Qualification of a whole blood intracellular cytokine staining assay to measure mycobacteria-specific CD4 and CD8 T cell immunity by flow cytometry[*]

Benjamin M. Kagina[a,b], **Nazma Mansoor**[a,c], **Eloi P. Kpamegan**[d], **Adam Penn-Nicholson**[a], **Elisa Nemes**[a], **Erica Smit**[a], **Sebastian Gelderbloem**[e], **Andreia P. Soares**[a,f], **Brian Abel**[a,g], **Alana Keyser**[a], **Mzwandile Sidibana**[a], **Jane E. Hughes**[a], **Gilla Kaplan**[h,i], **Gregory D. Hussey**[a,b], **Willem A. Hanekom**[a], and **Thomas J. Scriba**[a,*]

[a]South African Tuberculosis Vaccine Initiative (SATVI), Institute of Infectious Disease and Molecular Medicine and School of Child and Adolescent Health, University of Cape Town, South Africa

[b]Vaccines for Africa Initiative, Division of Medical Microbiology & Institute of Infectious Disease and Molecular Medicine, University of Cape Town, South Africa

[c]BD Biosciences, Johannesburg, South Africa

[d]Aeras, Rockville, MD, USA

[e]Aeras, Cape Town, USA

[f]Division of Medical Virology, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, South Africa

[g]Singapore Immunology Network, Agency for Science, Technology and Research, Singapore

[h]Public Health Research Institute, NJMS Rutgers, Newark, NJ, USA

[i]Bill & Melinda Gates Foundation, Seattle, WA, USA

Abstract

Background—Qualified or validated assays are essential in clinical trials. Short-term stimulation of whole blood and intracellular cytokine staining assay is commonly used to measure immunogenicity in tuberculosis vaccine clinical trials. Previously, the short-term stimulation process of whole blood with BCG was optimized. We aimed to qualify the intracellular cytokine staining process and assess the effects of long-term cryopreservation. Our hypotheses were that the assay is robust in the measurement of the mycobacteria-specific T cells, and long-term cryopreservation of fixed cells from stimulated whole blood would not compromise reliable measurement of mycobacteria induced CD4 T cell immunity.

[*]Corresponding author at: University of Cape Town, Health Sciences, Anzio Road, Observatory, 7925, South Africa. Tel.: +27 21 406 6427. thomas.scriba@uct.ac.za (T.J. Scriba).

Methods—Whole blood from healthy adults was collected in sodium heparinized tubes. The blood was left unstimulated or stimulated with mycobacterial antigens or mitogens for 12 h. Cells were harvested, fixed and multiple aliquots from each participant cryopreserved. Later, mycobacteria-specific CD4 and CD8 T cells expressing IFN-γ, TNF-α, IL-2 and IL-17 were quantitated by flow cytometry. Assay performance characteristics evaluated included limit of quantification and detection, reproducibility, precision, robustness, specificity and sensitivity. To assess the effects of long-term cryopreservation, fixed cells from the stimulated bloods were analysed one week post-cryopreservation and at 3-month intervals over a 3-year period.

Results—The limit of quantification for the different cytokines was variable: 0.04% for frequencies of IFN-γ- and IL-2-expressing T cells and less than 0.01% for TNF-α- and IL-17-expressing T cells. When measurement of the mycobacteria-specific T cells was assessed at levels above the detection limit, the whole blood intracellular cytokine assay showed high precision that was operator-independent. The assay was also robust: variation in staining conditions including temperature (4 °C or 20–23 °C) and time (45, 60 or 90 min) did not markedly affect quantification of specific T cells. Finally, prolonged periods of cryopreservation also did not significantly influence quantification of mycobacteria-specific CD4 T cells.

Conclusions—The whole blood intracellular cytokine assay is robust and reliable in quantification of the mycobacteria-specific T cells and is not significantly affected by cryopreservation of fixed cells.

1. Introduction

Control of the global tuberculosis (TB) pandemic, which still has significant morbidity with more than 8 million people getting the disease every year (World Health Organization, 2013), is an urgent priority. Epidemiological modelling suggests that elimination of TB can only be achieved with an effective vaccination strategy, coupled with better diagnosis and more effective treatment of persons infected with *Mycobacterium tuberculosis* (*M.tb*) and with TB disease (Abu-Raddad et al., 2009; Dye et al., 2013). Immunization with Bacille Calmette-Guerin (BCG), the only licensed TB vaccine, confers protection against severe forms of TB in infants, such as miliary TB and TB meningitis, however efficacy against adult and childhood pulmonary disease is variable and mostly poor (Trunz et al., 2006). A more efficacious vaccine is urgently needed and 16 new TB vaccines are currently in clinical testing (Evans et al., 2013).

Measurements of vaccine immunogenicity, or vaccine take, in preclinical development and human trials routinely include quantitation of antigen-specific Th1 and Th17 cells (Scriba et al., 2012; Day et al., 2013; Abel et al., 2010; Khader et al., 2007). A number of distinct immunological assays, among them IFN-γ ELISPOT, intracellular cytokine staining (ICS) assays and ELISA-based quantification of IFN-γ-release, have been used to measure immunogenicity of the different TB vaccine candidates (Ota et al., 2011; Abel et al., 2010; van Dissel et al., 2010). Results of outcomes generated by these different methodologies

could be usefully compared only if qualified or validated assays are used. Until this is achieved it is difficult to include immunogenicity criteria when gauging the potential value of a given vaccine against those of other candidates. There is a clear and urgent need to optimize, standardize, qualify and validate assays that are suitable for measuring antigen-specific T cell responses.

We previously optimized a whole blood (WB) ICS assay for measurement of antigen-specific Th1 responses in clinical studies (Hanekom et al., 2004), and have applied this WB-ICS assay to evaluate immunogenicity in many clinical studies of TB vaccines in infants and adults (Hawkridge et al., 2008; Scriba et al., 2012; Scriba et al., 2011; Scriba et al., 2010; Day et al., 2013; Abel et al., 2010; Tameris et al., 2013). Advantages of the WB-ICS assay over PBMC-based assays include a requirement for smaller blood volumes, and immediate antigen stimulation of fresh cells leading to less cell death and lower sensitivity of fixed cells to freezing and thawing procedures (Hanekom et al., 2008). This assay is thus suitable for resource-limited settings where TB, HIV/AIDS and malaria are endemic and are the ideal sites for clinical trials.

Here we describe qualification of the WB-ICS assay readout by flow cytometry. Assay qualification precedes assay validation. Validation is defined as an evaluation of the method on its fitness for the intended applications ((FDA) FaDa, 1996). The assay validation process involves evaluating performance characteristics of several parameters to ensure that the assay limitations are known when applied to measure predefined outcomes. Assay validation parameters include, among others, accuracy, specificity, limit of detection, reproducibility or inter- and intra-assay co-efficient of variation (CV), precision, linearity and robustness ((FDA) FaDa, 1996). Establishing accuracy (how close the measured value is to the true value ((FDA) FaDa, 1996)) in ICS assays is impractical because the true value is unknown. We aimed to qualify the WB-ICS assay and established some of the performance characteristics of the WB-ICS assay.

2. Methods

2.1. Study participants and blood collection

Healthy adults were recruited at the Institute of Infectious Disease and Molecular Medicine (IDM) of the University of Cape Town in South Africa, or at the South African TB Vaccine Initiative (SATVI) field site in Worcester, 110 km from Cape Town. All participants reported BCG vaccination at birth. The study protocol was approved by the University of Cape Town Research Ethics Committee, and all procedures adhered to the guidelines of the National Health Research Ethics Council. Good clinical practice guidelines were adhered to including written informed consent. Venous blood was collected from study participants in sodium heparin tubes and processed within 60 min.

