Cell Sorting-Assisted Microarray Profiling of Host Cell Response to Cryptosporidium parvum Infection†‡

Yi-Lin Yang,† Gregory A. Buck,‡ and Giovanni Widmer†§

Division of Infectious Diseases, Tufts Cummings School of Veterinary Medicine, North Grafton, Massachusetts 01536,† and Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, Virginia 23284§.

Received 3 September 2009/Returned for modification 5 October 2009/Accepted 16 December 2009

To study the transcriptional response of mammalian cells to infection with the intracellular apicomplexan parasite Cryptosporidium parvum, infected and uninfected cells were recovered from C. parvum-infected cell monolayers. This approach, which contrasts with a more conventional experimental design that compares infected to uninfected cell monolayers, enabled the identification of functional categories of genes that are differentially transcribed as a direct consequence of the presence of intracellular parasites. Among several categories of upregulated genes, glycoprotein metabolism was significantly overrepresented. To investigate whether these transcriptional changes affected the composition of the surface of infected cells, cells were probed with fluorescently labeled lectins. Among a panel of seven lectins, soybean agglutinin, which recognizes N-acetyl-D-galactosamine, generated the largest difference in fluorescence between infected and uninfected cells. The origin of the fluorescent signal emitted by infected cells was further investigated and attributed to the overexpression of glycoprotein on the surface of infected cells, as well as the presence of glycoprotein located in the proximity of intracellular parasites.

Cryptosporidium parvum, an apicomplexan protozoan species, infects the intestine and sometimes the respiratory tract and bile duct of various mammalian hosts. In immunocompromised humans the infection causes self-limited diarrhea, but in immunocompromised individuals, primarily those with AIDS, the infection can become chronic and potentially cause life-threatening syndromes (5). Infection with Cryptosporidium parasites occurs when food or water contaminated with oocysts is ingested, or possibly through the inhalation of oocysts. After entering the gastrointestinal tract of the host, four sporozoites released from the oocyst seek to invade the intestinal epithelial cells, where they multiply as meronts.

Global gene expression profiles in cell monolayers infected with C. parvum have been studied using microarrays (3, 8, 20). These studies found that genes belonging to the cell proliferation, apoptosis, signal transduction, and transcription categories are overrepresented among significantly regulated genes (8, 20, 39). The overexpression of the osteoprotegerin gene in host cells after infection also was reported (3).

Published microarray studies of C. parvum-infected cells are based on RNA extracted from cells in culture. The development of this parasite in culture is restricted in time, and the completion of the life cycle is only rarely observed (14, 32). High oocyst doses and long incubation times do not increase the proportion of infected cells (36) and may instead lead to the extensive perturbation of the monolayer caused primarily by apoptosis and mitosis (19, 21, 34, 38). Discerning which transcriptional changes occur directly in response to the infection, and which result from the perturbation of the monolayer (13, 36), is difficult. To eliminate the effect of partial infection and monolayer perturbation, we physically separated infected and uninfected cells to generate subpopulations that differed only with respect to the presence of intracellular parasites. We report on transcriptional and phenotypic changes identified by this approach, in particular the transcriptional upregulation and overexpression of cell surface glycoprotein in C. parvum-infected cells.

MATERIALS AND METHODS

Cell culture. Human ileocecal epithelial cells (HCT-8) (American Type Culture Collection no. CCL-244) were cultured in RPMI 1640 (Gibco/Invitrogen, Paisley, United Kingdom) supplemented with 10% heat-inactivated horse serum (Gibco), 1% penicillin-streptomycin, 2 mML-glutamine, and 1 mM sodium pyruvate (Sigma, St. Louis, MO). Cells were maintained in a 5% CO2 atmosphere at 37°C and 85% humidity. Monolayers were seeded in 96-well plates, 25- or 75-cm2 flasks at 2 × 104 cells/96-well plate, 1 × 105 cells/75-cm2 flask, or 3 × 106 cells/75-cm2, and grown to approximately 80% confluence.

Parasites. C. parvum oocysts were obtained from experimentally infected calves or immunosuppressed rodents. Oocysts were purified from stool as described previously (35). Oocysts for the inoculation of cell monolayers were surface sterilized with 10% commercial bleach for 7 min on ice and washed twice with phosphate-buffered saline (PBS). Oocytes for mock infections were heat inactivated at 80°C for 20 min.

