Identification of Genomic Responses to Collagen Binding by Trophozoites of *Entamoeba histolytica*

Anjan Debnath, Pradeep Das, Mohammed Sajid, and James H. McKerrow

1National Institute of Cholera and Enteric Diseases, Kolkata, India; 2Sandler Center for Basic Research in Parasitic Diseases, University of California, San Francisco

Attachment of *Entamoeba histolytica* trophozoites to collagen is a known stimulus for parasite activation, leading to subsequent tissue destruction and invasion. To identify cellular mechanisms of trophozoite activation, we assessed global variations in gene expression during collagen interaction with *E. histolytica*. A shotgun DNA microarray was constructed by use of 9600 random inserts from an *E. histolytica* genomic DNA library. Through differential hybridization, key differences between gene expression in collagen-activated trophozoites and that in nonactivated trophozoites were identified. Fourteen differentially regulated clones were reproducibly identified and selected for sequencing. Among the genes identified were those coding for (1) components of a signaling cascade that had been previously hypothesized to transmit responses to cell attachment, (2) adapter proteins for vesicle formation, and (3) proteins that are implicated in cytoskeletal reorganization and locomotion. Two known virulence-factor genes—those for cysteine proteinases and amebapore—also were up-regulated in response to collagen stimulation. These results provide important new clues about how a pathogen orchestrates responses to the host environment as well as a new tool for the analysis of other aspects of *Entamoeba* species infection and pathogenicity.

Amebiasis is the fourth leading cause of death and the third leading cause of morbidity due to protozoan infections worldwide and is responsible for ~70,000 deaths annually [1]. Amebiasis’s severe clinical form, an invasive enteric infection, is characterized by the capacity of *Entamoeba histolytica* trophozoites to destroy human tissues by penetrating the mucosa and submucosa of the large intestine. Tissue destruction leads to colitis and invasion of the bowel wall, which may, in turn, lead to the metastasis of amebae to other organs. Invasion requires (1) cell locomotion, (2) attachment to host epithelial cells and to the extracellular matrix (ECM), and (3) cytolysis and the destruction of the ECM.

Because collagen is a major component of the epithelial basal lamina and of the ECM of the intestine, *E. histolytica* encounters collagen during colonic invasion. In vitro, the incubation of trophozoites with human collagen type I leads to parasite activation, which includes the release of electron-dense granules (EDGs) containing factors considered to be important in the pathogenicity of this parasite [2–5]. The interaction of trophozoites with human collagen type I triggers biochemical cascades [6], but the mechanism of activation and the complement of downstream products released have not been elucidated.

Recently, it has become possible to decipher transcriptional programs of organisms by studying gene expression en masse [7]. Differences in cell types or states are correlated with changes in the mRNA levels of many genes [8]. DNA-microarray technology provides an opportunity to look simultaneously at changes in gene expression in thousands of genes under different physiological conditions [9].

Because the genome of *E. histolytica* has not yet been fully sequenced, it is currently not possible to construct gene-specific microarrays representing the complete *E.*
Entamoeba histolytica genome. However, given the relatively small size of the E. histolytica genome, valuable information on gene expression that is related to environmental signals can be initially assessed by microarray analyses. To begin to identify global changes in E. histolytica gene expression that are related to collagen activation, we produced an E. histolytica microarray using a variation of the standard array technology. Inserts (~1 kb) from an E. histolytica random-fragment genomic library were isolated to generate shotgun microarrays. To measure variation in gene expression during collagen interaction, the arrays were probed with differentially labeled cDNAs, which had been prepared from total RNA that was isolated from either collagen-activated or nonactivated cells. Polymerase chain-reaction (PCR) products from the arrays that showed differential hybridization were sequenced; these sequences identified several genes that are induced when trophozoites bind to collagen and provide a foundation for a more-comprehensive analysis of Entamoeba responses during host invasion.

**MATERIALS AND METHODS**

**Maintenance of cultures.** E. histolytica trophozoites from the HM1:iMSS strain were maintained in TYI-S-33 medium that was supplemented with penicillin (100 U/mL), streptomycin (100 μg/mL), and 15% adult bovine serum, according to the methods of Diamond [10]. All of the experiments were conducted with trophozoites that had been harvested during the logarithmic phase of growth.