2.2. Antigens used in WB-ICS

Bacillus Calmette Guerin (BCG) Danish 1331 (Statens Serum Institut, Copenhagen) was reconstituted in the vaccine vial with RPMI and used at a final concentration of 1.2×10^6 CFU/mL of blood, a dose previously optimized (Hanekom et al., 2004). Purified protein derivative (PPD, Statens Serum Institut, Copenhagen) was used at a final concentration of

10 μg/mL. In some experiments, peptide pools spanning ESAT-6 and CFP-10 or Ag85A mycobacterial proteins (15-mers, overlapping by 10 amino acids, each at 2 μg/mL; Peptide Protein Research Ltd.) were used. For positive controls, stimulation was done with either Staphylococcal enterotoxin B (SEB, Sigma) or Phytohemagglutinin (PHA, HA16, Sigma), at final concentrations of 10 μg/mL or 5 μg/mL, respectively, as previously described (Hanekom et al., 2004).

2.3. Short-term whole blood stimulation and cryopreservation

One milillitre of whole blood, pipetted into Sarstedt tubes, was either left unstimulated (negative control) or stimulated with mycobacterial antigens (as described above), or positive controls of SEB or PHA (either but not both positive controls), in the presence of co-stimulatory antibodies (anti-CD28 and anti-CD49d, each at 1 μg/mL; BD Biosciences). These co-stimulatory antibodies have been shown to increase cytokine expression in specific T cells (Hanekom et al., 2004). Blood was incubated at 37 °C for 12 h, and Brefeldin-A (Sigma-Aldrich, 10 μg/mL) was added during the last 5 h of incubation. The blood was then harvested with EDTA (Sigma, 2 μM) red blood cells were lysed and white blood cells fixed with FACS lysing solution (BD Biosciences). Fixed white cells were pelleted and multiple aliquots were cryopreserved with 10% DMSO (Sigma) in 40% foetal calf serum (BioWest) in RPMI at −190 °C in the vapor phase of liquid nitrogen.

2.4. Thawing and permeabilisation of cryopreserved fixed cells

Cryovials containing the stimulated, fixed and frozen white cells from whole blood were retrieved from liquid nitrogen tanks and thawed in a water bath at 37 °C for 2 min. Thawed cells were transferred from cryovials to labelled tubes containing 2 mL of phosphate buffered saline (PBS, BioWhittaker). Thereafter the cells were centrifuged at 215 g for 5 min. Next, the cells were permeabilised by adding 2 mL Perm/Wash solution (BD Biosciences) and incubated at room temperature for 10 min (unless specific incubation temperatures were investigated).

2.5. Intracellular cytokine staining (ICS) and flow cytometry

Thawed cryopreserved fixed cells were washed in PBS and immediately stained with cocktails of monoclonal antibodies for 60 min at 4 °C, unless otherwise indicated. Two different flow cytometry antibody panels were used: One monoclonal antibody panel was for multiparameter flow cytometry, utilizing a BD LSR II cytometer. For these experiments, the cells were thawed, permeabilised in BD Perm/Wash buffer and stained with previously optimized antibody-fluorochrome combinations to the following markers: CD3-PacBlue (BD Biosciences, clone MOPC-21), CD4-QDot605 (Invitrogen, S3.5), CD8-PerCPCy5.5 (BD Biosciences, SK1), IFN-γ-Alexa700 (BD Biosciences, B27), TNF-α-PeCy7 (eBioscience, Mb11), IL-2-FITC (BD Biosciences, 5344.111), IL-17-Alexa647 (eBioscience, SCPL1362) and the Ki67-PE (BD Biosciences, B1). Cytometer Setting and Tracking (CST) beads (BD Biosciences) were acquired before each experiment to ensure that cytometer parameters remained consistent across all experiments. Stained samples were acquired with a standard stopping gate set at 200,000 CD3 lymphocytes. Single stained and

negative compensation beads (BD Biosciences) were acquired for each experiment, before sample acquisition, and used to calculate the compensation matrix.

To measure effects of long-term cryopreservation on ICS outcomes in fixed white blood cells, a second monoclonal antibody panel comprising of CD4-APC (SK3) and IFN-γ-PE (25723.11; both from BD Biosciences) was acquired on a FACSCalibur (BD Biosciences). For these experiments, cells were thawed, permeabilized in BD Perm/Wash buffer and stained as indicated above before acquisition. At least 40,000 CD4 T cells were acquired.

2.6. IFN-γ ELISpot assay

We compared frequencies of IFN-γ expressing cells detected by WB-ICS and IFN-γ ELISpot assay from samples collected in a previously completed clinical trial of the candidate TB vaccine, MVA85A (Scriba et al., 2011). We analysed data from a subset of 36 healthy infants enrolled into the TB014 trial, who received a single intradermal vaccination of 5×10^7 pfu of MVA85A (clinicaltrials.gov NCT00679159). The WB-ICS assay was performed as described above. Whole blood and PBMC were stimulated in parallel with a single pool of peptides spanning the Ag85A protein (15-mers, overlapping by 10 amino acids, each at 2 μg/mL; Peptide Protein Research Ltd.). For ELISpot assay, medium alone served as negative control and PHA, (10 μg/mL) as a positive control. ELISpot plates, containing 3×10^5 peripheral blood mononuclear cells (PBMC) per well, were incubated with antigens for 18 h at 37 °C and developed according to the manufacturer's protocol (Mabtech), as previously described (Scriba et al., 2011). Assays were performed in duplicate wells and the average (with background subtracted) was used for analysis.

2.7. QuantiFERON-TB Gold In-Tube assay (QFT)

We also compared frequencies of IFN-γ expressing CD4 T cells detected by ICS upon blood stimulation with ESAT-6 and CFP-10 peptide pools with levels of antigen-specific IFN-γ release measured by the validated QFT (Mazurek et al., 2010). The QFT was performed according to the manufacturer's protocol (Cellestis).

2.8. WB-ICS assessments

To assess the lower limit of quantification (LOQ) of the WB-ICS assay, blood from two healthy adult volunteers was stimulated with BCG or left unstimulated. Stimulated cells were serially diluted with autologous unstimulated cells, ranging from 2 to 4096-fold. Unstimulated, BCG-stimulated and serially diluted samples were then stained with an 8-colour antibody panel, as described above, and cytokine expression by CD4 and CD8 T cells was analyzed on a LSR II cytometer. Expected frequencies were derived by dividing the frequencies of BCG-specific cytokine-expressing CD4 and CD8 T cells in the undiluted samples by the dilution factor.

To assess intra-assay variability, 3 to 5 aliquots of unstimulated and BCG stimulated whole blood generated in the same experiment (performed in 3 donors) were thawed, stained and acquired on 4 separate days.

To assess inter-operator variability, the same procedure outlined above to assess intra-assay variability was conducted by two different operators.

To assess robustness, two operators independently retrieved multiple aliquots of stimulated and fixed cells derived from 3 adult volunteers and stained the cells for 45, 60 and 90 min at either room temperature (20–23 °C) or at 4 °C.

To assess effects of long-term cryopreservation on the WB-ICS assay, whole blood from 11 healthy adults was stimulated with BCG, PPD, and SEB, or was left unstimulated, processed and cryopreserved. Fixed cells from stimulated WB were thawed, stained and acquired one week after cryopreservation and then every 3 months over a 4-year period.

2.9. Data analysis

All flow data analyses were performed using FlowJo software (version 9.0.2, TreeStar, Inc.). Statistical analysis was done with Microsoft Excel (2008) and GraphPad Prism (version 5a). Frequencies of cytokine-expressing CD4 and CD8 T cells are reported after subtracting frequencies of cytokine-expressing cells detected in unstimulated controls. Ki67-expression by both CD4 and CD8 T cells was too low for reliable quantification; therefore, the molecule was excluded from the analysis. To assess precision, reproducibility and robustness, we calculated the mean, standard error of the mean and the coefficient of variation of frequencies of specific cytokine-expressing CD4 and CD8 T cells. Variance component analysis was performed to estimate the inter-run variance and the intra-run variance. The overall precision of the assay is the variability calculated from the results of two operators.

