Differential Gene Expression Identifies Novel Markers of CD4+ and CD8+ T Cell Activation Following Stimulation by *Mycobacterium tuberculosis*

Jacqueline M. Cliff, Iryna N. J. Andrade, Rohit Mistry, Christopher L. Clayton, Mark G. Lennon, Alan P. Lewis, Ken Duncan, Pauline T. Lukey and Hazel M. Dockrell


doi: 10.4049/jimmunol.173.1.485

http://www.jimmunol.org/content/173/1/485

---

**Supplementary Material**

http://www.jimmunol.org/content/suppl/2004/06/17/173.1.485.DC1

**References**

This article cites 51 articles, 22 of which you can access for free at:
http://www.jimmunol.org/content/173/1/485.full#ref-list-1

**Subscription**

Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**Permissions**

Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**

Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Differential Gene Expression Identifies Novel Markers of CD4⁺ and CD8⁺ T Cell Activation Following Stimulation by Mycobacterium tuberculosis

Jacqueline M. Cliff, Iryna N. J. Andrade, Rohit Mistry, Christopher L. Clayton, Mark G. Lennon, Alan P. Lewis, Ken Duncan, Pauline T. Lukey, and Hazel M. Dockrell

T cell activation in response to antigenic stimulation is a complex process, involving changes in the expression level of a large number of genes. We have used cDNA array technology to characterize the differences in gene expression between human CD4⁺ and CD8⁺ T cells. PBMC from six healthy donors were stimulated with live Mycobacterium tuberculosis, and the gene expression profiles of each donor's CD4⁺ and CD8⁺ T cells were analyzed separately. ANOVA revealed 518 genes that were consistently differentially expressed between CD4⁺ and CD8⁺ T cells. These differentially expressed genes include a combination of well-known, previously characterized genes with a range of biological functions and unknown in silico predicted hypothetical genes. Where possible, the novel genes have been characterized using bioinformatics, and putative transcription factors, signaling molecules, transmembrane, and secreted factors have been identified. A subset of these differentially expressed genes could be exploited as markers of CD4⁺ and CD8⁺ T cell activation for use in vaccine trials. These observed differences in the gene expression profile of CD4⁺ and CD8⁺ T cells following activation by a human pathogen contribute to an increased understanding of T cell activation and differentiation and the roles these T cell subsets may play in immunity to infection. The Journal of Immunology, 2004, 173: 485–493.

W

hen a T cell is stimulated by encountering an APC expressing its cognate Ag, many biological processes occur, including metabolic activation, morphological changes, proliferation, and differentiation into distinct effector and memory cells. These processes are mediated in part by alterations in gene expression levels, and these changes are affected by co-stimulation via cytokine receptors and accessory and adhesion molecules, as well as the signals emanating from TCR triggering. CD4⁺ and CD8⁺ T cells are phenotypically similar, and both subsets respond to antigenic stimulation in a similar manner. Indeed, activated CD4⁺ and CD8⁺ T cells can both be induced to differentiate into either type 1 or 2 cells, secreting similar patterns of cytokines (1). The main difference between the two T cell subsets is their distinct roles in protective immunity, as effector CD8⁺ T cells recognize and destroy infected host cells by recognizing endogenous peptides presented by MHC class I, whereas effector CD4⁺ T cells recognize exogenous peptides presented via MHC class II and provide help to other immune cells. These differences in effector CD4⁺ and CD8⁺ T cells must be reflected in differences in the gene expression profile of the two T cell types.

High density cDNA and oligonucleotide microarray hybridization technologies can be used to measure the mRNA expression levels of thousands of genes simultaneously, and can be used to identify differentially expressed genes in different cell types or tissues. They have been successfully used in the analysis of various aspects of T cell function, including T cell activation in vivo in response to staphylococcal enterotoxin B (2) and determination of the gene expression profile of suppressor T cells in mice (3, 4). In human T cells, the genetic profiles of type 1 and 2 CD4⁺ and CD8⁺ T cells have been characterized (5, 6), and the effects of signaling through the IL-2R (7) and members of the CD28 family (8, 9) on TCR signaling have been determined.

To date, there has not been a systematic array-based analysis of the differences in gene expression profiles of human CD4⁺ and CD8⁺ T cells following physiological Ag-specific stimulation. This study addresses this issue by investigation of effector CD4⁺ and CD8⁺ T cells specific for a relevant human pathogen, Mycobacterium tuberculosis, which leads to 8 million new cases of tuberculosis (TB)³ disease and 2 million deaths each year (10). The efficacy of the currently available vaccine against M. tuberculosis, Mycobacterium bovis bacillus Calmette-Guerin (BCG), has been variable in trials with ~80% observed protection against adult pulmonary TB in the United Kingdom, but much lower protection demonstrated in large parts of the world, including Sub-Saharan Africa and Southeast Asia (11). The pivotal role of CD4⁺ T cells in protection against TB is clearly demonstrated by the susceptibility of HIV/AIDS patients to develop TB disease, and this correlates with the decline in CD4⁺ T cell count. In mice, Ab-mediated depletion of CD4⁺ T cells results in rapid reactivation of a persistent TB infection (12). Effector CD4⁺ T cells are recruited to infected lungs, where they secrete TNF-α and IFN-γ, leading to activation of the infected alveolar macrophages, enabling them to