**Activation of E. histolytica trophozoites with human collagen type I and Ca2+.** Trophozoites (~5 × 10^5) were incubated in modified TYI-S-33 medium (without serum or L-cysteine) containing human collagen type I (Sigma) (0.3 mg of human collagen type I/10^6 of trophozoites) and Ca^2+ (1 mmol/L), as described elsewhere [11]. After incubation for 2 h at 37°C, the cells were chilled (4°C) and were collected by centrifugation (230 g for 5 min). In another experiment, the same number of trophozoites were incubated in the absence of human collagen type I and Ca^2+ for 2 h at 37°C. After incubation, the cells were chilled (4°C) and were collected by centrifugation.

**Genomic shotgun library.** Genomic DNA was isolated from 10^8 E. histolytica trophozoites by use of the method of Huber et al. [12]. RNA contamination was removed by RNase digestion for 30 min at 37°C. Genomic DNA was sheared for 30, 60, 80, and 120 s, by use of a nebulizer (Invitrogen) that was connected to the laboratory’s compressed-air line. The sheared DNA was blunt-end repaired by use of T4 DNA and Klenow polymerases, and the ends were dephosphorylated by use of calf intestinal phosphatase, to create a suitable substrate for TOPO cloning (Invitrogen). DNA fragments were ligated into pCR4Blunt-TOPO vector (Invitrogen). The cloning reaction was performed for 5 min at room temperature, with undiluted and with 3- and 9-fold serially diluted blunt-end DNA; 3.3 μL of the precipitated TOPO cloning reaction was electroporated into 50 μL of One Shot TOP10 Electrocomp Escherichia coli (Invitrogen), and the transformants were incubated overnight on agar plates in the presence of X-gal and 100 μg/mL ampicillin.

**E. histolytica microarray.** Individual white colonies from the fresh random-fragment genomic DNA library were transferred to a 600-μL solution of Luria-Bertani medium (with 100 μg/mL ampicillin) in individual wells of a 96-well plate and were incubated for 18 h at 37°C; ~1 μL from each well was used for amplification of the inserts by PCR, using the T7 forward and M13 reverse primers. After analysis of the PCR products on 1% agarose gels, the products were consolidated onto 96-well plates. The PCR products from the 96-well plates were precipitated, were washed, and were transferred to 384-well plates and were printed on glass slides, as described elsewhere [8], by use of the ArrayMaker 2 (University of California, San Francisco). The concentration of each DNA product in 3× standard saline citrate (SSC; 1× SSC is 0.15 mol/L NaCl plus 0.015 mol/L sodium citrate) printing solution was 200 ng/μL, which produced acceptable spot quality. The robot positioned a cluster of 32 specialized, spring-loaded printing tips into adjacent wells of the DNA source plates to a depth of ~1 mm, filling the reservoir slot of each tip with ~1 μL of DNA solution. The tips were then lightly tapped at identical positions on each slide, leaving a small drop (<0.5 nL) of the DNA solution on the poly-L-lysine coated slide. After the DNA solution was deposited on every slide, the tips were washed with 0.5× SSC and were dried, and the process was repeated for the next set of DNA samples, with the new spots offset a small distance (200 μm) relative to previous spots, to produce a high-density grid. Microarrays were printed under controlled environmental conditions (i.e., a temperature of 19°C and a relative humidity of 50%). At the end of the print run, the slides were allowed to dry for ~12 h.

**RNA isolation, labeled cDNA preparation, and microarray hybridization.** Total RNA was isolated from ~5 × 10^6 trophozoites that had been incubated either with or without human collagen type I, by use of a TRIZOL kit (Invitrogen) and as directed by the manufacturer. Optical-density readings were taken at 260 and 280 nm, and the integrity of total RNA was checked in denatured formaldehyde gel, according to standard protocol [13]. Fluorescently labeled cDNA copies of the total RNA pool were prepared from both collagen-induced and uninduced E. histolytica. In brief, 30 μg of total RNA was reverse transcribed by use of StrataScript reverse transcriptase (Stratagene), 10 μg of oligo dT, 100 mmol/L dithiothreitol, and a ratio of amino-allyl dUTP to dTTP of 3:2, for 2 h at 42°C.
After hydrolysis of the RNA, the samples were purified by use of a Zymo-spin column (Zymo Research). The monofunctional NHS-ester Cy3 and Cy5 dyes were coupled with the cDNA, and the unincorporated/quenched Cy dyes were removed by use of a Zymo-spin column; Cy5-labeled sample was eluted into the tube containing the corresponding Cy3-labeled sample. The 2 cDNA pools to be compared were mixed and were applied to the array in a hybridization mixture containing 3× SSC, 0.25% SDS, and 25 mmol/L HEPEs (pH 7.0). Hybridization took place under a glass coverslip in a humidified slide chamber that was submerged in a 63°C water bath for 16 h. The arrays were washed at room temperature in 0.6× SSC and 0.03% SDS and were then briefly rinsed in 0.06× SSC.

