Table 1
Verification of array data by qRT-PCR

Locus represented by probe set	Length of time in vivo (days)	Array[a]		qRT-PCR[b,c]		Gene annotation
		Fold difference	SAM q value	Fold difference	p value (t-test)	
34.m00259	29	8.29	0	6.2	0.0006	Hypothetical protein
56.m00149	29	5.45	0	23.2	0.017	Calcium binding protein[d]
395.m00028	1	14.65	0.0165	2.2	0.01	EhMGL1[e]
347.m00051	1	2.57	0.0482	1.49	0.64	Tryptophanase[f]
613.m00022	1	−1.86	0.016	−2.37	0.01	URE3-BP
2.m00532	1	2.94	0.016	0.94	0.84	Cysteine protease 9[g]
147.m00095	1	1.07	0.51	1.02	0.87	RNA polymerase II L[h]
1.m00673	1	−1.62	0.09	−1.1	0.73	Lgl5[i]
176.m00112	29	1.22	0.68	1.2	0.55	TSA[j]
344.m00043	29	1.12	0.7	−1.03	0.96	RNA polymerase II subunit 13[h]
152.m00118	29	1.14	0.7	1.05	0.85	Cysteine protease 19[g]

[a] Change in transcript levels calculated from the results of Affymetrix array hybridization of cDNA derived from in vitro and in vivo amebae (fold increase and statistical significance).
[b] Values were compared after normalization using the geographic mean of empirically determined invariant genes (RNA pol II L, TSA, and cysteine protease 19) after correction for the efficiency of amplicon amplification.
[c] Change in transcripts determined by qRT-PCR for three mice (fold increase and statistical significance).
[d] EhCaBP has been previously described by Sahoo et al. [36].
[e] EhMGL1 has been previously described by Tokoro et al. [13].
[f] Tryptophanase has been previously described by Anderson et al. [30].
[g] Cysteine protease 9 and 19 have been previously described by Bruchhaus et al. [12].
[h] RNA polymerase II L and 13 have been previously described and used as internal controls by Beck et al. [33].
[i] Lgl5 is a novel member of the family of genes which encode the light subunit of the Gal/GalNAc lectin [5].
[j] TSA has been independently described by two groups [21].

Table 2
Cysteine protease transcripts

Probe set	Day 1 in vivo		Day 29 in vivo		In vitro ameba	Annotation of gene represented by probe set[a]
	Fold difference	SAM q value	Fold difference	SAM q value	Average normalized probe value	
10.m00362_at	**28.33**	**0**	**35.41**	**0**	**82**	**Cysteine proteinase 4[b] related**
2.m00545_at	**9.4**	**0**	**6.44**	**0.11**	**97**	**Cysteine protease 6[b] related**
2.m00532_at	**2.94**	**0.02**	**2.09**	**0.41**	**72**	**Cysteine protease 9[b]**
242.m00078_s_at[c]	**1.7**	**0.09**	**2.38**	**0.04**	**5403**	**Cysteine protease 1[b]**
79.m00156						
109.m00099_at	N/A	N/A	2.08	0.65	8	Cysteine protease 14-related[b]
222.m00084_at	1.41	0.44	−1.04	0.64	8	Cysteine protease 16[b]
152.m00118_at	1.24	0.44	1.14	0.7	853	Cysteine protease 19[b]
180.m00101_at	1.2	0.48	1.01	0.7	278	Cysteine protease 17[b]
65.m00146_at	1.11	0.51	1.77	0.53	29.5	OTU-like cysteine protease, putative
3.m00597_at	1.06	0.51	1.2	0.69	142	Calpain-like cysteine protease, putative
24.m00271_at	1.05	0.51	−1.11	0.6	23	Cysteine protease 10[b]
71.m00132_at	−1.08	0.39	−1.1	0.6	25.9	Cysteine protease, putative
2.m00493_at	−1.25	0.3	1.47	0.6	13.55	Peptidase, putative
191.m00117_at	−1.48	0.09	−1.2	0.47	7587	Cysteine protease 5[b] related
260.m00068_at	−1.49	0.12	1.09	0.7	910	Cysteine protease, putative
446.m00031_at	−1.62	0.09	−1.5	0.28	38	Cysteine protease 13[b]
97.m00133_at	−1.95	0.06	−1.62	0.4	38	Cysteine protease, putative cp112[d]
501.m00019_s_at	**−14.31**	**0.01**	**−4.88**	**0.02**	**2994**	**Cysteine protease 8[e]**

[a] Genes in bold demonstrate statistically significant change.
[b] In vitro expression previously determined by Bruchhaus et al. [21].
[c] Probe set also represents the 79.m00156 open reading frame which differs by one nucleotide from 242.m00078. This causes a change in the terminal 3′ of the open reading frame and hence an extension of 57 nucleotides.
[d] Gene described by Garcia-Rivera et al. [41].
[e] Probe set also represents the 358.m00031 open reading frame, which differs by three nucleotides from 501.m00019.

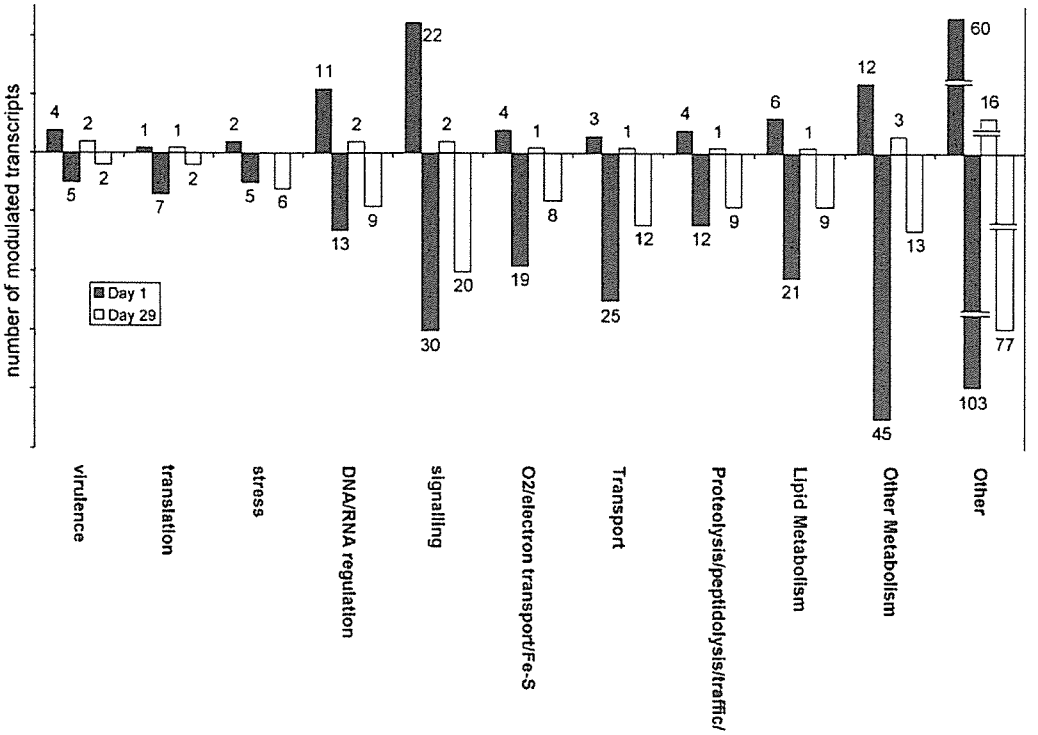


Fig. 3. Distribution of changed transcripts into functional categories. Differentially expressed genes were categorized by functional group. The changed transcripts were assigned to a putative function on the basis of gene annotation. Genes of unknown function were further annotated and categorized on the basis of Pfam/COG families and then loosely grouped (shown in more detail in Supplemental Table 5). The functional group is shown on the x-axis, and number of genes modulated in either a positive or negative direction is indicated on the y-axis. Transcripts changed at Day 1 are shown as solid bars and Day 29 modulated genes are shown as white bars. The exact number of changed transcripts is shown beside each column.

fold. The transcript from a CP4-like gene (CP10.m00362 gene) was expressed at low levels in vitro, but exhibited a 20–35-fold increase at Days 1 and 29. The CP6-like cysteine protease (CP2.m00545) was 10-fold increased at Day 1. In contrast, the CP8 transcript was significantly decreased in vivo (Table 2). The changes in transcript levels of specific CP genes suggest non-redundant functions for the individual proteases in the intestinal life of the parasite.

3.6. Laterally-transferred metabolic enzymes

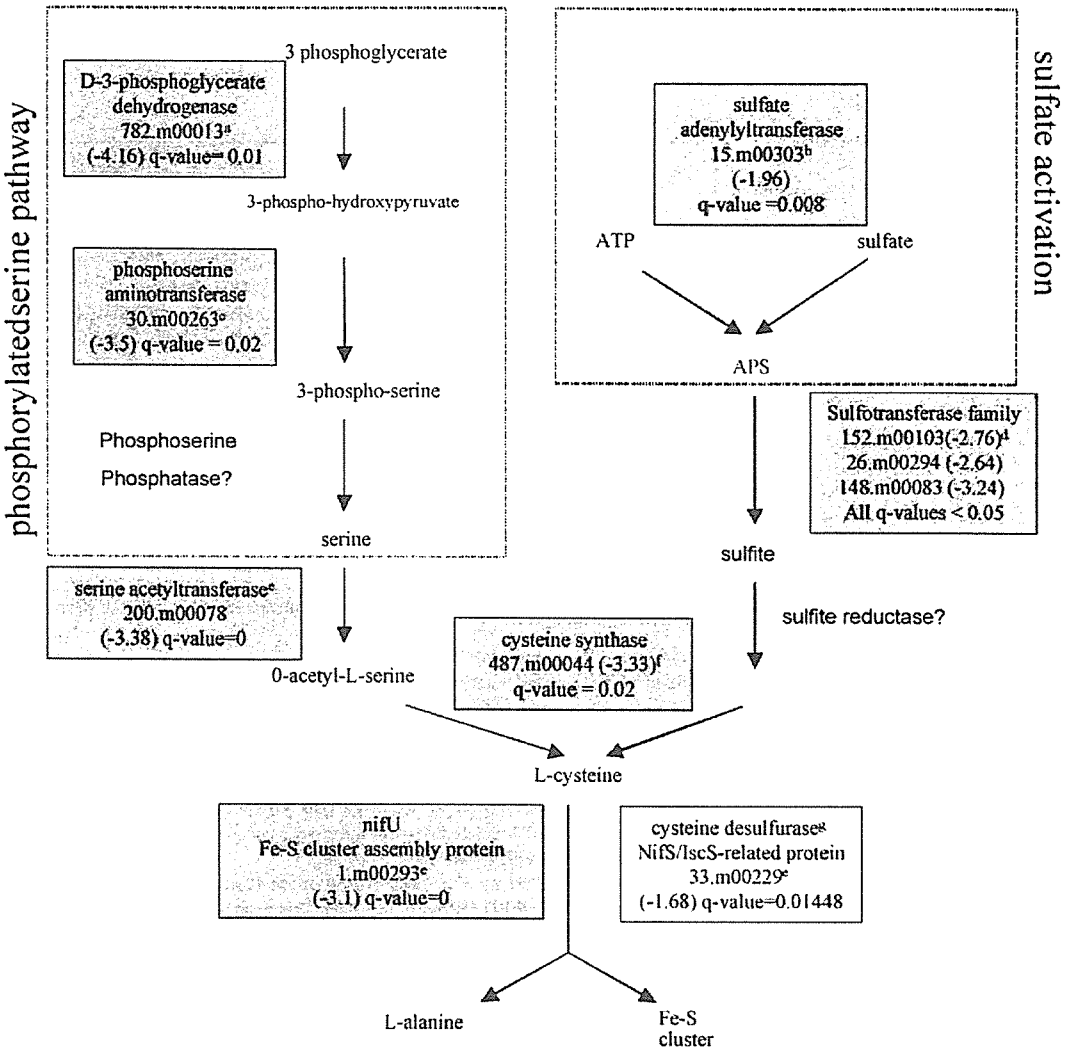
A number of important metabolic enzymes appear to have been acquired by lateral gene transfer from prokaryotes. These genes encode many of the enzymes involved in serine and cysteine biosynthesis, degradation of sulfur-containing amino acids (Fig. 4), and energy metabolism [13–16]. The results of the array hybridization indicated that 84 of 96 laterally transferred genes had detectable transcripts, with 11 decreased and 6 increased at least two-fold in vivo (Table 3).