³Abbreviations used in this paper: TB, tuberculosis; BCG, bacillus Calmette-Guerin; DAP10, DNAX activation protein 10; ESTs, expressed sequence tags; MOX, multiplicity of infection; PPD, purified protein derivative; qRT-PCR, quantitative RT-PCR.

Copyright © 2004 by The American Association of Immunologists, Inc.

0022-1767/04/$02.00
kill the mycobacteria (13). A CD8 \(^+\) T cell response is also necessary for immunity. \(\beta\)-microglobulin-deficient mice, which lack functional CD8 \(^+\) T cells, are highly susceptible to \textit{M. tuberculosis} (14), and these cells appear to play a crucial role in the latent phase of infection (15). In human TB patients, healthy contacts, and BCG-vaccinated donors, CD8 \(^+\) T cells reactive against a number of mycobacterial Ags have been identified (16, 17). Such CD8 \(^+\) T cells produce IFN-\(\gamma\) and TNF-\(\alpha\) (18) and are cytolytic for infected monocyte-derived macrophages (19). Furthermore, these CD8 \(^+\) T cells secrete granulysin, which, in conjunction with perforin, can directly kill the \textit{M. tuberculosis} bacillus within a macrophage (20). Cytotoxic CD4 \(^+\) T cells can also be generated by \textit{M. tuberculosis} stimulation. However, the kinetics of CD4 \(^+\) T cell-mediated killing are different from those of CD8 \(^+\) T cells, occurring after 16 h rather than 4 h following antigenic stimulation in vitro (21). The control of cell activation, which is mediated by the production of IL-10 and TNF-alpha, is therefore different, and may be reflected in differences in the expression level of certain genes.

\textit{M. tuberculosis} exerts an immunosuppressive effect on host T cell activation, which is mediated by the production of IL-10 and TGF-\(\beta\) by infected macrophages (22). At diagnosis, TB patients exhibit a depressed CD4 \(^+\) T cell proliferation and IFN-\(\gamma\) secretion response to purified protein derivative (PPD) stimulation in vitro compared with BCG-vaccinated control subjects (23). CD8 \(^+\) T cell responses are also reduced in TB patients, with a decrease in the amount of IFN-\(\gamma\) and TNF-\(\alpha\) produced (24) and also in cytolytic activity (25). Suppression of these functions of CD4 \(^+\) and CD8 \(^+\) T cells may therefore be to the pathogen’s advantage.

In the present study, we used human cDNA arrays to characterize the differences in the gene expression profile of effector CD4 \(^+\) and CD8 \(^+\) T cells, in response to \textit{M. tuberculosis} infection of human PBMC cultures. We identified 518 genes that were significantly differentially expressed, many of which have not been characterized previously as involved in lymphocyte activation.

### Materials and Methods

#### Generation of CD4 \(^+\) and CD8 \(^+\) T cell samples

Buffalo coats from healthy donors were purchased from the South Thames Blood Transfusion Service (London, U.K.), and PBMC were obtained by density centrifugation over Ficoll. The majority of the adult United Kingdom population has received BCG vaccination. The PBMC were cultured at 1 \times 10^6/ml in complete medium (RPMI 1640 supplemented with 5% AB serum) stimulated with live \textit{M. tuberculosis} strain H37Rv at a multiplicity of infection (MOI) of 1:1 (\textit{M. tuberculosis}:monocyte) for 6 days. Concurrently, autologous monocyte-derived macrophages were obtained by adherence, and cultured in complete medium for 6 days. On day 6, one-half of the stimulated PBMC was collected by centrifugation and restimulated by incubation with freshly infected (1:1 MOI) autologous monocyte-derived macrophages. Sixteen hours later, CD4 \(^+\) and CD8 \(^+\) T cells were positively selected from both the restimulated and the nonrestimulated PBMC samples, using Dynabeads (Dynal Biotech, Wirral, U.K.). In pilot studies using \textit{M. bovis} BCG stimulation in place of \textit{M. tuberculosis}, the average purity of CD4 \(^+\)-selected T cells as determined by flow cytometry was 98.7% CD4 \(^+\), 0.4% CD8 \(^+\), and for CD8 \(^+\)-selected T cells 97.8% CD4 \(^+\), and 1.2% CD8 \(^+\) (\(n = 3\)). The purity of the T cell samples used in the differential gene expression study was not tested because of safety considerations.