Cell monolayer infection, immunofluorescent labeling, and cell sorting. Subconfluent HCT-8 monolayers seeded in T75 flasks were infected with C. parvum oocysts at a 1:4 oocyst/host cell ratio and incubated for 24 h. This time point was chosen because at later time points monolayer perturbation increases. Taurocholic acid at a concentration of 0.05% was added with the oocysts to promote infection (11). After incubation, cell monolayers were recovered by treatment with Accutase (Innovative Cell Technologies, Inc.) and immunofluorescently labeled as described previously (31), except that polyclonal rabbit antibody specific for C. parvum sporozoites and oocysts was used as the primary antibody and the host cell membrane was not permeabilized. Labeled cells were resuspended at a concentration of 106 cells/ml in PBS and analyzed by flow cytometry (FACSCalibur; BD Biosciences, San Jose, CA). Fluorescence emitted by the secondary Alexa Fluor 488-conjugated anti-rabbit IgG antibody preadsorbed against human IgG was acquired in the FL1 channel and analyzed using CellQuest (BD Biosciences, San Jose, CA) software. Sorting was performed in
recovery sort mode, with the flow rate set to low. Approximately 5 × 10^5 infected cells, and the same number of uninfected cells, were sorted from the same monolayer into bovine serum albumin (BSA)-coated 50-ml tubes. Sorted uninfected and infected cells from the same monolayer were examined by fluorescence microscopy to confirm the efficiency of sorting (see the supplemental material). RNA was extracted from about 5 × 10^6 infected and uninfected cells using the RNAqueous kit (Ambion, Inc.).

RNA amplification and biotin labeling. RNA samples isolated from sorted cells were amplified using the WT-Ovation Pico RNA amplification system (NuGEN Technologies, Inc.). To ensure sufficient cDNA yield, two replicate amplifications were performed simultaneously from each RNA sample using 50 ng of total RNA samples as a substrate. Amplified cDNA was fragmented, labeled with Biotin using FL-Ovation cDNA Biotin Module V2 (NuGEN Technologies, Inc.), and hybridized to an Affymetrix Human Genome U133 Plus 2.0 GeneChip, which can analyze more than 47,000 transcripts. In all experiments, the manufacturers' protocols were followed.

Microarrays and data analysis. Transcriptional profiles in sorted cells identified using the Affymetrix Human Genome U133 Plus 2.0 array first were normalized using Robust Multichip Average (RMA) algorithms (17). A list of differentially expressed genes, their fold changes, and P values was compiled using the linear models for microarray data (LIMMA) package in BioConductor, which provides functions for fitting a linear model to the expression data for each gene and performing moderated t tests, to improve variance estimation for small sample sizes. Genes with absolute log2-fold changes of greater than 1.3 and a P value smaller than 0.05 were considered significantly regulated in response to infection.

The list of differentially regulated genes identified in U133 Plus 2.0 was subdivided into functional categories using the bioinformatics analysis resource DAVID (Database for Annotation, Visualization, and Integrated Discovery) of the National Institute of Allergy and Infectious Diseases (9, 15). Databases used in DAVID where the category terms are defined include Gene Ontology for GO terms, the Uniprot sequence feature, InterProt for protein functional domains, and the Swiss-Prot Protein Knowledgebase for the description of the function of a protein. The score of the Expression Analysis Systemic Explorer (EASE) (9), a conservative adjustment to the Fisher exact probability test, was used to measure the probability of observing a number of differentially regulated genes for a specific category given a category's representation on the array. The group of upregulated genes annotated with the keyword “glycoprotein” was analyzed by calculating the proportion of these genes in other categories using the kappa (κ) statistic, a chance-corrected measure of agreement between two sets of categorized data (7, 12). κ = 0 indicates no chance agreement, whereas κ = 1 indicates a perfect agreement, i.e., all genes annotated with keyword “glycoprotein” also belong to a second category.

Real-time RT-PCR. Twelve genes were selected for confirmatory reverse transcription-PCR (RT-PCR). These genes were among a collection of 778 differentially regulated genes identified in microarray analyses of RNA extracted from infected and uninfected HCT-8 cells and were selected to represent different biological categories. Primers were designed to amplify fragments of 200 to 350 bp from the selected transcripts (see Table S1 in the supplemental material) as biological categories. Primers were designed to amplify fragments of 200 to 350 bp from the selected transcripts (see Table S1 in the supplemental material) as biological categories.

In addition to the immunofluorescence analysis described below, and as an independent verification of the efficiency of cell sorting, RT-PCR was applied. The C. parvum 18S rRNA and β-tubulin loci were amplified in parallel with a fragment of the human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene. Portions of the same RNA samples extracted from sorted infected and uninfected HCT-8 cells and analyzed on microarrays were used. This experiment was conducted in duplicate using two pairs of RNA samples, each originating from infected and uninfected cells sorted from a monolayer.