3.7. Oxygen detoxification

Ameba colonize the anaerobic lumen of the gut. In this environment the proteins necessary for oxygen defense would likely not be required. *E. histolytica* is an aerotolerant anaerobe that uses an iron–sulfur ferredoxin to oxidize pyruvate as the end-electron acceptor. The levels of ferredoxin 2 [17] and iron–sulfur flavoproteins (328.m00064) transcripts were decreased, as was the mRNA of almost all genes encoding enzymes involved in the synthesis of iron–sulfur [Fe–S] clusters [18] including the sulfur assimilatory de novo cysteine biosynthetic pathway [14] (Fig. 4). One exception to the down-regulation of iron–sulfur transcripts was the mRNA of a protein with a strong match (99.8% aligned with a 770 Score, $E=0.0$) to the hybrid cluster protein (8.m00410; pfam HCP,cd01914.1). Hybrid cluster proteins incorporate both a [4Fe–4S] cubane cluster and a [4Fe–2S–2O] iron–sulfur cluster. Hybrid cluster proteins are thought to be involved in nitrate and/or nitrate respiration and have only been detected in facultative anaerobes cultivated in anaerobic conditions [19,20]. Also increased were thioredoxin (212.m00092) and glutamate synthase (78.m00152), both electron transport proteins lacking Fe–S clusters. The transcripts of other proteins involved in the detoxification of reactive oxygen metabolites were unchanged in vivo [21–24]. We concluded that adaptation to the intestine was associated with an overall decrease in the abundance of mRNAs encoding for oxygen detoxification proteins.

3.8. Energy metabolism

The lack of mitochondria necessitates that *E. histolytica* derive most of its ATP from glycolysis and fermentation. It was therefore notable that the transcripts of the genes encoding the glycolytic pathway were overall decreased at Day 1 but



[a] D-3-phosphoglycerate dehydrogenase has been previously described [14].
[b] sulfate adenylyltransferase has been previously described [14].
[c] phosphoserine aminotransferase has been previously described [14].
[d] the transcript of 152.m00103 is significantly increased at day 29.
[e] 200.m00078 (SAT1) has been previously described [14]. A second minor transcript. (253.m00083) encoding the same enzyme is significantly increased at day 1 and 29.
[f] 487.m00044 (CS2) has been previously described by Nozaki et al. [40].
[g] A second transcript annotated as cysteine desulfurase 46.m00208 was also down-regulated in vivo (-6.09; q-value=0). This gene may be a molybdenum cofactor sulfurase.
[h] 33.m00229 (NifS) and 1.m00293 (nifU) have been previously described [40].

Fig. 4. Modulation of transcripts encoding enzymes involved in iron–sulfur cluster biosynthesis. Transcripts significantly modulated in vivo are shown in shaded boxes. Gene name is followed by the locus represented by the microarray probe set, fold change in vivo, and significance (q-value) [14,40].

not Day 29 (Fig. 5). Diminution of mRNA levels for glycolytic enzymes with low catalytic efficiencies (that could therefore be pathway regulatory points) included one of the two pyruvate phosphate dikinase genes (−1.91 q-value = 0.05) and fructose-1,6-bisphosphate aldolase (−1.82 q-value = 0.0078) [25]. The PPi-dependent phosphofructose kinase (PFK) gene transcript was also significantly down-regulated (−2.03 q-value = 0.039). This enzyme, which accounts for most of the PFK activity of cultured *E. histolytica* trophozoites employs inorganic pyrophosphate as the phosphoryl donor to generate fructose-1,6-bisphosphate, and therefore increases the net energy generated by this pathway [26]. The transcript of a second gene encoding the ATP-requiring phosphofructose enzyme was expressed at low levels in culture, and was unchanged in vivo [27]. The potential down-regulation of the glycolytic pathway at Day 1 suggested by these results, could reflect a shift from rapid growth made possible by the abundant nutrients of tissue culture, to a more parsimonious existence in the resource poor colon [28]. An adaptation to slower growth is also suggested by the decrease in mRNA for the glucose ribose porter (53.m00214)

Table 3
Significantly changed transcripts from the laterally transfered genes

Probe set	Day 1 in vivo		Day 29 in vivo		Annotation of gene represented by probe set
	Fold difference	SAM q value	Fold difference	SAM q value	
395.m00028_s_at[a]	14.65	0.02	1.3	0.7	Methionine gamma-lyase
6.m00425_at	5.93	0.01	2.66	0.02	Endo-1,4-beta-xylanase
36.m00222_s_at[b]	3.01	0.01	1.42	0.6	Aminotransferase
78.m00152_s_at[c]	2.86	0.02	1.89	0.14	Glutamate synthase small subunit
347.m00051_at	2.64	0.04	1.93	0.6	Tryptophanase
189.m00093_x[d]_at	2.48	0	1.97	0.14	D-Hydantoinase
22.m00291_at	−1.15	0.36	−2.01	0.02	Aspartate ammonia-lyase
91.m00198_at[e]	−2.03	0.04	1.44	0.7	Pyrophosphate-dependent phosphofructokinase
25.m00254_at	−2.08	0.05	−1.65	0.35	Nicotinate phosphoribosyltransferase
1.m00704_at	−2.65	0	−1.2	0.53	Aspartate-ammonia ligase
289.m00068_at	−2.66	0.02	−2.83	0.05	Amidohydrolase
10.m00331_at	−2.96	0.01	−2.06	0.09	(2*r*)-Phospho-3-sulfolactate synthase
21.m00283_at[f]	−3.1	0	−1.18	0.57	Fe–S cluster assembly protein NifU
200.m00078_s_at[g]	−3.38	0	−5.95	0	Serine acetyltransferase
30.m00263_s_at[h]	−3.5	0.02	−1.22	0.53	Phosphoserine aminotransferase
782.m00013_s_at[i]	−4.16	0.01	−4.79	0	D-3-Phosphoglycerate dehydrogenase
103.m00185_at[j]	−10.53	0	−4.88	0.07	Fe-hydrogenase

[a] 395.m00028_s represents 202.m00088 with which it is identical and 132.m00106 from which it differs by one nucleotide. This gene has been previously described by Tokoro et al. as EhMGL1 [13].

[b] 36.m00222_s_at also represents 36.m00207 with which it differs by a 70 nt internal deletion.

[c] 78.m00152_s_at also represents 3.m00589 with which it differs by a 45 nt internal deletion.

[d] (_at) probe sets represent single genes, s_at. probe sets recognize two or more genes with which the probes are an exact match, _x_at probe sets may cross-hybridize in an unpredictable manner with sequences other than the main target.

[e] This gene has been previously described by Deng et al. [26].

[f] 21.m00283; this gene has been previously described [14].

[g] 200.m00078_s also represents 141.m00079 with which it is identical. This gene has been previously described by Nozaki et al. as SAT1 [14].

[h] 30.m00263_s_at represents 201.m00117 with which it is identical and 253.m00078 which differs from the other two genes by 9 nucleotides in 963 nts and whose ORF terminates 114 nts sooner than that of the other transcripts. This gene has been previously described [14].

[i] 782.m00013_s also represents 130.m00118 and 130.m00121. 130.m00118 differs by a one base pair insertion 12 nt from the terminal of 782.m00013 which changed the ORF 130.m00118 extend an additional 258 nt. 130.m00118 and 130.m00121 ORFs are identical. This gene has been previously described [14].

[j] 103.m00185 has been previously described as iron-dependent hydrogenase 2 by Nixon et al. [15].

and several tRNA synthetase genes, although no changes were observed for DNA replication and nucleotide synthesis genes.

An alternative hypothesis is that different pathways for energy generation are utilized in vivo. Zuo et al. had suggested that in the absence of glucose *E. histolytica* may catabolize amino acids; however, no consistent increase in the transcripts involved in this process was observed [29–30]. In contrast, an increase in the transcripts of two putative lipases and phospholipases suggested that this alternative energy source may be preferentially utilized in vivo.

3.9. Intracellular signaling proteins

Several potential DNA and RNA binding protein gene mRNA levels were modulated in response to the intestinal environment (Table 4). The EhEBP 1 transcription factor, an enhancer of Gal/GalNAc lectin *hgl5* subunit transcription, was down-regulated two-fold at Day 1 [31]. Four DNA binding proteins with 6–16-fold changes at Day 29 (111.m00140, 227.m00086, 143.m00093, 20.m00272) shared similarity at the protein level (71–83%). One of the other hypothetical transcripts increased at Day 29 (seven-fold) encoded a basic region leucine zipper domain. The pre-mRNA splicing factor PrP17 mRNA necessary for the splicing of introns greater than 200 nucleotides in *S. cerevisiae* and *S. pombe* was also increased [32]. A second potential RNA binding protein (91.m00190), similar to a *P. falciparum* protein RNA binding protein, was down-regulated two-fold at both Days 1 and 29.

Transmembrane receptor kinases (TMK) are key mediators of signal transduction in eukaryotic cells. *E. histolytica* is remarkable among protists for possessing over 80 putative TMKs. The *E. histolytica* TMK's have been grouped into six subfamilies (A–F) based on the kinase domain and motif structure of the extracellular domains [33]. Members of each of the subfamilies were discovered to be expressed in vitro and in vivo. Transcripts from eight TMKs were modulated in vivo (Table 5). Transcripts from 9 non-receptor *E. histolytica* kinases, and 10 phosphatases genes were also modulated. The transcript of phosphoinositide 3-kinase (9.m00406), implicated in phagocytosis, was one of the kinases decreased at Day 29 [34,35]. The mRNA of several calcium binding proteins were also changed in abundance. One of the up-regulated transcripts in vivo was EhCaBP [36], while transcripts of the potential calcium storage proteins Grainins 1 and 2 [37] were down-regulated. We concluded that the adaptation to the intestinal environment was accompanied, and likely in part mediated by, alterations in mRNA abundance of key signaling molecules.