Total RNA was extracted using RNeasy B (Biogenes, Dorset, U.K.). Typically, RNA samples were obtained from 3–6 \times 10^7 cells, resulting in an average total RNA yield of 15 \mu g. Any DNA contamination in the RNA samples was removed using DNase I (Ambion, Cambridge, U.K.) before use in differential gene expression or real-time quantitative RT-PCR (qRT-PCR) experiments. The quality of RNA samples was checked by gel electrophoresis.

#### Proliferation assay

PBMC from each donor were incubated at 1 \times 10^6/ml in round-bottom 96-well plates in complete medium. They were stimulated with \textit{M. tuberculosis} PPD (Statens Serum Institut, Copenhagen, Denmark), or with live \textit{M. bovis} BCG (GlaxoEvans strain). For the last 16 h of the 7-day assay, 1 \muCi of [\(^{3}H\)]thymidine (Amersham Biosciences, Bucks, U.K.) was added. Points shown represent the means of quadruplicate measurements for each donor.

#### IFN-\(\gamma\) ELISA

PBMC were incubated at 1 \times 10^6/ml and stimulated with PPD or \textit{M. bovis} BCG, and supernatants were collected on day 6. Quadruplicate supernatants were pooled and assayed in duplicate by sandwich ELISA, using a matched Ab pair and rIFN-\(\gamma\)-standard purchased from BD Pharmingen (San Diego, CA).

### Differential gene expression

Generation of nylon cDNA arrays, synthesis of \(^{3}P\)-labeled cDNA complex probes, and hybridization of complex probes to the arrays were performed, as previously described (26). cDNA arrays were synthesized from a normalcy-derived cDNA library from five healthy donors, which was purchased from Invitrogen Life Technologies (Renfrewshire, U.K.). The library was normalized so that the redundancy was reduced, and therefore the representation of low-level-expressed genes, such as cytokine and receptor genes, was enhanced (Invitrogen sales literature), allowing as many unique genes as possible to be analyzed. A total of 50,000 clones was picked from this library, and the inserts were amplified by PCR. The average product length was 1.2 kb. These PCR products were spotted in duplicate across three nylon membranes. Similarly, the human sequence-verified arrays were produced from nearly 8,000 clones (Incyte Genomics, Palo Alto, CA), comprising 6,000 characterized genes and 2,000 expressed sequence tags (ESTs), spotted in duplicate. Before use, the DNA loading of each array was checked by hybridization of \(^{3}P\)-labeled oligonucleotides, which were targeted to the vector sequence of the PCR products.

RNA samples (5 \mu g) were reverse transcribed using Superscript II (Invitrogen Life Technologies) and oligo(dT)\(_{18}\)VN primer, in the presence of \(\alpha\)-\(^{3}P\)dCTP (Amersham Biosciences). Unincorporated nucleotides were removed by purification on Sephadex G50 spin columns, followed by 40 cycles of 95°C for 1 min, 60°C for 1 min. Duplicate complex probes were synthesized from each RNA sample, and their quality was assessed by PAGE after denaturation. The complex probes were then hybridized to the arrays at 45°C in DIG Easy Hyb (Roche Diagnostics, East Sussex, U.K.) for 72 h, and the membranes were washed, as described (27). Complex probe hybridization intensities were quantified using a Storm phosphor imager (Amersham Biosciences), and in-house spot-finding software.

Clones from the normalized leukocyte library were sequenced using a pCMV sport vector-derivative primer (ATTTAGGTGACACTATAGAAG), using ABI Prism 3700 capillary sequencers with BigDye Terminator chemistry (Applied Biosystems, Foster City, CA). Clones used on the human sequence-verified arrays were sequenced first to ensure accuracy.

All data were analyzed on the log10 (base 10) scale, and the hybridization intensities on each individual array were normalized, by dividing the intensity for each gene by the geometric mean for that array. All genes that were not statistically significantly different from background were removed from the analysis: this was nearly 25% of the normalized leukocyte library array. For nearly 50% of the human sequence-verified arrays, there were four hybridization intensity values for each T cell sample from each donor, due to the duplicate reverse-transcription step and duplicate spots on each grid, allowing a determination of the variability in measurement for each gene within each donor. ANOVA (26) was performed using SAS software (version 8; SAS Institute, Cary, NC), to compare the intensities, within a donor, between CD4 \(^+\) and CD8 \(^+\) T cells for both the early and late stimulation conditions.

