (and TB7.7 for QFT-GIT) that are restricted to a region of the *M. tuberculosis* genome deleted from *M. bovis* BCG and which is not present in most environmental mycobacteria [26–29].

A practical benefit of IGRA tests is that they require only a single laboratory test with negative and positive controls, and only one visit. Moreover, the *in vitro* tests may discriminate true negative responses from anergy [25]. An updated version of the QFT-GIT has been launched [25, 30]: the QuantiFERON TB Plus includes an additional antigen tube to QFT-GIT, which contains peptides that are intended to specifically induce a CD8+ T-cell response in addition to the CD4+ T-cell response [31] detected with the original QFT-GIT assay [32]. The rationale for the inclusion of these new CD8-specific peptides is derived from the growing evidence that *M. tuberculosis*-specific CD8+ T-cells are more frequently detected in subjects with active TB disease compared to LTBI [33–36], they are associated with a recent exposure to TB [37–39].
and they decline when patients receive anti-TB treatment [38]. The first data on performance of QuantiFERON TB Plus in a multicentre European study were reported recently [39, 40]. Additional promising tests for LTBI detection should be mentioned. C-Tb is a skin test [41] that measures hypersensitivity to recombinant ESAT-6 and CFP-10 proteins following intradermal administration. The authors claim that it combines the strengths and advantages of TST and IGRA technologies: low cost and for ease of use, like the TST, and high specificity analogous to IGRAs. Another test, based on Rv3615c encoded outside the RD1 region, has shown potential as a new T-cell based immunodiagnostic [42]. However, it should be noted that the TST and IGRAs share limitations. They both suffer from low accuracy in immune-compromised patients and cannot distinguish between LTBI and active TB disease [25].

FIGURE 1 Outcome of Mycobacterium tuberculosis transmission and establishment of infection or disease based on the correlates of disease and correlates of risk. The outcome of a primary or secondary M. tuberculosis infection is not a simple two-state distribution represented by either active tuberculosis (TB) or latent TB infection, but rather represents a continuous spectrum of states that differ by the degree of the pathogen replication, host resistance and inflammatory markers. The identification of M. tuberculosis infection is complex, due to the absence of clinical signs, correlates of disease (COD), lung lesions detected by chest radiography or M. tuberculosis in the sputum culture. The latency state is characterised by an immunological equilibrium and by presumed control of the bacterial replication. As the infection advances, this balance is lost, resulting in increased bacterial burden and/or increased pathology. This state can be identified as subclinical or incipient TB disease, in which CODs may still be poorly informative. In contrast, correlates of risk (COR) may potentially allow the identification of those at risk, for preventive treatment. Indeed, upregulation of interleukin (IL-13 and type I and II interferon (IFN)-related gene expression, elevated activation markers on T-cells (e.g. expression of D-related human leukocyte antigen and loss of CD27 expression), as well as an elevated monocyte/lymphocyte (M/L) ratio, have been shown to be predictive of TB disease development. The progression of subclinical TB to clinical TB is likely to be associated with a further increase in bacterial burden and/or pathology. Therefore, active TB diagnosis is based on CODs, including chest radiography findings such as lung lesions indicative of disease, detection of M. tuberculosis in sputum and positive COR tests. Transmission of M. tuberculosis from active TB patients may lead to a primary or secondary M. tuberculosis infection. Primary M. tuberculosis infection is defined by IFN-γ release assay (IGRA)/tuberculin skin test (TST) conversion and absence of radiological lung lesions and sputum negative for M. tuberculosis. ↓: downregulation; ↑: upregulation; +: presence of a modulation based on current knowledge; −: absence of a modulation based on current knowledge.
The latter is a major issue in TB-endemic areas and leads to poor predictive value for the development of TB in persons with LTBI [25, 43, 44]. These limitations provide a compelling reason why a simple measure of IFN-γ response does not allow correct evaluation of risk of progression from latency to active disease. It follows that the discovery of biomarkers that can differentiate between active and quiescent bacterial replication in persons with LTBI, or host markers that identify those with LTBI who are at risk of developing active disease [11, 45, 46] is of great importance. Moreover, it is critical that newly identified markers and signatures are validated in different geographic settings, because human populations with different genetic backgrounds and environmental influences, as well as different circulating M. tuberculosis lineages may be associated with variable rates of progression to TB disease [47].