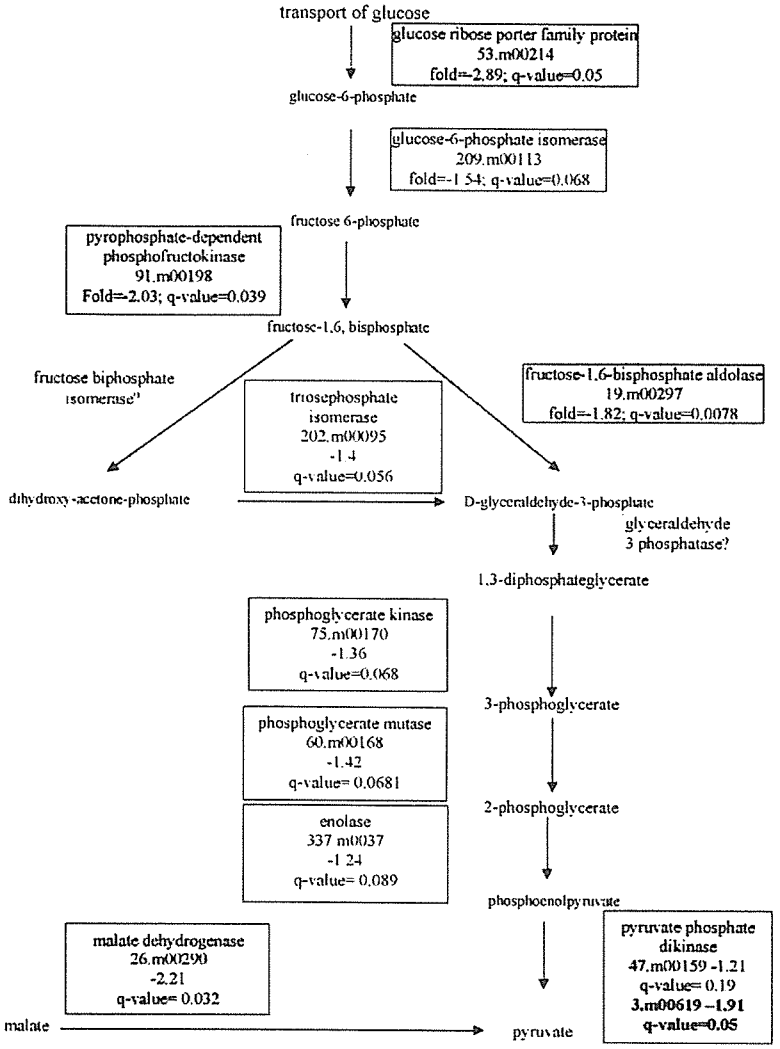


Fig. 5. Changes in expression of the genes encoding enzymes of the glycolytic pathway. Schematic representation of the *E. histolytica* glycolytic pathway. Transcripts significantly modulated in vivo are shown in shaded boxes. Gene name is followed by the locus represented by the microarray probe set, fold change in vivo and significance (q-value).

3.10. Other virulence genes

Of the proteins known to be important in *E. histolytica* virulence few changes in transcript levels were observed. This is perhaps not surprising because all of these factors were identified and characterized in vitro using cultured trophozoites. Among virulence factors that were altered was the Gal/GalNAc lectin, essential for ameba adherence and contact mediated cytolysis of host cells. The Gal/GalNAc lectin light and heavy subunits are each encoded by five member gene families, while the intermediate subunit is encoded by two different genes [5]. There were no statistically significant changes in the heavy and intermediate transcripts; however, two members of the light subunit gene family were statistically down-regulated (*lgl2* two-fold at Day 1 and *lgl3* four-fold at Day 1 and Day 29). In addition to these changes in known virulence factors, it is interesting to speculate if novel virulence factors may be among the hypothetical genes with altered in vivo expression.

3.11. Membrane transport

Nutrient acquisition may be an important aspect of virulence. Many of the genes with putative transport functions were down-regulated, including several members of the major facilitator superfamily, and two members of the ABC transporter family. Exceptions were increases (10-fold at Day 1 and 8-fold at Day 29) of a transcript encoding a potential amino-acid transporter (205.m00085), and the significant reversal of the down-regulation of one of the ABC transporter protein transcripts (41.m00219) (three-fold decrease at Day 1 but 1.3-fold increase at Day 29).

3.12. Endocytosis and vesicular trafficking

E. histolytica genes involved in motility and phagocytosis were also modulated. Rho family GTPases, a SNARE pfam domain-containing genes (141.m00087), Syntaxin A, and a

Table 4
DNA and RNA regulatory transcripts

Probe set	Day 1 in vivo		Day 29 in vivo		Annotation of gene represented by probe set
	Fold difference	SAM q value	Fold difference	SAM q value	
149.m00092_at	6.88	0.01	2.78	0.18	Hypothetical protein 149.t00006
169.m00139_at	4.56	0.01	1.98	0.41	Hypothetical protein 169.t00027
55.m00187_at	3.89	0.01	2.46	0.25	Hypothetical protein 55.t00025
146.m00110_at	2.87	0.04	1.25	0.7	Myb family DNA-binding protein
123.m00114_at	2.54	0	−1.55	0.28	Hypothetical protein 123.t00010
94.m00147_at	2.53	0.03	3.35	0.14	Hypothetical protein 94.t00018
20.m00305_at	2.44	0.49	4.05	0.02	RNA-binding protein, putative
197.m00084_x_at[a]	2.35	0.03	1.29	0.64	Hypothetical protein 197.t00015
524.m00023_at	2.3	0.01	1.38	0.69	Myb family DNA-binding protein
20.m00272_at	2.23	0.06	−6.25	0	Conserved hypothetical protein (DNA binding group)
3.m00577_at	2	0.01	6.55	0.09	High mobility group protein, putative
275.m00123_s_at[b]	1.92	0.04	2.34	0.38	Protein kinase, putative
98.m00143_at	1.86	0.05	−1.37	0.45	Zinc finger protein, putative
5.m00408_s_at[c]	1.82	0.25	−1.43	0.5	Hypothetical protein 5.t00017
4.m00655_at	1.75	0.19	7.36	0	Hypothetical protein 4.t00077
143.m00093_at	−1	0.42	−5.81	0.01	Conserved hypothetical protein (DNA binding group)
227.m00086_at	−1.09	0.42	−16.64	0	Conserved hypothetical protein (DNA binding group)
177.m00126_s_at[d]	−1.16	0.36	−4.09	0.04	Ribonuclease, putative
198.m00105_at	−1.26	0.16	−2.13	0.02	RNA-binding protein, putative
173.m00119_at	−1.37	0.12	−2.62	0.04	Putative CLK family kinase (*Dictyostelium discoideum*)
324.m00040_at[e]	−1.51	0.16	−4.21	0	Transcription initiation factor TFIID, putative
111.m00140_x_at[a]	−1.92	0.04	−11.04	0	Conserved hypothetical protein (DNA binding group)
62.m00167_at	−2.01	0.02	−1.4	0.28	Hypothetical protein 62.t00018
147.m00097_x_at[a,f]	−2.03	0.02	1.17	0.7	Enhancer binding protein-1
490.m00035_at	−2.07	0.03	−1.64	0.32	Hypothetical protein 490.t00001
268.m00065_x_at[a]	−2.19	0.01	−1.12	0.6	Predicted: similar to myb binding protein (P160) 1a-like (*Strongylocentrotus purpuratus*)
91.m00190_s_at[g]	−2.29	0.05	−2.48	0	RNA-binding protein, putative
88.m00155_s_at	−2.47	0.02	−2.08	0.11	Hypothetical protein 88.t00011
38.m00215_at	−2.53	0.04	−1.44	0.41	AAA family ATPase
14.m00334_at	−4.68	0.03	−1.47	0.41	Hypothetical protein 14.t00055

[a] (_at) probe sets represent single genes, s_at. probe sets recognize two or more genes with which the probes are an exact match, _x_at probe sets may cross-hybridize in an unpredictable manner with sequences other than the main target.
[b] 275.m00123 also represents 275.m00093 a truncated version that also contains an internal deletion of 24 nts.
[c] 5.m00408_s also represents 183.m00107 with which it is completely identical.
[d] 177.m00126_s_at also represents 100.m00129 with which it is 98% identical.
[e] This gene has been previously described by Luna-Arias et al. [42].
[f] Regulator of the transcription of the *hgl5* large subunit of the Gal/GalNAc lectin, Schaenman et al. [31].
[g] 91.m00190_s represents also the 204.m00092 ORF to which it is completely identical.

Vam3 homologue and a Ras-related gene (128.m00135) were down-regulated in vivo, while the transcripts of three Ras-related proteins (Ras-related protein 3, and two Rab family GTPases) and a Syntaxin C (yeast Pep13 homologue) (167.m00134) were increased at Day 1 (Table 5 and supplemental data). Although the Rab family GTPases mRNA levels remained high at Day 29 they were not statistically different from in vitro levels. Other genes, including those encoding potential Rho and Ras GTPase activating proteins, were down-regulated, and as previously mentioned, the phosphoinositide 3-kinase transcript was statistically decreased (Table 5 and supplemental data). These data suggest that the parasite alters the regulation of cytoskeletal-mediated processes in response to the host environment.

3.13. Unknown proteins

An intriguing result was the modulation of expression of four groups of hypothetical proteins (Table 6). Group 1 genes were similar to the AIG proteins implicated in plant defense from bacteria [38]. The group 1 transcripts were either un-regulated or up-regulated at Day 1, and with only two exceptions (AIG40.m00239 and AIG451.m00039) they were decreased at Day 29. Group 2 genes contained bacterial surface protein A-like sequences (BspA). Three members of the *E. histolytica* BspA family were modulated in vivo (BspA32 transcript was repressed at Days 1 and 29, BspA615.m00022 was decreased two-fold at Day 1, and BspA707.m00012 gene was up-regulated at Day 29) [33,39]. Group 3 genes encode small (~4–8 kDa) basic and

Table 5
Changes in transcripts potentially involved in intracellular signaling

Probe set	Day 1 in vivo		Day 29 in vivo		Annotation of gene represented by probe set
	Fold difference	SAM q value	Fold difference	SAM q value	
Transmembrane kinases					
A subgroup					
148.m00080_at[a]	4.66	0	2.31	0.5	TMK69
16.m00305_x_at[a,b]	3.01	0.01	1.47	0.68	TMK53
2.m00624_s_at	2.38	0.01	2.56	0.01	Protein kinase, putative
152.m00104_at[a]	−2	0.05	−3	0.02	TMK17
B1 subgroup					
67.m00091_x_at[a,b]	3.89	0.01	2.1	0.53	TMK95
174.m00087_x_at[b]	2.74	0.02	1.84	0.6	Protein kinase, putative
New subgroup					
251.m00091_at	−1.6	0.07	−2.67	0	Receptor protein kinase, putative
B2 subgroup					
15.m00355_at[a]	3.11	0.01	1.46	0.64	TMK105
C subgroup					
20.m00345_at[a]	2.81	0.02	1.67	0.5	TMK63
D1 subgroup					
5.m00482_at[a]	2.61	0.02	1.7	0.67	TMK56
136.m00104_at[a]	−2.76	0	−4.76	0	TMK03
Other kinases					
406.m00055_s_at[c]	2.02	0.03	2.77	0.25	Protein kinase
265.m00084_s_at[d]	1.45	0.51	−2.47	0	Protein kinase
16.m00336_s_at[e]	1.03	0.42	−2.03	0.04	Protein kinase
1.m00709_at	−1.01	0.42	−3.27	0.01	Polynucleotide kinase-3-phosphatase
9.m00406_s_at[f]	−1.08	0.36	−2.25	0	Phosphatidylinositol 3-kinase
77.m00170_at	−1.87	0.01	−2.35	0.02	Protein kinase
8.m00424_at	−2.16	0.04	−3.02	0	Protein kinase
116.m00121_at	−2.47	0.01	−11.01	0	Protein kinase
34.m00274_at	−3.48	0.02	−1.84	0.12	Protein kinase
Phosphatases					
98.m00146_at	2.08	0.04	−1.12	0.57	Ser/Thr protein phosphatase family protein (*Colwellia psychrerythraea* 34H)
92.m00163_x_at[b]	2.01	0.05	1.57	0.47	Conserved hypothetical protein
8.m00389_at	−1.08	0.42	−2.15	0.04	Protein phosphatase
58.m00145_at	−2	0.05	−1.03	0.65	Protein tyrosine phosphatase
171.m00101_at	−2.01	0.03	−1.22	0.53	Amidohydrolase
107.m00111 at	−2.09	0.05	−1.38	0.47	Protein phosphatase
393.m00036_at	−2.12	0.04	−1.24	0.47	Ser Thr protein phosphatase
69.m00153_at	−2.35	0.02	−2.37	0.04	Ser Thr protein phosphatase
289.m00068_at	−2.66	0.02	−2.83	0.05	Amidohydrolase
16.m00311_at	−3.19	0.02	−2.19	0.09	Hypothetical protein
Calcium binding					
3.m00563_at[g]	5.65	0	2.9	0.35	Calcium-binding protein Calcium Binding family
56.m00149_at[h]	4.31	0	4.33	0	Calcium-binding protein
94.m00131_at	4.12	0.01	1.27	0.64	SPRY domain protein
15.m00302_at	2.99	0.05	2.03	0.16	Hypothetical protein 15.t00011
247.m00081_s_at[i]	−2.03	0.01	−1.64	0.09	Grainin 2
32.m00201_s_a[j]	−2.12	0.01	−1.19	0.53	Conserved hypothetical protein
182.m00137_at[k]	−2.21	0.03	−2.73	0.08	Grainin 1
195.m00083_s_at[l]	−2.29	0.04	1.47	0.7	Plasma membrane calcium-transporting ATPase
101.m00121_at[g]	−2.51	0.01	−2.1	0.22	C2 domain protein Calcium Binding family
Small GTPase					
110.m00125_at	5.71	0	N/A	N/A	Hypothetical protein 110.t00013/GTPase-activator protein for Ras-like GTPase
288.m00067_at[m]	4.73	0	4.25	0.53	Ras-related protein 3
30.m00275_at	4.32	0.01	2.25	0.11	Rab family GTPase