To assess influence of time of cryopreservation on WB-ICS readout, intra-individual coefficient of variation was calculated across multiple experiments performed over time of cryopreservation.

Associations between WB-ICS, ELISpot and QuantiFERON readouts were assessed using a Spearman's correlation test.

3. Results

3.1. Establishment of the lower limit of quantification for the WB-ICS assay

Before conducting experiments to qualify the WB-ICS assay, we established the lower LOQ for the WB-ICS assay, by performing serial dilutions of stimulated WB with autologous unstimulated WB samples. Our aim was to determine the lowest frequencies of mycobacteria-specific IFN-γ, TNF-α, IL-2 and IL-17-expressing T cells that could be detected reliably using this assay. We used the gating strategy as shown in Supplementary Fig. 1. We calculated the percent deviation between expected and observed values of cytokine-expressing CD4 T cells. Observed frequencies of IFN-γ and IL-2-expressing CD4 T cells appeared to vary marginally (5%–18% and 4%–21% respectively) from their respective expected values above 0.04%, shown by the dashed horizontal lines (Fig. 1A and B). Marginal variation (3%–25% and 12%–29%) was observed for frequencies of TNF-α and IL-17 expressing CD4 T cells of 0.01% or above respectively (Fig. 1C and D). Similar

analyses of antigen-specific CD8 T cells showed that reliable enumeration with similar marginal variability between observed and expected frequencies was only possible for IFN-γ-expressing cells, above 0.04% (data not shown). TNF-α, IL-17 and IL-2 were not detected or were expressed at very low frequencies in CD8 T cells, elevating variability in the measurement (data not shown). These results indicate that accuracy of enumeration of cytokine-expressing cells is dependent on the levels of expression, and possibly, the antibody-fluorochrome combination.

Next, we aimed to assess non-specific frequencies of cytokine positive cells (background levels) detected in this assay. We analysed frequencies of IFN-γ-, TNF-α-, IL-2- and IL-17-expressing CD4 T cells detected in unstimulated, negative control samples from the 54 samples generated in the experiment comparing the WB-ICS with QFT assays. Median frequencies of CD4 T cells expressing IFN-γ, TNF-α, IL-2 and IL-17 in the unstimulated controls were 0.002 (range 0.0001–0.137), 0.002 (0.0001–0.071), 0.002 (0.0001–0.002) and 0.001 (0.0001–0.038), respectively (Fig. 1I). Medians and upper interquartile range (IQR) values for the four cytokines fell below the established limit of reliable quantification of the flow cytometer (Fig. 1I).

3.2. Intra-assay variability

Next, we analyzed coefficients of variation (CVs), a measure of precision within or between repeated measurements. CVs below 30% have been previously reported and are considered acceptable in ICS assays (Weinberg et al., 2000; Nomura et al., 2000). We used the gating strategy as shown in Supplementary Fig. 1. We calculated CVs from frequencies of BCG-specific CD4 and CD8 T cells, measured in quadruplet stimulation conditions for each sample (intra-assay variation). Each experiment was also repeated four times, on different days. CVs for IFN-γ, TNF-α-, IL-2- or IL-17-expressing CD8 and CD4 T cells were calculated. CVs for the frequencies TNF-α-, IL-2- and IL-17-expressing CD8 T cells were mainly above 30% (black, blue and green circles) while CVs for IFN-γ-expressing cells were all below 30% (red circles) (Fig. 1E and F). The four cytokine-expressing CD4 T cell subsets measured showed CVs below 30% (Fig. 1G and H). From these experiments, we concluded that measurement of BCG-specific IFN-γ, TNF-α-, IL-2- or IL-17-expressing CD4 T cells and IFN-γ-expressing CD8 T cells by WB-ICS assay was reliable. Therefore, our subsequent analyses focussed on these outcomes.

To measure intra-assay variability, CVs for each experiment and operator were computed separately. CVs below 15% were consistently observed for BCG-specific CD4 T cells expressing total IFN-γ, TNF-α and IL-2, and for total IFN-γ-expressing CD8 T cells (Table 1). CVs for specific CD4 T cells expressing IL-17 were greater than for the other cytokines due to the lower frequencies of these cells. Nevertheless, all the CVs were below 30% except for a single day for one operator, when a CV of 45.9% was observed. This was an unexpected result and further investigations revealed high IL-17 expression in the experimental control, likely compromising the precision of the ICS. Nevertheless, in general, these results show that intra-assay variability of the WB-ICS assay is acceptable.

3.3. Inter- and intra- operator variability

Ideally, a single operator should perform sample processing, acquisition and analyses to minimise variability in immunological assays. However, this is not always possible. We investigated inter-operator variability of the thawing, antibody staining and flow cytometry acquisition steps of the WB-ICS assay. Blood stimulation and processing until cryopreservation were performed by a single operator. Then, two operators thawed aliquots, applied identical reagents and antibody staining procedures to samples from 3 healthy donors and acquired the cells. Identical samples were stained and analysed on four different days. Data analysis, including gating was done by a single operator. Very little inter-operator variability was observed in frequencies of IFN-γ, TNF-α, IL-2 and IL-17-expressing CD4 T cells and IFN-γ-expressing CD8 T cells across the four experiments (Fig. 2). As expected, variability was greatest when low frequencies of specific T cells were detected, such as IL-17-expressing CD4 T cells (Fig. 2D). These findings were supported by analyses of intra-operator CVs, which were consistently below 30% (Table 1 and Fig. 3).

From these experiments, we concluded that the WB-ICS assay is not prone to major operator variability, provided that operators are suitably trained and adhere to standard operating procedures.

3.4. Robustness of the WB-ICS assay

Robustness is defined as the ability of an assay to withstand deliberate variation in assay conditions. We sought to determine effects of variation in duration and temperature of monoclonal antibody staining on assay outcome. Two operators independently retrieved multiple aliquots of stimulated and fixed cells derived from 3 adult volunteers and stained the cells for 45, 60 and 90 min at either room temperature (20–23 °C) or at 4 °C. These staining conditions did not significantly affect quantitation of frequencies of cytokine-expressing BCG-specific CD4 or CD8 T cells; all CVs were below 9% (Table 2). These data were also used to fit in a factorial design model. The data was fitted into the full model and the analysis revealed that the interactions were not significant (data not shown). From these analyses, we concluded that WB-ICS was robust to withstand the evaluated changes in the staining conditions.

3.5. Effects of long-term cryopreservation

An advantage of the WB-ICS assay is that blood is stimulated immediately and fixed lymphocytes are then cryopreserved, allowing later staining and analysis in sample batches. Our next experiments assessed effects of long-term cryopreservation on measurement of mycobacteria-specific IFN-γ-expressing CD4 T cells. Whole blood from 11 healthy adults was stimulated with BCG, PPD, and SEB, or was left unstimulated, then processed and cryopreserved. Cells were thawed, stained and acquired on a FACSCalibur one week after cryopreservation and then every 3 months over a 3-year period (Fig. 4A and B). Frequencies of CD4 T cells expressing IFN-γ either in the unstimulated or stimulated samples (minus unstimulated) were calculated for individual donors across different time intervals after cryo-preservation (Fig. 4C). Importantly, no increasing or decreasing trends in these frequencies were observed (Fig. 4C). CVs calculated for each stimulation condition for each individual donor over time (Fig. 4D) and across different donors at each time point (Fig. 4E)

were mostly less than 30%. Regression analysis suggested that frequencies of CD4 T cells expressing IFN-γ under all test conditions (Unstimulated, BCG, PPD and SEB) did not increase or decrease during the cryopreservation period. Overall, our data suggest that the duration of cryopreservation does not significantly influence the integrity of fixed cells to a degree that may affect measurement of specific T cells by the WB-ICS assay.