The characteristics of new prognostic tests that predict progression from latent to active TB disease have recently received much attention. In 2015, the WHO convened an expert group to discuss the target product profile (TPP) of such a test under the lead of FIND, based on currently available knowledge. For such a test to have utility in high TB-burden settings, it should ideally be based on a sample type more easily accessible than sputum. While a high positive predictive value (PPV; >95%) for progression from infection to active TB (figures 3 and 4) would be ideal, such a high value is probably unachievable in the short term, and a more realistic goal has to be considered. The performance characteristics presented here and in figures 3 and 4 represent expert opinion (a draft version of the TPP is published on the FIND website: www.finddx.org/wp-content/uploads/2016/05/TPP-LTBIprogression.pdf) that is undergoing validation in a survey of global stakeholders. A sensitivity and specificity of >90% (and minimally 75% for both) were considered an appropriate target to obtain acceptable positive and negative predictive values.

If we assume a 2-year cumulative TB incidence of 2% and an effectiveness of isoniazid preventive treatment of 50% (figures 3 and 4) for the optimal target the number of patients needed to treat (NNT) would be 13 with a PPV of 16%. Similarly, in the same scenario, for the minimal target the NNT would be 40 (figure 3) with a PPV of 6% (figure 4). If we evaluate the same parameters for the current generations of IGRAsthat based on performance characteristics of IGRAsthat predict progressing of TB as outlined in WHO LTBI guidelines [48, 49], the PPV is 2–3% and NNT is 85. In parallel to the development of the TPP, a framework for the validation of such tests is being formulated [50]. Of course it is important to recognise that the performance characteristics of such new prognostic tests (i.e. PPV and NPV) will depend on the underlying prevalence of M. tuberculosis infection, the incidence of TB disease and the risk group.

Recent research activities have aimed at developing correlates of risk of TB, which are likely to inform the future development of prognostic tests. These efforts have included unbiased approaches employing "omics" technologies, as well as more directed, hypothesis-driven approaches that assessed either a small set or even individual selected markers as candidate correlates of risk (COR) of TB (table 1). In the following sections we describe some of the most promising approaches.

**Blood transcriptomic correlates of risk**

The COR of TB in the most advanced stages of development is a whole-blood transcriptomic mRNA expression signature that was identified by mining RNA-sequencing data in a large prospective cohort of
adolescents with LTBI from South Africa [51]. The RNA signature was discovered in a set of 46 progressors and 107 controls, and is based on a classifier that computes a TB risk score from relative expression levels of 63 mRNA transcripts from 16 genes in whole blood [52]. The 16-gene transcriptomic COR signature predicted progression from infection to TB disease with a sensitivity of 66% and a specificity of 81% in the 12 months preceding diagnosis of incident TB disease in the test cohort.

FIGURE 4 Positive predictive value (PPV) to identify cases who have latent tuberculosis (TB) and will develop active TB identified by currently available, novel and envisioned diagnostic tests. If we assume a 2-year cumulative TB incidence of 2% and an effectiveness of isoniazid preventive treatment of 50%, the optimal PPV is 16%. Similarly, in the same scenario, for the minimal target the PPV would be 6%. If we evaluate the same parameters for the current generations of interferon-γ release assays (IGRAs) [based on performance characteristics of IGRA for predicting progression of TB as outlined in World Health Organization latent TB infection guidelines [48, 49]], the PPV is 2%–3%. In parallel to the development of the target product profile (TPP), a framework for the validation of such tests is being formulated [50]. TST: tuberculin skin test; COR: correlate of risk for TB development.
Measurement of the 16-gene transcriptomic COR was transferred from measurement of mRNA expression by RNA sequencing (HiSeq2000; Illumina, San Diego, CA, USA) to a high-throughput, microfluidic, real-time PCR platform to allow cheaper and simpler measurement of gene expression. The PCR-based 16-gene transcriptomic COR signature was validated by blind prediction in two independent cohorts of South African and Gambian progressors and controls from a prospective household TB contact study, GC6 to 74 (the Biomarkers for TB consortium) [52]. In the validation cohorts the sensitivity was 54% and the specificity 83%. Ongoing analyses are under way to explore the biological processes that underlie progression in the adolescent study. However, it is notable that the 16 genes that comprise the COR signature are all regulated by type I and II IFNs, suggesting that chronic peripheral activation of the IFN response, previously shown to be associated with active TB disease at the time of diagnosis [53–56], also precedes the onset of active TB disease.