Table 5 (*Continued*)

Probe set	Day 1 in vivo		Day 29 in vivo		Annotation of gene represented by probe set
	Fold difference	SAM q value	Fold difference	SAM q value	
71.m00153_at	3.32	0.03	3.23	0.08	Rho family GTPase
8.m00373_at	3.18	0	1.05	0.7	Ras guanine nucleotide exchange factor
106.m00140_at	3.04	0.03	1.49	0.6	Rap/Ran GTPase activating protein
4.m00593_at	2.95	0	2.86	0.28	Rab family GTPase
233.m00108_at	1.05	0.51	2.64	0.05	Alanine aminotransferase, RanBP1 domain
6.m00423_x_at[b]	−1.52	0.07	−4.16	0	Hypothetical protein 6.t00021 GTPase-activator protein for Ras-like GTPase
60.m00155_at	−1.61	0.04	−2.26	0	Hypothetical protein 60.t00020 GTPase-activator protein for Ras-like GTPase
47.m00168_x_at[b]	−1.7	0.07	−2.07	0.02	Ran-binding protein
24.m00315_s_at[n]	−2.07	0.01	−1.59	0.22	Rho GTPase activating protein
128.m00135_at	−2.16	0.03	−1.67	0.16	Ras family GTPase
265.m00063_x_at[b]	−2.24	0.01	−1.62	0.22	Rho guanine nucleotide exchange factor
1.m00710_at	−2.28	0.05	−2.82	0.02	Rho GTPase activating protein
334.m00040_at	−2.29	0.03	−1.26	0.53	Ras guanine nucleotide exchange factor
3.m00556_at	−2.42	0.02	−2.69	0.06	Hypothetical protein 3.t00019 small GTPase mediated signal transduction
131.m00145_at	−2.43	0.01	−1.83	0.05	Ras guanine nucleotide exchange factor
4.m00698_at	−2.57	0.03	−1.55	0.38	Rap/Ran GTPase activating protein
157.m00102_at	−3.2	0.01	−3.86	0	Rho GTPase activating protein
10.m00354_at	−5.05	0.01	−2	0.05	Hypothetical protein 10.t00038 GTPase-activator protein for Ras-like GTPase
151.m00103_at	−6.47	0	−4.41	0.01	Rho guanine nucleotide exchange factor

[a] Gene family previously described by Beck et al. [33].

[b] (_at) probe sets represent single genes, s_at. probe sets recognize two or more genes with which the probes are an exact match, _x_at probe sets may cross-hybridize in an unpredictable manner with sequences other than the main target.

[c] 406.m00055_s_at also represents 169.m00134 with which it is identical.

[d] 265.m00084_s at also represents 265.m00081with which it is identical.

[e] 16.m00336_s_at also represents 72.m00163 as they have an identical internal stretch of 808 nts.

[f] 9.m00406_s_at also represents 67.m00102 with which it is 82% identical.

[g] H. Moreno, personal communication.

[h] EhCaBP has been previously described by Sahoo et al. [36].

[i] 247.m00081_s_at is truncated by 69 nts but otherwise identical to 224.m00108 and 224.m00112 which are identical to each other. This gene has been previously described as grainin 2 by Nickel et al. [37].

[j] 32.m00210 represents 364.m00044 of which it is a truncated (575) 99% identical version.

[k] 182.m00137 was over 85% similar to grainin 1 which was previously described by Nickel et al. [37].

[l] 195.m00083_s_at represents 429.m00044 with which it is identical.

[m] 288.m00067 has been previously described as the Ras-related protein 3 by Kumagai et al. [43].

[n] 24.m00315_s also represents 13.m00346, these ORF's were identical but for an internal deletion of 60 nts.

cysteine rich proteins, which are over 97% identical at the protein level. These were significantly repressed at Day 29. Group 4 genes had some similarity to the TLDc domain of LysM, a protein with peptidoglycan-binding bacteriolytic and adhesive activities. Only two transcripts of the Group 4 TLDc gene family were expressed sufficiently highly to be included in our present analysis, with transcripts at Day 29 significantly increased when contrasted with the Day 1 in vivo levels.

3.14. Summary

This study is the first genome-wide analysis of transcription in *E. histolytica* trophozoites, and the first description of gene expression in trophozoites from the intestine. The most important findings were that >80% of ORFs are transcribed in trophozoites, and that in response to the host intestinal environment alterations occurred in the expression of genes implicated in metabolism, oxygen defense, cell signaling, virulence, antibacterial activity, and DNA-binding.

The inconsistent changes in the laterally-transferred metabolic enzymes suggest that the derivation of these genes did not lead to common regulation in response to the intestinal environment. Interestingly several members of the large gene families of *E. histolytica* were also regulated differently (e.g. the cysteine protease family Table 2) which suggested that, despite their similarity, the encoded proteins have distinct functions important in the adaptation of *E. histolytica* to the host environment.

It should be noted that there are several limitations to the approaches used in this report. Most importantly, changes in the level of mRNA abundance are only one level of control of gene expression. Ultimately the understanding of the parasite adaptation to the host will require deciphering the not

Table 6
Subgroups of unknown genes significantly changed between Days 1 and 29

Probe set	Day one in vivo		Day 29 in vivo		Annotation of gene represented by probe set
	Fold difference	SAM q value	Fold difference	SAM q value	
Group 1					
374.m00023_at	4.52	0.01	−1.84	0.2	AIG1 family protein, putative
451.m00039_at	3.62	0.05	3	0.53	AIG1 family protein, putative
374.m00021_at	2.79	0.01	−1.68	0.25	AIG1 family protein, putative
499.m00017_x_at[a]	2.43	0.02	−1.4	0.32	AIG1 family protein, putative
499.m00016_x_at[a]	2.1	0.03	−2.46	0.2	AIG1 family protein, putative
432.m00030_at	1.67	0.39	−34.8	0	AIG1 family protein, putative
432.m00029_at	1.6	0.16	−11.27	0	AIG1 family protein, putative
40.m00239_at	1.47	0.45	5.08	0.04	AIG1 family protein, putative
628.m00011_at	1.28	0.42	−3.15	0	AIG1 family protein, putative
477.m00021_at	1.08	0.42	−5.72	0.01	AIG1 family protein, putative
374.m00024_x_at[a]	1.07	0.51	−4.1	0	AIG1 family protein, putative
71.m00149_at	−1.86	0.09	−24.56	0	AIG1 family protein, putative
BspA-like[b]					
707.m00012_x_at[a]	2.08	0.16	7.08	0	BspA-like leucine rich repeat protein, putative
615.m00022_x_at[a]	−2.76	0.05	1.21	0.7	BspA-like leucine rich repeat protein, putative
7.m00417_a[c]	−2.84	0.02	−4.08	0.01	BspA-like leucine rich repeat protein, BspA32
Group 3					
496.m00027_x_at[a]	2.59	0.06	−3.72	0	Hypothetical protein 496.t00004
493.m00030_x_at[a]	2.57	0.12	−3.93	0	Hypothetical protein 493.t00001
586.m00015_s_at[d]	2.19	0.25	−4.06	0	Hypothetical protein
Group 4					
65.m00145_x_at[a]	2.25	0.02	14.75	0	Hypothetical protein 65.t00007
395.m00029_x_at[a]	−1.05	0.51	3.15	0.28	Conserved hypothetical protein

[a] (_at) probe sets represent single genes, s_at probe sets recognize two or more genes with which the probes are an exact match, _x_at probe sets may cross-hybridize in an unpredictable manner with sequences other than the main target.
[b] Gene family previously described by Davis et al. [39].
[c] Previously described by Beck et al. [33].
[d] 586.m00015_s represents also the 496.m00027 ORF which differs by two nucleotides.

only the transcriptome, but also the *E. histolytica* proteome and metabolome. The analysis of the transcriptome is in turn limited by the fact that it was not possible to independently represent every ORF on the array, because of the presence of families of genes with high sequence identity. In addition, the gene array hybridizations necessitated amplification of the amebic RNA, a step that introduces potential artifacts from non-linear amplification. However, the amplification used a technique that limits non-linear amplification. In addition, the array data was independently verified for 11/12 genes by quantitative real-time reverse transcriptase PCR from unamplified mRNA, suggesting that unequal amplification of mRNA was not a systematic problem with this analysis.

The description of the trophozoite transcriptome establishes the foundation for future interventional studies to delineate the contributions of the altered gene products to intestinal colonization and invasion. The long-term outcome of these studies promises to provide a molecular understanding of the parasite factors controlling the formation of amebic colitis, the most common and feared manifestation of amebiasis.

Acknowledgments

This work was supported by NIH Grant AI 39741 and Biotechnology Engagement Program Grant 17. The bioinformatics data analysis at VBI was funded by Department of Defense grant #DAAD 13-02-C-0018. We thank Dr P. Hoffman for advice and helpful discussion.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molbiopara.2006.02.007.

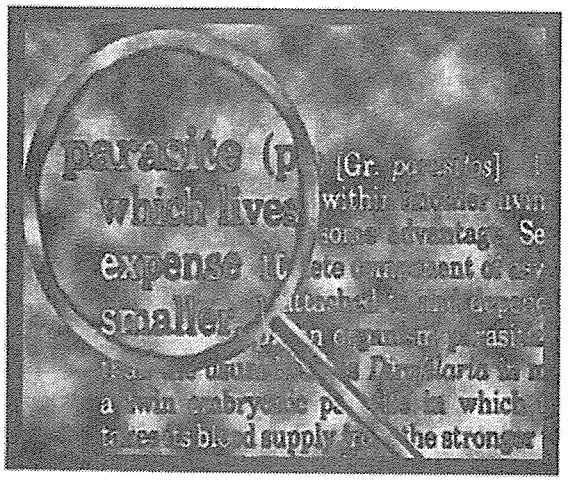
References

[1] Clark CG, Roger AJ. Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. Proc Natl Acad Sci U S A 1995;92:6518–21.
[2] Field J, Rosenthal B, Samuelson J. Early lateral transfer of genes encoding malic enzyme, acetyl-CoA synthetase and alcohol dehydrogenases from anaerobic prokaryotes to *Entamoeba histolytica*. Mol Microbiol 2000;38:446–55.
[3] Loftus B, Anderson I, Davies R, Alsmark UC, Samuelson J, et al. The genome of the protist parasite *Entamoeba histolytica*. Nature 2005;433:865–8.
[4] Stanley Jr SL. Amoebiasis. Lancet 2003;361:1025–34.
[5] Petri Jr WA, Haque R, Mann BJ. The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite *Entamoeba histolytica*. Annu Rev Microbiol 2002;56:39–64.
[6] Diamond LS, Harlow DR, Cunnick C. A new medium for axenic cultivation of *Entamoeba histolytica* and other Entamoeba. Trans R Soc Trop Med Hyg 1978;72:431–2.