Fluorescent lectin labeling and flow cytometry. HCT-8 cells seeded in T25 flasks, or 96-well plates were infected with C. parvum oocysts at a 1:1, 1:2, 1:4, or 1:50 oocyst/host cell ratio and incubated for 6, 19, 24, 30, 43, 48, 72, and 88 h as described above. Monolayers were released as described above, washed three times with PBS, and incubated with fluorescein-labeled lectins (Vector Laboratories, Burlingame, CA) at 37°C for 30 min (10) before or after the permeabilization of cell membranes with 100% methanol for 15 min. Lectin-labeled cells then were immunolabeled with primary rabbit antibody specific to C. parvum as described previously. Fluorescence originating from the fluorescein-labeled lectin and Alexa 667 fluorescence emitted by the secondary anti-rabbit IgG antibody conjugate (Molecular Probes, Invitrogen) were acquired in the FL1 and FL4 channels, respectively. FL1, FL4, and scatter signal amplifiers were set in log mode. Kolmogorov-Smirnov (K-S) statistics were computed with CellQuest software and used to compare the fluorescence intensity distributions of infected and uninfected control cells.

The seven fluorescein-labeled lectins used in this study include concanavalin A (ConA), wheat germ agglutinin (WGA), Ricinus communis agglutinin-I (RCA-I), peanut agglutinin (PNA), soybean agglutinin (SBA), Dolichos biflorus agglutinin (DBA), and Ulex europaeus agglutinin-I (UEA-I). To confirm the specificity of lectin binding, lectins were preincubated with the appropriate sugars for 30 min and then incubated with HCT-8 cells in the presence of sugars. The following sugars and concentrations were used: 0.2 M α-methyl-D-mannoside and 0.2 M α-methyl glucoside for ConA, 0.5 M N-acetyl-d-galactosamine (N-GalNAc) for WGA, and 0.5 M N-acetyl-d-galactosamine (N-GalNAc) for SBA.

Western blot analysis. C. parvum oocysts (6 × 10^7) were excised as described above, and the sporozoites were separated from oocysts and oocyst walls using a 2-μm Nucleopore Tack-Etch membrane (Whatman, Inc.). A sporozoite lysate equivalent to 2 × 10^6 sporozoites and the lysate of infected and uninfected HCT-8 cells (2 × 10^6/lanes) were resolved by SDS-PAGE under reducing conditions on a 12% polyacrylamide gel. Protein was electrophoretically transferred to a 0.2-μm-pore-size polyvinylidene fluoride membrane. The membrane was incubated with 0.6% hydrogen peroxide in methanol for 20 min at room temperature. The membrane then was blocked with 2% polyvinylpyrrolidone 10 in Tris-buffered saline supplemented with 0.1% Tween 20 for 1 h and incubated with 5 μg/ml horseradish peroxidase (HRP)-conjugated SBA (EY Lab, Inc., San Mateo, CA). The washed membrane was incubated in the dark with ECL Western blotting detection reagent (GE Healthcare) and luminal-based enhanced chemiluminescence substrate. Luminescence was visualized on autoradiography film.

RESULTS

Analysis of global gene expression. RNA was extracted from C. parvum-infected and uninfected cells originating from the same monolayer and analyzed with microarrays. In two independent experiments with Affymetrix HG-U133 Plus 2.0 arrays, 778 genes were found to be differentially regulated 24 h postinfection (P < 0.05, moderated t statistics) (see Table S2 in the supplemental material). A total of 417 genes were upregulated and 361 genes were downregulated. The up- and down-regulated genes were analyzed using the one-tailed Fisher exact probability test for gene enrichment analysis. The biological categories found to be overrepresented among upregulated genes (P ≤ 1.69 × 10^{-8}, EASE score) include the N-linked glycosylation site, the keyword “glycoprotein” (18), the protein domain UDP-glucuronosyl/UDP-glucosyltransferase, and the GO molecular function term “glucuronosyltransferase activity” (Table 1). The biological categories found to be overrepresented among downregulated genes (P ≤ 5.08 × 10^{-4}, EASE score) include cytoskeleton, intracellular, cytoplasm, and intracellular part (Table 2).

