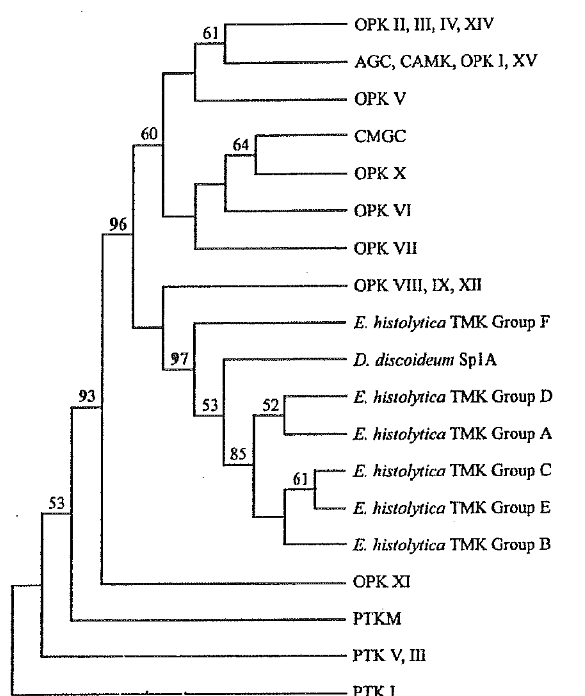

FIG. 1. Phylogenetic and sequence analysis of the kinase domain of *E. histolytica* transmembrane kinases. CMGC is the cyclin-dependent kinase, mitogen-activated protein kinase, glycogen synthase kinase, and CDK-like kinase family. CAMK is the calmodulin-dependent kinase family. AGC is the family of protein kinases A, G, and C. TMKs were aligned to Hanks's alignment (20) using CLUSTALX (61). The aligned sequences were bootstrapped using Seqboot, Protpars, and Consense (15). Significant bootstrap values are shown in bold, and bootstrap values above 50 are shown. The GenBank numbers for the TMKs are in Tables S1 and S2. SplA is GenBank accession no. U32174 (40).

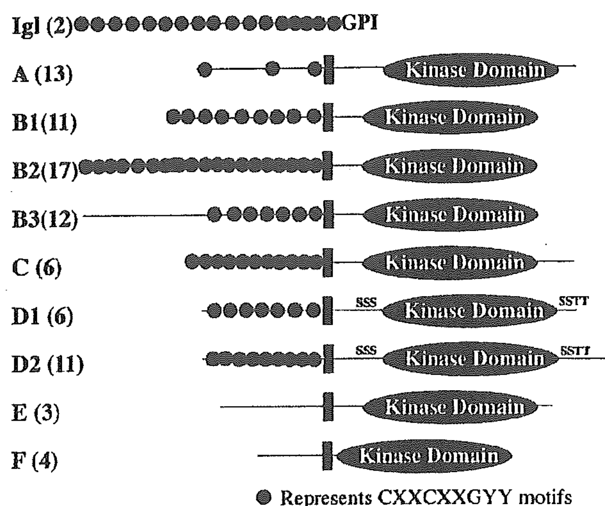

FIG. 2. Diagram of Igl and the TMKs. This diagram shows the approximate sizes of the different proteins and the distribution of the CXXCXXGYY motifs in the extracellular domain (indicated by a black circle). The CXXCXXGYY motifs are part of a larger motif, CXXCXXG(Y)(Y/F)(L/V/F/Y/M)-Polar-Polar, which also can begin with CXC instead of CXXC. GPI represents a putative GPI anchor. A black rectangle indicates a transmembrane domain, and a black oval indicates a putative kinase domain. The serine (SSS)- and serine/threonine (SSTT)-rich regions found in groups D1 and D2 are shown. Numbers in brackets indicate numbers of known family members.

Production of anti-ΔTMK96 rabbit serum production. The kinase region of Tmk96 (ΔTMK96) was PCR amplified with the primers 5′-CAATTTAGAGAAGGAATTCCT-3′ (5′ primer) and 5′-TCACATTAATTGAAGATGTTTTAAAACAACA-3′ (3′ primer). This 1,000-bp fragment was cloned into TOPO NT/T7 (Invitrogen), in frame with an amino-terminal six-His tag via TA cloning. Bacteria were grown at 37°C to an optical density at 600 nm of 0.5 and induced with isopropyl-β-D-thiogalactopyranoside for 4 h, and the recombinant protein was purified with nickel agarose beads (Qiagen). Antibodies were raised to this purified recombinant protein by a 90-day protocol including three inoculations of New Zealand White rabbits with recombinant ΔTMK96 (Covance, Princeton, NJ), and the antibodies were purified from serum with a protein A column.

Western blots using ΔTMK96. Soluble proteins were extracted from amebae by harvesting 5×10^7 trophozoites by incubation on ice for 10 min, followed by centrifugation (200 × *g* at 4°C for 5 min). The amebae were lysed in 10 mM sodium phosphate buffer with protease inhibitor cocktail I (Sigma, St. Louis, MO) per the manufacturer's directions. Membranes were then cleared by centrifugation (100,000 × *g* at 4°C for 1 h). Whole-cell lysates were prepared by sonication of 10^6 amebae in three 5-min pulses on ice. Large intact particles were eliminated by centrifugation (20,000 × *g* at 4°C for 30 min). All samples were then separated on 10% polyacrylamide gels and then electrotransferred to polyvinylidene difluoride membranes (Millipore, Bedford, MA). Nonspecific binding was blocked by preincubation with Tris-buffered saline with 5% bovine serum albumin and 0.05% Tween 20 (TBST). In order to detect proteins on the blot, either anti-ΔTMK96 rabbit serum (a dilution of 1:5,000) or preimmune serum (1:2,500) was added in TBST for 1 h at room temperature. Interactions were detected by the addition of peroxidase-conjugated goat anti-rabbit IgG (Sigma) and development with ECL (Amersham) per the manufacturer's directions.

Confocal microscopy. *E. histolytica* trophozoites (10^6) were bound to glass coverslips in a 24-well plate for 30 min at 37°C in TYI-S-33 medium. Adherent amebae were washed twice in phosphate-buffered saline (PBS) and fixed in 3% paraformaldehyde for 30 min at room temperature. To make the plasma membrane permeable, amebae were treated with 0.2% Triton X-100 in PBS for 1 min. Nonspecific binding was blocked by incubation with 20% goat serum and 5% bovine serum albumin (Sigma) in PBS for 1 h at 37°C. After incubation with either the anti-ΔTMK96 rabbit polyclonal antibody (200 μg/ml) or anti-Gal/GalNAc lectin antibody (6 μg/ml) for 1 h at 37°C, the coverslips were washed twice before Cy3-conjugated goat anti-rabbit secondary antibodies (Jackson Laboratories, Bar Harbor, ME) were added at a 1:160 dilution for 1 h at 37°C. The coverslips were washed twice and mounted on slides with Gel/Mount (Biomeda, Foster City, CA). Confocal images were visualized using a Zeiss LSM 510 laser scanning microscope (Carl Zeiss, Inc., Thornwood, NY).

RESULTS

Phylogenetic analysis of the *E. histolytica* TMK family. Previously we reported a family of CXXC-containing proteins with sequence similarity to Igl1 and Igl2 (8). With the completion of the genome project, it became apparent that the majority of these CXXC-containing proteins were members of an approximately 80-gene family with intracellular kinase domains. The extracellular domains shared sequence similarity to Igl and to VSPs of *Giardia*. The kinase domain of these proteins in most cases contained all five conserved kinase motifs: the key conserved glycine-rich motif, the K residue, and the HRDL, DFG, and APE motifs (Table 2). However, a few lacked the glycine-rich motif (Tmk01, Tmk24, Tmk45, and Tmk73), which helps coordinate the second phosphate of ATP (14). The glycine, K, and DFG motifs are implicated in the binding and orientation of ATP. The HRDL motif is involved in catalysis. The APE motif is responsible for anchoring the substrate and thereby influencing the specificity of interaction with the protein that is being phosphorylated (14, 20, 60). Despite the conservation of the essential kinase motifs, overall the kinase domains were

TABLE 3. Consensus sequence comparison of subdomains of malian tyrosine kinases, serine/threonine kinases, and the *E. histolytica* transmembrane kinase groups

Kinase	Consensus sequence in subdomain[a]				
	I	II	III	IV	V
Tyrosine	LGxGxFGxVxxG	VA(I/V)K	ExxxM(R/K)	GxCooo	GxL(F/Y)lx
Ser/Thr	LGxGxFGxVxxx	VAIK	Exxxx(K/R/H)	Gxxxxx	GXX(Y/F)(L/I)H
A	IAEGAMG-VY-G	VA(I/V)K	ExxxLK	GSVTYI	GSLgeYOR
B1	IGEGsFGIVYVG	VAIK	Ee(S/N)EDK	GAVFIP	GSIQDIMNK
B2	oGEGsFGoVY(K/o)G	VAIK	EVxMLDK	GavfIP	GSLxxooxK
B3	oGEGsFGoV(Y/F)KG	VAIK	EoxmLdK	Gavoop	GSoxxoo(x)K
C	IGEGAFGMVFRG	VAIK	EQEKEF	GAVYTE	GS(M/L)SKOWxK
D1	IGGGTFGIVYRA	VAVK	EAEMMER	GSVoTx	GSLRKFMK
D2	IGxGTFGoVY(K/R)(G/A)	VA(V/I)K	ExxooEx	Gsvxxx	GSLRK(Y/f)ok
E	IGEG(S/T)FgiV(Y/F)xG	VAIK	ExFxxE	Gxxxox	GSoKxoY
F	ogxg(S/T)xgxvwxa	oaVK	evxlmK	Gsgxdx	gxossoL

[a] Subdomain is according to Hanks's alignment of kinases (20). Underlined in tyrosine and serine/threonine kinases indicate highly conserved residues. Letters in uppercase are conserved. Letters in lowercase indicate moderate conservation. x, any residue; o, hydrophobic residue; —, sequence of varying length.

divergent from other known kinases. The TMK family branched closely with the other protein kinase group IX (OPK IX), which are the TGFβ receptor and activin family of serine/threonine kinases, OPK VIII, which are a Raf family of serine/threonine kinases, and OPK XII, which are the casein kinase I family of kinases (Fig. 1). It was not possible to determine by sequence analysis alone if the TMKs were serine/threonine or tyrosine kinases. Sp1A from *Dictyostelium discoideum* was the most closely related kinase identified by sequence similarity in the NCBI database that has been functionally characterized. The Sp1A kinase has been shown to be a dual-specificity kinase that phosphorylates both tyrosine and serine/threonine residues (40). The Sp1A kinase was found to be phylogenetically within the TMK family (Fig. 1). Unlike the TMK family, SplA is a cytoplasmic kinase rather than a transmembrane kinase. Two TMKs (Tmk58 and Tmk89) clearly did not group with the TMK family but grouped closely with other protein kinase groups VI and VII (data not shown), suggesting that these may represent serine/threonine transmembrane kinases that have evolved separately from the TMK family.

The TMK family members were grouped based on sequence and phylogenetic analysis of their cytoplasmic kinase domains. A specific signature motif between the conserved HRDL and APE motifs was identified for each family (Table 2). The motif was 15 to 24 amino acids C-terminal to the histidine in the HRDL motif. The motif sequences were CC(I/V)KITDFGTSR (group A), KLTDFGS(A/S)R (group B), C(A/G)KLTDFGTC (group C), PITAKVTDFGTS (group D1), V(T/V)(C/X)KV(T/S)DFGTS (group D2), AKLSDFGTSR (group E), and VKVSDFGLS with a conserved tryptophan two resi-

TABLE 4. Significantly changed genes during growth

Growth period[a]	Probe(s)[b]	Protein	Group	Change	*P* value
Early (12 h p.i.) to mid-log (48 h p.i.)	1, 91, 99	Actin		Increased 2.17-fold	<0.05
	190	Lgl3	Gal/GalNAc lectin	Increased 1.37-fold	<0.05
Late log (96 h p.i.) to nonadherent (144 h p.i.)	1, 91, 99	Actin		Decreased 1.88-fold	<0.01
	321	EHCP1&2	Cysteine proteinase	Decreased 1.79-fold	<0.05
	320	EHCP20	Cysteine proteinase	Increased 1.37-fold	<0.05
	88	Hgl family	Gal/GalNAc lectin	Decreased 1.58-fold	<0.01
	189	Lgl1	Gal/GalNAc lectin	Decreased 2.87-fold	<0.05
	280	Lgl4	Gal/GalNAc lectin	Decreased 2.28-fold	<0.01
	232	Lgl5	Gal/GalNAc lectin	Decreased 4.26-fold	<0.05
	224	ENT3593	Unknown—surface	Increased 1.61-fold	<0.05
	145	TSA	Thiolredoxin	Decreased 1.65-fold	<0.05
	239, 277	SREHP	Surface protein	Decreased 3.08-fold	<0.05

[a] Three hybridizations (biological replicates) were analyzed in two dye experiments comparing 12 h to 48 h, 48 h to 96 h (no significant differences were found), and 96 h, to 148 h postinoculation (p.i.).
[b] Probe numbers are from Table S1 in the supplemental material.