[7] Houpt ER, Glembocki DJ, Obrig TG, Moskaluk CA, Lockhart LA, et al. The mouse model of amebic colitis reveals mouse strain susceptibility to infection and exacerbation of disease by CD4+ T cells. J Immunol 2002;169:4496–503.
[8] Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 2001;29:365–71.
[9] Wu Z, Irizarry R, Gentleman R, Murillo F, Spencer F. A model based background adjustment for oligonucleotide expression arrays. Technical report, 2003, John Hopkins University, Department of Biostatistics Working Papers, Baltimore, MD, 2004.
[10] Ginzinger DG. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. Exp Hematol 2002;30:503–12.
[11] Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;3 [RESEARCH0034].
[12] Bruchhaus I, Loftus BJ, Hall N, Tannich E. The intestinal protozoan parasite *Entamoeba histolytica* contains 20 cysteine protease genes, of which only a small subset is expressed during in vitro cultivation. Eukaryot Cell 2003;2:501–9.
[13] Tokoro M, Asai T, Kobayashi S, Takeuchi T, Nozaki T. Identification and characterization of two isoenzymes of methionine gamma-lyase from *Entamoeba histolytica*: a key enzyme of sulfur-amino acid degradation in an anaerobic parasitic protist that lacks forward and reverse transsulfuration pathways. J Biol Chem 2003;278:42717–27.
[14] Nozaki T, Ali V, Tokoro M. Sulfur-containing amino acid metabolism in parasitic protozoa. In: Baker J, Muller R, Rollinson D, editors. Advances in Parasitology. Academic Press; 2005. p. 1–100.
[15] Nixon JE, Field J, McArthur AG, Sogin ML, Yarlett N, et al. Iron-dependent hydrogenases of *Entamoeba histolytica* and *Giardia lamblia*: activity of the recombinant entamoebic enzyme and evidence for lateral gene transfer. Biol Bull 2003;204:1–9.
[16] Rosenthal B, Mai Z, Caplivski D, Ghosh S, de la Vega H, et al. Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica*. J Bacteriol 1997;179:3736–45.
[17] Bruchhaus I, Roeder T, Lotter H, Schwerdtfeger M, Tannich E. Differential gene expression in *Entamoeba histolytica* isolated from amoebic liver abscess. Mol Microbiol 2002;44:1063–72.
[18] Ali V, Shigeta Y, Tokumoto U, Takahashi Y, Nozaki T. An intestinal parasitic protist, *Entamoeba histolytica*, possesses a non-redundant nitrogen fixation-like system for iron–sulfur cluster assembly under anaerobic conditions. J Biol Chem 2004;279:16863–74.
[19] Filenko NA, Browning DF, Cole JA. Transcriptional regulation of a hybrid cluster (prismane) protein. Biochem Soc Trans 2005;33:195–7.
[20] van den Berg WA, Hagen WR, van Dongen WM. The hybrid-cluster protein ('prismane protein') from *Escherichia coli*. Characterization of the hybrid-cluster protein, redox properties of the [2Fe–2S] and [4Fe–2S–2O] clusters and identification of an associated NADH oxidoreductase containing FAD and [2Fe–2S]. Eur J Biochem 2000;267:666–76.
[21] Akbar MA, Chatterjee NS, Sen P, Debnath A, Pal A, et al. Genes induced by a high-oxygen environment in *Entamoeba histolytica*. Mol Biochem Parasitol 2004;133:187–96.
[22] Samarawickrema NA, Brown DM, Upcroft JA, Thammapalerd N, Upcroft P. Involvement of superoxide dismutase and pyruvate:ferredoxin oxidoreductase in mechanisms of metronidazole resistance in *Entamoeba histolytica*. J Antimicrob Chemother 1997;40:833–40.
[23] Tekwani BL, Mehlotra RK. Molecular basis of defence against oxidative stress in *Entamoeba histolytica* and *Giardia lamblia*. Microbes Infect 1999;1:385–94.
[24] Bruchhaus I, Tannich E. Induction of the iron-containing superoxide dismutase in *Entamoeba histolytica* by a superoxide anion-generating system or by iron chelation. Mol Biochem Parasitol 1994;67:281–8.
[25] Saavedra E, Encalada R, Pineda E, Jasso-Chavez R, Moreno-Sanchez R. Glycolysis in *Entamoeba histolytica*. Biochemical characterization of recombinant glycolytic enzymes and flux control analysis. FEBS J 2005;272:1767–83.
[26] Deng Z, Huang M, Singh K, Albach RA, Latshaw SP, et al. Cloning and expression of the gene for the active PPi-dependent phosphofructokinase of *Entamoeba histolytica*. Biochem J 1998;329:659–64.
[27] Chi AS, Deng Z, Albach RA, Kemp RG. The two phosphofructokinase gene products of *Entamoeba histolytica*. J Biol Chem 2001;276:19974–81.
[28] Wolfe AJ. The acetate switch. Microbiol Mol Biol Rev 2005;69: 12–50.
[29] Zuo X, Coombs GH. Amino acid consumption by the parasitic, amoeboid protists *Entamoeba histolytica* and *E. invadens*. FEMS Microbiol Lett 1995;130:253–8.
[30] Anderson IJ, Loftus BJ. *Entamoeba histolytica*: observations on metabolism based on the genome sequence. Exp Parasitol 2005;110: 173–7.
[31] Schaenman JM, Gilchrist CA, Mann BJ, Petri Jr WA. Identification of two *Entamoeba histolytica* sequence-specific URE4 enhancer-binding proteins with homology to the RNA-binding motif RRM. J Biol Chem 2001;276:1602–9.
[32] Sapra AK, Arava Y, Khandelia P, Vijayraghavan U. Genome-wide analysis of pre-mRNA splicing: intron features govern the requirement for the second-step factor. Prp17 in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. J Biol Chem 2004;279:52437–46.
[33] Beck DL, Boettner DR, Dragulev B, Ready K, Nozaki T, et al. Identification and gene expression analysis of a large family of transmembrane kinases related to the Gal/GalNAc lectin in *Entamoeba histolytica*. Eukaryot Cell 2005;4:722–32.
[34] Ghosh SK, Samuelson J. Involvement of p21racA, phosphoinositide 3-kinase, and vacuolar ATPase in phagocytosis of bacteria and erythrocytes by *Entamoeba histolytica*: suggestive evidence for coincidental evolution of amebic invasiveness. Infect Immun 1997;65:4243–9.
[35] Meza I, Clarke M. Dynamics of endocytic traffic of *Entamoeba histolytica* revealed by confocal microscopy and flow cytometry. Cell Motil Cytoskeleton 2004;59:215–26.
[36] Sahoo N, Labruyere E, Bhattacharya S, Sen P, Guillen N, et al. Calcium binding protein 1 of the protozoan parasite *Entamoeba histolytica* interacts with actin and is involved in cytoskeleton dynamics. J Cell Sci 2004;117:3625–34.
[37] Nickel R, Jacobs T, Urban B, Scholze H, Bruhn H, et al. Two novel calcium-binding proteins from cytoplasmic granules of the protozoan parasite *Entamoeba histolytica*. FEBS Lett 2000;486:112–6.
[38] Reuber TL, Ausubel FM. Isolation of Arabidopsis genes that differentiate between resistance responses mediated by the RPS2 and RPM1 disease resistance genes. Plant Cell 1996;8:241–9.
[39] Davis PH, Zhang Z, Chen M, Zhang X, Chakraborty S, et al. Identification of a family of BspA like surface proteins of *Entamoeba histolytica* with novel leucine rich repeats. Mol Biochem Parasitol 2005.
[40] Nozaki T, Asai T, Kobayashi S, Ikegami F, Noji M, et al. Molecular cloning and characterization of the genes encoding two isoforms of cysteine synthase in the enteric protozoan parasite *Entamoeba histolytica*. Mol Biochem Parasitol 1998;97:33–44.
[41] Garcia-Rivera G, et al. *Entamoeba histolytica*: a novel cysteine protease and an adhesin form the 112 kDa surface protein. Mol Microbiol 1999;33(3):556–68.
[42] Luna-Arias JP, et al. The TATA-box binding protein of *Entamoeba histolytica*: cloning of the gene and location of the protein by immunofluorescence and confocal microscopy. Microbiology 1999;145(Part 1):33–40.
[43] Kumagai M, et al. Molecular cloning and characterization of a protein farnesyltransferase from the enteric protozoan parasite *Entamoeba histolytica*. J Biol Chem 2004;279(3):2316–23.

Volume 55, Issue 4 December 2006 ISSN 1383-5769

PARASITOLOGY INTERNATIONAL



www.elsevier.com/locate/parint

Available online at www.sciencedirect.com
ScienceDirect
Parasitology International 55 (2006) 307–311
www.elsevier.com/locate/parint

Short communication

Genetic diversity of glucose phosphate isomerase from *Entamoeba histolytica*☆

Elham Razmjou [a], Ali Haghighi [b], Mostafa Rezaian [a], Seiki Kobayashi [c], Tomoyoshi Nozaki [d,*]

[a] *Department of Medical Parasitology and Mycology, School of Public Health and Institute of Public Health Research, Tehran University of Medical Sciences, Iran*
[b] *Department of Parasitology and Mycology, Shaheed Beheshti University of Medical Sciences, Iran*
[c] *Department of Tropical Medicine and Parasitology, School of Medicine, Keio University, Japan*
[d] *Department of Parasitology, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan*

Received 28 June 2006; received in revised form 30 July 2006; accepted 3 August 2006
Available online 18 September 2006

Abstract

To investigate the molecular basis of zymodeme analysis in the enteric protozoan parasite *Entamoeba histolytica*, genes encoding glucose phosphate isomerase (GPI) were isolated from four representative *E. histolytica* strains belonging to zymodeme II, IIα-, XIV, or XIX. Two alleles were obtained from each strain; six alleles with eight polymorphic nucleotide positions were identified among the four strains. Two of these eight polymorphic nucleotides resulted in non-conserved amino acid substitutions. Three GPI isoenzymes with distinct predicted isoelectric points were identified, which agrees well with the observed electrophoretic patterns of GPI from these strains. Amino acid comparisons of GPI from *E. histolytica* and other organisms revealed that all amino acid residues implicated for substrate binding and catalysis were conserved. Biochemical characterization of recombinant *E. histolytica* GPI confirmed that it possessed kinetic parameters similar to GPI from other organisms. The electrophoretic mobility of three GPI isoenzymes was examined by starch gel electrophoresis. Thus, we have established the molecular basis of the classical isoenzymes patterns that have been used for grouping *E. histolytica* isolates and for differentiation of *E. histolytica* from non-pathogenic *Entamoeba dispar*.