For the 107 upregulated genes with the functional annotation “glycoprotein,” the chance-corrected agreement between “glycoprotein” and eight other categories with a minimum threshold of 60 genes was determined using κ scores (Table 3). The highest score (κ = 0.92) was found for the term “N-linked glycosylation site,” indicating an almost perfect overlap between the set of genes annotated with these two terms. The κ scores for “glycoprotein” and other annotations were lower, ranging from 0.70 for “transmembrane” to 0.61 for “membrane part” (Table 3). A κ of 0.7 and above indicates strong agreement, whereas κ values larger than 0.3 are indicative of biological meaning (16). Examples of terms found at the lower end of the κ scale include “nucleus,” “cytoplasm,” and “transport.” Taken together, the data not only suggest that glycopro-
Transcriptional changes in sorted and unsorted cells (39) were compared (see Table S3 in the supplemental material). Out of 22,277 transcripts included on the U133A 2.0 GeneChip, 181 transcripts and 77 transcripts were identified as significantly upregulated in sorted and unsorted cells, respectively ($P < 0.01$, Student’s $t$ test). Out of the significantly upregulated genes, 171 and 51 transcripts identified in sorted and unsorted cells that are annotated with functional terms were subdivided into functional categories using DAVID (see Tables S4 and S5 in the supplemental material). Based on the criterion of $P < 0.01$ by EASE score, the functional categories with the highest abundance of upregulated genes were identified. In sorted cells genes annotated with “glycoprotein” and “N-linked glycosylation site” were overrepresented among the upregulated genes ($P \leq 6.7 \times 10^{-5}$, EASE score). Upregulated genes annotated with the terms “glycoprotein” and “N-linked glycosylation site” were identified in unsorted cells, but these were not significantly overrepresented ($P > 0.05$, EASE score) among the upregulated genes.

**Confirmatory reverse transcription-PCR analyses.** From the most up- and downregulated genes ($P < 0.05$) identified in microarray analyses of sorted cells, and from genes that were identified in a previous microarray study (39), 12 genes belonging to different functional categories were randomly selected for RT-PCR validation (Table 4). Portions of the same RNA samples analyzed on microarrays were used as a template for RT-PCR as described previously (39). The concentration of each mRNA transcript in each sample was estimated using a standard curve and normalized against GAPDH (Fig. 1) or succinate dehydrogenase complex subunit A (SDHA) mRNA (not shown). Although RT-PCR results showed considerable variability among replicates, RT-PCR results from seven upregulated genes and four downregulated genes were in agreement with the microarray data. The upregulation of MAX5 predicted by microarrays was, however, not apparent from the RT-PCR results. Normalization against GAPDH or SDHA gave the same results.

**Lectin screening.** Since microarray results indicated that glycoprotein-related genes were overrepresented among upregulated genes in infected cells, seven fluorescent lectins recognizing different oligosaccharides were used to examine which types of glycoproteins were overexpressed. At 48 h postinfection, unsorted cells from *C. parvum*-infected and control monolayers were reacted with fluorescein-conjugated lectin as described in Materials and Methods. Lectins that recognize glucose (ConA), mannose (ConA), galactose (PNA, RCA-1), fucose (UEA-1), N-GalNAc (DBA, SBA, RCA-1), and N-GlcNAc (WGA) were used. Out of seven lectins, SBA generated the largest difference when the fluorescence of infected and uninfected cell monolayers was analyzed using K-S statistics (Table 5). Consistently with the microarray data, this result suggests the overexpression of N-GalNAc glycoproteins on *C. parvum*-infected cells. The specificity of lectin binding was confirmed by preincubating lectins with the matching sugar. As

---

**TABLE 1. Gene ontology classification of upregulated genes**

<table>
<thead>
<tr>
<th>Gene ontology term</th>
<th>Count</th>
<th>$%$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-linked glycosylation site</td>
<td>95</td>
<td>22.78</td>
<td>$4.33 \times 10^{-11}$</td>
</tr>
<tr>
<td>Glycoprotein</td>
<td>107</td>
<td>25.66</td>
<td>$4.56 \times 10^{-11}$</td>
</tr>
<tr>
<td>UDP-glucuronosyl/UDP-glucosyltransferase</td>
<td>9</td>
<td>2.16</td>
<td>$3.12 \times 10^{-10}$</td>
</tr>
<tr>
<td>Glucuronosyltransferase</td>
<td>9</td>
<td>2.16</td>
<td>$4.66 \times 10^{-10}$</td>
</tr>
<tr>
<td>Pentose and glucuronate interconversions</td>
<td>10</td>
<td>2.40</td>
<td>$1.13 \times 10^{-9}$</td>
</tr>
<tr>
<td>Glucuronosyltransferase activity</td>
<td>9</td>
<td>2.16</td>
<td>$1.60 \times 10^{-8}$</td>
</tr>
<tr>
<td>Metabolism of xenobiotics by cytchrome P450</td>
<td>13</td>
<td>3.12</td>
<td>$2.61 \times 10^{-8}$</td>
</tr>
<tr>
<td>Oxidative stress induced gene expression via Nrf2</td>
<td>9</td>
<td>2.16</td>
<td>$5.57 \times 10^{-8}$</td>
</tr>
<tr>
<td>Porphyrin and chlorophyll metabolism</td>
<td>9</td>
<td>2.16</td>
<td>$2.23 \times 10^{-6}$</td>
</tr>
<tr>
<td>Microsome</td>
<td>12</td>
<td>2.88</td>
<td>$2.40 \times 10^{-6}$</td>
</tr>
</tbody>
</table>

* a Number of genes in this category.
* b Percentage of genes with annotation among all upregulated genes.
* c EASE $P$ value.