3.6. Comparison of WB-ICS assay with IFN-γ ELISpot and QFT assays

Next, we compared frequencies of antigen-specific, IFN-γ expressing CD4 T cells detected by WB-ICS with levels of antigen-specific IFN-γ expression measured by ELISpot assay or by QFT. Frequencies of Ag85A-specific CD4 T cells expressing IFN-γ detected in MVA85A-vaccinated infants by WB-ICS assay were highly correlated with frequencies of IFN-γ expressing PBMC, detected by ELISpot assay (Fig. 5A).

Frequencies of ESAT-6/CFP-10-specific CD4 T cells expressing IFN-γ detected by WB-ICS assay in healthy, *M.tb*-infected or uninfected adults also correlated very strongly with levels of released IFN-γ detected by QFT, a whole blood assay validated for diagnosis of latent *M.tb* infection (Mazurek et al., 2010) (Fig. 5B). Of note, the magnitude of WB-ICS assay-detected specific T cells correlated well with the QFT-detected responses, even at very low frequencies. Interestingly, very strong correlations were also observed between frequencies of TNF-α or IL-2-expressing ESAT-6/CFP-10-specific CD4 T cells, detected by WB-ICS assay, and levels of soluble IFN-γ detected by QFT (Fig. 5C and D).

These findings from our WB-ICS assay concur with those from the validated QFT, and further support the utility of the WB-ICS assay for accurate quantitation of antigen-specific cytokine-expressing T cells.

3.7. Sensitivity and specificity of the WB-ICS assay

Finally we were interested in assessing the sensitivity and specificity of the WB-ICS assay. In our study setting (Western Cape Province of South Africa), the population is highly exposed to *M.tb* and environmental non-tuberculous mycobacteria, and BCG vaccination coverage is high (Corrigal, 2005). As a result it is very rare to find persons without mycobacteria-specific T cell responses precluding evaluation of assay specificity. However, since the antigens ESAT-6 and CFP-10 measure T cell responses induced only by *M.tb*, and not BCG or most environmental mycobacteria, assay sensitivity and specificity could be assessed by comparing data from QFT negative and positive persons as shown in Fig. 5. The highest sensitivity and specificity pairing were observed for frequencies of IFN-γ-, TNF-α-, and IL-2-expressing CD4 T cells at 0.014%, 0.020% and 0.013%, respectively (Fig. 6A, B & C), suggesting that these frequencies could be used as assay positivity thresholds. We then applied these threshold frequencies to classify the QFT status from the 54 participants. Sensitivity and specificity values for IFN-γ-, TNF-α-, and IL-2-expressing CD4 T cells were 90% and 92%; 90% and 88%; and 93% and 92%, respectively. ESAT-6/CFP-10-specific IL-17-expressing CD4 T cells were more infrequent than Th1 cells, even in QFT-positive persons. This precluded reliable derivation of a positivity threshold for IL-17-expressing CD4 T cells and subsequent calculations of sensitivity and specificity.

4. Discussion

Intact mycobacteria-specific T cell responses are necessary for human immunity against mycobacteria. However, exactly which functions such T cells should possess to protect against TB disease, is not known. As a consequence, assessment of vaccine-induced immune responses can only be used to infer vaccine immunogenicity or vaccine take. Regardless, regulatory agencies require data generated by qualified and/or validated assays for assessment of vaccine-induced immune responses to ultimately license candidate vaccines.

We report that the WB-ICS assay is a highly reproducible, reliable and robust method for quantifying antigen-specific cytokine-expressing CD4 and CD8 T cells. Variability in measured frequencies of specific T cells by WB-ICS, both within and across experiments was within accepted limits; CVs remained below 30%, even when performed by multiple operators and across different days.

The WB-ICS assay was originally optimized a decade ago, and has since been applied to many clinical studies (Hawkridge et al., 2008; Scriba et al., 2012; Scriba et al., 2011; Scriba et al., 2010; Day et al., 2013; Abel et al., 2010; Tameris et al., 2013). During the development and optimization process, many variables that may affect assay performance were tested, including blood volume, delay in initiation of antigen stimulation following blood collection and type of polypropylene tube used for stimulation (Hanekom et al., 2004). Here, we build on these earlier optimization steps by investigating precision, robustness and effects of long-term cryopreservation to validate the WB-ICS assay readout for reliable measurement of mycobacteria-specific T cell frequencies. This assay is ideal for application in clinical field trials, where samples cannot be analysed in real time but rather in batches at a later time point, after cryopreservation. The sensitivity and accuracy of the WB-ICS assay are further supported by the strong correlation between *M.tb* antigen-specific frequencies of IFN-γ, TNF-α and/or IL-2-expressing CD4 T cells and the quantitative value of IFN-γ detected by the validated diagnostic QFT. Good agreement was also observed when compared with the highly quantitative IFN-γ ELISpot assay. Furthermore, sensitivity and specificity curves generated with ESAT-6/CFP-10-specific T cell data derived from QFT positive and negative participants showed that the WB-ICS assay could reliably discriminate *M.tb* infection from non-infection, even at very low frequencies of mycobacteria-specific Th1 cells. Importantly, long-term cryopreservation of fixed cells from stimulated whole blood did not result in significant increases in assay variability. Our results suggest that the performance characteristics of the WB-ICS assay facilitate reliable measurement of antigen-specific T cells.

Whole blood-ICS assay experiments to establish sensitivity and specificity can be challenging in settings with high prevalence of sensitization to the antigen of interest. We applied frequencies of ESAT-6/CFP-10-specific T cells measured in QFT positive and negative participants to calculate the sensitivity and specificity of the WB-ICS, and derive assay positivity thresholds. Further, accuracy determination in WB-ICS assay is impractical because the true frequency of the outcome being quantified is often unknown. Comparing the WB-ICS assay with other validated assays such as ELISpot and QFT in quantifying specific outcomes could be considered as an alternative evaluation of the accuracy. The

WB-ICS assay allowed highly sensitive (sensitivities of 90% to 93%) detection of cytokine-expressing ESAT-6/CFP-10-specific CD4 T cells, when compared with QFT assay. WB-ICS assay outcomes were highly correlated with QFT and ELISpot assay outcomes, even when including WB-ICS results below the lower LOQ formally suggested by the serial dilution experiments. An important factor that may have led to the underestimate of WB-ICS LOQ was the frequency of non-specific cytokine-positive T cells in the unstimulated sample. Because unstimulated samples were used to perform serial dilutions of stimulated cells, the lower LOQ for each cytokine was largely determined by the frequencies of cytokine expressing T cells in the unstimulated cells in this particular donor (0.04% for IFN-γ and IL-2, 0.01% for TNF-α and IL-17 CD4 T cells, and 0.04% for IFN-γ CD8 T cells).

To assess how inter-individual variability may have influenced the LOQ, we measured Th1 cell frequencies in unstimulated samples from a larger sample size (n = 54). These mostly fell below the LOQ, suggesting that the sensitivity of quantification shown in Fig. 1 is likely underestimated. This possible underestimation of the LOQ is further supported by the positivity threshold frequencies of IFN-γ (0.014%) and IL-2 (0.02%), both below the cytometer's LOQ. However, a positivity threshold frequency of TNF-α (0.02%) was greater than the LOQ.

A major advantage of the WB-ICS assay is that cells can be fixed and cryopreserved after antigen stimulation is completed, thus eliminating any potentially detrimental effects of cryopreservation on the function of antigen presenting and effector T cells. Further, because stimulated cells are fixed before cryopreservation, maintenance of cell integrity during cryopreservation and thawing procedures is less important. A disadvantage of cryopreserving fixed cells from stimulated whole blood is that the choice of antigen for stimulation cannot be changed, limiting flexibility of the assay.