**Data acquisition and analysis.** Arrays were scanned by use of a GenePix 4000 scanner (Axon Instruments); images were acquired by scanning with the Cy3 and Cy5 channels at a resolution of 5 μm. Data were quantified by use of GenePix Pro 4.0 software (version 1.9; Axon Instruments). Local background was subtracted from the value of each spot on the array. Only spots in which almost >70% of the pixels had a signal of at least twice the SD of the local background were considered in subsequent data analysis. The data were stored and normalized by use of the NOMAD microarray database (available at: http://ucsf-nomad.sourceforge.net). The Cy5/Cy3 fluorescence ratios and log10-transformed ratios were calculated from the normalized values. Data from related genes in different experiments were analyzed by use of such data-mining tools as CLUSTER analysis and were visualized by use of TREEVIEW [14]. Genes whose expression changed ≥2-fold relative to those in collagen-uninfected *E. histolytica* were selected for further analysis.

**Sequence analysis.** Individual clones were sequenced by use of dye-terminator cycle sequencing reactions (Perkin-Elmer) and an automated ABI Prism 377 DNA sequencer. These sequences were compared with others through searches of the DNA- and protein-sequence databases at the National Center for Biotechnology Information (Bethesda, MD), by use of the BLAST program.

**Reverse-transcription PCR (RT-PCR).** For a few clones (AEC1293, AEC3950, AEC3223, AEC4732, AEC2484, and AEC3233), cases of differential transcription were confirmed by RT-PCR, which was performed with gene-specific primers. For cDNA synthesis, 1 μg of RNase-free DNase-treated total RNA was reverse transcribed for 1 h at 42°C, in 50 μL of 500 μmol/L dNTP, 10 μmol/L dithiothreitol, 1.2 μmol/L dT(15) primer, 40 U of RNAsin, and 75 U of Superscript Reverse Transcriptase II (Invitrogen). Control reactions in which the reverse transcriptase was eliminated in the first-strand reaction were performed for each sample. An aliquot of synthesized cDNAs was amplified by PCR, with the same master mix. All PCRs were performed within the linear range of amplification of the corresponding mRNA species. The linear range of the PCR amplification was verified by the quantification of the cDNA-PCR product that was obtained after amplification for 15–30 cycles (data not shown). In all PCRs, after initial denaturation for 5 min at 94°C, the targeted genes were amplified by 25 amplification cycles (30 s at 94°C, 30 s at 55°C or 50°C, and 1 min at 72°C), followed by final extension for 5 min at 72°C. The products were analyzed on 1% agarose gel, stained with ethidium bromide, and quantitated by use of Quantity One software (Bio-Rad). As a control for a constitutively expressed housekeeping gene, amplification reactions were also conducted with primers for actin, so that the amounts of RNA used in our reactions could be normalized. The primer sequences used to amplify the clones were as follows: for the AEC1293 clone, 5′-CCAGAAGCTTTGATGCAGCT-3′ and 5′-TTGATCTCTGTTGGCACCA-3′; for the AEC3950 clone, 5′-CTAATGGGGCGGAAAAG-3′ and 5′-GTATCTCATGCTGCGGAAAAGG-3′; for the AEC3223 clone, 5′-GAGAAAGGTGTTGATGTGAG-3′ and 5′-AGTGGATCAGTAGAAAGCA-3′; for the AEC4732 clone, 5′-TCCATGATGATAGAAGGATT-3′ and 5′-TGATGATCATGAACACT-3′; for the AEC2484 clone, 5′-AGGAACGGCCAAATCCGTA-3′ and 5′-GCGCCTCCTACCGTTGAT-3′; and for the AEC3233 clone, 5′-GTTTTTGGATTTGATTTTGGTGC-3′ and 5′-ACCAATAAATTCAGA-3′. The primer sequences for *E. histolytica* actin were 5′-GAGGATATGCTTGCACACT-3′ and 5′-ATAGCTGGTCCAGGATCT-3. The primers were designed to yield PCR products that were 300–600 bp in length.