TABLE 3—*Continued*

Consensus sequence in subdomain[a]						
VIa	VIb	VII	VIII	IX	X	XI
GMxYlx	o(I/V)HRDLaar N(c/l)oo	O-KixDFGl	PIW(K/R) WXAPEOO	SDVW(S/A)(F/Y) GoooWE–Pyp	—	CWxxxxxRPxF
(G/A)ox (Y/F)oH	ooHRDoKPxNOOO	KooDFGo	xxxx(Y/F) xAPEoo	XDoW(G/S) oGoooe–PFx	—	xxxxxxxRxxx
GMxFLH	IoHLDLKPDNLLVN	KITDFGT	GTPoYxAPExY	(g/s)DV(Y/F)(s/a) (F/Y)AotaWEoFYqxEPYK	—	CWkQxxxeRpxF
GISYLH	iLHRDLIKPDNFLVV	KLTDFGS	GTPKYMAPEvL	SDIYSFSITMLQiiTWQDPFP	—	SWQQEPKERITI
GIxYLH	ILHRDLIKPDNoLoo	KLTDFGS	GTPxYMAPEoL	ADIySFaotmY(e/q) xoxWxxPyP	—	(s/c)WxqxPxxRox
GIxYLH	ILHRDLIKpdNoLoo	KLTDFGS	GTPxYMAPEoL	ADI(y/f)sFaotmyEo oxWxxa(y/f)p	—	(s/c)WxxxPxxRox
ALX(Y/F) LH	IIHRDVKGENoLo (Y/F)	KLTDFGT	GTPTYMAPECL	vDVYAYGIVLYET(y/f) xExxAYxxDERFNQPWM	—	CwxQxxxxRPxF
GMEYLH	ILHRDLKTDNVLVY	KVTDFGT	GTPVYMAPEI (s/t)	SDVYSFAICMLEIWLGRDPYDP	—	(a/s)WxHxPS (E/D)RPTF
GMxYLH	IoHrDLKtDNVLvo	KV(T/S) DFGT	GTPmYMAPEoh	SDVYSFAICoLEIWoxxxPY (d/p)	—	CWxxxPxxRPxF
GMxFLH	IIHRDLKPDNoLoo	KLSDFGT	GTPxYMAPEoo	oDV(F/Y)s(F/Y)AoVx (Y/F)EoosrkoPYs	—	CWaxDPxxRPxF
GMoYLH	IoHxdLksxNoLVx	KvSDfgl	gtlxWxAPEoL	xDvYS(F/Y)gIomWEoot rxxP(Y/F)x	—	CWxxxpexrpxf

dues N-terminal to the APE motif (group F) (Table 2 and Fig. 2). Tmk58 and Tmk89, which did not group with the rest of the TMKs, contained an ITDFGLAKK motif. Group G TMKs lacked a conserved motif in the kinase domain and also lacked one or more of the conserved kinase domain motifs. Sequence similarity between family members was not limited to the kinase domain. Additional motifs were found in some families that can be used to identify group members between the cytoplasmic and kinase domain (Table 2 footnotes). Additionally, group D TMKs had a serine-rich region N-terminal to their kinase domain and a serine/threonine region C-terminal to their kinase domain. Serine/threonine rich regions are often found in kinases and typically regulate kinase activity (30). The kinase domains of many of the *E. histolytica* TMKs were a mix of serine/threonine and tyrosine kinase signature motifs, further complicating efforts to predict their kinase activity (Table 3). However, the sequence similarity to SplA of *D. discoideum* suggests that they likewise may be dual-specificity kinases.

An overview of the sequence motifs of the different subfamilies of kinases is shown in Table 2 (all domains except kinase domains), Table 3 (kinase domains), and Fig. 2. The TMKs had extracellular domains of 36 to 2,117 amino acids (with 11 to 98% similarity between family members), a transmembrane domain, and a kinase domain. Groups B, C, and D shared significant sequence similarity in their extracellular domain to Igl1 (Fig. 2). Most extracellular domains were rich in CXC/CXXC/CXXXC motifs. Almost all of the TMKs had at least one CXXCXXGYY motif beginning approximately 25 amino acids N-terminal to the transmembrane domain. Like Igl1 most of the TMKs had many additional CXXCXXGYY motifs (Table 2 and Fig. 2). A very similar C(D/E)XCXXG(Y/F)(Y)(G) motif was found in Igl, VSP of *Giardia* (1, 38), and laminin LE domains (27, 57, 63). In VSP of *Giardia* the CXXC motifs have not been crystallized but have been shown to bind zinc (38), to have their N terminus at the host parasite interface (38), and have nonreactive cysteines (42). In laminin the LE domains form a rod-like structure of mini-globular folds (57, 63). This is consistent with the CXXCXXGYY motifs forming a linear array of mini-globular folds. By analogy to the laminin LE domains, the TMKs may have a rod-like structure.

Gene expression analysis of *E. histolytica* during growth. To characterize the expression profile of these genes, we constructed an oligoarray of the TMK genes, Gal/GalNAc lectin genes, and other putative surface virulence genes. Few changes in gene expression were seen when early phase (12 h) was compared with mid-log phase (48 h) and late log phase (96 h) was compared with nonadherent phase (144 h) (Table 4). When mid-log-phase (48 h) and late-log-phase (96 h) amebic cultures were compared, very little change was seen (data not shown). *actin* was clearly growth regulated (Table 4), as were the Gal/GalNAc lectin *hgl* genes, many but not all of which decreased significantly during late log phase (Table 4 and data not shown). This decrease is consistent with previous observations of *hgl1*, *hgl2*, and *hgl3* gene expression (48). Expression of known genes was, in general, consistent with RT-PCR, Western, and/or Northern analysis (data not shown). We concluded that *E. histolytica* did not appear to growth-phase regulate expression of most putative virulence genes when grown under lab culture conditions.

Expression analysis of the TMKs. The expression profile of 67 of the TMK genes during mid-log phase (72 h) is shown in Fig. 3. At this time point, 19 genes (group A, *tmk61*, *tmk65*, and *tmk72*; group B2, *tmk02*, *tmk08*, and *tmk74*; group B3, *tmk21* and *tmk28*; group C, *tmk39* and *tmk63*; group D1, *tmk40* and *tmk56*; group D2, *tmk19*, *tmk44* and *tmk46*; group E, *tmk22* and *tmk54*; group F, *tmk59*; group G, *tmk06*) (Fig. 3) showed hybridization values in trophozoites significantly greater than that for the cyst-specific transcript *jacob* ($P < 0.05$).

Analysis of TMK expression by RT-PCR. Real-time PCR was conducted on *jacob*, *tmk31*, *tmk80*, *tmk96*, and *tmk98*, all of

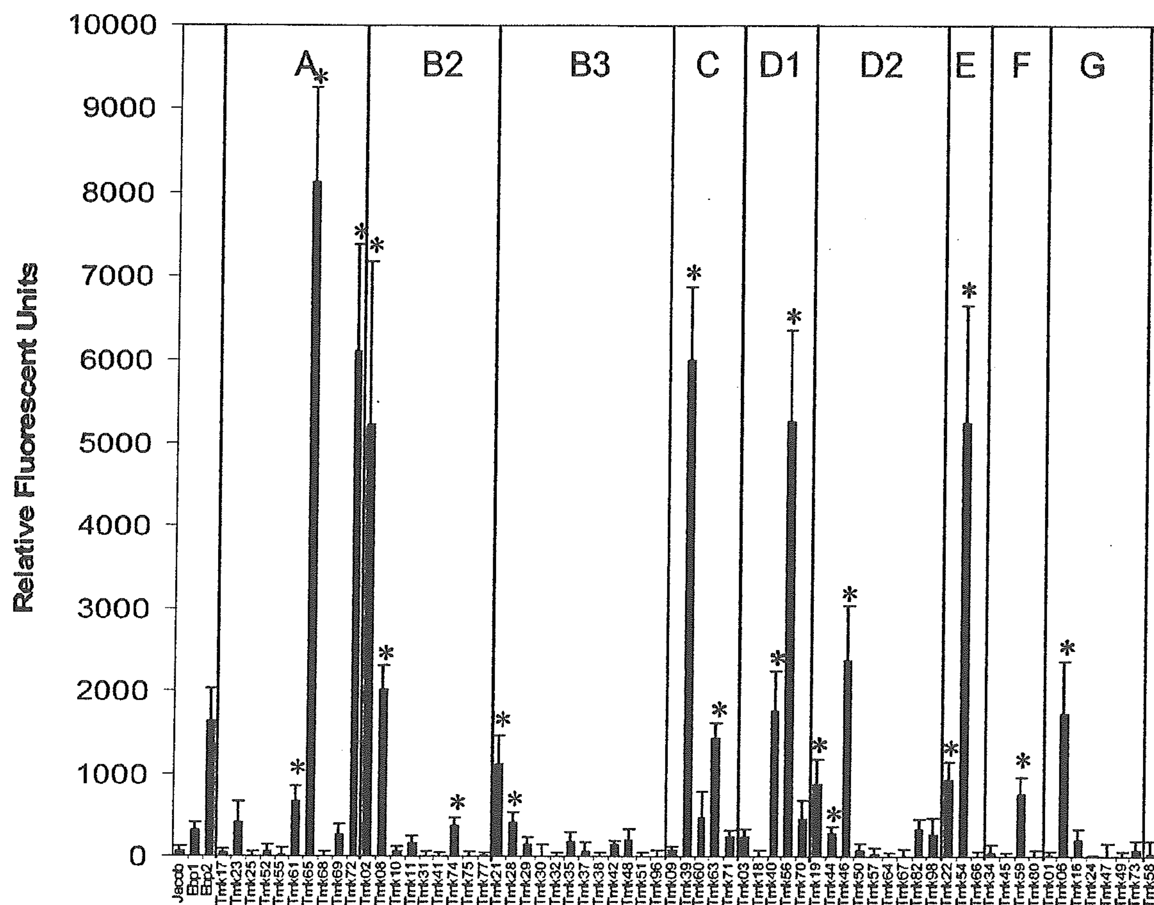

FIG. 3. Expression of TMKs during log-phase culture as determined by oligoarrays. The TMKs were classified into groups: A, B2, B3, C, D1, D2, E, F, G, and "other" (*tmk58*). Expression of *jacob* (encoding a cyst protein not expected to be expressed in trophozoites), *ebp1*, and *ebp2* genes is shown for reference. TMK genes expressed at a higher level than *jacob* ($P < 0.05$) are indicated by an asterisk and were as follows: group A, *tmk61* (391.t00004-AAFB01000774, 279.t00010-AAFB01000993), *tmk65* (62.t00013-AAFB01000240), and *tmk72* (302.t00003-AAFB01000819); group B2, *tmk02* (70.t00014-AAFB01000264), *tmk08* (10.t00040-AAFB01000051), and *tmk74* (6.t00088-AAFB01000031); group B3, *tmk21* (42.t000019-AAFB01000175), and *tmk28* (66.t00027-AAFB01000251); group C, *tmk39* (359.t00009-AAFB01000933), and *tmk63* (20.t00067-AAFB01000094); group D1, *tmk40* (65.t00015-AAFB01000247), and *tmk56* (5.t00091-AAFB01000028); group D2, *tmk19* (135.t00017-AAFB01000458), *tmk44* (159.t00012-AAFB01000511), and *tmk46* (131.t00015-AAFB01000449); group E, *tmk22* (12.t00043-AAFB01000464) and *tmk54* (75.t00011-AAFB01000285); group F, *tmk59* (304.t00008-AAFB01000821); and group G, *tmk06* (274.t00010-AAFB01000764). Error bars represent the standard error of the mean of three hybridizations (biological replicates).

which were predicted by oligoarray not to be expressed, and *tmk19*, *tmk21*, *tmk63*, *tmk65*, *tmk71*, *tmk75*, *sa*, *hgl*, and *actin*, all of which were predicted to be expressed. Gene expression was monitored sequentially in trophozoites in laboratory culture over a 12-day period. To allow comparison between time points, results were normalized to the average of three RNA polymerase II genes (Fig. 4). All real-time PCR results of the TMK genes were consistent with the oligoarray results. Significant variations in expression during laboratory culture were observed for *tmk19*, *tmk63*, and *tmk79* ($P < 0.01$) (Fig. 4D, 4F, and 4I).