Keywords: Glucose phosphate isomerase; *Entamoeba histolytica*; Zymodemes

The protozoan parasite *Entamoeba histolytica* is an etiological agent of amoebiasis, causing an estimated 50 million cases of amebic colitis, dysentery, and extraintestinal abscesses [1] and 40,000–100,000 deaths annually [2]. *E. histolytica*, together with *Giardia intestinalis*, is categorized as a type I amitochondriate protist, which lacks both typical mitochondria and hydrogenosomes [3] and is also deficient in all features of aerobic metabolism. *E. histolytica* produces energy by glycolysis and fermentation under an anaerobic or microaerophilic environment [3,4]. In *E. histolytica*, all enzymes involved in the conversion of glucose into pyruvate by glycolysis have been identified [5,6]. Molecular identification and enzymological characterization of 10 enzymes including hexokinase, inorganic pyrophosphate-dependent phosphofructokinase, and phosphoglucomutase were previously accomplished [7–11]. A recent study by Saarvedra et al. indicated that phosphoglycerate mutase, fructose-1,6-phosphate aldolase, glyceraldehyde-3-phosphate dehydrogenase, and pyruvate phosphate dikinase are flux control steps because they possess the lowest catalytic efficiencies [11]. Four of about a dozen enzymes in the glycolytic pathway, namely hexokinase, phosphoglucomutase, glucose-6-phosphate isomerase (GPI), and malic enzyme, have been used to differentiate *E. histolytica* strains as well as pathogenic *E. histolytica* from non-pathogenic *Entamoeba dispar* species. Although GPI, which is involved in the reversible conversion of glucose-6-phosphate to fructose-6-phosphate, has been used as a gold standard to differentiate the four major groups of *E. histolytica* strains (zymodemes II, IIα-, XIV, and XIX) [12], the molecular basis of this

Abbreviations: GPI, glucose phosphate isomerase; *EhGPI*, *Entamoeba histolytica GPI* gene; EhGPI, *Entamoeba histolytica* glucose phosphate isomerase; ORF, open reading frame; PCR, polymerase chain reaction; rEhGPI, recombinant *E. histolytica* GPI.

☆ Nucleotide sequence data reported in this paper are available in the EMBL, GenBank, and DDJB databases under the accession numbers BAE53444–BAE53451.

* Corresponding author. Tel.: +81 27 220 8020; fax: +81 27 220 8025.
E-mail address: nozaki@med.gunma-u.ac.jp (T. Nozaki).

1383-5769/$ - see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.parint.2006.08.001

differentiation has not been established. In this study, we cloned and characterized GPI from representative isolates that belong to four distinct zymodemes (HM1:IMSS cl6, SAW1627, SAW755CR clB, and KU2 for zymodemes II, IIα-, XIV, and XIX, respectively).

A putative *EhGPI* gene of the HM1:IMSS cl6 reference strain was identified in the *E. histolytica* genome database (http://www.tigr.org/tdb/). It contains an open reading frame (ORF) of 1641 nucleotides encoding a protein of 546 amino acids with a predicted molecular mass of 61.4 kDa and pI of 6.91. To examine polymorphisms of GPI sequences in *E. histolytica* strains, the *EhGPI* ORF was amplified by PCR using genomic DNA from the four above-mentioned *E. histolytica* strains with sense (5′-CCTGGATCCG*ATG*TTACCAACTCTTCCTGA-3′) and antisense (5′-CCAGGATCC*TTA*GTTTTTTCTCATATC TTTAACA-3′) primers designed based on the sequence information (engineered *Bam*HI sites are underlined, and the start and stop codons are italicized). Conditions and parameters for PCR were previously described [13]. PCR products were digested with *Bam*HI and cloned into *Bam*HI-digested pET-15b (Novagen) in the same orientation as the T7 promoter to produce a GPI fusion protein with a histidine tag at the amino terminus. Six plasmid clones were randomly chosen for each *E. histolytica* strain, and their inserts were sequenced on both strands. Nucleotide sequences of four GPI clones derived from HM1:IMSS cl6 were identical to the sequence in the database (XP 650595 and AAT92031) and the sequence recently reported [11] and were designated "allele 1". The remaining two clones contained four nucleotide substitutions and were designated "allele 2" (Fig. 1A). Similarly, two different alleles of GPI genes were identified from each of the three representative strains of other zymodemes (IIα-, XIV, and XIX) (Fig. 1A). These allelic gene sequences (BAE53444–BAE53451) revealed heterogeneity at eight nucleotide positions. Altogether, six distinct alleles, designated "alleles 1–6" (Fig. 1A), were identified in the four isolates. Two of these nucleotide polymorphisms resulted in non-conserved amino acid substitutions at positions 32 and 449, respectively, and distinct predicted isoelectric points. Isoelectric points of three allelic GPIs (designated isotypes 1–3) varied (isotype 1, 6.91; 2, 7.15; 3, 6.73) (Fig. 1A). However, the total number of amino acids and the predicted molecular mass of all GPIs obtained from these four strains were identical. The predicted pIs of the two allelic gene products in these strains agree well with the observed mobility of GPI isoenzymes in conventional starch gel electrophoresis, assuming that *E. histolytica* GPI forms homo- and heterodimers as previously shown in other organisms (e.g., [14]).

To further examine features of *E. histolytica* GPIs (EhGPIs), they were compared with homologs from archaea, bacteria, fungi, protists, plants, and metazoa, revealing 30–64% identity to EhGPI (representative species are shown in Fig. 1B). EhGPIs showed 44–60% identity to GPIs from *Escherichia coli*, *Aspergillus oryzae*, *Homo sapiens*, and *Spinacia oleracea*. EhGPIs showed 41–54% identity to GPIs from other parasitic protists including *Leishmania mexicana*, *Trypanosoma brucei*,

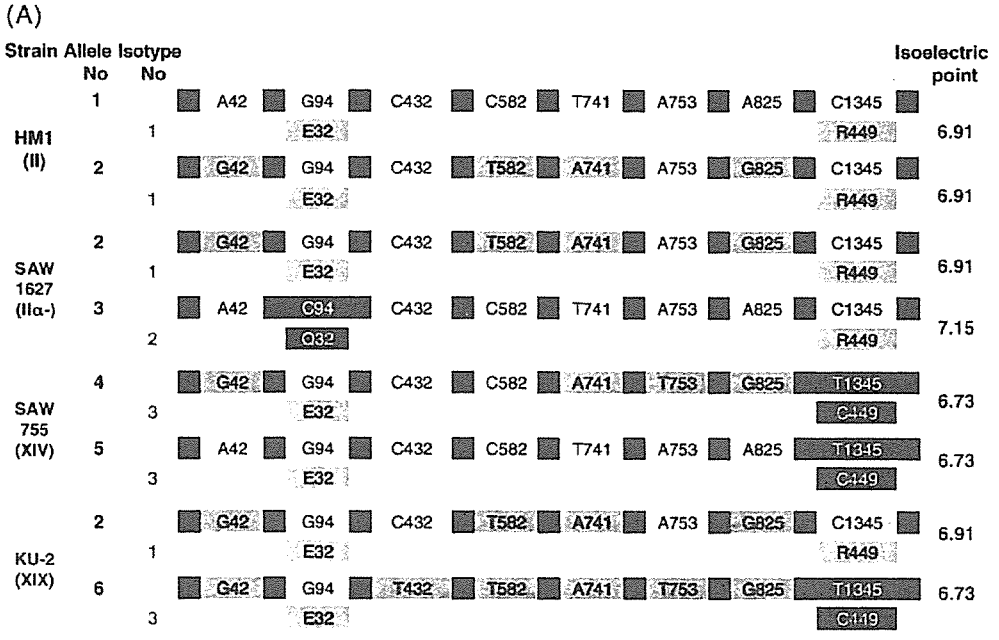


Fig. 1. Comparison of GPI nucleotide and protein sequences. (A) Schematic representation of nucleotide (top) and amino acid (bottom) sequences of two GPI alleles in four representative *E. histolytica* strains. Only nucleotides and amino acids showing heterogeneity are indicated, with predicted pI values. (B) Multiple alignments of deduced amino acid sequences of GPIs from *E. histolytica* and other organisms. Protein primary structures were aligned using the CLUSTAL W program version 1.83 (http://www.ddbj.nig.ac.jp/search/clustalw-e.html) with the BLOSUM matrix. Sequences are *Entamoeba histolytica* (Eh), *Homo sapiens* (Hs), *Escherichia coli* (Ec), *Trypanosoma cruzi* (Tc), *T. brucei* (Tb), *Leishmania mexicana* (Lm), *Aspergillus oryzae* (Ao), *Plasmodium falciparum* (Pf), *Toxoplasma gondii* (Tg), *Spinacia oleracea* (So), *Giardia intestinalis* (Gi), and *Trichomonas vaginalis* (Tv). Asterisks indicate identical amino acids. Dots and colons indicate conserved amino acids substitutions. Dashes indicate computer-generated gaps. Residues deduced from crystal structures (16–18) to be important for substrate binding and/or catalysis of GPIs are shaded in black, and the two signature patterns of the GPI family, [DENS]-X-[LIVM]-G-G-R-[FY]-S-[LIVMT]-X-[STA]-[PSAC]-[LIVMA]-G and [GS]-X-[LIVM]-[LIVMFYW]-XXXX-[FY]-[DN]-Q-X-G-V-E-X-X-K, are highlighted by gray boxes.

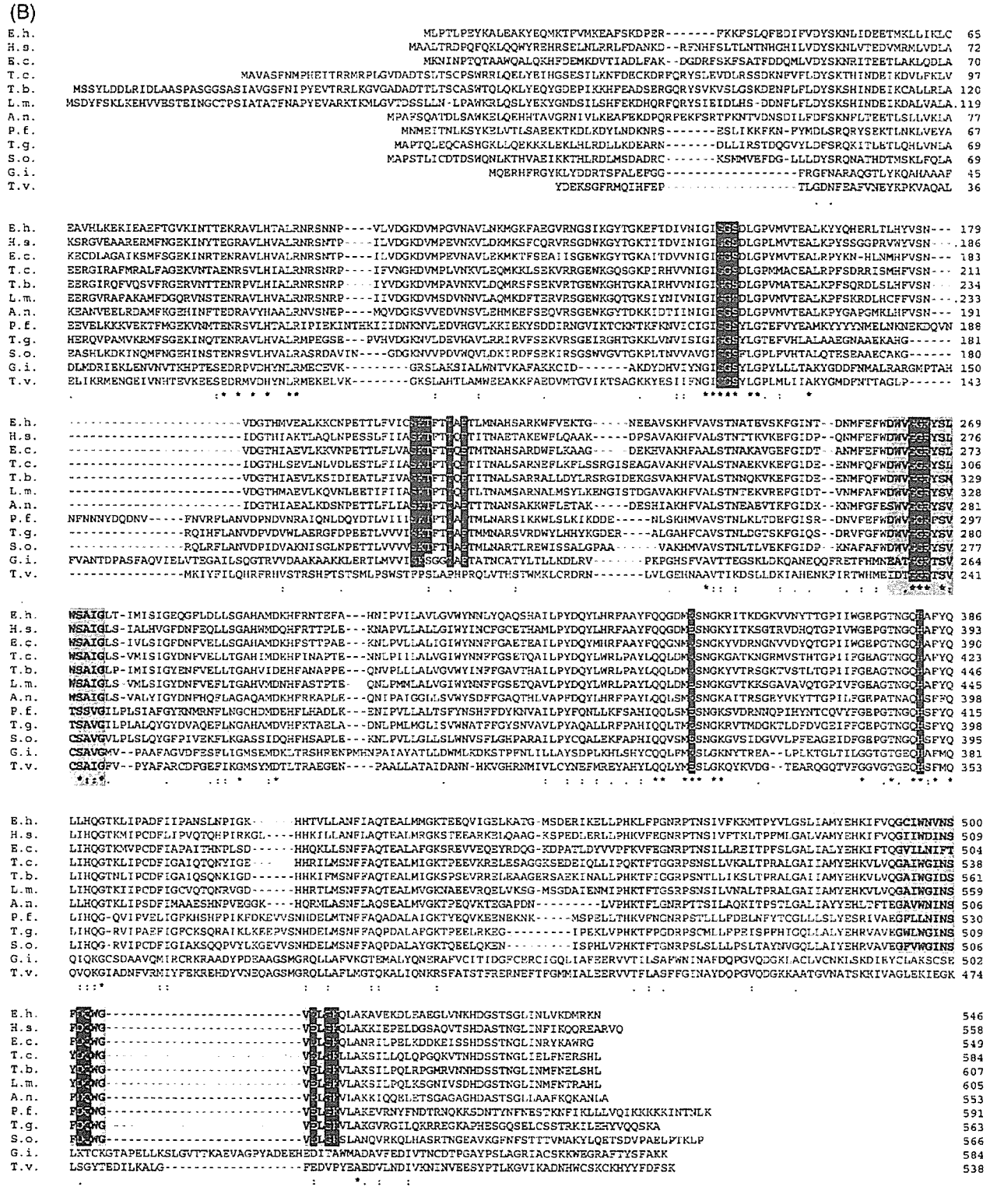


Fig. 1 (*continued*).