**RESULTS AND DISCUSSION**

ECM degradation is a necessary step in the invasion of the intestinal wall by *E. histolytica* trophozoites. Tissue invasion is accomplished by a complex response that comprises the up-regulation of a number of ameba proteins. Knowledge of this response is key to our understanding ameba pathogenicity [15].

To identify transcriptional responses to collagen attachment, we produced a shotgun genomic DNA microarray for *E. histolytica*, by use of 80 s–sheared 1.5–4.0-kb genomic DNA (figure 1A and 1B). When 3-fold–diluted, blunt-end, dephosphorylated DNA was cloned in pCR4Blunt-TOPO vector, ~2500 white transformants were obtained per cloning reaction. The transformants were analyzed for inserts by PCR, and the average size of the inserts was found to be 1 kb (figure 2). The array was constructed by printing 9600 PCR-amplified inserts from twenty-five 384-well plates, by use of an arraying robot with a similarly sized print-tip cluster of 32 tips, spaced for 384-well plates.

The size of the genome of *E. histolytica* is predicted to be 20 Mb. The average size of the insert applied to the array was 1 kb, and the microarray was constructed by printing 9600 PCR-amplified inserts from an *E. histolytica* genomic-DNA library. Theoretically, the microarray covered 9.6 Mb—almost one-half—of the *E. histolytica* genome. Because the *E. histolytica*
Genomic Responses by Entamoeba histolytica to Collagen

Figure 1. Shearing of Entamoeba histolytica genomic DNA, by use of a nebulizer. A, Lane 1, DNA molecular weight marker; lane 2, Genomic DNA of E. histolytica isolated from 10⁸ trophozoites, according to the method described in Materials and Methods. B, Lane 1, DNA molecular weight marker; lanes 2–5, 1.5–4.0-kb sized sheared genomic DNA of E. histolytica.

The genome has large number of repetitive sequences, some inserts will have either identical or nearly identical sequences. On the basis of these calculations, we estimate that our microarray may cover one-third to one-half of the E. histolytica genome. Moreover, due to the random nature of the shotgun-cloning procedure, there is an unbiased representation of sequences in the genomic library, such that genes that are rarely transcribed are as equally represented as highly expressed ones—overall genome representation is usually good in shotgun libraries, with comparatively little variation across the genome. Thus, even if randomly selected clone inserts are used as probes, there should be good coverage and relatively little redundancy. Also, not only can coding regions be studied, but intergenic regions can be studied too [16].

The faithfulness of the shotgun DNA microarray for the reporting of collagen-specific gene expression was apparent from the fact that, due to the size-selection process that was undertaken prior to shotgun cloning, the fragments were similar in length. Also, variation in the efficiency of priming events was minimized, because an identical primer pair was used for all reactions. Three separate hybridizations from 3 independent cDNA preparations showed virtually identical differential-hybridization patterns and, in total, 225 clones of 9600 white colonies printed on the DNA array were actually up-regulated (≥2-fold) due to collagen activation (figure 3). A scatterplot of the relative expression ratios for all detected signals is shown in figure 4; unregulated data points are found on the main diagonal, and regulated data points are away from the diagonal. Clustering the genes enhanced their similarities and made it easy to correlate those genes that might have a relationship with the specific event (figure 5). The numerical values for the signal intensities at the spots showing overexpression due to collagen activation and that were selected for future sequencing are shown in table 1. The sequencing results of the initial 10 clones showed good homology with functionally important proteins. Although many other clones also showed high Cy5/Cy3 ratios, only an accessible number of clones was selected for sequencing, which will readily lend itself to more-detailed functional analysis. The identification, by homology, of the genes up-regulated due to collagen activation in E. histolytica is shown in table 2. Few genes that were shown by the DNA microarray to be overexpressed during collagen interaction were validated by RT-PCR, affirming the reliability of our microarray (figure 6).