Expression analysis of the TMKs during erythrophagocytosis. One million trophozoites were grown in 50 ml of medium with or without a vast excess of erythrocytes (24 million/ml of medium) for 24 h. We did not observe significant changes in TMK gene expression during erythrophagocytosis (data not shown).

Detection of expression of TMK family members with polyclonal antibodies. The kinase domain of Tmk96 was expressed in *Escherichia coli* and used to generate polyclonal antibodies. Multiple trophozoite proteins were detected with polyclonal anti-kinase domain antisera (Fig. 5A). No bands were observed with preimmune sera (data not shown). The recognition of multiple proteins by the antisera raised against the Tmk96 kinase domain was not surprising given that the kinase domain is conserved between different TMKs. The predicted TMKs vary in size from 482 to 2,577 amino acids, and Tmk96 is in the largest subfamily (group B) of TMKs with multiple members expressed. Of the TMK genes that we detected with expression significantly above *jacob* the size range was from 686 to 2,577 amino acids (78 kDa to 294 kDa not accounting for potential posttranslational modifications). The smallest proteins observed on Western blots were of a mass consistent with that of a typical kinase domain (270 amino acids) and may represent

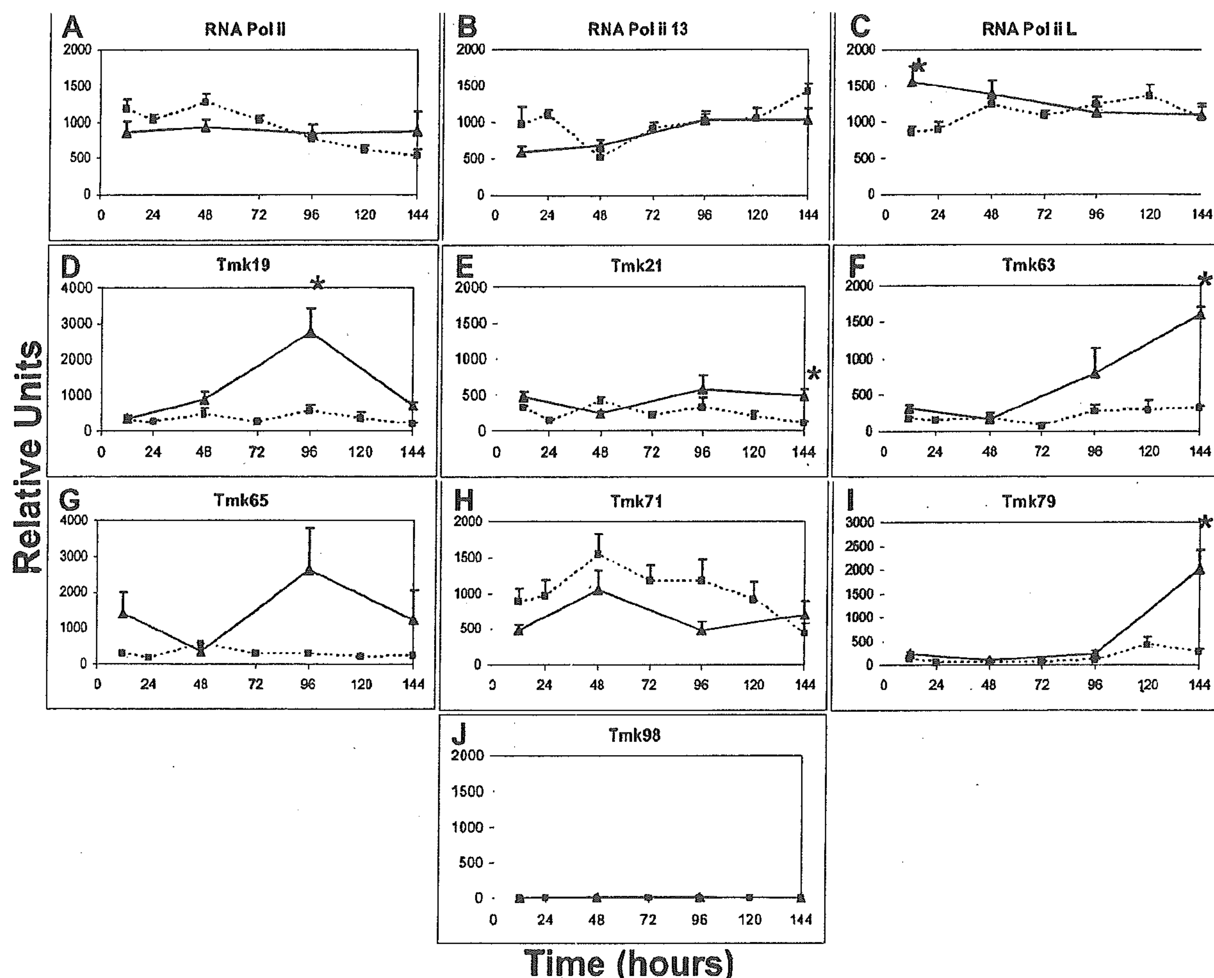

FIG. 4. Expression of TMK family genes during culture as determined by real-time PCR. Quantitative real-time PCR was performed on (A) the RNA polymerase II gene (*rna pol ii*) (27.t00035-AAFB01000114), (B) the RNA polymerase II 13 gene (*rna pol ii 13*) (344.t00001-AAFB01000903), (C) the RNA polymerase II L gene (*rna pol ii l*) (147.t0005-AAFB01000482), (D) *tmk19* (135.t00017-AAFB01000458), (E) *tmk21* (42.t00019-AAFB01000175), (F) *tmk63* (20.t00067-AAFB01000094), (G) *tmk65* (62.t00013-AAFB01000240), (H) *tmk71* (268.t00007-AAFB01000754), (I) *tmk79* (71.t00002-AAFB01000266), and (J) *tmk98* (361.t00001-AAFB01000937). Two sequential growth curves are shown. For growth curve A (triangles and solid line), samples were collected at 12, 48, 96, and 144 h postinoculation. For growth curve B (squares and dashed line), samples were collected at 12, 24, 48, 72, 96, 120, and 144 h postinoculation. Triplicate samples were collected at each time point. Culture B was established by transferring 300,000 amebae from culture A at 144 h. The standard errors of three biological samples, with each sample analyzed in duplicate, are shown. To allow comparison between time points, data for the TMKs were normalized to the average of RNA polymerase II, RNA polymerase II L, and RNA polymerase II 13. The average expression of these three genes was defined as 1,000 units of expression.

TMKs that do not have an extracellular domain or whose extracellular domain has been cleaved off.

Localization of TMK kinase domains to plasma membrane microdomains. Permeabilized amebae showed a focal plasma membrane staining pattern with anti-TMK kinase domain antibodies (Fig. 5B). No staining was seen with nonpermeabilized cells or with permeabilized cells stained with preimmune sera (data not shown). We concluded that the kinase domain was on the cytoplasmic side of the plasma membrane. The TMKs therefore appeared to be typical type I transmembrane proteins with an amino-terminal signal sequence and a predicted transmembrane domain preceding the kinase domain. The focal staining pattern contrasted with the uniform plasma membrane staining pattern seen with anti-Gal/GalNAc lectin antisera (Fig. 5C).

DISCUSSION

The most important finding of this work is the identification of a large family of over 80 transmembrane kinases in *E. histolytica*. Although eukaryotic-type transmembrane kinases are found in organisms from bacteria to humans (4, 64, 68), large families of TMKs have been previously described only in multicellular organisms (20). For example, we were not able to find any TMKs in the partially completed genome of the ameba *D. discoideum* (17), and only a few TMKs were found in the apicomplexan *P. falciparum* (4). The TMKs were also found in *Entamoeba invadens*, *Entamoeba dispar*, and *Entamoeba moshkovskii*, indicating that this family is conserved within *Entamoeba* (data not shown) (66). Multiple TMKs were observed to be expressed at the mRNA and protein levels,

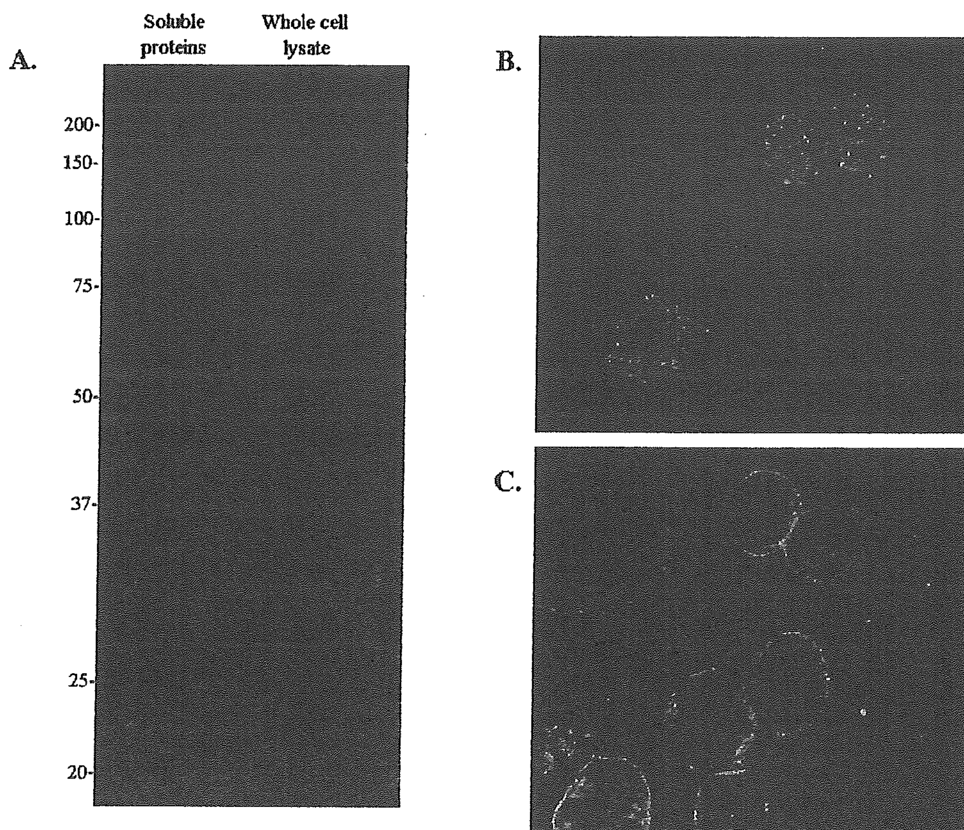

FIG. 5. Recognition of *Entamoeba histolytica* surface proteins by anti-ΔTMK96 rabbit serum. (A) Soluble trophozoite proteins as well as whole-cell lysates were analyzed by Western blots with anti-ΔTMK96 rabbit serum (1:5,000 dilution). Preimmune serum did not recognize any trophozoite proteins on Western blots (data not shown). (B) Confocal microscopy of permeabilized trophozoites with anti-ΔTMK96 rabbit serum. (C) Confocal microscopy of permeabilized trophozoites with anti-Gal/GalNAc lectin antibodies. No staining was seen with preimmune rabbit serum or in nonpermeabilized trophozoites with the anti-ΔTMK96 serum (data not shown). Magnification, ×400.

consistent with nonoverlapping biological functions for individual members of the TMK family.

The existence of multiple different extracellular domains of the TMKs suggests that each interacts with the host environment and signals into the parasite in distinct ways. The extracellular domains of the group B, C, and D TMKs had sequence similarity to Igl of *E. histolytica* (8), laminin LE domains (27, 57, 63), and VSPs of *Giardia lamblia* (1, 38). The sequence similarity is largely limited to a repeated CXXCXXGYY motif. In laminin the LE domains function as mini-globular folds arranged in tandem to form a rod-like structure. If the CXX CXXGYY motifs in the TMKs take on a similar conformation, then these motifs may function to help the extracellular domain of the TMKs project off the surface of the cell in a pilus-like manner. This may make them available for interaction with host cell factors. In each TMK subfamily there was one or more members expressed; however, most family members appeared not to be expressed under the conditions of laboratory culture. By RT-PCR we did observe that the expression of some TMKs varied between growth curves. This indicates that the expression of these genes may be dynamic. Whether the TMKs share with the *Giardia* VSPs the process of antigenic variation under different biological conditions remains to be determined.