Trypanosoma cruzi, *Toxoplasma gondii*, and *Plasmodium falciparum*, while their similarity to representative amitochondriate and microaerophilic/anaerobic protists *Trichomonas vaginalis* and *G. intestinalis* was lower (31–32% identity). This finding suggests distinct GPI ancestors for these amitochondriate parasites, which was also supported by the

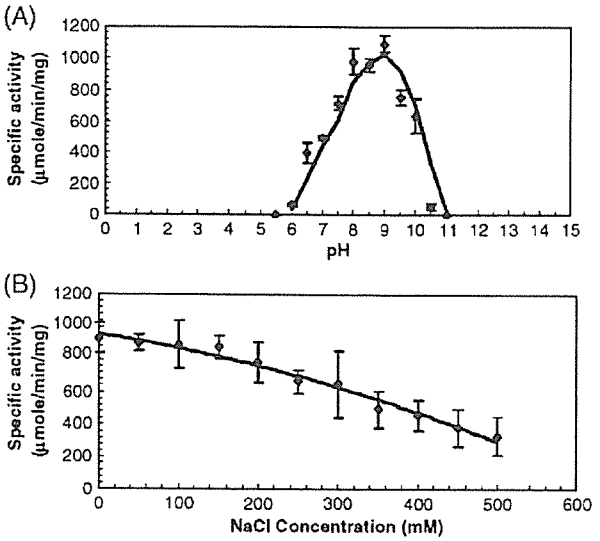


Fig. 2. (A) pH dependence of *E. histolytica* GPI. The pH optimum of the enzyme was determined in the direction of glucose-6-phosphate formation from fructose-6-phosphate between pH 5.5 and 11 at 25 °C in a coupling assay. The 100-μl assay mixture was comprised of 100 mM either 2-Morpholinoethanesulfonic acid (pH 5.5, 6.0, and 6.5), *N*-(2-hydroxyethyl)piperazine-*N'*-(2-ethanesulfonic acid) (pH 7.0, 7.5, and 8.0), *N*-[Tris(hydroxymethyl)methyl]-3-aminopropanesulfonic acid (pH 8.5, 9.0, and 9.5), or 3-(cyclohexylamino)-1-propanesulfonic acid (pH 10, 10.5, and 11); 2 mM fructose-6-phosphate; 0.5 mM NADP; 0.1 U glucose-6-phosphate dehydrogenase; and 0.1 μg purified EhGPI protein. Error bars represent standard errors from three independent experiments. (B) The effect of NaCl on EhGPI activity. The activity of EhGPI was determined in a 100-μl assay mixture comprised of 100 mM TAPS, pH 9.0; 2 mM fructose-6-phosphate; 0.5 mM NADP; 0.1 U glucose-6-phosphate dehydrogenase; and 0.1 μg purified EhGPI protein, with the addition of 0–500 mM NaCl. Error bars represent standard errors from five independent experiments.

previous phylogenetic study [15]. Two conserved signature patterns of GPI [16,17] are totally conserved in EhGPIs (Fig. 1B). In addition, all conserved residues that were shown by various crystal structures or mutagenesis studies [18–20] to be crucial for substrate recognition, binding, or catalysis are highly conserved in EhGPI (Fig. 1B). For instance, the following amino acid residues are completely conserved: Thr205, Thr208, Ser153, Ser203, and Lys204 (numbering based on EhGPI), crucial for binding to the phosphate group; His382, Lys510, Glu210, Ile150, Thr378, Glu507, Arg266, Asp502, Gln503, and Glu351, essential for catalysis; and Gly151, Gly152, Gly264, and Gly265, involved in substrate specificity. EhGPIs lack the long amino-terminal extension present in GPIs from *L. mexicana*, *T. cruzi*, and *T. brucei* that is involved in targeting the enzyme to glycosomes [21]. This is consistent with the premise that EhGPI is cytosolic. EhGPI also lacks an internal insertion (between amino acids 227 and 228) present in GPIs from *L. mexicana*, *T. cruzi*, and *T. brucei*, a 27-amino-acid amino-terminal extension in *G. intestinalis* GPI, and a long carboxyl-terminal extension present in GPI from *T. vaginalis* and *G. intestinalis* [22].

To investigate biochemical features of EhGPIs, the recombinant EhGPIs corresponding to three representative proteins [EhGPI isotypes 1–3, which represent allele 1 of HM1:IMSS cl6 (BAE53444), allele 3 of SAW1627 (BAE53447), and allele 4 of SAW755CR clB (BAE53448)] were produced using a prokaryotic expression system and purified as previously described [23]. Upon SDS-polyacrylamide gel electrophoresis, the recombinant EhGPI (rEhGPI) proteins appeared as an apparently homogeneous single band of approximately 63 kDa, which agrees with the predicted size of a monomeric EhGPI protein with an additional stretch of 21 amino acids at the amino terminus. The formation of glucose-6-phosphate from fructose-6-phosphate in the reverse reaction catalyzed by rEhGPI was measured by monitoring the reduction of $NADP^+$ spectrophotometrically at 340 nm [24]. rEhGPI isotype 1 was active over a wide pH range (7.0–10.0) with optimal pH of 8.0–9.0 (Fig. 2A), consistent with the previous study [11]. GPI activity was inhibited by addition of a monovalent cation (e.g., Na^+) (Fig. 2B), similar to GPI from *Aspergillus niger* [25]. rEhGPI followed Michaelis–Menten kinetics. Kinetic constants under the standard assay conditions (containing 25–300 μM substrates; see Fig. 2 legend for details) were determined with Lineweaver-Burk plots. rEhGPI showed a K_m value of 122±16 μM for fructose-6-phosphate and a specific activity of 786±59 μmol min^{-1} mg $protein^{-1}$ in the reverse reaction, which are of the same order of magnitude as those determined in the previous study (V_{max}, 620 μmol min^{-1} mg $protein^{-1}$; K_m, 480 μM; measured at pH 8 and 37 °C) [11]. The K_m and specific activity of rEhGPI were comparable to GPI from other organisms, e.g., *A. niger* [25], *L. mexicana* [21,26], *T. cruzi* [27], *T. brucei* [26,28], and rabbit [28].

To establish the relationship between individual GPI isotypes identified in this study and zymodemes revealed by starch gel electrophoresis in four representative *E. histolytica* strains, we examined both native and recombinant GPI isoenzymes using starch gel electrophoresis. The electrophoretic mobilities of native and recombinant GPI isoenzymes were similar despite marginal differences in molecular mass and isoelectric point due to the amino-terminal addition of the histidine tag (plus 2.4 kDa; 6.91 vs. 6.88, 7.15 vs. 7.04, and 6.73 vs. 6.74 for EhGPI isotypes 1–3, respectively) (data not shown), confirming that isoenzyme profiles displayed by starch gel electrophoresis reflect two allelic GPI isotypes. Although zymodeme XIX showed the three expected bands on the starch gel, the mixture of recombinant GPI isotypes 1 and 3 gave rise to only two bands corresponding to two homodimers (isotypes 1/1 and 3/3) (data not shown). This is consistent with the hypothesis that a heterodimer (composed of isotypes 1 and 3) is not efficiently formed unless two GPI isotypes are simultaneously synthesized in vivo. These results may also indicate that the homodimers are structurally stable.

These data demonstrate for the first time the molecular basis of classical isoenzyme patterns that have been used for the characterization of *E. histolytica* isolates by zymodeme analyses, considered the "gold standard" to differentiate *E. histolytica* strains as well as pathogenic from non-pathogenic species [12,29]. Our results, together with the allelic heterogeneity of hexokinase and phosphoglucomutase [7–10], indicate that isoenzyme profiles are primarily, if not solely, attributable to sequence diversities at the primary sequence level. Our data further enable the establishment of PCR-based zymodeme analysis that

can replace laborious and time-consuming zymodeme analysis using starch gel electrophoresis.

Acknowledgments

The authors express their appreciation to Vahab Ali, Gunma University; Yasuo Shigeta and Yumiko Saito-Nakano, National Institute of Infectious Diseases, Japan; Bahram Kazemi, Shaheed Beheshti University of Medical Sciences; and Bijan Farzami, Tehran University of Medical Sciences for technical assistance and helpful discussions. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan to T.N.; a grant for Research on Emerging and Re-emerging Infectious Diseases from the Ministry of Health, Labor, and Welfare; a grant for the Project to Promote the Development of Anti-AIDS Pharmaceuticals from the Japan Health Sciences Foundation to T.N.; a fellowship from Japan Society for the Promotion of Science to A.H.; a scholarship from The Ministry of Health, Treatment, and Medical Education of Iran to E.R.; and financial support from Tehran University to M.R.

References

[1] Gilchrist CA, Petri WA. Virulence factors of *Entamoeba histolytica*. Curr Opin Microbiol 1999;2:433–7.

[2] WHO/UNESCO Report. A consultation with experts on amoebiasis. Epidemiol Bull 1997;18:13–4.

[3] Müller M. Enzymes and compartmentation of core energy metabolism of anaerobic protists—a special case in eukaryotic evolution. In: Coombs GH, Vickerman K, Sleigh MA, Warren A, editors. Evolutionary Relationships among Protozoa. The Netherlands: Kluwer Academic Publishers; 1998. p. 109–32.

[4] Reeves RE. Metabolism of *Entamoeba histolytica* Schaudinn. Adv Parasitol 1903;23:105–42.

[5] Loftus B, Anderson I, Davies R, Alsmark UC, Samuelson J, Amedeo P, et al. The genome of the protist parasite *Entamoeba histolytica*. Nature 2005;433:865–8.

[6] Anderson IJ, Loftus BJ. *Entamoeba histolytica*: observations on metabolism based on the genome sequence. Exp Parasitol 2005;110:173–7.

[7] Ortner S, Plaimauer B, Binder M, Scheiner O, Wiedermann G, Duchene M. Molecular analysis of two hexokinase isoenzymes from *Entamoeba histolytica*. Mol Biochem Parasitol 1995;73:189–98.