An emerging body of knowledge suggests that different cel-
lular activities, including adhesion, growth, motility, and secretion, are regulated by cytoskeletons. The interaction between *E. histolytica* trophozoites and collagen results in rapid polymerization of the amorphous cytoplasmic ground substance into an organized cytoskeletal superstructure that is made of microtubules, microfilaments, and intermediate filaments. However, cytoskeletons from trophozoites that are not activated by collagen are not structured [18]. In *E. histolytica*, cytoskeletons also play a role in pseudopod formation, adhesion-plaque formation, and phagocytosis. The interplay of several cytoskeletal proteins, as well as of receptors and members of the signal-transduction pathway, participate in pseudopod formation and phagocytosis. Therefore, the genomic fragment of clone AEC1293 encodes >1 gene that is involved in cytoskeletal rearrangement and in signal-transduction pathways (particularly, the Src homology 3 [SH3] domain-containing protein), leading to invasion of basal epithelial layers via both locomotion and the phagocytosis of human cells (myosin IB).

It is well known that cell adhesion involves interaction of the cell surface with the ECM, remodeling of the plasma membrane, and reassembly of the cytoskeleton [19]. Cell at-

**Figure 4.** Scatterplot comparing the hybridization of collagen-induced and uninduced cDNA probes with the arrayed DNA. The numerical median values for all the spots on the array for the red channel (635 nm) were subtracted from the background value at 635 nm and were plotted on the X-axis; similarly, the numerical median values for all the spots on the array for the green channel (532 nm) were subtracted from the background value at 532 nm and were plotted on the Y-axis.
Genomic Responses by Entamoeba to Collagen

Figure 5. Clustered display of data from a portion of the entire microarray, indicating collagen-induced gene expression from 3 different experiments (EhCS1p1-056, EhCS1p2-055, and EhCS1p2-060). Data from related genes in 3 separate hybridizations from 3 independent cDNA preparations, prepared from both nonactivated and collagen-activated trophozoites, were analyzed by use of such data-mining tools as CLUSTER analysis and were visualized by use of TREEVIEW. Increasingly positive log ratios due to collagen activation are colored red; highlighted sections are among the clones selected for sequencing.

tachment and spreading induce the phosphorylation of several focal adhesion proteins, including the cytoskeletal protein paxillin and a 125-kDa focal adhesion kinase (pp125FAK) [20]. Previous studies have provided evidence for the recruitment of pp60Src and paxillin by pp125FAK in the collagen type I– and Ca2+-induced signaling cascade [21]; pp60Src binds to tyrosine-phosphorylated motifs on paxillin via SH3 domains [22]. Among the collagen-induced gene transcripts that were identified by use of the present study’s microarray, there were sequences that showed homology with mRNA for SH3 domain–containing proteins. SH3 domains are conserved protein modules in many proteins and are involved in signal transduction as well as in cytoskeletal processes [23]. SH3 domains have proposed roles in directing the assembly of NADPH oxidase subunits [24], in modulating both the activity of the phosphatidylinositol-3’-kinase [25] and the GTPase activity of dynamin [26], and in localizing proteins to distinct subcellular sites. A previous study has concluded that E. histolytica trophozoite interaction elicits the rearrangement of cytoskeleton [27] and triggers the tyrosine kinase cascades, resulting in adhesion that is important in the pathogenicity of this parasite. This conclusion is consistent with the results reported here and validates the usefulness of microarray-hybridization analysis for the identification of differences between 2 pools of total RNA. The SH3-domain transcript that is involved in signal transduction and cytoskeletal rearrangement due to collagen activation is, in fact, up-regulated.

Intersectin, a novel adapter protein containing both EH and SH3 domains, may regulate the formation of clathrin-coated vesicles [28]. The present study’s microarray has identified the participation, during activation of E. histolytica with human collagen type I, of an intersectin-like protein that is involved in the formation of clathrin-coated vesicles. Previous findings also have suggested that clathrin plays a role in the regulated secretion of EDGs in E. histolytica during interaction with collagen type I [29].