**TABLE 2. Gene ontology classification of downregulated genes**

<table>
<thead>
<tr>
<th>Gene ontology term</th>
<th>Count</th>
<th>$%$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytoskeleton</td>
<td>31</td>
<td>8.76</td>
<td>$1.04 \times 10^{-4}$</td>
</tr>
<tr>
<td>Intracellular</td>
<td>159</td>
<td>44.92</td>
<td>$1.39 \times 10^{-4}$</td>
</tr>
<tr>
<td>Alternative splicing</td>
<td>105</td>
<td>29.66</td>
<td>$1.94 \times 10^{-4}$</td>
</tr>
<tr>
<td>Cytoplasm</td>
<td>108</td>
<td>30.51</td>
<td>$2.66 \times 10^{-4}$</td>
</tr>
<tr>
<td>Intracellular part</td>
<td>151</td>
<td>42.66</td>
<td>$5.08 \times 10^{-4}$</td>
</tr>
<tr>
<td>Chromosomal rearrangement</td>
<td>12</td>
<td>3.39</td>
<td>$7.19 \times 10^{-4}$</td>
</tr>
<tr>
<td>Microtubule</td>
<td>11</td>
<td>3.11</td>
<td>$1.11 \times 10^{-3}$</td>
</tr>
<tr>
<td>Proto-oncogene</td>
<td>11</td>
<td>3.11</td>
<td>$1.31 \times 10^{-3}$</td>
</tr>
<tr>
<td>Cytoskeleton</td>
<td>14</td>
<td>3.95</td>
<td>$1.35 \times 10^{-3}$</td>
</tr>
<tr>
<td>Splice variant</td>
<td>73</td>
<td>20.62</td>
<td>$2.40 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

* a Number of genes in this category.
* b Percentage of genes with annotation among all upregulated genes.
* c EASE $P$ value.

**TABLE 3. Function of upregulated genes annotated with glycoprotein keyword**

<table>
<thead>
<tr>
<th>Gene ontology term</th>
<th>Kappa score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycoprotein</td>
<td>1.00</td>
</tr>
<tr>
<td>N-linked glycosylation site</td>
<td>0.92</td>
</tr>
<tr>
<td>Transmembrane</td>
<td>0.70</td>
</tr>
<tr>
<td>Transmembrane region</td>
<td>0.67</td>
</tr>
<tr>
<td>Signal</td>
<td>0.67</td>
</tr>
<tr>
<td>Intrinsic to membrane</td>
<td>0.64</td>
</tr>
<tr>
<td>Membrane</td>
<td>0.63</td>
</tr>
<tr>
<td>Signal peptide</td>
<td>0.63</td>
</tr>
<tr>
<td>Membrane part</td>
<td>0.61</td>
</tr>
<tr>
<td>Transport</td>
<td>0.00</td>
</tr>
<tr>
<td>Cytoplasm</td>
<td>0.00</td>
</tr>
<tr>
<td>Nucleus</td>
<td>0.00</td>
</tr>
</tbody>
</table>

* a Examples of terms showing no association with the keyword “glycoprotein” in upregulated genes are indicated in italic.
shown in Table 5, this treatment significantly reduced the K-S statistic. Because this comparison was performed with unpermeabilized cells, only glycoprotein expression on the cell surface was measured.

**Flow cytometry and microscopy of SBA fluorescence.** Having identified SBA as the lectin giving the largest K-S statistic difference between infected and uninfected HCT-8 monolayers, we investigated the course of glycoprotein expression during parasite development in cell monolayers and the cellular location of the SBA-reactive glycoprotein. HCT-8 cells infected with *C. parvum* oocysts at a 1:2 oocyst/host cell ratio and incubated for 6, 19, 24, 30, 43, and 48 h were analyzed using flow cytometry (Fig. 2). The K-S values expressing the difference in FL-1 fluorescence from fluorescein-labeled SBA between infected and uninfected monolayers for the four treatments are shown in Fig. 3. Fluorescence histograms were acquired for cells labeled with SBA prior to and after permeabilization. These treatments enabled us to measure fluorescence originating from cell surface glycoprotein only (unpermeabilized cells) and fluorescence emitted by any reactive glycoprotein regardless of its cellular location (permeabilized cells), including the intracellular protein of cellular or parasite origin. Significantly, starting at 19 h postinfection, K-S values for the SBA-labeled samples were positive regardless of cell permeabilization, as shown by a rightward shift of the FL1 histograms from infected monolayers (Fig. 2). These data indicate that N-GalNAc oligosaccharides bound by SBA are located outside and inside of the host cell membrane and that glycoprotein expression in both locations is affected by the infection. The lectin fluorescence of cells that were permeabilized prior to incubation with lectin was stronger than that observed with unpermeabilized cells, indicating the binding of lectin to intracellular and extracellular glycoproteins, possibly due to the presence of the glycoprotein of parasite origin. As shown in Fig. 3, the