Application of correlates to clinical trials and prevention strategies

Further development of this PCR-based 16-gene transcriptomic COR included assessment of diagnostic performance against published microarray data from TB patients and healthy controls. In a set of HIV-uninfected South African adults including 130 prevalent TB cases and 230 controls, combined from four published studies [53–56], the COR discriminated between active TB and healthy controls with 87% sensitivity and 97% specificity. Therefore, although the prognostic 16-gene transcriptomic COR was not originally developed for diagnostic purposes, these performance characteristics suggest the signature has excellent potential as a nonsputum-based triage test to trigger investigation for undiagnosed TB disease. Thus, the 16-gene transcriptomic COR has potential to identify individuals at high risk of undiagnosed TB disease at the time of screening and in addition, if active disease has been excluded, to identify individuals

<table>
<thead>
<tr>
<th>TABLE 1 Candidate correlates of tuberculosis (TB) risk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biomarker</strong></td>
</tr>
<tr>
<td><strong>Validated correlates of TB risk</strong></td>
</tr>
<tr>
<td>Commercial or traditional tests for LTBI diagnosis</td>
</tr>
<tr>
<td>Molecular tests</td>
</tr>
<tr>
<td>Cell activation markers Blood cell counts</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Unvalidated correlates of TB risk</strong></td>
</tr>
<tr>
<td>Molecular tests</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Antigen-specific T-cells</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cell differentiation markers Serum/plasma cytokine tests Antigen-specific antibodies Responses to latency antigens</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CD8+ T-cell response</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

LTBI: latent TB infection; RD: region of difference; IGRA: interferon (IFN)-γ release assays; PPD: purified protein derivative; TST: tuberculin skin test; HLA-DR: D-related human leukocyte antigen; IL: interleukin; AIRE: autoimmune regulator; BCG: bacille Calmette–Guérin; Ag: antigen; Th: T-helper cell; IP: IFN-γ-inducible protein; Ig: immunoglobulin; PBMCs: peripheral blood mononuclear cells; HBHA: heparin-binding haemagglutinin.
at high risk for progression to incident TB disease within 12–18 months of screening. These performance characteristics would be ideal for a non-sputum-based triage test to identify persons for definitive, sputum-based investigation; and to identify those persons who would benefit maximally from targeted preventive therapy, thus avoiding unnecessary treatment of persons with LTBI who would remain healthy. Prevalence of COR status in young HIV-uninfected South Africans is ~12%, and is not different between IGRA-negative and IGRA-positive persons (Penn-Nicholson A and Scriba TJ, Cape Town, South Africa; personal communication). Given these criteria, screening with the 16-gene transcriptomic COR would allow several-fold reduction in the number of people requiring COR-targeted preventive therapy in high TB burden settings, compared to IGRA-targeted prevention strategies (figures 2 and 3).

The ideal COR-targeted preventive therapy regimen would be short, safe, sterilising and able to be rolled out rapidly with high coverage in serial mass campaigns. A clinical trial to test the efficacy of a 3-month short-course, once-weekly, high-dose, directly-observed regimen of isoniazid and rifapentine (3HP) for 16-gene transcriptomic COR-positive persons without active TB disease started in October 2016 (clinicaltrials.gov/NCT02735590). In parallel, efforts are ongoing to transfer COR testing to a point-of-care device that would allow mass COR screening by local healthcare providers. Translation to a point-of-care test may require a more parsimonious signature and efforts are underway to reduce the 16-gene transcriptomic COR to a smaller set of mRNA transcripts (Scriba T J et al., Cape Town, South Africa; personal communication). Whole-blood transcriptomic signatures for diagnosis of TB disease comprising either three [57] or four genes [58] have recently been validated, providing proof of concept that very small sets of genes have potential for promising diagnostic utility. If successful, the CORTIS trial will provide proof of concept for the potential of community-wide COR screening campaigns to target curative and preventive therapy for those individuals at highest risk of TB disease, and impact the global epidemic through early diagnosis and interruption of transmission (table 1).