We also have identified the expression of genes that are related to the vav oncogene. Vav proteins are guanine nucleotide exchange factors for Rho family GTPases that activate pathways leading to actin cytoskeletal rearrangements and transcriptional alterations. Vav proteins contain several protein-binding domains that can link cell-surface receptors to downstream signaling proteins. There is some evidence that, in E. histolytica, the activity of collagen type I and of Ca2+-dependent EDG-containing collagenase [11], the release of soluble proteolytic activities, and adhesion are dependent on cytoskeletal integrity [5, 30]. The results from our microarray suggest that the interaction of E. histolytica trophozoites and the major ECM component, type I collagen, elicits the rearrangement of the cytoskeleton through Vav proteins.

In the present study, the gene encoding a protein that is homologous with the human GRB2 (growth factor receptor–binding protein) has been found to be up-regulated because of collagen interaction. GRB2 remains associated with activated tyrosine-phosphorylated epidermal growth factor receptors and platelet-derived growth factor receptors via its SH2 domain. It has been found that the stimulation of trophozoites with collagen type I and Ca2+-dependent EDG-containing collagenase [11], the release of soluble proteolytic activities, and adhesion are dependent on cytoskeletal integrity [5, 30]. The results from our microarray suggest that the interaction of E. histolytica trophozoites and the major ECM component, type I collagen, elicits the rearrangement of the cytoskeleton through Vav proteins.
Asparaginyl-tRNA synthetases are involved in translation, ribosomal structure, and biogenesis; they catalyze the formation of asparaginyl-tRNA in a 2-step reaction by activating asparagine. The presence of the transcript for asparaginyl-tRNA synthetase during collagen interaction may relate to the gene encoding the asparagine-rich surface protein found in E. histolytica. The corresponding functional gene does not exist in E. dispar, suggesting that this protein might be important to pathogenicity in E. histolytica [33].

Cysteine proteinases that are released by E. histolytica trophozoites play a key role in the invasion and inflammation of the gut. Most of the proteinase activity that is detected in E. histolytica lysates can be attributed to the expression of 4 cysteine proteinase genes: acp1, acp2, acp3, and ehcp5 [34, 35]. The microarray results from the present study show an overexpression of genes encoding EhCP1 (acp3) and EhCP2 (acp2) due to collagen interaction, suggesting a role for these proteinases in invasion. The hypothesis that gives them the most direct role is that they are involved in the cleaving of collagen, a major component of the ECM. Although increased expression of the EhCP2 enzyme does not augment liver-abscess formation in gerbils [36], it might be involved in degrading collagen in the colonic wall.

Another putative virulence-factor gene that was transcriptionally up-regulated by collagen interaction showed homology with the E. histolytica gene for amebapore (a pore-forming peptide). Amebapore is an effector molecule for E. histolytica-mediated cytolyis [37, 38]. The present study’s microarray also identified increased expression of E. histolytica ehapt2 retrotransposon due to collagen activation. It has been found recently that the presence of ehapt2 elements is a species-specific property—such elements are present in E. histolytica and are absent in E. dispar [39]. It is likely that the increased expression of ehapt2 influences the expression of at least some of the genes, because the location of ehapt2 is in close proximity to various gene-translation initiation or termination codons [40]. Collagen interaction also led to the enhanced expression of E. histolytica putative GTPase, putative RNA-binding protein genes, and putative DNA ligase genes. Because transcription in E. histolytica is performed in a monocistronic manner [41], most likely the expression of a number of E. histolytica genes is tightly regulated, and the increased expression of ehapt2 resulted in the overexpression of closely located genes, namely, GTPase and putative RNA-binding protein genes, as well as DNA ligase genes.

Ribosomal proteins catalyze ribosomal assembly and stabilize tRNA tertiary structure, adapting the structure of the ribosome for optimal function [42]. 40S ribosomal proteins S19 and S24 were found, in the present study, to be up-regulated during collagen activation in E. histolytica. It has appeared that the eukaryotic S19 (a core protein that is associated with the 18S rRNA of the 40S small subunit) had no counterpart in prokaryotes, mitochondria, and chloroplast, and therefore this protein was thought to be a recent addition to the eukaryotic ribosomal protein repertoire [43]. However, on the basis of the results from our microarray, it seems that E. histolytica may possess the gene encoding the 40S ribosomal protein S19; this finding will shed new light on the understanding of the evolution of this parasite. No data are currently available con-

Table 1. Summary of the collagen-responsive gene-expression data.