![Image](vol78n10-1043f1.jpg)

**FIG. 1.** Validation of gene expression by quantitative RT-PCR. Twelve mRNA transcripts were RT-PCR amplified from two infected and two mock-infected RNA samples used in microarray hybridizations. For each sample, expression values were normalized against GAPDH mRNA levels, and relative expression values were obtained by comparing RNA extracted from corresponding infected and uninfected monolayers. Mean relative expression (n = 2 ± standard deviation (SD)) from each transcript normalized against GAPDH transcription and quantified using RT-PCR (gray bar) was compared to the mean relative expression (n = 2 ± SD) of the transcript detected using microarray analyses (black bar). For gene symbols see Table 4.

### Table 4. Twelve genes selected for RT-PCR analysis

<table>
<thead>
<tr>
<th>ID</th>
<th>Symbol</th>
<th>Description</th>
<th>FC</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>209590_at</td>
<td>MXRA5</td>
<td>Matrix-remodelling associated 5</td>
<td>2.53</td>
<td>0.004</td>
</tr>
<tr>
<td>208506_at</td>
<td>HIST1H3F</td>
<td>Histone cluster 1, H3f</td>
<td>2.51</td>
<td>0.032</td>
</tr>
<tr>
<td>206561_s_at</td>
<td>AKR1B10</td>
<td>Aldo-keto reductase family 1, member B10 (aldose reductase)</td>
<td>2.33</td>
<td>0.018</td>
</tr>
<tr>
<td>234673_s_at</td>
<td>HHL2A</td>
<td>HERV-HLTR-associating 2</td>
<td>2.21</td>
<td>0.009</td>
</tr>
<tr>
<td>204532_s_at</td>
<td>UGT1A9</td>
<td>UDP glucuronosyltransferase 1 family, polypeptide A9</td>
<td>1.50</td>
<td>0.015</td>
</tr>
<tr>
<td>218368_s_at</td>
<td>TNFRSF12A</td>
<td>Tumor necrosis factor receptor superfamily, member 12A</td>
<td>1.27</td>
<td>0.222</td>
</tr>
<tr>
<td>213420_s_at</td>
<td>KRT4</td>
<td>Keratin 4</td>
<td>1.23</td>
<td>0.277</td>
</tr>
<tr>
<td>213506_at</td>
<td>F2RL1</td>
<td>Coagulation factor II (thrombin) receptor-like 1</td>
<td>1.15</td>
<td>0.309</td>
</tr>
<tr>
<td>209189_at</td>
<td>EPS15</td>
<td>Epidermal growth factor receptor pathway substrate 15</td>
<td>1.15</td>
<td>0.009</td>
</tr>
<tr>
<td>213240_s_at</td>
<td>HHLA2</td>
<td>HERV-HLTR-associating 2</td>
<td>2.21</td>
<td>0.009</td>
</tr>
<tr>
<td>234673_at</td>
<td>MXRA5</td>
<td>Matrix-remodelling associated 5</td>
<td>2.53</td>
<td>0.004</td>
</tr>
<tr>
<td>206561_s_at</td>
<td>AKR1B10</td>
<td>Aldo-keto reductase family 1, member B10 (aldose reductase)</td>
<td>2.33</td>
<td>0.018</td>
</tr>
<tr>
<td>234673_s_at</td>
<td>HHL2A</td>
<td>HERV-HLTR-associating 2</td>
<td>2.21</td>
<td>0.009</td>
</tr>
</tbody>
</table>

* ID: Affymetrix probe identity.
* Symbol: As assigned by the HUGO Gene Nomenclature Committee.
* FC: Log2 transcriptional fold change between sorted infected and uninfected cells.
* p: Calculated using the moderated t test of Limma.