This highly reproducible, reliable and robust WB-ICS assay is ideal for detection and characterisation of antigen-specific cytokine expressing CD4 and CD8 T cells. This is essential for determining which attributes of antigen-specific T cell responses may contribute to vaccine efficacy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Abel B, Tameris M, Mansoor N, Gelderbloem S, Hughes J, Abrahams D, Makhethe L, Erasmus M, de Kock M, van der Merwe L, et al. The novel tuberculosis vaccine, Aeras-402, induces robust and polyfunctional CD4+ and CD8+ T cells in adults. Am J Respir Crit Care Med. 2010; 181:1407. [PubMed: 20167847]

Abu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini IM Jr, Dye C, Halloran ME. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. Proc Natl Acad Sci U S A. 2009; 106:13980. [PubMed: 19666590]

Corrigal, J. [accessed 16 September 2014] Western Cape Provincial EPI Vaccination Survey. 2005. www.capegateway.gov.za/Text/2007/6/cd_volume_7_childhood_diseases_overview.pdf

Day CL, Tameris M, Mansoor N, van Rooyen M, de Kock M, Geldenhuys H, Erasmus M, Makhethe L, Hughes EJ, Gelderbloem S, et al. Induction and regulation of T-cell immunity by the novel

tuberculosis vaccine M72/AS01 in South African adults. Am J Respir Crit Care Med. 2013; 188:492. [PubMed: 23306546]

Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. Annu Rev Public Health. 2013; 34:271. [PubMed: 23244049]

Evans TG, Brennan MJ, Barker L, Thole J. Preventive vaccines for tuberculosis. Vaccine. 2013; 31 (Suppl 2):B223. [PubMed: 23598486]

(FDA) FaDA. Guidance for Industry: Q2b Validation of Analytical Procedures: Methodology. Food and Drug Administration (FDA); 1996.

Hanekom WA, Hughes J, Mavinkurve M, Mendillo M, Watkins M, Gamieldien H, Gelderbloem SJ, Sidibana M, Mansoor N, Davids V, et al. Novel application of a whole blood intracellular cytokine detection assay to quantitate specific T-cell frequency in field studies. J Immunol Methods. 2004; 291:185. [PubMed: 15345316]

Hanekom WA, Dockrell HM, Ottenhoff TH, Doherty TM, Fletcher H, McShane H, Weichold FF, Hoft DF, Parida SK, Fruth UJ. Immunological outcomes of new tuberculosis vaccine trials: WHO panel recommendations. PLoS Med. 2008; 5:e145. [PubMed: 18597551]

Hawkridge T, Scriba TJ, Gelderbloem S, Smit E, Tameris M, Moyo S, Lang T, Veldsman A, Hatherill M, Merwe Lvd, Fletcher HA, Mahomed H, Hill AVS, Hanekom WA, Hussey GD, McShane H. Safety and immunogenicity of a new tuberculosis vaccine, MVA85A, in healthy adults in South Africa. The Journal of Infectious Diseases. 2008; 198:544. [PubMed: 18582195]

Khader SA, Bell GK, Pearl JE, Fountain JJ, Rangel-Moreno J, Cilley GE, Shen F, Eaton SM, Gaffen SL, Swain SL, et al. IL-23 and IL-17 in the establishment of protective pulmonary CD4+ T cell responses after vaccination and during *Mycobacterium tuberculosis* challenge. Nat Immunol. 2007; 8:369. [PubMed: 17351619]

Mazurek GH, Jereb J, Vernon A, LoBue P, Goldberg S, Castro K. Updated guidelines for using interferon gamma release assays to detect *Mycobacterium tuberculosis* infection — United States, 2010. MMWR Recomm Rep. 2010; 59:1. [PubMed: 20577159]

Nomura LE, Walker JM, Maecker HT. Optimization of whole blood antigen-specific cytokine assays for CD4+ T cells. Cytometry. 2000; 40:60. [PubMed: 10754518]

Ota MO, Odutola AA, Owiafe PK, Donkor S, Owolabi OA, Brittain NJ, Williams N, Rowland-Jones S, Hill AV, Adegbola RA, et al. Immunogenicity of the tuberculosis vaccine MVA85A is reduced by coadministration with EPI vaccines in a randomized controlled trial in Gambian infants. Sci Transl Med. 2011; 3:88ra56.

Scriba TJ, Tameris M, Mansoor N, Smit E, van der Merwe L, Isaacs F, Keyser A, Moyo S, Brittain N, Lawrie A, Gelderbloem S, Veldsman A, Hatherill M, Hawkridge A, Hill AV, Hussey GD, Mahomed H, McShane H, Hanekom WA. Modified vaccinia Ankara-expressing Ag85A, a novel tuberculosis vaccine, is safe in adolescents and children, and induces polyfunctional CD4+ T cells. Eur J Immunol. 2010; 40:279. [PubMed: 20017188]

Scriba TJ, Tameris M, Mansoor N, Smit E, van der Merwe L, Mauff K, Hughes EJ, Moyo S, Brittain N, Lawrie A, et al. Dose-finding study of the novel tuberculosis vaccine, MVA85A, in healthy BCG-vaccinated infants. J Infect Dis. 2011; 203:1832. [PubMed: 21606542]

Scriba TJ, Tameris M, Smit E, van der Merwe L, Hughes EJ, Kadira B, Mauff K, Moyo S, Brittain N, Lawrie A, Mulenga H, de Kock M, Makhethe L, Janse van Rensburg E, Gelderbloem S, Veldsman A, Hatherill M, Geldenhuys H, Hill AV, Hawkridge A, Hussey GD, Hanekom WA, McShane H, Mahomed H. A phase IIa trial of the new tuberculosis vaccine, MVA85A, in HIV- and/or *Mycobacterium tuberculosis*-infected adults. Am J Respir Crit Care Med. 2012; 185:769. [PubMed: 22281831]

Tameris MD, Hatherill M, Landry BS, Scriba TJ, Snowden MA, Lockhart S, Shea JE, McClain JB, Hussey GD, Hanekom WA, Mahomed H, McShane H. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo controlled phase 2b trial. Lancet. 2013; 381:1021. [PubMed: 23391465]

Trunz BB, Fine P, Dye C. Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. Lancet. 2006; 367:1173. [PubMed: 16616560]

van Dissel JT, Arend SM, Prins C, Bang P, Tingskov PN, Lingnau K, Nouta J, Klein MR, Rosenkrands I, Ottenhoff TH, et al. Ag85b-Esat-6 adjuvanted with IC31 promotes strong and long-lived *Mycobacterium tuberculosis* specific T cell responses in naive human volunteers. Vaccine. 2010; 28:3571. [PubMed: 20226890]

Weinberg A, Zhang L, Brown D, Erice A, Polsky B, Hirsch MS, Owens S, Lamb K. Viability and functional activity of cryopreserved mononuclear cells. Clin Diagn Lab Immunol. 2000; 7:714. [PubMed: 10882680]