<table>
<thead>
<tr>
<th>Clone identification number</th>
<th>F635 median minus B635 (635/532)</th>
<th>F532 median minus B532 (635/532)</th>
<th>Ratio of medians (635/532)</th>
<th>Median of ratios (635/532)</th>
<th>Log ratio (635/532)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEC1293</td>
<td>6616</td>
<td>2664</td>
<td>2.483</td>
<td>2.399</td>
<td>1.312</td>
</tr>
<tr>
<td>AEC2390</td>
<td>7263</td>
<td>953</td>
<td>7.621</td>
<td>6.644</td>
<td>2.930</td>
</tr>
<tr>
<td>AEC3223</td>
<td>32,062</td>
<td>5156</td>
<td>6.218</td>
<td>6.227</td>
<td>2.637</td>
</tr>
<tr>
<td>AEC4732</td>
<td>20,702</td>
<td>3463</td>
<td>5.978</td>
<td>5.481</td>
<td>2.580</td>
</tr>
<tr>
<td>AEC0490</td>
<td>6609</td>
<td>1776</td>
<td>3.721</td>
<td>3.725</td>
<td>1.896</td>
</tr>
<tr>
<td>AEC0221</td>
<td>2234</td>
<td>653</td>
<td>3.421</td>
<td>3.237</td>
<td>1.774</td>
</tr>
<tr>
<td>AEC0807</td>
<td>8286</td>
<td>1816</td>
<td>4.563</td>
<td>4.099</td>
<td>2.190</td>
</tr>
<tr>
<td>AEC2484</td>
<td>17,580</td>
<td>3605</td>
<td>4.877</td>
<td>4.905</td>
<td>2.286</td>
</tr>
<tr>
<td>AEC3233</td>
<td>4860</td>
<td>1626</td>
<td>2.989</td>
<td>3.156</td>
<td>1.580</td>
</tr>
<tr>
<td>AEC6195</td>
<td>12,175</td>
<td>2452</td>
<td>4.965</td>
<td>4.577</td>
<td>2.312</td>
</tr>
</tbody>
</table>