The most significant feature of the TMKs is the kinase domain that, with the exception of two TMKs, is distinct from other known kinases. It is not possible, based on sequence analysis, to predict activity, as most have similarity to both the serine/threonine and tyrosine kinases. Interestingly, a closely related kinase, SplA from *D. discoideum*, is a dual-specificity kinase with both tyrosine and serine/threonine kinase activity (40). All of the essential kinase motifs were conserved, suggesting that these are functional kinases. We were not able to demonstrate kinase activity when the kinase domain was expressed in *E. coli*. Since all of the functional residues were conserved, the most likely explanation for this is that the kinase is not functional in *E. coli* or that the kinase domain is not able to phosphorylate the substrates we have used. It has been previously shown that some kinases were not functional when expressed in *E. coli* (18). Additionally, even if the kinase is functional, identification of a substrate is often the rate-limiting step in characterizing a kinase (5, 23). Further experimentation will be necessary to identify the substrate or interacting partners of the TMKs and determine if the TMKs are serine/threonine and/or tyrosine kinases. Phylogenetic and sequence analysis shows that there are six subfamilies of kinases with distinct motifs within the kinase domains. Some families had additional conserved motifs outside of the kinase domain. This

would imply that the subfamilies may represent functionally different families of kinases in sensing (differences in extracellular domains) and signaling (differences in kinase domains).

The focal staining pattern of the TMKs distinctly contrasts with the uniform plasma membrane staining pattern seen with the Gal/GalNAc lectin of *E. histolytica* or VSPs in *G. lamblia*, both of which lack cytoplasmic kinase domains (38, 42, 45). This localization suggests that the TMKs form a focal multimolecular signaling complex in the plasma membrane (31, 33).

In conclusion, the work presented here may begin to explain how *E. histolytica* is able to persist in the host for long periods of time despite immune surveillance, as well as sensing and responding to host stimuli. The large families of TMKs described here could serve in both biological sensing and antigenic variation. The distinct extracellular and kinase domains of the TMKs suggest that each TMK may sense or interact with different host factors and cause a distinct signaling event in response to that environmental cue.

ACKNOWLEDGMENTS

We thank Aaron J. Mackey and William R. Pearson for guidance with the bioinformatics analyses and Brendan Loftus and Neil Hall for access to the *E. histolytica* genome sequencing project data at the TIGR and Sanger sequencing centers. Barbara Mann provided the Spl sequence.

This study was supported by NIH grant AI26649 to W.A.P. B.D. was supported by the Biomolecular Research Facility of the University of Virginia. T.N. was supported by a grant for Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan (16017307, 16044259, 15590378), and a grant from the Japan Health Sciences Foundation.

REFERENCES

1. **Adam, R. D., A. Aggarwal, A. A. Lal, V. F. Cruz, T. McCutchan, and T. E. Nash.** 1988. Antigenic variation of a cysteine-rich protein in *Giardia lamblia*. J. Exp. Med. **167**:109–118.
2. **Ali, V., T. Hashimoto, Y. Shigeta, and T. Nozaki.** 2004. Molecular and biochemical characterization of D-phosphoglycerate dehydrogenase from *Entamoeba histolytica*. A unique enteric protozoan parasite that possesses both phosphorylated and nonphosphorylated serine metabolic pathways. Eur. J. Biochem. **271**:2670–2681.
3. **Ali, V., Y. Shigeta, and T. Nozaki.** 2003. Molecular and structural characterization of NADPH-dependent d-glycerate dehydrogenase from the enteric parasitic protist *Entamoeba histolytica*. Biochem. J. **375**:729–736.
4. **Anamika, N. Srinivasan, and A. Krupa.** 2005. A genomic perspective of protein kinases in *Plasmodium falciparum*. Proteins **58**:180–189.
5. **Brinkworth, R. I., R. A. Breinl, and B. Kobe.** 2003. Structural basis and prediction of substrate specificity in protein serine/threonine kinases. Proc. Natl. Acad. Sci. USA **100**:74–79.
6. **Bruchhaus, I., B. J. Loftus, N. Hall, and E. Tannich.** 2003. 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 **2**:501–509.
7. **Chavez-Rios, R., and M. Vargas-Mejia.** 2000. Isolation and identification of L10 gene from *Entamoeba histolytica* homologous to Wilms' tumor suppressor. Arch. Med. Res. **31**:S305–S306.
8. **Cheng, X. J., M. A. Hughes, C. D. Huston, B. Loftus, C. A. Gilchrist, L. A. Lockhart, S. Ghosh, V. Miller-Sims, B. J. Mann, W. A. Petri, Jr., and H. Tachibana.** 2001. Intermediate subunit of the Gal/GalNAc lectin of *Entamoeba histolytica* is a member of a gene family containing multiple CXXC sequence motifs. Infect. Immun. **69**:5892–5898.
9. **Cheng, X. J., H. Tsukamoto, Y. Kaneda, and H. Tachibana.** 1998. Identification of the 150-kDa surface antigen of *Entamoeba histolytica* as a galactose- and *N*-acetyl-D-galactosamine-inhibitable lectin. Parasitol. Res. **84**:632–639.
10. **Das, S., and F. D. Gillin.** 1991. Chitin synthase in encysting *Entamoeba invadens*. Biochem. J. **280**:641–647.
11. **de la Vega, H., C. A. Specht, C. E. Semino, P. W. Robbins, D. Eichinger, D. Caplivski, S. Ghosh, and J. Samuelson.** 1997. Cloning and expression of chitinases of *Entamoebae*. Mol. Biochem. Parasitol. **85**:139–147.
12. **Diamond, L. S., D. R. Harlow, and C. C. Cunnick.** 1978. A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. Trans. R. Soc. Trop. Med. Hyg. **72**:431–432.
13. **Donelson, J. E.** 2003. Antigenic variation and the African trypanosome genome. Acta Trop. **85**:391–404.
14. **Drennan, D., and A. G. Ryazanov.** 2004. Alpha-kinases: analysis of the family and comparison with conventional protein kinases. Progress Biophys. Mol. Biol. **85**:1–32.
15. **Felsenstein, J.** 1989. PHYLIP—phylogeny inference package (version 3.2). Cladistics **5**:164–166.
16. **Frisardi, M., S. K. Ghosh, J. Field, K. Van Dellen, R. Rogers, P. Robbins, and J. Samuelson.** 2000. The most abundant glycoprotein of amebic cyst walls (Jacob) is a lectin with five Cys-rich, chitin-binding domains. Infect. Immun. **68**:4217–4224.
17. **Glockner, G., L. Eichinger, K. Szafranski, J. A. Pachebat, A. T. Bankier, P. H. Dear, R. Lehmann, C. Baumgart, G. Parra, J. F. Abril, R. Guigo, K. Kumpf, B. Tunggal, E. Cox, M. A. Quail, M. Platzer, A. Rosenthal, A. A. Noegel**, and the *Dictyostelium* Genome Sequencing Consortium. 2002. Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*. Nature **418**:79–85.
18. **Gong, D., Z. Gong, Y. Guo, and J. K. Zhu.** 2002. Expression, activation, and biochemical properties of a novel Arabidopsis protein kinase. Plant Physiol. **129**:225–234.
19. **Grass, S., and J. W. St. Geme III.** 2000. Maturation and secretion of the non-typable *Haemophilus influenzae* HMW1 adhesin: roles of the N-terminal and C-terminal domains. Mol. Microbiol. **36**:55–67.
20. **Hanks, S. K., and A. M. Quinn.** 1991. Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. Methods Enzymol. **200**:38–62.
21. **Haque, R., I. K. M. Ali, R. B. Sack, B. M. Farr, G. Ramakrishnan, and W. A. Petri, Jr.** 2001. Amebiasis and mucosal IgA antibody against the *Entamoeba histolytica* adherence lectin in Bangladeshi children. J. Infect. Dis. **183**:1787–1793.
22. **Haque, R., P. Duggal, I. K. M. Ali, M. B. Hossain, D. Mondal, R. B. Sack, B. M. Farr, T. H. Beaty, and W. A. Petri, Jr.** 2002. Innate and acquired resistance to amebiasis in Bangladeshi children. J. Infect. Dis. **186**:547–552.
23. **Hardie, D. G.** 1999. Protein phosphorylation: a practical approach, 2nd ed. Oxford University Press, Oxford, United Kingdom.
24. **Hegde, P., R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J. Earle-Hughes, E. Snesrud, N. Lee, and J. Quackenbush.** 2000. A concise guide to cDNA microarray analysis. BioTechniques **29**:548–562.
25. **Honma, K., H. K. Kuramitsu, R. J. Genco, and A. Sharma.** 2001. Development of a gene inactivation system for *Bacteroides forsythus*: Construction and characterization of a BspA mutant. Infect. Immun. **69**:4686–4690.
26. **Huber, M., L. Garfinkel, C. Gitler, D. Mirelman, M. Revel, and S. Rozenblatt.** 1988. Nucleotide sequence analysis of an *Entamoeba histolytica* ferredoxin gene. Mol. Biochem. Parasitol. **31**:27–33.
27. **Jaye, M., W. S. Modi, G. A. Ricca, R. Mudd, I. M. Chiu, S. J. O'Brien, and W. N. Drohan.** 1987. Isolation of a cDNA clone for the human laminin-B1 chain and its gene localization. Am. J. Hum. Genet. **41**:605–615.
28. **Kane, M. D., T. A. Jatkoe, C. R. Stumpf, J. Lu, J. D. Thomas, and S. J. Madore.** 2000. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucleic Acids Res. **28**:4552–4557.
29. **Kraemer, S. M., and J. D. Smith.** 2003. Evidence for the importance of genetic structuring to the structural and functional specialization of the *Plasmodium falciparum var* gene family. Mol. Microbiol. **50**:1527–1538.
30. **Krupa, A., G. Preethi, and N. Srinivasan.** 2004. Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in Ser/Thr and Tyr kinases. J. Mol. Biol. **339**:1025–1039.
31. **Laughlin, R., G. C. McGugan, R. P. Powell, B. H. Welter, and L. A. Temesvari.** 2004. Involvement of raft-like plasma membrane domains of *Entamoeba histolytica* in pinocytosis and adhesion. Infect. Immun. **72**:5349–5357.
32. **Leippe, M., and H. J. Müller-Eberhard.** 1994. The pore-forming peptide of *Entamoeba histolytica*, the protozoan parasite causing human amoebiasis. Toxicology **87**:5–18.
33. **Lucero, H. A., and P. W. Robbins.** 2004. Lipid rafts-protein association and the regulation of protein activity. Arch. Biochem. Biophys. **426**:208–224.
34. **Mai, Z., and J. Samuelson.** 1998. A new gene family (*ariel*) encodes asparagine-rich *Entamoeba histolytica* antigens, which resemble the amebic vaccine candidate serine-rich *E. histolytica* protein. Infect. Immun. **66**:353–355.
35. **Mann, B. J., B. E. Torian, T. S. Vedvick, and W. A. Petri, Jr.** 1991. Sequence of a cysteine-rich galactose-specific lectin of *Entamoeba histolytica*. Proc. Natl. Acad. Sci. USA **88**:3248–3252.
36. **McCoy, J. J., B. J. Mann, T. S. Vedvick, Y. Pak, D. B. Heimark, and W. A. Petri, Jr.** 1993. Structural analysis of the light subunit of the *Entamoeba histolytica* galactose-specific adherence lectin. J. Biol. Chem. **268**:24223–24231.
37. **McCoy, J. J., A. M. Weaver, and W. A. Petri, Jr.** 1994. Use of monoclonal anti-light subunit antibodies to study the structure and function of the *Entamoeba histolytica* Gal/GalNAc lectin. Glycoconjugate J. **11**:432–436.
38. **Nash, T. E.** 2002. Surface antigenic variation in *Giardia lamblia*. Mol. Microbiol. **45**:585–590.