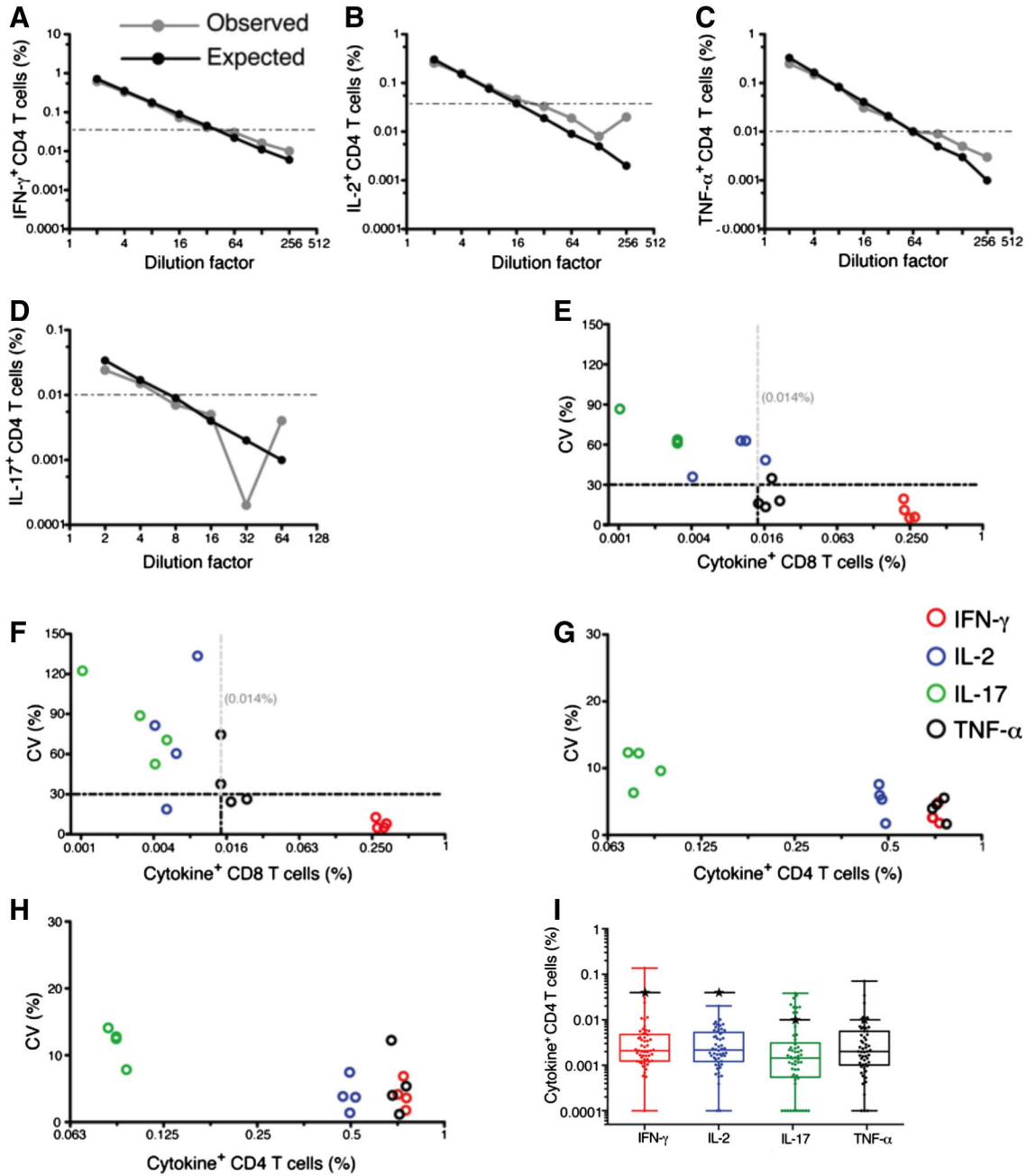World Health Organization. Global Tuberculosis Report. 2013

Fig. 1.
Variability of the WB-ICS assay for detecting cytokine-expressing CD4 and CD8 T cells. Establishing the lower limit of quantification of the WB-ICS assay (A–D). Unstimulated cells were spiked with a series of 2-fold diluted BCG-stimulated cells, then stained and analysed. Expected frequencies of cytokine positive cells were derived by dividing the frequencies of BCG-specific cytokine-expressing CD4 and CD8 T cells in the undiluted samples by the dilution factor. The plots show the expected and observed frequencies of BCG-specific IFN-γ (A), TNF-α (B), IL-2 (C) and IL-17 (D) CD4 T cells for each dilution factor. Establishing variability in detection of cytokine-expressing CD4 and CD8 T cells (E–