#### Real-time qRT-PCR validation

The cDNA equivalent of 50 ng of total RNA was used in each reaction. RNA was reverse transcribed using Superscript II, in duplicate for each sample. The PCR were performed in SYBR Green PCR Master Mix (Applied Biosystems), containing 400 nM of each primer. These reactions were performed by an ABI Prism 7500 Sequence Detection System, with a 10-min 95°C denaturation step, followed by 40 cycles of 95°C for 15 s, 60°C for 1 min. The fluorescence emitted from the dsDNA PCR product was measured at the end of each cycle. At the end of all the cycles, a dissociation analysis was performed to ensure the presence of only one specific PCR product. The starting mRNA copy number was determined by comparison with a standard curve. In parallel, PCR were performed on reverse-transcriptase-negative control samples, to ensure there was no genomic DNA contamination. For each sample and for each gene, the results were
normalized by dividing by the calculated β-actin copy number, to allow comparisons between samples.

The PCR primers for each gene were: IFN-induced protein 35 kDa, CAGCAGGCCTAGCAGTCTTC and CTTCCGACTCTGGTGGAAC; brefeldin A-inhibited guanine-nucleotide exchange protein 2, TCCCATGATTGCTATAAAACG and GGTGCGAGTTAGGCTGACG; furin, TCTGGTACCCAGCCATCTG and CTGTGCACCAACCCAGCATC; GATTGGCTATAAACG and GGTGGCAGTTAGCCGTTACAAG; PPD; and CD40 ligand.

**In silico analysis of differentially expressed genes**

The predicted full-length protein sequences obtained from the National Center for Biotechnology Information were used in the transmembrane hidden Markov model program based at the Center for Biological Sequence Analysis (29) to predict whether the genes are likely to be transmembrane proteins. Signal peptides were identified using SignalP to determine whether proteins are likely to be secreted (30). Their intracellular location was predicted using PSORT II (31). Previously characterized domains and motifs were identified using screening the PROSITE (32), Pfam (33), PRINTS (34), and SMART (35) databases.

**Results**

**Mycobacteria-specific responses of donors**

CD4⁺ and CD8⁺ T cell samples were prepared from six healthy donors. Initially, the memory T cell responsiveness to mycobacterial Ags was tested, by incubation of PBMC with either live *M. bovis* BCG or PPD from *M. tuberculosis* H37Rv. Each donor responded in an Ag-specific manner, as measured by both proliferation assay and IFN-γ secretion (Fig. 1), reflecting memory T cell activation due to prior BCG vaccination of the donors. The magnitude of the responses varied considerably between donors, which is typical in BCG-vaccinated subjects (36).

![FIGURE 1. Donor T cell responses to mycobacterial Ags. PBMC were cultured with live *M. bovis* BCG at 1:1 MOL (mycobacterium:monocyte) or with 10 µg/ml PPD for 6 days. A, T cell proliferation was assayed by [³H]thymidine incorporation. B, IFN-γ secretion was assayed by ELISA.](http://www.jimmunol.org/)

**Identification of differentially expressed genes**

PBMC from the six donors were incubated with live *M. tuberculosis* for either 7 days to identify late activation genes, or 6 days, followed by restimulation for a further 16 h with freshly infected autologous monocyte-derived macrophages, to identify early activation genes. An early activation time point could not be generated by 16-h stimulation directly ex vivo, as the mycobacteria-specific memory T cell frequency is too low without prior clonal expansion in vitro. Array-based experiments only test genes present on the arrays, and most arrays only contain previously characterized genes and ESTs. So as to comprehensively cover the gene transcript profile of the activated T cells, we chose to screen a normalized leukocyte cDNA library. An in-house GlaxoSmithKline human sequence-verified cDNA array was also screened.

Many genes were found to be differentially expressed by CD4⁺ and CD8⁺ T cells at both the early and late activation time points. As expected, there was considerable variation between donors. ANOVA was performed to identify genes that were consistently differentially expressed by T cells from all six donors. Fold differences between expression levels in CD4⁺ and CD8⁺ T cells were calculated from the geometric means of the hybridization intensities from the six donors for both the early and the late activation time points. A total of 518 genes was at least 1.3-fold more highly expressed in either CD4⁺ or CD8⁺ T cells at least one time point, with a p value <0.05 (see Supplemental tables I and II for full lists of these genes). Functional categories were ascribed to these genes using a combination of literature searches, the Locuslink, Online Mendelian Inheritance in Man and Unigene databases at the National Center for Biotechnology Information (37), and the GeneCards database (available at [http://](http://))

The on-line version of this article contains supplemental material.

![FIGURE 2. Proportions of functional classes of differentially expressed genes. PBMC were infected with live *M. tuberculosis* for 6 days, followed by 16-h restimulation (early activation genes), or were incubated for 7 days with *M. tuberculosis* (late activation genes) before harvest. CD4⁺ and CD8⁺ T cells were isolated from the PBMC, and total RNA was reverse transcribed to make ³²P-labeled cDNA probes that were used to screen the cDNA arrays. All statistically significantly (p < 0.05, n = 6) differentially expressed genes at either time point were classified according to biological function. X, Unknown: this category includes in silico predicted hypothetical genes, as well as identified expressed genes with no functional information.](http://www.jimmunol.org/)
bioinformatics.weizmann.ac.il/cards). Genes involved in many different biological and molecular processes were identified using this cDNA array library screening methodology (Fig. 2).