(86)

39. **Nicholas, K. B., H. B. Nicholas, Jr., and D. W. Deerfield II.** 1997. GeneDoc: analysis and visualization of genetic variation. EMBnet News **4**:14.
40. **Nuckolls, G. H., Osherov, N., Loomis, W. F., and J. A. Spudich.** 1996. The *Dictyostelium* dual-specificity kinase splA is essential for spore differentiation. Development **122**:3295–3305.
41. **Pacheco, J., M. Shibayama, R. Campos, D. L. Beck, E. Houpt, W. A. Petri, Jr., and V. Tsutsumi.** 2004. *In vitro* and *in vivo* interaction of *Entamoeba histolytica* Gal/GalNAc lectin with various target cells: an immunocytochemical analysis. Parasitol Int. **53**:35–47.
42. **Papanastasiou, P., T. Bruderer, Y. Li, C. Bommeli, and P. Kohler.** 1997. Primary structure and biochemical properties of a variant-surface protein of *Giardia*. Mol. Biochem. Parasitol. **86**:13–27.
43. **Petri, W. A., Jr., M. D. Chapman, T. Snodgrass, B. J. Mann, J. Broman, and J. I. Ravdin.** 1989. Subunit structure of the galactose and *N*-acetyl-D-galactosamine-inhibitable adherence lectin of *Entamoeba histolytica*. J. Biol. Chem. **264**:3007–3012.
44. **Petri, W. A., Jr., R. Haque, and B. J. Mann.** 2002. The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite *Entamoeba histolytica*. Annu. Rev. Microbiol. **56**:39–64.
45. **Petri, W. A., Jr., R. D. Smith, P. H. Schlesinger, C. F. Murphy, and J. I. Ravdin.** 1987. Isolation of the galactose binding adherence lectin of *Entamoeba histolytica*. J. Clin. Investig. **80**:1238–1244.
46. **Purdy, J. E., B. J. Mann, E. C. Shugart, and W. A. Petri, Jr.** 1993. Analysis of the gene family encoding the *Entamoeba histolytica* galactose-specific adhesin 170-kDa subunit. Mol. Biochem. Parasitol. **62**:53–60.
47. **Radonic, A., S. Thulke, I. M. Mackay, O. Landt, W. Siegert, and A. Nitsche.** 2004. Guideline to reference gene selection for quantitative real-time PCR. Biochem. Biophys. Res. Commun. **313**:856–862.
48. **Ramakrishnan, G., B. D. Ragland, J. E. Purdy, and B. J. Mann.** 1996. Physical mapping and expression of gene families encoding the *N*-acetyl D-galactosamine adherence lectin of *Entamoeba histolytica*. Mol. Microbiol. **19**:91–100.
49. **Ravdin, J. I., C. F. Murphy, R. A. Salata, R. L. Guerrant, and E. L. Hewlett.** 1985. The *N*-Acetyl-D-galactosamine-inhibitable adherence lectin of *Entamoeba histolytica*. Partial purification and relation to amoebic virulence in vitro. J. Infect. Dis. **151**:804–815.
50. **Reverchon, S., C. Rouanet, D. Expert, and W. Nasser.** 2002. Characterization of indigoidine biosynthetic genes in *Erwinia chrysanthemi* and role of this blue pigment in pathogenicity. J. Bacteriol. **184**:654–665.
51. **Rifkin, M. R., and F. R. Landsberger.** 1990. Trypanosome variant surface glycoprotein transfer to target membranes: a model for the pathogenesis of trypanosomiasis. Proc. Natl. Acad. Sci. USA **87**:801–805.
52. **Roberts, D. J., A. G. Craig, A. R. Berendt, R. Pinches, G. Nash, K. Marsh, and C. I. Newbold.** 1992. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature **357**:689–692.
53. **Saito-Nakano, Y., M. Nakazawa, Y. Shigeta, T. Takeuchi, and T. Nozaki.** 2001. Identification and characterization of genes encoding novel Rab proteins from *Entamoeba histolytica*. Mol. Biochem. Parasitol. **116**:219–222.
54. **Saito-Nakano, Y., T. Yasuda, K. Nakada-Tsukui, and T. Nozaki.** 2004. Rab5-associated vacuoles play a unique role in phagocytosis of the enteric protozoan parasite *Entamoeba histolytica*. J. Biol. Chem. **279**:49497–49507.
55. **Schaenman, J. M., C. A. Gilchrist, B. J. Mann, and W. A. Petri, Jr.** 2001. Identification of two *Entamoeba histolytica* sequence-specific URE4 enhancer-binding proteins with homology to the RNA-binding motif RRM. J. Biol. Chem. **276**:1602–1609.
56. **Stanley, S. L., Jr., A. Becker, C. Kunz-Jenkins, L. Foster, and E. Li.** 1990. Cloning and expression of a membrane antigen of *Entamoeba histolytica* possessing multiple tandem repeats. Proc. Natl. Acad. Sci. USA **87**:4976–4980.
57. **Stetefeld, J., U. Mayer, R. Timpl, and R. Huber.** 1996. Crystal structure of three consecutive laminin-type epidermal growth factor-like (LE) modules of laminin γ1 chain harboring the nidogen binding site. J. Mol. Biol. **257**: 644–657.
58. **Tannich, E., F. Ebert, and R. D. Horstmann.** 1991. Primary structure of the 170 kDa surface lectin of pathogenic *Entamoeba histolytica*. Proc. Natl. Acad. Sci. USA **88**:1849–1853.
59. **Tannich, E., F. Ebert, and R. D. Horstmann.** 1992. Molecular cloning of cDNA and genomic sequences coding for the 35-kilodalton subunit of the galactose-inhibitable lectin of pathogenic *Entamoeba histolytica*. Mol. Biochem. Parasitol. **55**:225–228.
60. **Taylor, S. S.** 1989. cAMP-dependent protein kinase. J. Biol. Chem. **264**: 8443–8446.
61. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**:4876–4882.
62. **Tokoro, M., T. Asai, S. Kobayashi, T. Takeuchi, and T. Nozaki.** 2003. Identification and characterization of two isoenzymes of methionine gamma-lyase from *Entamoeba histolytica*: a key enzyme of sulphur-amino acid degradation in an anaerobic parasitic protist that lacks forward and reverse trans-sulfuration pathways. J. Biol. Chem. **278**:42717–42727.
63. **Tunggal, P., N. Smyth, M. Paulsson, and M. Ott.** 2000. Laminins: structure and genetic regulation. Microsc. Res. Tech. **51**:214–227.
64. **Udo, H., J. Munoz-Dorado, M. Inouye, and S. Inouye.** 1995. *Myxococcus xanthus*, a gram-negative bacterium, contains a transmembrane protein serine/threonine kinase that blocks secretion of beta-lactamase by phosphorylation. Genes Dev. **9**:972–983.
65. **van Dellen, K., S. K. Ghosh, P. W. Robbins, B. Loftus, and J. Samuelson.** 2002. *Entamoeba histolytica* lectins contain unique 6-Cys or 8-Cys chitin-binding domains. Infect. Immun. **70**:3259–3263.
66. **Wang, Z., J. Samuelson, C. G. Clark, D. Eichinger, J. Paul, K. van Dellen, N. Hall, I. Anderson, and B. Loftus.** 2003. Gene discovery in the *Entamoeba invadens* genome. Mol. Biochem. Parasitol. **129**:23–31.
67. **Zhai, Y., and M. H. Saier, Jr.** 2000. The amoebapore superfamily. Biochim. Biophys. Acta **1469**:87–99.
68. **Zhang, W., L. Li, W. Jiang, G. Zhao, Y. Yang, and J. Chiao.** 2000. A novel transmembrane serine/threonine protein kinase gene from a rifamycin SV-producing *Amycolatopsis mediterranei* U32. Eur. J. Biochem. **267**:3744–3752.

ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Experimental Parasitology 110 (2005) 244–252

Experimental Parasitology

www.elsevier.com/locate/yexpr

The diversity of Rab GTPases in *Entamoeba histolytica* ☆

Yumiko Saito-Nakano [a], Brendan J. Loftus [b], Neil Hall [b], Tomoyoshi Nozaki [c,d,*]

[a] *Department of Parasitology, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo 162-8640, Japan*
[b] *The Institute for Genomic Research, Rockville, MD 20850, USA*
[c] *Department of Parasitology, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan*
[d] *Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, 2-20-5 Akebono-cho, Tachikawa, Tokyo 190-0012, Japan*

Received 31 January 2005; received in revised form 31 January 2005; accepted 16 February 2005
Available online 7 April 2005

Abstract

Rab proteins are ubiquitous small GTP-binding proteins that form a highly conserved family and regulate vesicular trafficking. Recent completion of the genome of the enteric protozoan parasite *Entamoeba histolytica* enabled us to identify an extremely large number (>90) of putative Rab genes. Multiple alignment and phylogenic analysis of amebic, human, and yeast Rab showed that only 22 amebic Rab proteins including *Eh*Rab1, *Eh*Rab2, *Eh*Rab5, *Eh*Rab7, *Eh*Rab8, *Eh*Rab11, and *Eh*Rab21 showed significant similarity to Rab from other organisms. The 69 remaining amebic Rab proteins showed only moderate similarity (<40% identity) to Rab proteins from other organisms. Approximately one-third of Rab proteins including Rab7, Rab11, and RabC form 15 subfamilies, which contain up to nine isoforms. Approximately 70% of amebic Rab genes contain single or multiple introns, and this proportion is significantly higher than that of common genes in this organism. Twenty-five Rabs possess an atypical carboxyl terminus such as CXXX, XCXX, XXCX, XXXC, and no cysteine. We propose annotation of amebic Rab genes and discuss biological significance of this extraordinary diversity of *Eh*Rab proteins in this organism.
© 2005 Elsevier Inc. All rights reserved.

Index Descriptors and Abbreviations: DNA, deoxyribonucleic acid; *Eh*Rab, *Entamoeba histolytica* Rab; GTP, guanosine 5′-triphosphate

Keywords: *Entamoeba histolytica*; Membrane traffic; Phylogeny; Rab GTPase

1. Introduction

Small GTP-binding proteins are ubiquitous molecular switches involved in a variety of important cellular processes including cell proliferation, cytoskeletal assembly, and intracellular membrane trafficking in all eukaryotes. This superfamily is classified into five families: Ras, Rho/Rac, Rab, Sar/Arf, and Ran families based on its primary sequences (Bourne et al., 1990; Takai et al., 2001). Rab GTPases constitute the largest group of this superfamily and known as essential regulators of vesicular transport pathways (Novick and Zerial, 1997). In general, but not necessarily always, the complexity of Rab genes correlate with multicellularity of organisms. The higher the number of cells consisting of an organism, the higher the number of Rab genes encoded in its genome is. For example, unicellular yeast *Saccharomyces pombe*, *Saccharomyces cerevisiae*, and a nematode *Caenorhabditis elegans*, a fruit fly *Drosophila melanogaster*, or human *Homo sapiens*, which consists of one, $\sim 10^3$, 10^9, or 10^{13} cells, have 7, 11, 29, 29, or 60 Rab genes, respectively (Pereira-Leal and Seabra, 2001). In multicellular organisms, several Rab proteins are expressed in a highly coordinated (i.e., tissue-, organ- or developmental

☆ Nucleotide sequence data reported in this paper have been submitted to the DDBJ/EBI/GenBank Data Bank with Accession Nos. AB197055 to AB197121.

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

0014-4894/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2005.02.021

stage-specific) fashion (Seabra et al., 2002; Zerial and McBride, 2001).

The enteric protozoan parasite *E. histolytica* is a etiological agent of amoebiasis, causing an estimated 50 million cases of amebic colitis, dysentery, and extraintestinal abscesses (Petri, 2002) and 40,000–100,000 deaths annually (WHO/PAHO/UNESCO, 1997). The size of the recently completed genome of *E. histolytica* (Loftus et al., 2005) is 20 Mb, which is about 1.6 times of *S. cerevisiae* and one fifth of *C. elegans*. We and other groups have previously reported 16 *Eh*Rab genes, identified by cDNA isolation using degenerate PCR primers and homology-based search (Juarez et al., 2001; Rodriguez et al., 2000; Saito-Nakano et al., 2001, 2004; Temesvari et al., 1999). However, the complexity of Rab genes in *E. histolytica* remains totally unknown. In this paper, we describe identification of additional 75 *Eh*Rab genes by thorough search of the latest genome database. To our knowledge, this is the first demonstration of a uni- or multi-cellular organism possessing more than 90 Rab genes. Based on our analysis, we propose annotation of amebic Rab genes.