### Table 5. Comparison of lectin fluorescence emitted by *C. parvum*-infected and uninfected HCT-8 monolayers

<table>
<thead>
<tr>
<th>Designation</th>
<th>Common name</th>
<th>Kolmogorov-Smirnov statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConA</td>
<td>Concanavalin A</td>
<td>0.13 ± 0.08</td>
</tr>
<tr>
<td>DBA</td>
<td>Dolichos biflorus</td>
<td>0.05 ± 0.02</td>
</tr>
<tr>
<td>PNA</td>
<td>Peanut</td>
<td>0.12 ± 0.02</td>
</tr>
<tr>
<td>RCA-I</td>
<td>Ricinus communis</td>
<td>0.14 ± 0.10</td>
</tr>
<tr>
<td>SBA</td>
<td>Soybean</td>
<td>0.33 ± 0.16</td>
</tr>
<tr>
<td>UEA-I</td>
<td>Ulex europaeus</td>
<td>0.11 ± 0.04</td>
</tr>
<tr>
<td>WGA</td>
<td>Wheat germ</td>
<td>0.14 ± 0.05</td>
</tr>
<tr>
<td>ConA + glucose + mannose</td>
<td>0.04 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>WGA + N-GlcNAc</td>
<td>0.07 ± 0.05</td>
<td></td>
</tr>
<tr>
<td>SBA + N-GalNAc</td>
<td>0.09 ± 0.09</td>
<td></td>
</tr>
<tr>
<td>SBA + N-GlcNAc</td>
<td>0.18</td>
<td></td>
</tr>
</tbody>
</table>

* A measure of the differences between fluorescence histograms. Means ± standard deviations of three to five experiments are shown.
* Sugars were preincubated with lectin prior to incubating lectin with cells.
permeabilization of cells after they were reacted with SBA gave K-S distances similar to those of unpermeabilized cells, suggesting that methanol permeabilization by itself does not affect SBA binding. Consistently with the observed upregulation of host glycoprotein expression in response to the infection and with the putative presence of parasite glycoprotein, K-S values for SBA fluorescence paralleled parasite immunofluorescence over time (Fig. 3). As typically observed with *C. parvum*-infected monolayers (1, 36), parasite fluorescence peaked after 2 days of culture, a trend that also was observed with the SBA K-S values. Confirmatory experiments with HCT-8 monolayers grown in 96-well plates infected with *C. parvum* oocysts at a 1:50 ratio and incubated for 48 and 72 h were performed. Foci of intracellular parasites were visualized with immunofluorescence, and the infected monolayer was double labeled with SBA. As expected from the flow cytometry analyses, fluorescence from SBA and antibody was found to colocalize on the monolayer (not shown). To ensure that lectin and antibody bound to the intended target, controls also were included in all experiments: infected monolayers labeled with SBA preincubated with N-GalNAc or labeled with *C. parvum*-specific antibody in the absence of secondary antibody showed, as expected, no fluorescence. Infected cells labeled separately with either fluorescent SBA or *C. parvum*-specific antibody also were examined to confirm that green (fluorescein) and red (Alexa 568) fluorescent signals were specific and that no bleed-over from red to green and vice versa took place (not shown).

We used microscopy to investigate the cellular location of SBA fluorescence. Consistently with the results presented in Fig. 2 and 3, at 48 h postinfection binding of SBA was observed primarily on *C. parvum*-infected cells (Fig. 4). Based on the observation that particularly in the later phase of the infection some SBA fluorescence is intracellular (Fig. 2 and 3), we further investigated the cellular location of SBA relative to that of the parasite (Fig. 5). Infected cells double labeled with SBA and *C. parvum*-specific antibody after permeabilization were used to visualize the location of glycoprotein recognized by SBA relative to that of the intracellular parasite. These microscopic analyses showed a striking colocalization of parasite antigen and SBA fluorescence. Glycoprotein expression in infected and uninfected cells was further evaluated on a Western blot (Fig. 6). Cell lysates from uninfected host cells, *C. parvum*.
infected cells, and *C. parvum* sporozoites were blotted and reacted with HRP-conjugated SBA. The blot indicates that SBA binds to the glycoprotein of host and parasite origins, which is consistent with the data presented in Fig. 2 to 5, and with the view that the permeabilization of infected cells increased SBA fluorescence. Taken together, these results indicate that, in addition to recognizing glycoproteins on the cell surface, SBA also binds intracellular glycoproteins located in close proximity to the parasite. It is conceivable that intracellular fluorescence originates from glycoprotein of parasite origin.

**DISCUSSION**

Because the majority of host cells in a *C. parvum*-infected monolayer remain uninfected, it is difficult in a direct comparison of infected and uninfected monolayers to identify the transcriptional responses of the host cells to the infection, and to differentiate such transcriptional changes from secondary changes resulting from the disturbance of the monolayer. The approach used here, i.e., the analysis of infected and uninfected cells originating from one monolayer, was intended to eliminate from the microarray data changes in the transcriptome that do not occur directly in response to the intracellular parasites. Microarray data presented here reveal that, among others, genes related to cell proliferation and apoptosis, which were identified using unsorted cells (8, 20, 39), were not overrepresented among the upregulated genes identified in sorted cells (see Table S5 in the supplemental material). Genes annotated with glycoprotein and the N-linked glycosylation site also were discovered in unsorted infected cells, but these genes were not significantly overrepresented.