There were similar numbers of genes involved in some biological pathways, such as the immune system, signal transduction, and gene transcription, which were more highly expressed in either CD4⁺ or CD8⁺ T cells (Table I). Other types of genes, in particular those involved in protein biosynthesis, metabolism, and those encoding structural proteins, tended to be more highly expressed in CD4⁺ T cells, especially at the late activation time point. This may reflect the higher proliferative rate of CD4⁺ T cells. The overall pattern of differential gene expression was different at the early and late activation time points (Supplemental tables I and II).

Validation of cDNA array screening technique

To confirm that the gene expression changes were valid, real-time qRT-PCR experiments were performed on biologically independent samples. Six genes identified as being differentially expressed by the cDNA array hybridization experiments were analyzed, and the expression patterns as measured by qRT-PCR of these genes were similar to those observed by cDNA array (Fig. 3). These six genes include genes more highly expressed in either CD4⁺ or CD8⁺ T cells, at either early (furin, hypothetical protein LOC285381) or late (ARFGEF2 (brefeldin A-inhibited guanine nucleotide-exchange protein 2), KIAA0100, FLJ38447 fis, IFN-induced protein 35) activation time points, and either previously characterized genes or unknown, hypothetical genes. These results clearly demonstrate the biological as well as the technical reproducibility of our system.

Kinetics of gene expression variance

The 518 genes that were differentially expressed were clustered into six groups, based on their expression patterns in CD4⁺ and CD8⁺ T cells at the early and late activation time points (supplementary figure 1 and supplemental table III). Some genes were consistently more highly expressed in either CD4⁺ or CD8⁺ T cells, whereas for other genes the differences were only observed at either the early or the late activation time point. Overall, more genes were more highly expressed in CD4⁺ T cells (335) than in CD8⁺ T cells (184).

Differential expression of immune system genes

Genes involved in various aspects of immune function were differentially expressed between CD4⁺ and CD8⁺ T cells (Fig. 4). These included secreted molecules, such as IL-16, which was expressed more highly by CD8⁺ T cells, and cyclophilin A, which was more highly expressed by CD4⁺ T cells. Various receptor

Table I. Numbers of genes more highly expressed in CD4⁺ and CD8⁺ T cells, at early and late activation time points

<table>
<thead>
<tr>
<th>Category</th>
<th>Early Activation Genes</th>
<th>Late Activation Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Higher expression in CD4⁺ T cells</td>
<td>Higher expression in CD8⁺ T cells</td>
</tr>
<tr>
<td>Immune</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Signal transduction</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Transcription factor</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>Nuclear - other</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Intracellular trafficking</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Proteolysis</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Protein biosynthesis</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Metabolism</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Hormone</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Structural</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Solute transporter</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Channel</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Unknown</td>
<td>44</td>
<td>12</td>
</tr>
<tr>
<td>ESTs</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>146</td>
<td>77</td>
</tr>
</tbody>
</table>

*Experimental details are given in Fig. 2.*
molecules were also differentially expressed, including the IL-7R expressed by CD8\(^+\) T cells and CD99 expressed by CD4\(^+\) T cells. Interestingly, the chemokine receptors CCR5, CX3CR1, and the Duffy blood group Ag were more highly expressed in CD4\(^+\) T cells than CD8\(^+\) T cells. Several proteins involved in signal transduction were differentially expressed, including the \textit{vav} 1 oncogene, which was up-regulated in CD8\(^+\) T cells, and IL-2-inducible T cell kinase, which was up-regulated in CD4\(^+\) T cells. Three MHC class II genes, DQ\(\beta\), DR\(\beta\), and DO\(\beta\), were all preferentially up-regulated by the activated CD4\(^+\) T cells (Fig. 4). Data from the HSV grids showed that
In silico characterization of unknown genes

Of the genes found to be differentially expressed, 19% have no known function, and most of these are in silico predicted genes based on analysis of the genomic DNA sequence. A further 6% of the differentially expressed clones are ESTs that do not map to known genes (Fig. 2). To gain some insight into their potential role in CD4+ T cells, we performed additional analysis of the genomic DNA sequence. A further 6% of the differentially expressed clones are ESTs that do not map to known genes (Fig. 2). To gain some insight into their potential role in CD4+ T cells and the results of these predictions are summarized in Table II.