2. Methods

Nucleotide and protein sequences of human and yeast Rab were retrieved from GenBank. For accession numbers of these proteins, see (Pereira-Leal and Seabra, 2001). To obtain a maximum number of Rab genes from *E. histolytica*, we first searched for putative homologues of all yeast and human small GTPases including Rab, Ras/Rap, Rho/Rac, Arf/Sar, and Ran against a translated protein subset of the *E. histolytica* genome database (http://www.tigr.org/tdb/e2k1/eha1/) using BLASTP algorism. Additionally, we used *E. histolytica*-specific members of Rab, Ras, and Rho/Rac, e.g., RabA (Temesvari et al., 1999), Ras4 (Kumagai et al., 2004), and RacD (Lohia and Samuelson, 1996), as an inquiry. Altogether, we obtained 146 possible small GTPases. All the possible *Entamoeba* Rab protein sequences were re-examined with BLASTP analysis using individual amebic protein as an inquiry sequence against the human database at National Center of Biotechnology Information (NCBI). Among 146 possible small GTPases, 55 sequences were tentatively assigned as Sar/Arf, Ras/Rap, Rho/Rac, or Ran, as they showed highest identities to these small GTPase subfamilies from human. One should be cautious that due to the method we used to retrieve sequences, it is possible additional small GTPases may exist in *E. histolytica*. The 91 remaining small GTPases were individually verified for the presence of conserved GTP-binding consensus sequences (GDXXVGKT, DTAGQE, and GNKXD) and additional five conserved regions that are specific only to Rab family (IGVDF, KLQIW, RFRSIT, YYRGA, and LVYDIT) (Pereira-Leal and Seabra, 2000) by manual inspections. In the present study, these 91 putative Rab are further analyzed. Highly conserved domains stretching from the first to the third GTP-binding consensus regions of 91 *E. histolytica*, 7 *S. cerevisiae*, and 35 human Rab proteins were aligned using the CLUSTAL W program version 1.81 (Thompson et al., 1994) with default parameters. After alignments were manually corrected and non-aligned gaps were removed, 110 unambiguously aligned sites were selected and used for phylogenetic analysis by the neighbor-joining method (Saitou and Nei, 1987). Phylogenetic trees were drawn using the TreeView software (http://taxonomy.zoology.gla.ac.uk/rod/rod.html).

3. Results and discussion

3.1. Identification and annotation of 91 EhRab proteins

We identified 91 putative Rab proteins in the *E. histolytica* genome database (Loftus et al., 2005). They included 16 previously reported Rab proteins consisting of putative amebic homologues of human Rab1, Rab2, Rab5, Rab7, Rab8, Rab11, and Rab proteins that showed limited homology (<40% identity) to Rab from other organisms and were annotated as *Eh*RabA to *Eh*RabI (Juarez et al., 2001; Rodriguez et. al., 2000; Saito-Nakano et al., 2001, 2004; Temesvari et al., 1999). Twenty-two of 91 amebic Rab showed >40% identity to human Rab1/yeast Ypt1p, human Rab2/Rab4/Rab14, Rab5/Ypt5p, Rab7/Ypt7p, Rab8/Sec4p, Rab11/Ypt31p, or human Rab21 and were considered to be their amebic homologues. We accordingly designated them as *Eh*Rab1A-1B, *Eh*Rab2A-2C, *Eh*Rab5, *Eh*Rab7A-7I, *Eh*Rab8-8B, *Eh*Rab11A-11D, and *Eh*Rab21, where the alphabet after the number represents an individual isotype, in case where their overall identity to human or yeast homologues was >40% (Tables 2–6). Among the remaining 69 *E. histolytica*-specific Rab proteins, 30 Rab proteins showed >40% mutual identity to one or more of these 30 amebic Rab and clustered in nine subgroups designated as *Eh*RabC, *Eh*RabD, *Eh*RabF, *Eh*RabI, EhRabK, *Eh*RabL, *Eh*RabM, *Eh*RabN, and *Eh*RabP subfamilies (Tables 7–10). The residual 39 Rab, including previously reported 3 *Eh*Rab, i.e., *Eh*RabA, *Eh*RabB, and *Eh*RabH, represent novel and solitary amebic Rab proteins showing low homology to Rab from human, yeast, and to other members of the amebic Rab. Accordingly we designated these newly identified solitary Rab as *Eh*RabX1 to *Eh*RabX36 (Fig. 1 and Table 1).

3.2. EhRab proteins showing significant homology to human or yeast counterparts

Entamoeba histolytica possesses at least nine isotypes of Rab7 (*Eh*Rab7A–*Eh*Rab7I, Fig. 1, Table 5), which has

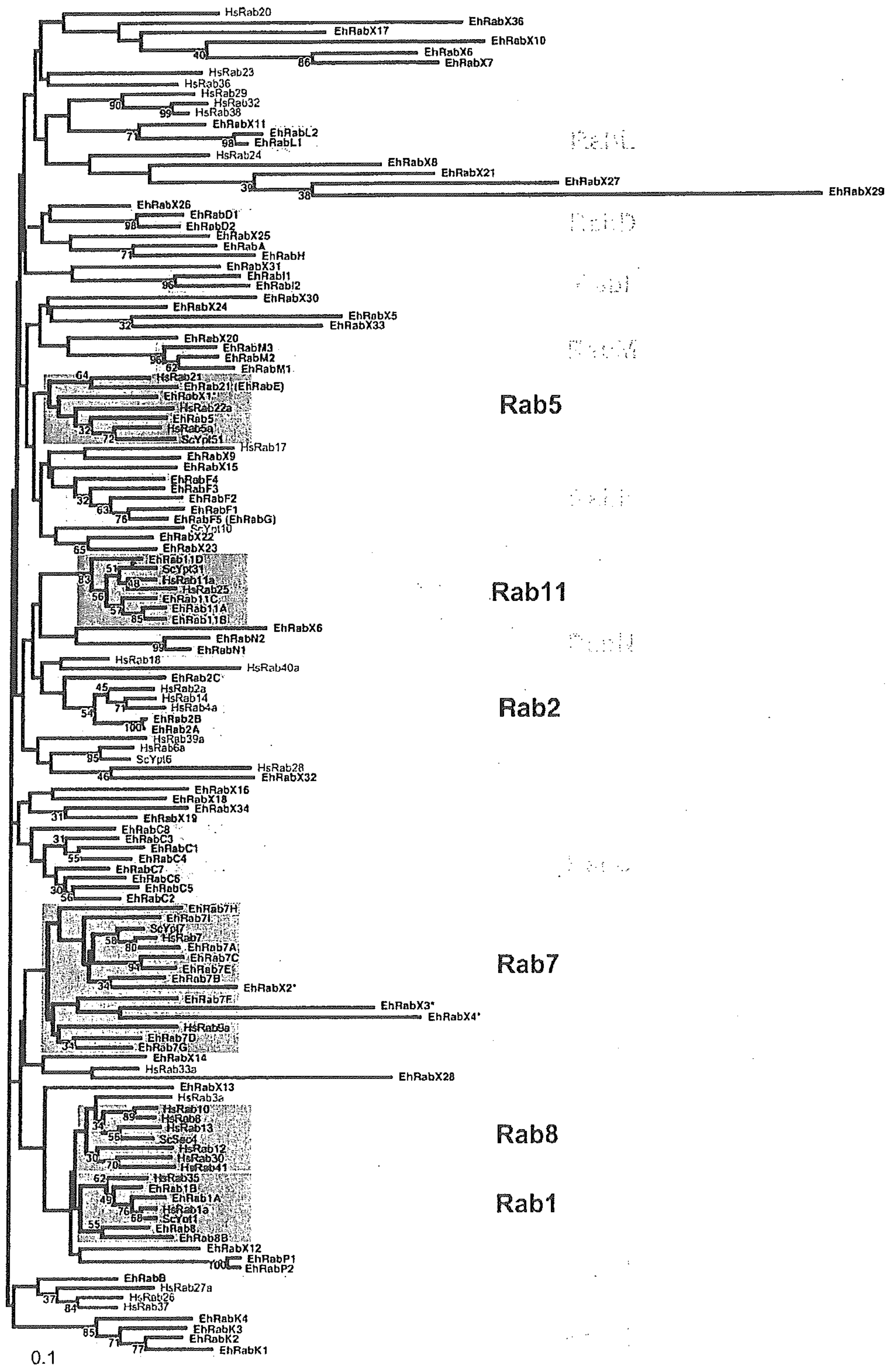

Fig. 1. A phylogenic tree of Rab proteins from *E. histolytica*, human, and yeast. The numbers on the nodes represent the bootstrap proportions (%) of 1000 pseudo samples; only bootstrap proportions >30% are shown. *E. histolytica* Rab proteins are indicated in bold. Tentative subfamilies that revealed significant homology (>40% identity) to human or yeast counterpart are shaded by dark boxes, while *Entamoeba*-specific subfamilies are highlighted by light boxes. Asterisks indicate Rab proteins that were not considered as isotypes of the subfamily based on <40% identity to other members of the subfamily. The scale bar indicates 0.1 substitutions at each amino acid position.

Table 1
A list of 91 *Eh* Rab proteins and their features of nucleotide and amino acid sequences

Name	Previous name	Accession No.	No. of intron	Missing regions	C-terminal peptides	References
Eh Rab1A	*Eh* Rab1	AB054578	1		CXXX	3,6
Eh Rab1B		AB197055	3		XXCC	
Eh Rab2A		AB197071	1		—	
Eh Rab2B		AB197072	1		—	
Eh Rab2C		AB197073	1		—	
Eh Rab5		AB054582	1		XXCC	5
Eh Rab7A	*Eh* Rab7	AB054583	—		XCXC	1,5,6
Eh Rab7B		AB186363	1		XCXC	6
Eh Rab7C		AB186364	1		XXCC	6
Eh Rab7D		AB186365	—		XXCC	6
Eh Rab7E		AB186366	1		XXCC	6
Eh Rab7F		AB186367	1		XXCC	
Eh Rab7G		AB186368	—		XXCC	
Eh Rab7H		AB186369	1		XXCC	
Eh Rab7I		AB197056	3		XXCC	
Eh Rab8		AF363067	1		XXCC	4
Eh Rab8B		AB197057	1		XXCC	
Eh Rab11A	*Eh* Rab11	AB186370	1		XXCC	1
Eh Rab11B		AB054587	—		XXCC	3
Eh Rab11C		AB054588	—		XXCC	3
Eh Rab11D		AB197058	1		XCXC	
Eh Rab21	*Eh* RabE	AB054581	1	switch I	XXCC	3
Eh RabA		AF030184	2		XCXC	1
Eh RabB		AF127375	—		XXCC	2
Eh RabC1	*Eh* RabC	AB054579	—		XXCC	3,6
Eh RabC2		AB186371	—		XXCC	
Eh RabC3		AB197059	—		XXCC	
Eh RabC4		AB197060	—		XXCC	
Eh RabC5		AB197061	1		XXCC	
Eh RabC6		AB197062	2		XXCC	
Eh RabC7		AB197063	2		XXCC	
Eh RabC8		AB197064	1		XXCC	
Eh RabD1	*Eh* RabD	AB054580	—		XXCC	3
Eh RabD2		AB197065	—		XXCC	
Eh RabF1	*Eh* RabF	AB054584	1		CXXX	3
Eh RabF2		AB197068	3		XXCC	
Eh RabF3		AB197067	2		XXCC	
Eh RabF4		AB197066	1		XXCC	
Eh RabF5	*Eh* RabG	AB054585	1		XXXC	3
Eh RabH		AB054586	—	switch II	XCXC	3
Eh RabI1	*Eh* RabI	AB054589	2	switch I, II	XCXC	3
Eh RabI2		AB197069	1		XCXC	
Eh RabK1		AB197077	—		XXCC	
Eh RabK2		AB197079	1		XXCC	
Eh RabK3		AB197078	3		XXCC	
Eh RabK4		AB197120	1	switch I	XXCC	
Eh RabL1		AB197081	1		XCCX	
Eh RabL2		AB197080	1		XCCX	
Eh RabM1		AB197083	4	box 2, 3	XXCC	
Eh RabM2		AB197084	—		XXCC	
Eh RabM3		AB197082	2		XXCC	
Eh RabN1		AB197085	2		XCXX	
Eh RabN2		AB197086	1		XXXC	
Eh RabP1		AB197087	—		CXXX	
Eh RabP2		AB197088	1		CXXX	
Eh RabX1		AB197070	3		XXCC	
Eh RabX2		AB197074	—		XXCC	
Eh RabX3		AB197075	3	switch I, II	—	
Eh RabX4		AB197076	—	switch I	XXCX	
Eh RabX5		AB197089	1	switch II	XXXC	
Eh RabX6		AB197090	1	switch II	CXXX	
Eh RabX7		AB197091	1	switch I, II	CXXX	