Prediction of TB disease months before the onset of clinical symptoms in HIV-infected persons using host biomarker profiles was first demonstrated in a unique longitudinal cohort as part of the Amsterdam cohort studies [59-61]. HIV-infected, intravenous drug users had regular blood collections at municipal health services and the archive of these samples allowed selection of PBMCs 6 months to 1 year before the clinical diagnosis of TB. Gene expression profiling using dual-color reverse transcriptase multiplex ligation-dependent probe amplification [62] revealed a predictive signature that discriminated those who did develop TB in the next period from those that remained free of disease for at least 2 years following sample collection [60, 61]. Expression of interleukin (IL)-13 in the absence of AIRE (APECED, autoimmune regulator) was predictive for TB disease development within the next year in this high-risk population. Interestingly, when analysing individuals from the IL-13/AIRE group who progressed to TB in the following couple of months, there was also abundant expression of type I IFN-related genes, indicating detection of early disease markers in those individuals. Although this needs to be repeated in larger, longitudinal and independent cohorts, preferably also HIV-uninfected groups, it is an important proof of principle, illustrating the power of predictive TB biomarkers (table 1).

**Immune activation-based COR tests**

T-cell activation-based CORs of TB have also been the subject of investigation. A recent study, performed in BCG-vaccinated infants who participated in the recent phase IIb efficacy trial of MVA85A [63], compared a number of immunological outcomes as potential correlates of risk at trial baseline and after administration of the MVA85A boosting vaccine versus placebo. 53 infants who developed active TB disease and 205 controls who remained healthy were evaluated. An intriguing result was that infants who developed TB disease during follow-up had significantly higher levels of activated CD4+ T-cells, expressing D-related human leukocyte antigen at study baseline than infants who remained healthy [64, 65]. Importantly, this result was validated independently in the cohort of *M. tuberculosis*-infected adolescents discussed earlier, in whom elevated CD4+ T-cell activation was also found to correlate with risk of TB [64]. Elevated T-cell activation may thus reflect the same immunopathogenesis that underlies the observed upregulated IFN response in TB progressors from the adolescent study (table 1).

Taken together, the infant, adolescent, and the GC6-74 household contact studies all imply that persons at risk of developing active TB have evidence of inflammation and/or immune activation. The source of this immune activation is not definitive and requires further study. In adolescents and household contacts, *M. tuberculosis* replication may be the primary driver of increased inflammation, but this cannot be the cause in the infants, who were all IGRA-negative at sample collection. A notable result in the infant study was that T-cell responses to cytomegalovirus were significantly higher in infants who progressed to TB disease [64]. This raises the hypothesis that underlying co-infections may drive immune activation [66, 67], which may be associated with risk of TB. Genetic predisposition to immune activation and other environmental factors may also be involved.
Two other candidate COR of TB were identified in the 4–6-month-old infant cohort from the phase IIb MVA85A trial [63]. Higher levels of Ag85A-binding immunoglobulin (Ig) antibodies were observed in controls than in progressors. Ag85A-specific IgG was thus associated with reduced risk of developing TB disease [64]. In addition, elevated frequencies of BCG-specific IFN-γ secreting T-cells, measured by ELISPOT assay at trial baseline, were also associated with reduced risk of developing TB disease. The latter result was surprising in light of the previous immune correlates study in BCG-vaccinated infants, which aimed to determine whether frequencies or cytokine co-expression patterns of BCG-specific CD4 or CD8 cells were associated with subsequent risk of TB disease [68]. This study, which measured CD4+ and CD8+ T-cell expression of IFN-γ, TNF-α, IL-2 and IL-17 at 10 weeks of age, found no associations between frequencies or cytokine-expression patterns of BCG-specific CD4+ and CD8+ T-cells and subsequent risk of TB [68].