From these predictions, we characterized three of the nine proteins as having transmembrane domains, the hypothetical proteins FLJ10936, DKFZp564A022, and FLJ11240, which showed different expression patterns. All three of these proteins were predicted to be located in the endoplasmic reticulum. Another two proteins, hypothetical protein MGC2508 and hypothetical protein LOC51255, are likely to be transcription factors, as they have high probabilities of nuclear location and contain predicted zinc finger domains. These genes have opposing expression patterns. The hypothetical protein MGC8407 was more highly expressed in CD4+ T cells at both time points, and this protein contains a tyrosine kinase domain and is therefore possibly involved in signal transduction. Three genes were identified that are likely to encode secreted proteins. These are 37-kDa leucine-rich repeat protein and hypothetical protein MGC10993, which are more highly expressed in CD4+ T cells, and hypothetical gene supported by AL449243, which was expressed more highly by CD8+ T cells.

Discussion

This is the first comprehensive description of the differences in the gene expression profile of human activated CD4+ and CD8+ effector T cells following Ag-specific stimulation. In this study, we used live M. tuberculosis to generate effector T cells in PBMC cultures from BCG-vaccinated donors. Live mycobacteria are known to elicit stronger CD8+ T cell activation than dead mycobacteria or soluble Ags (19), possibly by creating pores in the vacuole membrane, allowing Ags to enter the cytoplasm and thereby enter the MHC class I presentation pathway. CD4+ T cells are also strongly activated in this system, due to the binding of mycobacterial peptides to MHC class II molecules in the infected macrophage phagolysosome. In addition, the monocyte/macrophage APCs may receive signals from live M. tuberculosis, such as via Toll-like receptors, some of which may be M. tuberculosis specific, and this may affect the activation of both CD4+ and CD8+ T cells. M. tuberculosis infection also leads to immunosuppression and immunomodulation, including down-regulation of MHC class II and induction of IL-10 and TGF-β, which inhibit CD4+ and CD8+ T cell activation (22, 38, 39). In this way, we believe our stimulation conditions more closely reflect the activation of both CD4+ and CD8+ T cells within an infected lung granuloma than other culture systems. We were able to identify a large number of genes that were differentially expressed between CD4+ and CD8+ T cells, and these came from a variety of different gene families, underlining the fact that immune function is the result of interplay of a large number of biological processes.

The recall response of memory T cells to in vitro stimulation is highly variable between human donors. ANOVA was performed between paired CD4+ and CD8+ effector T cells from individual donors, to identify those genes that were consistently differentially expressed. The time point at which gene expression levels are measured is clearly important. We observed large differences in the genes that were more highly expressed at either an early (16-h restimulation) or late (7-day stimulation) time point, although over one-half of the genes showed a consistent differential gene expression pattern.

A low dynamic range in the fold difference in expression level was observed for a combination of reasons. The T cells were stimulated in an Ag-specific manner, so the collected samples contained a mixed population of M. tuberculosis-specific activated cells and nonactivated resting T cells. mRNA from the nonactivated T cells will have reduced the signal observed derived from the activated cells, thereby dampening the observed fold differences. Therefore, a 1.3-fold overall observed fold difference in the mixed cell populations reflects much greater variations occurring within the responder cell populations. Also, the extent of the donor variability means that for each gene, some donors will have much higher fold differences between CD4+ and CD8+ T cells, while others will have lower differences. As both CD4+ and CD8+ T cells are taken from each donor, they are compared using the variability within donors. Within donor variability will be smaller than between donor variability; thus, we have a more statistically powerful experiment, which enables us to detect smaller fold differences as being significant than one would expect if one had to use between donor variability.

### Table II. In silico characterization of novel genes

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Reference Sequence</th>
<th>Length</th>
<th>Secreted</th>
<th>Location</th>
<th>Domains</th>
<th>Expression Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyp Prot* FLJ11240</td>
<td>NM_018368</td>
<td>540</td>
<td>No</td>
<td>ER (66.7%)</td>
<td>9-Transmembrane actin binding</td>
<td>Up CD4+ early</td>
</tr>
<tr>
<td>37-kDa Leucine-rich repeatProt</td>
<td>NM_005824</td>
<td>313</td>
<td>Yes</td>
<td>Nuclear (73.9%)</td>
<td>Zinc finger C3HC4 type</td>
<td>Up CD4+ early</td>
</tr>
<tr>
<td>Hyp Prot MGC10993</td>
<td>NM_030577</td>
<td>311</td>
<td>Yes</td>
<td>Nuclear (70%)</td>
<td>C2H2 type</td>
<td>Up CD4+ early</td>
</tr>
<tr>
<td>Hyp Prot LOC51255</td>
<td>NM_016494</td>
<td>150</td>
<td>No</td>
<td>Nuclear (65%)</td>
<td>Tyrosine kinase</td>
<td>Up CD4+ late</td>
</tr>
<tr>
<td>Hyp Prot MGC8407</td>
<td>NM_024046</td>
<td>501</td>
<td>No</td>
<td>Nuclear (91%)</td>
<td>4-Zinc finger C2H2 type</td>
<td>Up CD8+ early</td>
</tr>
<tr>
<td>Hyp Prot MGC2508</td>
<td>NM_023427</td>
<td>170</td>
<td>No</td>
<td>Nuclear (91.3%)</td>
<td>4-Zinc finger C2H2 type</td>
<td>Up CD8+ early</td>
</tr>
<tr>
<td>Hyp Prot supported by AL449243</td>
<td>NM_033318</td>
<td>107</td>
<td>Yes</td>
<td>No</td>
<td>6-Transmembrane DUF92</td>
<td>Up CD8+ early</td>
</tr>
<tr>
<td>Hyp Prot FLJ10936</td>
<td>NM_018279</td>
<td>396</td>
<td>No</td>
<td>ER (66.7%)</td>
<td>3-Transmembrane ring_2_zinc finger</td>
<td>Up CD8+ late</td>
</tr>
<tr>
<td>Hyp Prot DKFZp564A022</td>
<td>NM_030954</td>
<td>258</td>
<td>No</td>
<td>ER (44.4%)</td>
<td>3-Transmembrane ring_2_zinc finger</td>
<td>Up CD8+ late</td>
</tr>
</tbody>
</table>