(*continued on next page*)

Table 1 (*continued*)

Name	Previous name	Accession No.	No. of intron	Missing regions	C-terminal peptides	References
Eh RabX8		AB197092	—	switch I, II, box 2, 3	CCXX	
Eh RabX9		AB197093	3		XXCC	
Eh RabX10		AB197094		switch I, II	CCXXX	
Eh RabX11		AB197095	2		XXCC	6
Eh RabX12		AB197096	2		XXCC	
Eh RabX13		AB197097	1		XXCC	
Eh RabX14		AB197098	1		XXCC	
Eh RabX15		AB197099	1		XXCC	
Eh RabX16		AB197100	—		XXCC	
Eh RabX17		AB197101	—	switch I	CXXX	
Eh RabX18		AB197102	1		CXXX	
Eh RabX19		AB197103	2		XXCC	
Eh RabX20		AB197104	—		XXCC	
Eh RabX21		AB197105	1	switch II	CXXXXX	
Eh RabX22		AB197106	2		XXCC	
Eh RabX23		AB197107	2		XXCC	
Eh RabX24		AB197108	1	switch I	XXCC	
Eh RabX25		AB197109	—		XXCC	
Eh RabX26		AB197110	1		XXCC	
Eh RabX27		AB197111	1	box 2, 3	CXXX	
Eh RabX28		AB197112	—	box 1, 3	CXXX	
Eh RabX29		AB197113	1	switch I, II, box 2	CXXX	
Eh RabX30		AB197114	1		CXXX	
Eh RabX31		AB197115	1		XXCC	
Eh RabX32		AB197116	2	box 2	CXXXXX	
Eh RabX33		AB197117	2	switch I	CCXXX	
Eh RabX34		AB197118	1		CXXX	
Eh RabX35		AB197119	—		XXCC	
Eh RabX36		AB197121	—	switch I	—	

Reference 1, Temesvari et al. (1999); 2, Rodriguez et al. (2000); 3, Saito-Nakano et al. (2001); 4, Juarez et al. (2001); 5, Saito-Nakano et al. (2004); and 6, Okada et al. (2005).

no precedent in unicellular eukaryotes as well as fly, worm, and mammals (Pereira-Leal and Seabra, 2001). In human, Rab7 and related Rab9 were shown to be involved in late endocytic trafficking (Rodman and Wandinger-Ness, 2000). We previously showed in *E. histolytica* Rab7A plays a role in the biogenesis of the unique organelle to this organism ("prephagosomal vacuole") and also in the processing, storage, and transport to phagosomes of digestive proteins (Saito-Nakano et al., 2004). It was also shown that *Eh*Rab7A was concentrated in endosome-enriched fraction (Temesvari et al., 1999). At least five *Eh*Rab7 isotypes (*Eh*Rab7A-7E) were identified by proteomic analysis of latex bead-containing phagosomes, suggesting the involvement of multiple Rab7 isotypes in phagosome biogenesis (Okada et al., 2005). Although three other Rab proteins, *Eh*RabX2-X4 were clustered in a same clade in the phylogenetic reconstruction, they were not considered as Rab7 isotypes based on the fact that they showed <40% identities to any Rab7 isotypes from *E. histolytica*, yeast, or human (Table 5).

In general Rab11/Ypt31p function in the recycling of membrane proteins on recycling endosomes (Rodman and Wandinger-Ness, 2000) and form a cluster together with human Rab25 (Pereira-Leal and Seabra, 2001), which showed similar subcellular localization to Rab11 (Casanova et al., 1999). *E. histolytica* possesses four

Table 2
Amino acid identities among Rab1 and closely related homologues from *E. histolytica*, human, and yeast

	EhRab1A	EhRab1B	HsRab1a	ScYpt1	HsRab35	EhRab8	EhRab8B	HsRab8a	ScSec4
EhRab1A	100	56	60	55	47	47	40	46	45
EhRab1B		100	58	56	53	49	49	54	49
HsRab1a			100	71	54	51	42	52	50
ScYpt1				100	49	51	42	48	47
HsRab35					100	43	41	45	42
EhRab8						100	54	50	47
EhRab8B							100	42	42
HsRab8a								100	50
ScSec4									100

Eh, *E. histolytica*; Hs, *Homo sapiens*; and Sc, *Saccharomyces cerevisiae*.

Table 3
Amino acid identities among Rab2 and closely related homologues from *E. histolytica* and human

	EhRab2A	EhRab2B	EhRab2C	HsRab2a	HsRab14	HsRab4
EhRab2A	100	89	46	51	46	40
EhRab2B		100	45	51	46	42
EhRab2C			100	37	38	35
HsRab2a				100	56	51
HsRab14					100	59
HsRab4						100

Values less than 40% are shown in reverse.

Rab11 homologues (*Eh*Rab11A-11D) including *Eh*Rab11A, which was proposed to play a role during encystation (McGugan and Temesvari, 2003) and previously identified *Eh*Rab11B and *Eh*Rab11C (Saito-Nakano et al., 2001; Temesvari et al., 1999) (Table 6).

Despite significant similarity to human and yeast homologues (51 and 43% identity, respectively), which are localized to endosomes (Zerial and McBride, 2001), Rab5 was shown to be excluded from endocytic pathway in *E. histolytica* (Saito-Nakano et al., 2004). We also previously showed that *Eh*Rab5 is involved, together with *Eh*Rab7A, in the initial phase of biogenesis of prephagosomal vacuoles, and is essential for efficient phagocytosis (Saito-Nakano et al., 2004). In contrast to yeast, where three isotypes with apparently redundant function exist, no additional *Eh*Rab5 isotype is present in this organism (Fig. 1). Previously reported *Eh*RabE (Saito-Nakano et al., 2001) showed 47% identity to human Rab21, which localizes to early endosomes like Rab5 and Rab22 (Simpson et al., 2004; Table 4). We redesignated *Eh*RabE as *Eh*Rab21 based on the criteria described above.

We identified four putative Rab1/Rab8 homologues (Table 2), which include previously described *Eh*Rab1 (AB054578), showing 60% identity to human Rab1a, and *Eh*Rab8 (AF363067), which shows 42% identity to human Rab8a (Juarez et al., 2001; Saito-Nakano et al., 2001) and was proposed to be involved in the targeting of vesicles to plasma membrane (Juarez et al., 2001). We designated the other two Rab proteins as *Eh*Rab1B and *Eh*Rab8B as the former shows highest similarity to human Rab1a and the latter shows highest identity to *Eh*Rab8.

We identified three putative Rab2/Rab4/Rab14 homologues (Table 3). Although human Rab2, Rab4, and Rab14 belong to one subfamily (Pereira-Leal and Seabra, 2001), they reveal distinct localizations and functions Rab2, Rab4, or Rab14 is localized in the ER-to-Golgi, early endosomes, or Golgi-to-endosomes (Junutula et al., 2004; Mohrmann et al., 2002; Tisdale, 1999).

3.3. EhRab proteins with no obvious homologues from other organisms

Sixty-nine of 91 *Eh*Rab proteins showed <40% identity to human and yeast Rab, are presumed to be unique to this organism. Thirty of these ameba-specific Rabs also form nine subfamilies (Tables 7–10) and designated as e.g., *Eh*RabC3, where the alphabet represents a subfamily and the number after the alphabet represents an individual isotype. The largest *E. histolytica*-specific group (only after *Eh*Rab7 subfamily described above), *Eh*RabC subfamily, consists of eight isotypes, *Eh*RabC1-C8 (Table 7). *Eh*RabC1-C3 were identified from isolated phagosomes (C1, Okada et al., 2005; C2 and C3, Okada and Nozaki, unpublished), suggesting that these three members of *Eh*RabC subfamily are involved in phagosome biogenesis. Four isotypes of *Eh*RabF (*Eh*RabF1-F4) as well as *Eh*RabG were grouped in the same subfamily, *Eh*RabG was renamed as *Eh*RabF5 (Tables 1 and 8). Other Rab proteins were also grouped in subfamilies: *Eh*RabD1-D2, *Eh*RabI1-I2,, *Eh*RabK1-K4, *Eh*RabL1-L2, *Eh*RabM1-M3, *Eh*RabN1-N2, and *Eh*RabP1-P2 (Tables 9 and 10).

Table 4
Amino acid identities among Rab5 and closely related homologues from *E. histolytica*, human, and yeast

	EhRab5	HsRab5a	ScYpt51	HsRab22a	EhRabX1	HsRab21	EhRab21/E
EhRab5	100	51	43	44	30	40	36
HsRab5a		100	48	47	36	38	29
ScYpt51			100	39	36	37	32
HsRab22a				100	37	39	30
EhRabX1					100	33	30
HsRab21						100	47
EhRab21/E							100

Values less than 40% are shown in reverse.

Table 5
Amino acid identities among Rab7 and closely related homologues from *E. histolytica*, human, and yeast

	EhRab7A	EhRab7B	EhRab7C	EhRab7D	EhRab7E	EhRab7F	EhRab7G	EhRab7H	EhRab7I	EhRabX2	EhRabX3	EhRabX4	HsRab7a	HsRab9	ScYpt7	EhRab5	HsRab5a
EhRab7A	100	45	44	39	45	41	36	32	35	37	16	21	56	46	54	27	29
EhRab7B		100	46	49	49	38	36	35	46	39	15	25	49	42	46	27	29
EhRab7C			100	46	64	42	43	40	42	36	15	18	51	46	46	32	28
EhRab7D				100	46	38	51	40	36	36	35	12	47	45	42	34	29
EhRab7E					100	40	43	33	39	35	15	19	46	42	46	29	30
EhRab7F						100	36	33	35	30	10	23	42	36	45	30	30
EhRab7G							100	37	38	31	18	21	40	40	37	31	31
EhRab7H								100	36	25	24	15	35	32	32	38	28
EhRab7I									100	32	15	18	41	33	37	35	33
EhRabX2										100	11	19	38	33	36	26	21
EhRabX3											100	18	18	17	22	22	18
EhRabX4												100	22	19	22	21	15
HsRab7a													100	48	62	34	34
HsRab9														100	47	31	27
ScYpt7															100	37	28
EhRab5																100	51
HsRab5a																	100

Values less than 40% are shown in reverse. EhRab5 and HsRab5a are included as control.

3.4. Peculiarity of nucleotide and protein sequences

Approximately 70% (64) of amebic *Rab* genes contained introns, and 23% (22) of amebic *Rab* genes contained 2–4 introns (Table 1). Considering that among 9938 predicted open reading frames of the entire *Entamoeba* genome only 25.2% of genes have introns and 6% of these genes contained multiple introns (Loftus et al., 2005), and that the average size of amebic *Rab* genes is 702 bp, which is shorter than that of all genes (1173 bp), indicating that amebic *Rab* genes are extremely intron rich.

Rab GTPases typically have three GTP-binding consensus, Rab-specific effector (also called "Switch I"), and α2 helix ("Switch II") regions (Pereira-Leal and Seabra, 2000). The "Switch I and II" regions are located on the surface of the molecule as demonstrated in structural studies, and crucial for the interaction with regulatory proteins (effectors) such as guanine nucleotide exchange factors and GTPase-activating proteins (Stenmark and Olkkonen, 2001). While the GTP-binding consensus regions (boxes 1–3) are conserved in the small GTPase superfamily, the "Switch I and II" regions are unique to Rab GTPases. About 20% of amebic Rab proteins lack one or more of these functional regions (Table 1). We have a line of evidence supporting that these *Eh*Rab genes are functional GTPase. First, *Eh*Rab21, *Eh*RabH, and *Eh*RabI1, which lack one or both of Switch I and II regions, are expressed as mRNA in the axenic trophozoites, which excludes a possibility of pseudogenes (Saito-Nakano et al., 2001). In addition, *Eh*Ras4, which lacks the Switch I region and one of the GTP-binding consensus regions, showed a GTP-binding activity in vitro (Kumagai et al., 2004). These data suggest that amebic Rab that lack these conserved regions are likely functional, and that it is not easy, in general, to predict functionality of these important regions and domains of the amebic small GTPases by the primary protein sequences.