NOTE. Clones showing overexpression due to collagen activation were selected after scanning the microarrays by use of a GenePix 4000 scanner and obtaining the numerical expression data for each spot by use of GenePix Pro 4.0 software (version 1.9). F635 is the fluorescence at 635 nm for the Cy5 channel, and F532 is the fluorescence at 532 nm for the Cy3 channel; B635 and B532 are background values at 635 nm and 532 nm, respectively. Log ratio, log transform of the ratio of the medians; median of ratios, median of pixel-by-pixel ratios of pixel intensities, with the median background value subtracted; ratio of medians, ratio of the median intensities of each feature for each wavelength (635 nm and 532 nm), with the median background value subtracted.
<table>
<thead>
<tr>
<th>Clone identification number</th>
<th>Putative identity</th>
<th>Function</th>
<th>No (%). of nucleotide or amino acid residues matched/compared</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEC1293</td>
<td>SH3 domain–containing protein SH3P18</td>
<td>Signal transduction and rearrangement of cytoskeleton</td>
<td>28/51 (54)</td>
<td>AAC50593</td>
</tr>
<tr>
<td>1e-11</td>
<td>Intersectin 1</td>
<td>Clathrin-coated vesicle formation</td>
<td>25/47 (53)</td>
<td>O42287</td>
</tr>
<tr>
<td>7e-10</td>
<td>VAV (Homo sapiens)</td>
<td>Actin cytoskeletal rearrangement, leading to release of soluble proteolytic activities and adhesion</td>
<td>25/50 (50)</td>
<td>CAA34383</td>
</tr>
<tr>
<td>1e-07</td>
<td>Growth factor receptor–bound protein 2 (GRB2 adapter protein) (SH2/SH3 adapter GRB2) (ASH protein)</td>
<td>Reorganization of vimentin filament network, for growth, motility, and secretion of proteinases</td>
<td>Not found in database</td>
<td>Q07883</td>
</tr>
<tr>
<td>1e-06</td>
<td>Unconventional myosin IB (Entamoeba histolytica)</td>
<td>Locomotion and phagocytosis</td>
<td>29/53 (54)</td>
<td>AAC47535</td>
</tr>
<tr>
<td>AEC3950</td>
<td>Asparaginyl-tRNA synthetase (Mus musculus)</td>
<td>Protein synthesis, especially Entamoeba histolytica asparagine-rich surface protein</td>
<td>131/238 (55)</td>
<td>XP_128957</td>
</tr>
<tr>
<td>1e-74</td>
<td>Cysteine proteinase 1 precursor</td>
<td>Collagen or other host-protein degradation</td>
<td>183/197 (92)</td>
<td>Q01957</td>
</tr>
<tr>
<td>1e-122</td>
<td>Cysteine proteinase 2 precursor</td>
<td>Collagen or other host-protein degradation</td>
<td>183/197 (92)</td>
<td>M94162</td>
</tr>
<tr>
<td>AEC4732</td>
<td>Entamoeba histolytica gene for pore-forming peptide</td>
<td>Cytolysis</td>
<td>114/120 (95)</td>
<td>X70851</td>
</tr>
<tr>
<td>5e-46</td>
<td>Entamoeba histolytica putative GTPase and putative RNA-binding protein genes, partial cds; and ehapt2 retrotansposon</td>
<td>Gene activation?</td>
<td>168/199 (84)</td>
<td>AY141200</td>
</tr>
<tr>
<td>AEC0490</td>
<td>Entamoeba histolytica putative mitochondrial type and endoplasmic reticulum-like hsp70 gene</td>
<td>Stress response</td>
<td>116/125 (92)</td>
<td>AY141198</td>
</tr>
<tr>
<td>AEC0221</td>
<td>40S ribosomal protein S24</td>
<td>Ribosomal assembly, stabilization of rRNA tertiary structure</td>
<td>49/84 (58)</td>
<td>Q962Q6</td>
</tr>
<tr>
<td>2e-18</td>
<td>Ribosomal protein S19</td>
<td>Ribosomal assembly, stabilization of rRNA tertiary structure</td>
<td>49/84 (58)</td>
<td>1909359A</td>
</tr>
<tr>
<td>AEC0807</td>
<td>40S ribosomal protein S24</td>
<td>Ribosomal assembly, stabilization of rRNA tertiary structure</td>
<td>49/84 (58)</td>
<td>1909359A</td>
</tr>
<tr>
<td>AEC3233</td>
<td>40S ribosomal protein S24</td>
<td>Ribosomal assembly, stabilization of rRNA tertiary structure</td>
<td>49/84 (58)</td>
<td>1909359A</td>
</tr>
<tr>
<td>AEC5195</td>
<td>40S ribosomal protein S24</td>
<td>Ribosomal assembly, stabilization of rRNA tertiary structure</td>
<td>49/84 (58)</td>
<td>1909359A</td>
</tr>
</tbody>
</table>

**NOTE.** Subsets of individual clones that are contained within the microarray and that hybridized to collagen-activated *E. histolytica* cDNAs are shown with their sequence similarity entries in protein databases (BLAST X) or nucleic acid databases (BLAST N). 1e-11 and similar notations represent the statistical probability of the match. The parameter *E*-value (expect value) was used to describe the number of hits one can expect to see by chance alone when one searches a database of a particular size. The lower the *E*-value (or the closer it is to 0), the more significant is the match [17].
cerning a putative functional role played by the ribosomal protein S24.

Heat-shock proteins (HSPs) are synthesized by cells not only in response to heat, but also to various other stressful stimuli. The cytoprotective capacity of HSPs may be attributed, in part, to their ability to recognize proteins that are not yet in their native conformation or that are denatured [44].

The cluster of genes identified in the present study by microarray analysis provides insight into collagen-induced signaling in amebae and supports the hypothesis that a genetic program is stimulated by interaction with the ECM. Until the genome-sequencing project is complete, the random array produced for the present study provides a tool for the identification of other biologically important variations in Entamoeba gene expression.

References

25. Pleiman GM, Hertz WM, Cambier JC. Activation of phosphatidyli-