* Hyp Prot, hypothetical protein.

ER, endoplasmic reticulum.
Some genes that were found to be up-regulated in either CD4+ or CD8+ T cells have been described previously. For example, the NK cell-activating receptor NK2D is known to be expressed on CD8+ T cells, where it pairs with the adaptor molecule DNAX activation protein 10 (DAP10), thereby providing costimulation via a B7/CD28-independent pathway (40). In the present study, both NK2D and DAP10 were identified as being significantly more highly expressed in CD8+ T cells. The ligands for NK2D, MHC class I chain-related gene A and B, are expressed on APCs in response to infection or stress, and it is known that *M. tuberculosis* induces MHC class I chain-related gene A expression on dendritic cells and epithelial cells (41). Also, our finding of enhanced IL-7R expression on CD8+ T cells fits in with current understanding of the role of IL-7 in driving memory CD8+, but not CD4+, T cell proliferation, and survival (42). Although expression of MHC class II molecules is constitutive on professional APCs, the expression on T cells is tightly regulated (43). Clearly, CD4+ T cells up-regulated MHC class II expression more than CD8+ T cells in response to *M. tuberculosis* stimulation. CD4+ T cells also up-regulated their expression of the chemokine receptors CCR5, CX3CR1, and the Duffy group Ag more than CD8+ T cells; this would lead to differential homing of the T cells during an in vivo infection. Different requirements for the secretion of biologically active IL-16 by CD4+ and CD8+ T cells have been described previously, with CD4+ T cells displaying slower kinetics and a dependence on costimulation (44). In our system, we observed increased IL-16 mRNA expression in the CD8+ T cells (Fig. 4); hence, our stimulation conditions may favor IL-16 production by CD8+ over CD4+ T cells.

In contrast, other genes were found to be differentially expressed, which had not previously been identified as being expressed in CD4+ or CD8+ T cells. For example, the serine endoprotease furin is a widely expressed transmembrane protein located in various cellular compartments, including the trans-Golgi network and endosomal compartments. It is known to be expressed in endocrine and neuroendocrine cells in the immature secretory granules, but not the mature granules (45); it is possible that its higher expression in activated CD8+ T cells is important for the maturation of the cytolytic granules in these cells. We also observed the enhanced expression of an IFN-inducible leucine zipper protein, IFN-induced protein 35 in CD8+ T cells compared with CD4+ T cells. This protein forms a heterodimer with the AP-1 transcription factor (46), and may thereby participate in the regulation of cellular responses to the IFN-γ secreted in response to *M. tuberculosis*. Interestingly, this gene was found to be up-regulated in Th1 cells compared with Th2 cells in a different study (5).

We had expected to find increased expression of cytolytic effector molecules, such as perforin or granzymes, in the CD8+ T cells. There are several possible explanations as to why this was not the case. The statistical analysis of the data from nearly 60,000 cDNA clones, but only six separate donors, is complex, and it may be that our cutoff limits were too stringent to identify these genes. Alternatively, the mRNA levels may not be that different, but protein levels may be controlled via another mechanism. Another explanation is that *M. tuberculosis* also induces the activation of cytolytic CD4+ T cells (21, 47). Indeed, perforin mRNA was assayed using the human sequence-verified arrays, and high levels were detected in both CD4+ and CD8+ samples (data not shown). Interestingly, Ksp37, which has previously been described as a protein selectively produced by cytotoxic lymphocytes (48), was actually more highly expressed by CD4+ than CD8+ T cells in this system. As expected, the CD8sg gene was more highly expressed in CD8+ than CD4+ T cells; however, this failed to reach significance due to variability in the replicate hybridization intensities for one donor. Granulysin was more highly expressed by CD8+ than CD4+ T cells, with a geometric fold difference of 1.25. The IFN-γ gene was highly expressed in all T cell samples analyzed, and there was a close correlation between the hybridization intensity in CD4+ T cells and the donor’s response to PPD in PBMC cultures.