Several other studies that also addressed T-cell activation status have shown that phenotypic markers of T-cell activation and/or differentiation can discriminate between distinct clinical presentations of TB infection. During antigen-driven differentiation T-lymphocytes pass through several stages (early, late and terminally differentiated effector cells), and each stage can be characterised by a set of cell surface markers. Since the differentiation process depends on antigenic stimulation, markers of T-cell differentiation may serve as indicators of \textit{M. tuberculosis} replication or antigen load. T-cell expression of CD27 has been observed to show promise as a COR. CD27 is a member of the TNF receptor superfamily, which is constitutively expressed by the naive T-cells and early effector lymphocytes, but downregulated during late stages of effector cell differentiation. Therefore, late effector lymphocytes exhibit low to no CD27 expression [69–70]. Several studies have demonstrated that significantly higher proportions of \textit{M. tuberculosis}-specific IFN-γ-producing CD4+ T-cells do not express CD27 (CD27−IFN-γ CD4+) in persons with active TB disease, compared with healthy controls. In addition, it has been shown that frequencies of CD27−IFN-γ CD4+ cells strongly correlated with the degree of pulmonary destruction [71] and TB treatment success [73, 76]. \textit{Schütze et al.} [77] reported that HIV-infected patients without TB disease have higher proportions of CD27−IFN-γ CD4+ cells than HIV-negative persons without TB disease and suggested that accumulation of \textit{M. tuberculosis}-specific CD27−CD4+ cells may reflect the degree of \textit{M. tuberculosis} replication, thus possibly identifying subclinical \textit{M. tuberculosis} infection. The loss of CD27 expression on specific CD4+ T-cells was shown to precede the development of active TB disease in one HIV-infected patient who had recently HIV-seroconverted, lending further support to this hypothesis [77] (table 1). The accuracy of assays based on the modulation of CD27 may be increased by combining several tests either based on CD27 or on cytokine expression [78]; however, further studies are needed to confirm better accuracy compared to IGRA.

IFN-γ inducible protein-10 (IP-10) is a chemokine secreted by multiple cell types, including monocytes, endothelial cells and fibroblasts, in response to IFN-γ. It acts as a chemoattractant for monocytes/macrophages, T-cells, natural killer cells, and dendritic cells and promotes T-cell adhesion [79, 80]. IP-10 concentrations, measured after \textit{M. tuberculosis} antigen exposure or in unstimulated blood, are elevated in patients with TB either with or without HIV co-infection either in blood [81–84] or urine [85, 86]. High levels of plasma IP-10 was associated with incident TB in both HIV-uninfected [87] and HIV-infected subjects [88] (table 1).

HBHA and other \textit{M. tuberculosis} latency antigens

Heparin-binding haemagglutinin (HBHA) antigen is expressed at the surface of a variety of mycobacterial species and promotes binding to host epithelial cells [89]. This feature increases \textit{M. tuberculosis} pathogenicity and facilitates the extrapulmonary dissemination of mycobacteria. Detection of T-cell responses specific to HBHA has been associated with LTBI, suggesting a protective role for the HBHA-specific responses [90–93]. This is supported by the presence of HBHA-specific CD8+ T-cells in LTBI and the proposed role of such cells in maintaining the latent state [94, 95]. However, as previously stated, LTBI is considered a heterogeneous entity which may include spontaneous cure, stable and persistent asymptomatic infection and, as discussed earlier, early progression towards active disease [10]. \textit{Corbierre et al.} [96] have suggested that measuring IFN-γ responses to two different mycobacterial antigens, ESAT-6 and HBHA, may allow stratification of LTBI subjects into several groups. In this model, those who score negative to both antigens represent successful elimination of the \textit{M. tuberculosis} infection, those who score positive to both antigens are in a stage of true latency able to control \textit{M. tuberculosis} replication and those who are positive only to ESAT-6 have actively replicating \textit{M. tuberculosis} (as in active TB disease). The latter is informed by the demonstration that ESAT-6 is produced at high levels during mycobacterial growth [97]. Based on this model, the authors report that loss of the \textit{in vitro} IFN-γ response to HBHA preceded development of TB in a dialysis patient [96].

Other \textit{M. tuberculosis} latency antigens have also been reported [98]. Adaptation of \textit{M. tuberculosis} to a state of non- or slow replicating persistence during LTBI, with downregulated metabolic activity and concomitant altered gene expression patterns have been suggested [99, 100]. In this state, \textit{M. tuberculosis} has been shown to
induce the expression of the 48-gene-encoding DosR regulon [101]. T-cell responses to M. tuberculosis DosR-encoded latency antigens were particularly dominant in TST+ LTBI individuals, compared to successfully treated or active TB patients. T-cell proliferation and IFN-γ production driven by these latency antigens was consistently seen in LTBI in many diverse genetic and geographic populations [102–104]. Responses to one of these antigens, Rv2628, discriminated recent from remote LTBI in a cross-sectional study [104]. Moreover, the ratio of responses between M. tuberculosis DosR-encoded antigens and ESAT6, in analogy with the results found for HBHA and ESAT6, correlated with low risk of TB disease [105]. These results suggest that the evaluation of this response may be useful for identification of LTBI subjects who are more likely to develop active TB within 2 years and may benefit from preventive treatment (QFT-IT-positive and Rv2628-negative) (table 1). In addition, it has been shown that the specific response to Rv2628 is found at the site of TB disease [106], emphasising the concept of TB as a spectrum of different stages of infection [10].