[8] Ortner S, Clark CG, Binder M, Scheiner O, Wiedermann G, Duchene M. Molecular biology of the hexokinase isoenzyme pattern that distinguishes pathogenic *Entamoeba histolytica* from nonpathogenic *Entamoeba dispar*. Mol Biochem Parasitol 1997;86:85–94.

[9] Ortner S, Binder M, Scheiner O, Wiedermann G, Duchene M. Molecular and biochemical characterization of phosphoglucomutases from *Entamoeba histolytica* and *Entamoeba dispar*. Mol Biochem Parasitol 1997;90:121–9.

[10] Kroschewski H, Ortner S, Steipe B, Scheiner O, Wiedermann G, Duchene M. Differences in substrate specificity and kinetic properties of the recombinant hexokinases HXK1 and HXK2 from *Entamoeba histolytica*. Mol Biochem Parasitol 2000;105:71–80.

[11] Saavedra E, Encalada R, Pineda E, Jasso-Chavez R, Moreno-Sanchez R. Glycolysis in *Entamoeba histolytica*. Biochemical characterization of recombinant glycolytic enzymes and flux control analysis. FEBS J 2005;272:1767–83.

[12] Sargeaunt PG, Williams JE, Grene JD. The differentiation of invasive and non-invasive *Entamoeba histolytica* by isoenzyme electrophoresis. Trans R Soc Trop Med Hyg 1978;72:519–21.

[13] Nozaki T, Asai T, Sanchez LB, Kobayashi S, Nakazawa M, Takeuchi T. Characterization of the gene encoding serine acetyltransferase, a regulated enzyme of cysteine biosynthesis from the protist parasite *Entamoeba histolytica* and *Entamoeba dispar*: regulation and possible function of the cysteine biosynthetic pathway in *Entamoeba*. J Biol Chem 1999;274:32445–52.

[14] Van Beneden RJ, Powers DA. Structural and functional differentiation of two clinally distributed glucose phosphate isomerase allelic isozymes from the teleost *Fundulus heteroclitus*. Mol Biol Evol 1989;6:155–70.

[15] Huang J, Mullapudi N, Lancto CA, Scott M, Abrahamsen MS, Kissinger JC. Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. Genome Biol 2004;5:R88.

[16] Sun YJ, Chou CC, Chen WS, Wu RT, Meng M, Hsiao CD. The crystal structure of a multifunctional protein: phosphoglucose isomerase/autocrine motility factor/neuroleukin. Proc Natl Acad Sci U S A 1999;96:5412–7.

[17] Hansen T, Oehlmann M, Schonheit P. Novel type of glucose-6-phosphate isomerase in the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol 2001;183:3428–35.

[18] Lee JH, Chang KZ, Patel V, Jeffery CJ. Crystal structure of rabbit phosphoglucose isomerase complexed with its substrate D-fructose 6-phosphate. Biochemistry 2001;40:7799–805.

[19] Jeffery CJ, Hardre R, Salmon L. Crystal structure of rabbit phosphoglucose isomerase complexed with 5-phospho-D-arabinonate identifies the role of Glu357 in catalysis. Biochemistry 2001;40:1560–6.

[20] Cordeiro AT, Godoi PH, Silva CH, Garratt RC, Oliva G, Thiemann OH. Crystal structure of human phosphoglucose isomerase and analysis of the initial catalytic steps. Biochim Biophys Acta 2003;1645:117–22.

[21] Cordeiro AT, Michels PA, Delboni LF, Thiemann OH. The crystal structure of glucose-6-phosphate isomerase from *Leishmania mexicana* reveals novel active site features. Eur J Biochem 2004;271:2765–72.

[22] Henze K, Horner DS, Suguri S, Moore DV, Sanchez LB, Muller M, et al. Unique phylogenetic relationships of glucokinase and glucose phosphate isomerase of the amitochondriate eukaryotes *Giardia intestinalis*, *Spironucleus barkhanus* and *Trichomonas vaginalis*. Gene 2001;281:123–31.

[23] Ali V, Shigeta Y, Tokumoto U, Takahashi Y, Nozaki T. An intestinal parasitic protist *Entamoeba histolytica* possesses a non-redundant NIF-like system for iron–sulfur cluster assembly under anaerobic conditions. J Biol Chem 2004;279:16863–74.

[24] Jeong JJ, Fushinobu S, Ito S, Jeon BS, Shoun H, Wakagi T. Characterization of the cupin-type phosphoglucose isomerase from the hyperthermophilic archaeon *Thermococcus litoralis*. FEBS Lett 2003;535:200–4.

[25] Ruijter GJ, Visser J. Characterization of *Aspergillus niger* phosphoglucose isomerase. Use for quantitative determination of erythrose 4-phosphate. Biochimie 1999;81:267–72.

[26] Nyame K, Do-Thi CD, Opperdoes FR, Michels PA. Subcellular distribution and characterization of glucose phosphate isomerase in *Leishmania mexicana mexicana*. Mol Biochem Parasitol 1994;67:269–79.

[27] Concepcion JL, Chataing B, Dubourdieu M. Purification and properties of phosphoglucose isomerases of *Trypanosoma cruzi*. Comp Biochem Physiol B Biochem Mol Biol 1999;122:211–22.

[28] Marchand M, Kooystra U, Wierenga RK, Lambeir AM, Van Beeumen J, Opperdoes FR, et al. Glucose phosphate isomerase from *Trypanosoma brucei*. Cloning and characterization of the gene and analysis of the enzyme. Eur J Biochem 1989;184:455–64.

[29] Sargeaunt PG. Zymodemes of *Entamoeba histolytica*. In: Ravdin JI, editor. Amoebiasis: human infection by *Entamoeba histolytica*. New York: John Wiley and Jone Inc; 1988. p. 370–87.

Molecular and Biochemical Parasitology, in press

An *Entamoeba* sp. strain isolated from rhesus monkey is virulent but genetically different from *Entamoeba histolytica*☆

Hiroshi Tachibana[a,*], Tetsuo Yanagi[b], Kishor Pandey[b], Xun-Jia Cheng[a], Seiki Kobayashi[c], Jeevan B. Sherchand[d], Hiroji Kanbara[b]

[a] *Department of Infectious Diseases, Tokai University School of Medicine, Isehara, Kanagawa 259-1193, Japan*

[b] *Department of Protozoology, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan*

[c] *Department of Tropical Medicine and Parasitology, Keio University School of Medicine, Shinjuku-ku, Tokyo 160-8582, Japan*

[d] *Department of Microbiology and Parasitology, Infectious and Tropical Diseases Center, Tribhuvan University Teaching Hospital, Kathmandu, Nepal*

Abbreviations: GPI, glucose phosphate isomerase; HXK, hexokinase; ITS, internal transcribed spacer; PAS, periodic acid-Schiff; PRX, peroxiredoxin

☆ *Note:* Nucleotide sequence data reported in this paper are available in the DDBJ, EMBL, and GenBank™ databases under the accession numbers **AB282657** to **AB282673**.

* Corresponding author. Tel.: +81 (463) 93-1121; fax: +81 (463) 95-5450.

E-mail address: htachiba@is.icc.u-tokai.ac.jp (H. Tachibana)

Abstract

An *Entamoeba* sp. strain, P19-061405, was isolated from a rhesus monkey in Nepal and characterized genetically. The strain was initially identified as *E. histolytica* using PCR amplification of peroxiredoxin genes. However, sequence analysis of the 18S rRNA gene showed a 0.8% difference when compared to the reference *E. histolytica* HM-1:IMSS human strain. Differences were also observed in the 5.8S rRNA gene and the internal transcribed spacer regions 1 and 2, and analysis of the serine-rich protein gene from the monkey strain showed unique codon usages compared to *E. histolytica* isolated from humans. The amino acid sequences of two hexokinases and two glucose phosphate isomerases also differed from those of *E. histolytica.* Isoenzyme analyses of these enzymes in the monkey strain showed different electrophoretic mobility patterns compared with *E. histolytica* isolates. Analysis of peroxiredoxin genes indicated the presence of at least seven different types of protein, none of which were identical to proteins in *E. histolytica*. When the trophozoites from the monkey strain were inoculated into the livers of hamsters, formation of amebic abscesses was observed 7 days after the injection. These results demonstrate that the strain is genetically different from *E. histolytica* and is virulent. Revival of the name *Entamoeba nuttalli* is proposed for the organism.

Keywords: Entamoeba histolytica; *Entamoeba nuttalli*; Rhesus monkey; rRNA gene; Virulency

1. Introduction

The enteric protozoan *Entamoeba histolytica* causes an estimated 50 million cases of amebic colitis and liver abscess in human, resulting in 100,000 deaths annually [1]. *Entamoeba dispar* is morphologically indistinguishable from *E. histolytica* but is nonpathogenic [2]. It has been reported that *E. histolytica/E. dispar* is commonly found in the feces of non-human primates, such as macaques and baboons [3-6]. Since *E. histolytica* infection in non-human primates is problematic not only for animal health but also as a possible source of transmission to humans, it is important to discriminate between *E. histolytica* and *E. dispar*. Recent studies have demonstrated that *E. dispar* infection is prevalent in captive and wild non-human primates, including baboons, macaques and chimpanzees [7-11]. We have demonstrated that the prevalence of *E. dispar*, but not *E. histolytica*, is 43% in a wild colony of *Macaca fuscata fuscata* in Japan [10], 66% in captive *Macaca* monkeys [8], and 56% in captive chimpanzees [9]. In contrast, a limited number of *E. histolytica* infections in non-human primates have been confirmed [12-14]. However, it is unknown whether *E. histolytica* isolated from non-human primates is identical to that isolated from humans.

In the present study, we analyzed several genes of an *Entamoeba* strain isolated from a rhesus monkey, *Macaca mulatta*, and compared it with a reference strain isolated from humans. We report here that the monkey strain is virulent but genetically different from *E. histolytica*.

2. Materials and methods

2.1. Isolation and culture conditions

A fecal sample containing *Entamoeba* cysts was obtained from a rhesus monkey in

Pashupati Nath Temple, Kathmandu, Nepal on June 2005. The sample was suspended in water for 24 h to remove *Blastocystis* spp. and then cultured in modified Tanabe-Chiba medium, which consists of an agar slant and an upper liquid medium in a screw-cap test tube, at 37°C. The slant contained 1% agar in Ringer's solution supplemented with 0.1% L-asparagine, and the liquid medium contained phosphate-buffered saline (pH 7.6) supplemented with one-eighth the volume of horse serum. A microspatula of rice powder was added before use. Grown trophozoites of the *Entamoeba,* designated as the P19-061405 strain, were brought to Japan and cultured in Robinson's medium at 37°C [15]. The trophozoites were treated with a cocktail of antibiotics and then cultured monoxenically with live *Crithidia fasciculata* in BI-S-33 medium supplemented with 15% adult bovine serum at 37°C [16]. The monoxenically cultured trophozoites were used in most experiments in the present study. Finally, trophozoites of the strain were cultured axenically in the BI-S-33 medium and then cloned by limiting dilution, followed by examination using microscopy. Trophozoites of *E. histolytica* strains HM-1:IMSS, HK-9, Rahman, and NOT-12 were cultured axenically in BI-S-33 medium and were used as reference strains. Trophozoites of the *E. histolytica* SAW1453 strain were cultured xenically in Robinson's medium. Trophozoites of the *E. dispar* SAW1734RclAR strain were grown monoxenically with sterilized *Crithidia fasciculata* in YIGADHA-S medium supplemented with 15% adult bovine serum at 37°C and were also used as a reference strain [17].

2.2. PCR amplification

Genomic DNA of trophozoites was isolated using a DNeasy tissue kit (Qiagen).