A better understanding of the protective CD4+ and CD8+ T cell immune response to intracellular infection, derived from this type of array-based analysis, will facilitate the development of correlates of protection to be used as surrogate markers in field trials of new vaccines. New prophylactic and immunotherapeutic TB vaccines are urgently needed, due to the limited efficacy of *M. bovis* BCG in some areas and because the length of antibiotic therapy makes effective TB treatment difficult for many low income countries. New vaccine candidates are being tested in animal models, including live attenuated *M. tuberculosis*, recombinant vaccines, DNA vaccines, and protein subunit vaccines (49), but correlates of protection must be identified that demonstrate the induction of protective immunity against *M. tuberculosis* and therefore the efficacy of the candidate vaccine before large-scale clinical trials can be initiated (50). The production of IFN-γ in whole blood cultures in response to mycobacterial Ags has been proposed as a possible correlate, and has been used to demonstrate the difference between immune responses in Malawi and the United Kingdom induced following BCG vaccination (36). However, in vaccine studies in cattle and mice, protection is not related to the magnitude of the IFN-γ response, suggesting that additional mechanisms must be present or absent to provide protection against disease (50, 51). Any correlates of protection identified in the present culture system with *M. tuberculosis* are also likely to be applicable to vaccine development for viral infections, where both CD4+ and CD8+ T cell immune responses are required.

It is likely that a combination of assays, measuring effector CD4+ and CD8+ T cell function, will be required in conjunction with in vitro analysis of the killing of mycobacteria (52) to assess candidate vaccine efficacy. Ideally, ELISA-based assays based on secreted proteins would be developed. At the present time, 64 of the unknown genes identified in this study have full-length protein reference sequences. Of these, 5 are predicted to be secreted proteins, according to SignalP. In time, the remaining 38 unknown genes and the 30 ESTs will become fully annotated, allowing the prediction of more secreted proteins. Currently, CTL activity is usually measured using the 51Cr release assay, which is labor intensive, fairly insensitive, and involves radioactivity, and is thus unsuitable for large-scale field trials. The identification of 48 previously uncharacterized genes that are more highly expressed in activated effector CD8+ T cells than CD4+ T cells provides new leads in the discovery of biological pathways unique to CD8+ T cells and involved in their cytotoxic response. Eighty-six uncharacterized genes were found to be more highly expressed in CD4+ T cells, and some of these may also play a crucial role in immunity to intracellular infection.

This study has identified many genes that are up-regulated in CD4+ T cells following stimulation with *M. tuberculosis*, by comparison with CD8+ T cells, and vice versa. Some of these genes may be important during the immune response to TB, as they were found to be differentially expressed in the blood of active TB patients compared with Mantoux-positive latently infected people in a parallel study.6 The G protein-coupled receptor EDG6 was more highly expressed in CD4+ T cells than CD8+ T cells following in

vitro stimulation with *M. tuberculosis*, while the Src homology 3 domain kinase-binding protein 1 was more highly expressed in CD8+ T cells; both of these proteins are involved in lymphocyte signal transduction. Both of these genes are much more highly expressed in the blood of active TB patients than in latently infected controls, and this probably reflects the greater mycobacterial burden in these people, leading to a greater circulation of activated T cells in the periphery. Conversely, some genes were more highly expressed by latently infected people than in active TB patients, and these genes may therefore be contributing to the immune protection. These include the oyster-sterol-binding protein 10, which was more highly expressed by CD4+ T cells, and the signal transduction molecule DAP10 and the lysosomal siadilase Neu1, which were more highly expressed by CD8+ T cells.

We now propose to develop some of these genes into surrogate markers of CD4+ and CD8+ T cell activation. Likely candidates include those that were most highly differentially expressed, such as the NKG2D receptor and its signal transduction partner DAP10. The extent of expression of these proteins in vitro in response to mycobacterial stimulation might reflect the magnitude of the memory CD8+ T cell response induced following vaccination. Also, secreted proteins such as IL-16 and KSP37, and the putatively secreted hypothetical proteins MGCI0993 and AL449243 will be analyzed to determine whether their production following in vitro stimulation correlates with the induction of protective immunity by vaccination. These surrogate markers might then be used in field trials of new immunotherapeutic and prophylactic vaccines.

Acknowledgments

We are grateful to Debbie Smith, Joanne Cox, and Paul Murdoch for scientific advice, and to Sara Atkinson for providing the *M. tuberculosis*. We thank the Transcriptome Analysis and Quantitative Expression Department and the Core Sequencing group at GlaxoSmithKline Medicines Research Centre (Stevenage, U.K.).

References