H). Aliquots of fixed cells from a single BCG-stimulated (and unstimulated) whole blood sample were thawed and stained independently by two operators on 4 different days. For each operator on each day, 2–5 aliquots of cells were stained and acquired to calculate the coefficient of variation (CV) for frequencies of BCG-specific IFN-γ-, TNF-α-, IL-2- or IL-17-expressing CD8 T cells (E, F) and CD4 T cells (G, H) for one operator are shown. The vertical dashed lines are a 30% CV mark. Frequencies of cytokine-expressing CD8 T cells with CVs less than 30% (IFN-γ) and more than 30% (TNF-α-, IL-2- and IL-17) (panels E and F) are shown. For each repeat, frequencies detected in the unstimulated sample were subtracted from frequencies in the respective BCG-stimulated sample. (I) Frequencies of non-specific IFN-γ-, TNF-α-, IL-2- or IL-17-expressing CD4 T cells, detected in unstimulated cells, in 54 adults. Horizontal lines within the boxes represent medians, boxes represent the IQR and whiskers represent the range. The black horizontal lines with stars represent the established limit of reliable quantification of the flow cytometer. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
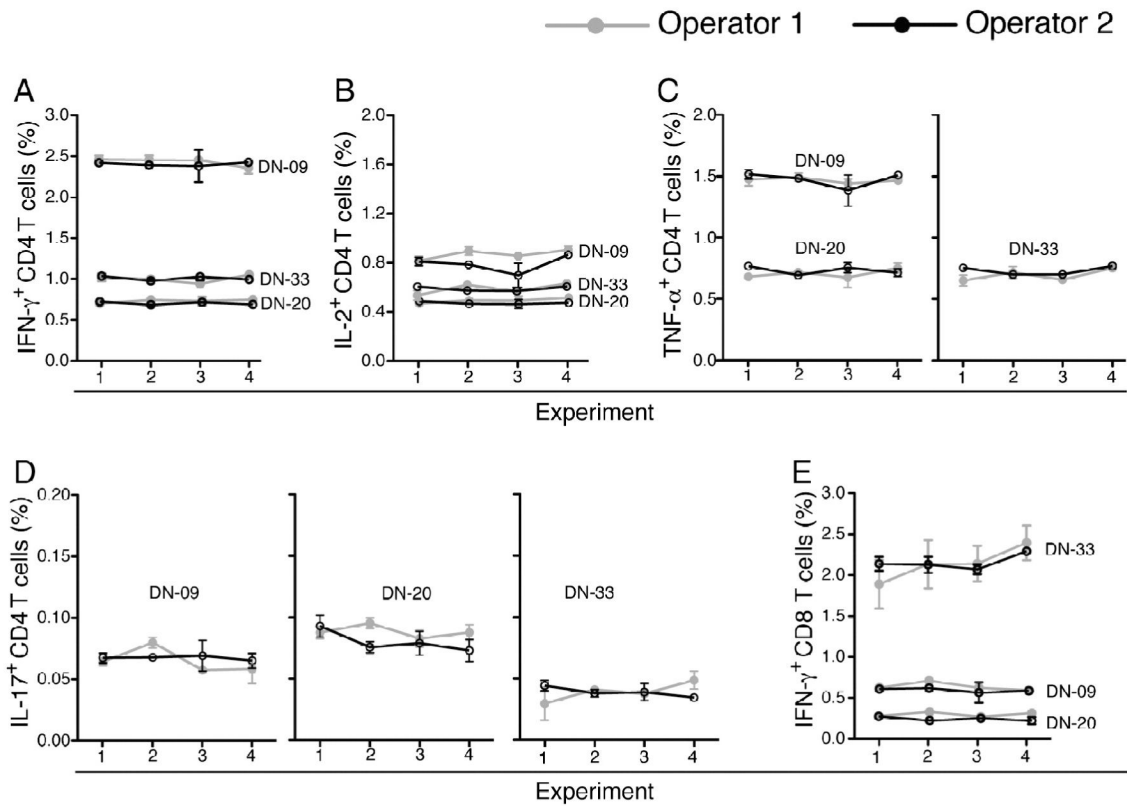
Fig. 2.
Assay reproducibility. Four aliquots of fixed cells from 3 BCG-stimulated whole blood samples were thawed, stained and acquired by two operators. Analysis of flow cytometry data to determine frequencies of BCG-specific CD4 T cells expressing IFN-γ (A), IL-2 (B), TNF-α (C) and IL-17 (D), and IFN-γ-expressing CD8 T cells (E) were completed by a single analyst. Frequencies detected in unstimulated cells were subtracted from the respective BCG-induced frequencies. Mean and standard deviation of the 3 samples for each experiment are shown.
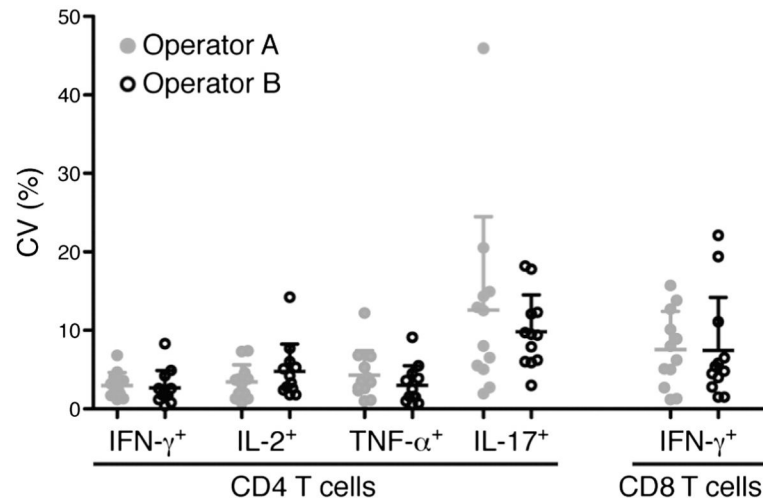
Fig. 3.
Precision and intra-operator variability. CVs of frequencies of BCG-specific IFN-γ, IL-2, TNF-α and IL-17-expressing CD4 T cells or IFN-γ-expressing CD8 T cells, measured by 2 operators (grey and black dots). Each operator independently stained and acquired 3 sets of samples in 4 independent experiments. Frequencies detected in unstimulated cells were subtracted from respective BCG-induced frequencies for each sample and experiment. Then, the CVs for each operator and cytokine were computed. Lines on the plots represent means and SD.
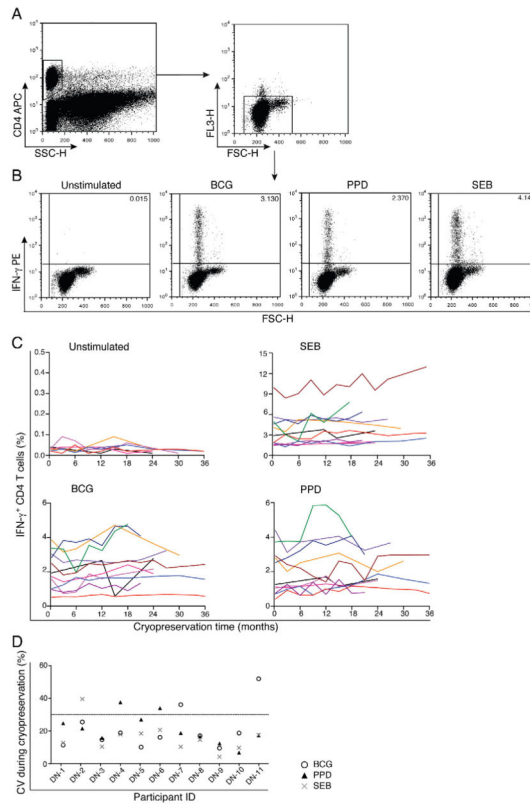
Fig. 4.
Effects of long-term cryopreservation on measurement of IFN-γ-expressing CD4 T cells. Blood from 11 donors was either left unstimulated or was stimulated and cells from each stimulated condition were split into 12 aliquots and cryopreserved. A single set of aliquots was thawed, stained and acquired on a FACSCalibur after one week of cryopreservation, and then every 3 months, up to 36 months. (A and B) Gating strategy used to measure IFN-γ-expressing CD4 T cells. (A) CD4 T cells were first selected by gating against side scatter (SSC). Then an exclusion gate comprising forward scatter (FSC) against the empty PerCP channel was applied to remove autofluorescent, compromised cells. (B) Representative dot plots of IFN-γ expression in CD4 T cells from unstimulated, BCG, PPD and SEB-stimulated conditions from a single donor after cryopreservation of stimulated whole blood. (C) Frequencies of IFN-γ-expressing CD4 T cells from 11 healthy donors analysed at different time intervals after cryopreservation. (D) Coefficient of variation of BCG-, PPD- and SEB-stimulated IFN-γ-expressing CD4 T cells, calculated for individual donors across different time intervals after cryopreservation.
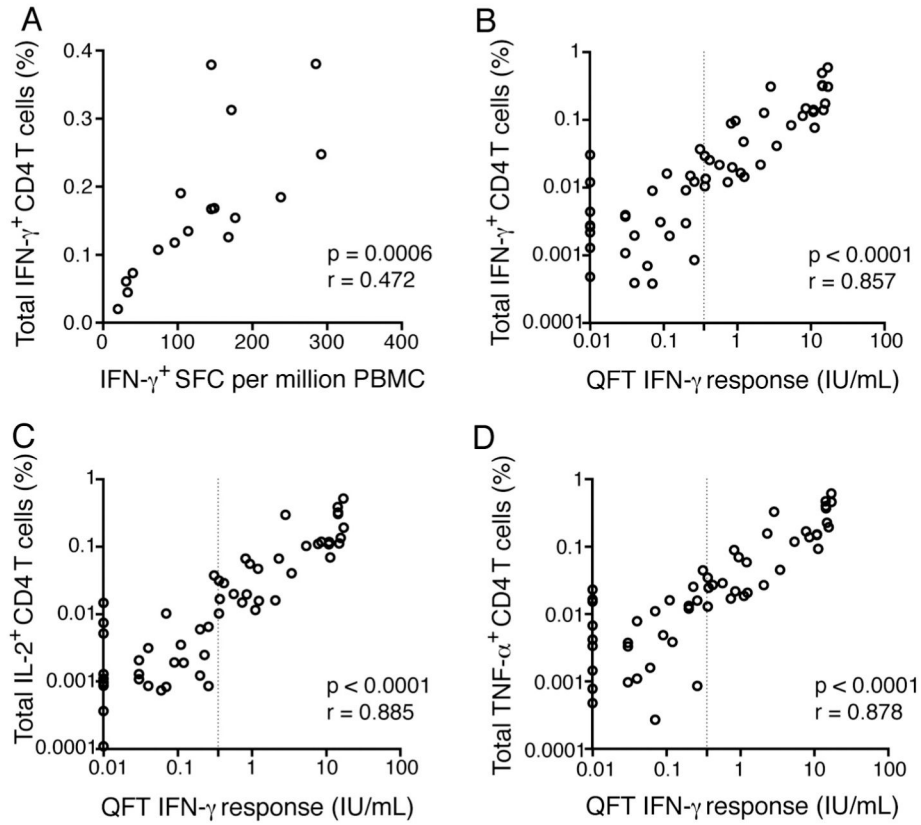
Fig. 5.
Comparison of frequencies of Th1 cytokine-expressing CD4 T cells, measured by WB-ICS, with IFN-γ release detected by IFN-γ ELISpot assay or by QuantiFERON-TB Gold In-Tube assay (QFT). (A) Ag85A-specific T cell responses measured in 4–6 month old infants 28 days after intradermal vaccination with the MVA85A vaccine (Scriba et al., 2011), by WB-ICS and IFN-γ ELISpot assay. (B–D) Correlation analyses between IFN-γ- (B), IL-2- (C) or TNF-α-expressing (D) ESAT-6/CFP-10-specific CD4 T cell frequencies detected by WB-ICS assay and plasma levels of IFN-γ measured by QFT. Assays were run in parallel on blood collected from 57 healthy adults with or without *M.tb* infection (QFT-positive or negative). The vertical line represents the QFT cut-off (0.35 IU/mL of IFN-γ) for diagnosis of *M.tb* infection. For graphical representation of data on a log scale, the background-normalized QFT response (TB-Ag minus unstimulated) from 5 donors was artificially adjusted from 0.0 IU/mL to 0.01 IU/mL. Significance and strength of the associations were calculated using Spearman's correlation test.
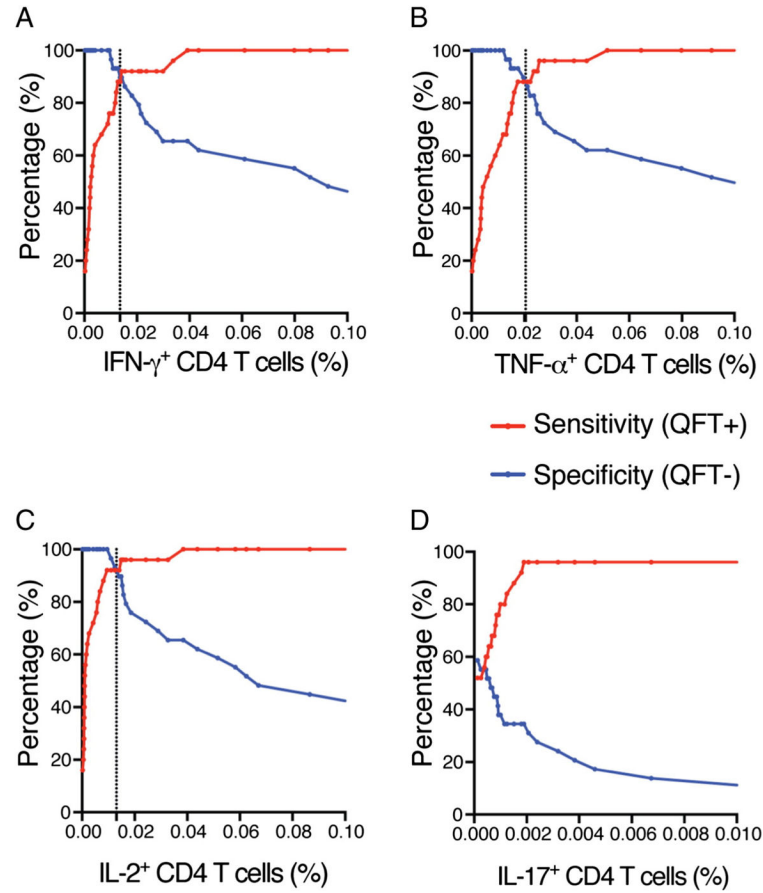
Fig. 6.
Sensitivity and specificity analysis of WB-ICS assay in QFT negative and positive adults.
Curves showing paired sensitivity and specificity values at variable WB-ICS assay cut-offs
for frequencies of ESAT-6/CFP-10-specific CD4 T cells expressing IFN-γ (A), TNF-α (B),
IL-2 (C), and IL-17 (D). Data are from 54 healthy adults with or without *M.tb* infection
(QFT-positive or negative). The vertical lines represent paired sensitivity and specificity
with the highest values, indicating the assay positivity threshold.