Although there are clear advantages in analyzing sorted cells, this approach involves additional manipulations that may affect the RNA. Prior to being sorted, suspensions of live cells were immunolabeled, which raised the possibility that mRNA integrity could be compromised during incubation. Because we performed these manipulations prior to sorting, the possibility...
of differentially affecting the transcriptome of infected and uninfected cells was not an immediate concern, although the possibility that the infection influences RNA integrity cannot be ruled out. Moreover, after being sorted, all suspensions were treated in an identical manner such that differences in the transcriptome could unambiguously be attributed to the infection. The observation that intact rRNA was recovered from sorted cells, and that control cells that had passed through the sorter remained viable (not shown), demonstrated that these manipulations were compatible with the goals of the experiments. Consistently with this view, a relatively high Pearson’s correlation coefficient (0.72) between replicate microarray experiments was obtained. It is conceivable that RNA integrity is differentially affected in infected and uninfected cells. This possibility cannot be excluded on the basis of the experimental design and the controls included here.

Because the number of cells recovered by flow cytometry was relatively small (∼5 × 10⁵), we used linear enzymatic amplification (6, 25, 30, 33) to generate sufficient RNA for microarray hybridization. Before and after RNA amplification, portions of the same RNA samples as those analyzed with microarrays were used as the template for RT-PCR. Relative transcriptional changes of FOS, KRT4, F2RL1, and TNFRSF12A genes before and after RNA amplification were in good agreement (not shown), indicating that RNA amplification is compatible with the experiments.

Consistently with the intracellular location of C. parvum, the punctate immunofluorescence pattern observed on live sorted cells (see Fig. S1 in the supplemental material) suggests that the antibody detected parasite antigen that is shed by sporozoites prior to host cell invasion and is deposited on the cell membrane (4, 23, 29, 37). When sorting on the basis of this fluorescent signal, the assumption is made that the presence of sporozoite antigen is indicative of a particular cell being infected. Although more direct evidence of infection would be preferable, to date no direct antigenic marker of C. parvum infection has been found.

A total of seven lectins were tested to characterize glycoproteins that are upregulated on C. parvum-infected HCT-8 cells. SBA generated the largest K-S difference between infected and uninfected cell populations. DBA and RCA-I, which, like SBA, recognize N-GalNAc, did not generate the same difference in fluorescence. RCA-I recognizes β-galactosidase (β-gal)-linked GalNAc, and DBA recognizes α-linkage-β-GalNAc. In contrast, SBA does not distinguish between different anomeric derivatives of GalNAc. It has been reported that the difference in N-GalNAc specificity between DBA, RCA-I, and SBA lectins can affect their labeling patterns on the same tissue sample, indicating that anomeric specificity affects lectin binding (2, 24, 26–28). The observation that SBA, but not DBA or RCA-I, binds to infected cells indicates that β-gal-linked GalNAc or α-link-

FIG. 5. Location of SBA fluorescence in relation to parasite. Postpermeabilized infected cells were double labeled with FITC-SBA (green) and C. parvum-specific antibody (red). The localization of parasites on host cells is difficult to discern under visible light, and their positions are indicated with black arrows.
age-d-GalNAc are not overexpressed on C. parvum-infected cells. The structure of GalNAc-decorated glycoproteins and the importance of the regulated expression of this glycoprotein in C. parvum infection are unknown. Whether this effect occurs only in cultured monolayers or is representative of changes in the intestine requires further investigation. Similarly, determining whether the observed overexpression of membrane glycoproteins is a component of the immune response (22) deserves further investigation.

In conclusion, we applied microarray analysis and RT-PCR to identify transcriptional changes in C. parvum-infected and uninfected cells recovered from the same monolayer. This approach revealed that glycoprotein-related genes are upregulated in infected cells, which was confirmed by reacting infected and control cells with fluorescent lectins. In addition to glycoprotein located on the host cell surface, a portion of glycoprotein localizes in close proximity to intracellular parasites. Whether this glycoprotein is of host cell or parasite origin remains to be determined.

ACKNOWLEDGMENTS

Financial support from the NIAID (grant AI055347) is gratefully acknowledged. The Center for Neuroscience Research is supported by grant P30 NS047243. Y.-L.Y. was supported by an Agnes Varis Graduate Fellowship.

We thank Chris Parkin from the Tufts Center for Neuroscience Research Computational Genomics Core for assistance with data analysis. Our thanks to Chuck Shoemaker, TCSVM, John Leong, and Brian Akerley, University of Massachusetts Medical Center, for feedback and suggestions and to Xiaochuan Feng and Claudia Abeijon for assistance with Western blotting.

REFERENCES