Elevated proportion of peripheral monocytes as COR for identifying contacts of TB patients at highest risk of developing active TB

Sabin and colleagues [107–109] showed in the 1920s that the ratio between monocyte and lymphocyte numbers in peripheral blood, or monocyte/lymphocyte (M/L) ratio, was associated with progressive outcome of mycobacterial infections in rabbits. Rediscovering this experimental study, a series of prospective cohort evaluations have recently been performed in adults, pregnant women and infants in sub-Saharan Africa [110–112]. In each study an elevated M/L ratio was associated with risk of TB disease before the appearance of symptoms. The biological underpinning of this finding may be related to activation of myelopoiesis in the bone marrow induced by elevated expression of IFN-γ, which is involved in the antimycobacterial response, thus leading to increased myeloid cellularity in the blood [113, 114]. Since lymphopoiesis is not activated in a similar manner, and peripheral lymphoid cells, including T- and B-cells, are recruited to the site of M. tuberculosis replication, this may lead to a decrease in peripheral blood lymphoid cells. The higher M/L ratio is thus likely to be a product of both these processes. Interestingly, in addition to this hypothesis, it was shown that in healthy donors a high M/L ratio was associated with antimycobacterial activity and with transcriptomic profiles characterised by an enrichment of IFN-associated transcripts in monocytes. These data suggest that the M/L ratio may be associated with subclinical disease by acting as a marker of monocyte function [115] (table 1).

Distinct host immune responses to BCG vaccination can obscure identification of COR

Another TB COR study was recently completed in 10-week-old BCG-vaccinated infants from South Africa, who during 2 years of follow-up either progressed to TB disease (progressors) or remained healthy (controls) [116]. Comprehensive comparisons of global gene expression and cellular responses to BCG between progressors and controls did not lead to identification of COR in these infants. However, gene expression profiles revealed two distinct clusters of infants, each of which contained both progressors and controls. One cluster had enrichment of differentially expressed genes in biological pathways including IFN responses and T-cell activation, while the other cluster had enrichment of myeloid cell and glucose metabolism pathways. Importantly, within cluster 1 infants, those who progressed to TB had elevated monocyte to T-cell ratios and frequencies of BCG-specific CD4+ T-cells expressing IFN-γ, among other T-helper type 1 cytokines. It was striking that the former finding is consistent with the M/L ratio associations with risk of TB [115, 116], while the latter is consistent with the observed immune activation and frequencies of IFN-γ-expressing Ag85B-specific T-cells identified as COR in the 4–6-month-old infant cohort from the phase Ib MVA85A trial.

Conclusions

Although only a handful of studies have identified biomarkers of progression to TB, a picture is emerging from the common biological processes, including immune activation, IFN responses and changes in peripheral blood myeloid and lymphoid cells which have been clinically validated and are associated with risk of TB. Promising biomarkers and biomarker signatures are being discovered that may allow accurate prediction of progression from infection to active TB disease. Several of these biomarkers are being taken further towards test development, evaluation and possible use in clinical settings. However, despite this great progress, no biomarker signature has yet been identified that fully meets the performance goals for a prognostic test of progression from latent infection to TB disease set forward by the global health community led by the WHO and other stakeholders. Further intensified research efforts are required for ongoing biomarker discovery, evaluation and test development. In addition, it is critical that different markers and signatures are validated in different geographic settings (and thus different host genetic backgrounds and circulating pathogen lineages) to ensure that accuracy in different high-burden settings is sufficiently acceptable to enable a “screen and treat” strategy. Utility of such biomarkers is also highly dependent on translation of the test to a near-patient platform to allow testing in settings where patients at highest risk for progression present for care (e.g. HIV clinics in TB-endemic populations). Without further
progress on the development of tests that better predict the development of active TB, we will not be able to eradicate the seedbeds of TB and achieve the end-TB strategy of the WHO.

Acknowledgements

We are deeply grateful to Samuel G. Schumacher (FIND, Geneva, Switzerland) for contributing to the preparation of the figures.

References

Host RNA signatures for diagnostics: an example from paediatric tuberculosis