Table 1

Intra-assay coefficient of variation using the WB-ICS assay: Two operators (A and B) independently processed, stained and acquired fixed unstimulated and BCG-stimulated cells derived from 3 adult healthy volunteers (DN-09, DN-20 and DN-33). Each operator performed each experiment (on 4 different days) using 2–5 separate aliquots of identical cells. Thereafter, flow cytometry data analyses were done by a single analyst and the CVs of frequencies of specific total CD4 T cells expressing IFN-γ, IL-2, TNF-α or IL-17 and frequencies of specific total CD8 T cells expressing IFN-γ, were calculated for each operator, donor and experiment (day). Unstimulated responses were subtracted from the respective BCG-stimulated responses for each sample and experiment.

| | Intra-assay CV (%) | | | | | | | | | | Number of repeats per run/day |
| | IFN-γ⁺ CD4 T cells | | IL-2⁺ CD4 T cells | | TNF-α⁺ CD4 T cells | | IL-17⁺ CD4 T cells | | IFN-γ⁺ CD8 T cells | | |
Operator	A	B	A	B	A	B	A	B	A	B	A (B)
DN-09	1.9	1.2	4.6	4.2	3.5	2.5	5.0	5.9	6.2	6.5	5 (5)
	2.3	1.7	3.8	2.4	2.3	1.0	5.5	3.0	1.2	1.5	3 (3)
	1.2	8.3	1.3	14.2	2.6	9.1	2.7	18.2	2.7	22.1	3 (3)
	3.0	0.4	3.2	1.8	1.0	0.6	20.5	9.4	1.3	5.4	3 (3)
DN-20	4.1	1.8	3.8	1.8	4.0	1.6	12.9	9.5	5.0	5.8	5 (4)
	1.7	2.6	1.3	6.0	1.1	3.9	8.0	6.0	8.0	11.1	3 (3)
	6.8	4.9	7.4	7.7	12.2	5.5	14.3	12.1	12.7	4.8	3 (3)
	3.6	2.6	3.7	5.3	5.3	4.5	12.5	12.3	5.1	19.4	3 (3)
DN-33	4.7	1.8	7.3	2.6	6.8	0.7	45.9	9.7	15.7	4.0	5 (5)
	1.3	0.8	2.0	3.3	6.7	1.5	6.5	7.9	13.8	4.5	3 (3)
		4.2		5.1		3.6		17.8		2.8	2 (3)
	3.2	2.0	2.0		3.4		14.9		8.9		3 (2)

Table 2

Robustness of the WB-ICS assay. Two operators (A and B) independently processed, stained and acquired fixed cells from either unstimulated or BCG-stimulated samples derived from 3 adult healthy volunteers. Variability in the temperature and duration of monoclonal antibody cocktail staining was assessed. Each operator performed the experiments independently. Thereafter, flow data analyses were done by one analyst and the mean frequencies of specific total CD4 T cells expressing IFN-γ, IL-2, TNF-α or IL-17 and frequencies of specific total CD8 T cells expressing IFN-γ, were calculated for each operator and each staining condition. Unstimulated responses were subtracted from the respective BCG-stimulated responses for each sample and experiment.

ICS antibody staining temperatures	Operator	ICS antibody staining duration (min)	Mean cytokine+ CD4 T cells (%)				Mean cytokine+ CD8 T cells (%)
			IFN-γ	IL-2	TNF-α	IL-17	IFN-γ
4 °C	A	45	1.070	0.487	0.823	0.057	0.791
4 °C	B	45	1.104	0.530	0.819	0.056	0.728
20–23 °C	A	45	1.110	0.556	0.909	0.060	0.737
20–23 °C	B	45	1.135	0.565	0.899	0.067	0.718
4 °C	A	60	1.181	0.550	0.888	0.056	0.739
4 °C	B	60	1.150	0.587	0.886	0.057	0.780
20–23 °C	A	60	1.073	0.526	0.847	0.064	0.678
20–23 °C	B	60	1.151	0.566	0.892	0.064	0.716
4 °C	A	90	1.277	0.535	0.896	0.066	0.825
4 °C	B	90	1.174	0.591	0.927	0.069	0.834
20–23 °C	A	90	1.146	0.576	0.975	0.062	0.816
20–23 °C	B	90	1.186	0.611	0.925	0.071	0.730
Overall mean			1.146	0.557	0.891	0.062	0.758
Standard deviation			0.056	0.034	0.044	0.005	0.050
CV			4.927	6.080	4.987	8.407	6.593