It has been well established that Rab proteins are lipid-modified at the carboxyl terminus by isoprenylation with Rab geranylgeranyl transferase (geranylgeranyl transferase II) (Takai et al., 2001). Prenylation motifs located at the carboxyl terminus of Rab generally consist of two cysteine residues such as XXCC, XCCX, CCXX, and XCXC (Takai et al., 2001). While the majority of amebic Rab proteins contain two cysteine residues at the carboxyl terminus, about 13% (13) possess a single cysteine, such as CXXX, which was often found in Ras and Rho/Rac subfamilies (Table 1). Five amebic Rab proteins totally lack cysteine at the carboxyl terminus. It has been reported that a number of human Rab also lack the typical carboxyl termini; 8 Rab possess no cysteine and other 8 Rab contain a single cysteine (Pereira-Leal and Seabra, 2001). Interestingly, seven amebic Rab proteins contain novel types of carboxyl termini containing a single cysteine: XCXX, XXCX, XXXC or

Table 6
Amino acid identities among Rab11 and closely related homologues from *E. histolytica*, human, and yeast

	EhRab11A	EhRab11B	EhRab11C	EhRab11D	HsRab11a	ScYpt31	HsRab25
EhRab11A	100	63	58	55	57	42	51
EhRab11B		100	56	58	54	43	48
EhRab11C			100	51	51	42	42
EhRab11D				100	51	40	42
HsRab11a					100	44	57
ScYpt31						100	42
HsRab25							100

Table 7
Amino acid identities among RabC isotypes from *E. histolytica*

	EhRabC1	EhRabC2	EhRabC3	EhRabC4	EhRabC5	EhRabC6	EhRabC7	EhRabC8
EhRabC1	100	44	55	49	41	42	41	37
EhRabC2		100	46	41	57	53	48	48
EhRabC3			100	49	44	45	40	34
EhRabC4				100	40	40	41	39
EhRabC5					100	49	44	43
EhRabC6						100	52	50
EhRabC7							100	49
EhRabC8								100

Values less than 40% are shown in reverse.

Table 8
Amino acid identities among RabF and related homologues from *E. histolytica* and human

	EhRabF	EhRabF5/G	EhRabF2	EhRabF3	EhRabF4	EhRabX21	HsRab17	EhRabX15
EhRabF	100	59	44	44	40	33	33	31
EhRabF5/G		100	43	45	41	38	31	33
EhRabF2			100	45	48	32	29	32
EhRabF3				100	47	38	32	27
EhRabF4					100	39	35	35
EhRabX21						100	29	28
HsRab17							100	26
EhRabX15								100

Values less than 40% are shown in reverse.

Table 9
Amino acid identities among RabK isotypes in *E. histolytica*

	EhRabK1	EhRabK2	EhRabK3	EhRabK4
EhRabK1	100	60	47	39
EhRabK2		100	45	41
EhRabK3			100	37
EhRabK4				100

Values less than 40% are shown in reverse.

Table 10
Pairwise amino acid identities between isotypes that belong to RabD, RabI, RabL, RabM, RabN, and RabP subfamilies

Subfamilies	
EhRabD1:EhRabD2	60
EhRabI1:EhRabI2	52
EhRabL1:EhRabL2	62
EhRabM1:EhRabM2	56
EhRabM2:EhRabM3	65
EhRabM1:EhRabM3	54
EhRabN1:EhRabN2	75
EhRabP1:EhRabP2	85

CXXXXX, which have not been reported in human (Pereira-Leal and Seabra, 2001).

Finally, the diversity and complexity of Rab proteins we report here may reflect the vigorous dynamism of the membrane transport and the committed reliance on Rab proteins in the determination of specificity of vesicular trafficking in this unicellular protozoan parasite. In addition, the presence of unique structural exceptions including lack of conserved functional boxes and domains, and non-conventional carboxyl terminus found in some Rab proteins may suggest that novel Rab modifications and functions exist in this organism. Further analysis of such peculiarities may help in the elucidation of functional constrains during evolution and the ameba-specific biological mechanisms, which are likely associated with its unique virulent competence.

Acknowledgments

We are grateful to Tetsuo Hashimoto, Tsukuba University, for many helpful comments and discussions. We thank Mami Okada, Mai Nudejima, and Yumiko Tsukamoto for their technical assistance. This work was supported in part by a grant for Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency, Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan to Y.S-N. (15790219) and T.N. (16017307, 16044250, and 15590378), and a grant for the Project to Promote Development of Anti-AIDS Pharmaceuticals from the Japan Health Sciences Foundation to T.N.

References

Bourne, H.R., Sanders, D.A., McCormick, F., 1990. The GTPase superfamily: a conserved switch for diverse cell functions. Nature 348, 125–132.

Casanova, J.E., Wang, X., Kumar, R., Bhartur, S.G., Navarre, J., Woodrum, J.E., Altschuler, Y., Ray, G.S., Goldenring, J.R., 1999. Association of Rab25 and Rab11a with the apical recycling system of polarized Madin–Darby canine kidney cells. Molecular Biology of the Cell 10, 47–61.

Juarez, P., Sanchez-Lopez, R., Stock, R.P., Olvera, A., Ramos, M.A., Alagon, A., 2001. Characterization of the *Eh*Rab8 gene, a marker of the late stages of the secretory pathway of *Entamoeba histolytica*. Molecular and Biochemical Parasitology 116, 223–228.

Junutula, J.R., De Maziere, A.M., Peden, A.A., Ervin, K.E., Advani, R.J., van Dijk, S.M., Klumperman, J., Scheller, R.H., 2004. Rab14 is involved in membrane trafficking between the Golgi complex and endosomes. Molecular Biology of the Cell 15, 2218–2229.

Kumagai, M., Makioka, A., Takeuchi, T., Nozaki, T., 2004. Molecular cloning and characterization of a protein farnesyltransferase from the enteric protozoan parasite *Entamoeba histolytica*. Journal of Biological Chemistry 279, 2316–2323.

Loftus, B., Anderson, I., Davies, R., Alismark, U.C.M., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J. Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Guillén, N. Gilchrist, C., Stroup, S.E., Bhattacharya, S., Lohia, A., Foster, P.G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., Petri, Jr., W.A., Clark, C.G., Embley, T.M., Barrell, B., Fraser, C.M., Hall, N., 2005. The genome of the protist parasite *Entamoeba histolytica*. Nature 433, 865–868.

Lohia, A., Samuelson, J., 1996. Heterogeneity of *Entamoeba histolytica* rac genes encoding p21rac homologues. Gene 173, 205–208.

McGugan, G.C.J., Temesvari, L.A., 2003. Characterization of a Rab11-like GTPase, *Eh*Rab11, of *Entamoeba histolytica*. Molecular and Biochemical Parasitology 129, 137–146.

Mohrmann, K., Gerez, L., Oorschot, V., Klumperman, J., van der Sluijs, P., 2002. Rab4 function in membrane recycling from early endosomes depends on a membrane to cytoplasm cycle. Journal of Biological Chemistry 277, 32029–32035.

Novick, P., Zerial, M., 1997. The diversity of Rab proteins in vesicle transport. Current Opinion in Cell Biology 9, 496–504.

Okada, M., Huston, C.D., Mann, B.J., Petri, W.A.J., Kita, K., Nozaki, T., 2005. Proteomic analysis of phagocytosis in the enteric protozoan parasite *Entamoeba histolytica*. Eukaryotic Cell, in press.

Pereira-Leal, J.B., Seabra, M.C., 2000. The mammalian Rab family of small GTPases: definition of family and subfamily sequence motifs suggests a mechanism for functional specificity in the Ras superfamily. Journal of Molecular Biology 301, 1077–1087.

Pereira-Leal, J.B., Seabra, M.C., 2001. Evolution of the Rab family of small GTP-binding proteins. Journal of Molecular Biology 313, 889–901.

Petri Jr., W.A., 2002. Pathogenesis of amebiasis. Current Opinion in Microbiology 5, 443–447.

Rodman, J.S., Wandinger-Ness, A., 2000. Rab GTPases coordinate endocytosis. Journal of Cell Science 113, 183–192.

Rodriguez, M.A., Garcia-Perez, R.M., Garcia-Rivera, G., Lopez-Reyes, I., Mendoza, L., Ortiz-Navarrete, V., Orozco, E., 2000. An *Entamoeba histolytica* rab-like encoding gene and protein: function and cellular location. Molecular and Biochemical Parasitology 108, 199–206.

Saitou, N., Nei, M., 1987. The neighbour-joining method: a new method to reconstruct phylogenetic trees. Molecular Biology of Evolution 4, 406–425.

Saito-Nakano, Y., Nakazawa, M., Shigeta, Y., Takeuchi, T., Nozaki, T., 2001. Identification and characterization of genes encoding novel Rab proteins from *Entamoeba histolytica*. Molecular and Biochemical Parasitology 116, 219–222.

Saito-Nakano, Y., Yasuda, T., Nakada-Tsukui, K., Leippe, M., Nozaki, T., 2004. Rab5-associated vacuoles play a unique role in phagocytosis of the enteric protozoan parasite *Entamoeba histolytica*. Journal of Biological Chemistry 279, 49497–49507.

Seabra, M.C., Mules, E.H., Hume, A.N., 2002. Rab GTPases, intracellular traffic and disease. Trends in Molecular Medicine 8, 23–30.

Simpson, J.C., Griffiths, G., Wessling-Resnick, M., Fransen, J.A., Bennett, H., Jones, A.T., 2004. A role for the small GTPase Rab21 in the early endocytic pathway. Journal of Cell Science 117, 6297–6311.

Stenmark, H., Olkkonen, V.M., 2001. The Rab GTPase family. Genome Biology 2, REVIEWS3007.1-3007.7.

Takai, Y., Sasaki, T., Matozaki, T., 2001. Small GTP-binding proteins. Physiological Reviews 81, 153–208.

Temesvari, L.A., Harris, E.N., Stanley, S.L.J, Cardelli, J.A., 1999. Early and late endosomal compartments of *Entamoeba histolytica* are enriched in cysteine proteases, acid phosphatase and several Ras-related Rab GTPases. Molecular and Biochemical Parasitology 103, 225–241.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the. sensitivity of progressive multiple sequence alignment through sequence weighting, positive-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

Tisdale, E.J., 1999. A Rab2 mutant with impaired GTPase activity stimulates vesicle formation from pre-Golgi intermediates. Molecular Biology of the Cell 10, 1837–1849.

WHO/PAHO/UNESCO. 1997. A consultation with experts on amebiasis. Epidemiological Bulletin 18, 13–14.

Zerial, M., McBride, H., 2001. Rab proteins as membrane organizers. Nature Reviews Molecular Cell Biology 2, 107–117.

ELSEVIER

Archives of Medical Research 37 (2006) 277–279

Archives of Medical Research

REVIEW ARTICLE

Diversity of Clinical Isolates of *Entamoeba histolytica* in Japan

Tomoyoshi Nozaki,[a] Seiki Kobayashi,[b] Tsutomu Takeuchi,[b] and Ali Haghighi[c]

[a]*Department of Parasitology, Gunma University Graduate School of Medicine, Showa-machi, Maebashi, Gunma, Japan*
[b]*Department of Tropical Medicine and Parasitology, Keio University School of Medicine, Shinanomachi, Shinjuku-ku, Tokyo, Japan*
[c]*Department of Parasitology and Mycology, School of Medicine, Shaheed Beheshti University of Medical Sciences, Tehran, Iran*

Received for publication September 26, 2005; accepted September 27, 2005 (ARCMED-D-05-00393).

In Japan, amebiasis is domestically transmitted by two major populations: male homosexuals and mentally handicapped persons, which is remarkably different from most other developed countries where *Entamoeba dispar* infection is predominantly observed. Here we briefly summarize epidemiology of amebiasis in Japan. We also review our current understanding of the diversity of *Entamoeba histolytica* clinical isolates in Japan, based on polymorphic genetic markers, clinical representations, and *in vivo* virulence, using an animal model. © 2006 IMSS. Published by Elsevier Inc.

Key Words: Amebiasis, Epidemiology, Heterogeneity, Sexually transmitted disease, Mentally handicapped persons.

Peculiarities of Amebiasis Endemic in Japan

The most unusual characteristic of amebiasis in Japan is that imported cases comprise only a minor proportion of all cases discovered in the country and cases attributable to domestic transmission dominate (1). We have two major populations infected by amebiasis: male homosexuals and mentally handicapped persons in institutions. We have approximately 500–600 cases of amebiasis reported to the Ministry of Health, Labour and Welfare, including three to four deaths annually. Approximately 90% of the reported cases are male. About 80% of cases have neither a history of traveling in endemic countries nor are they mentally handicapped. Thus, most of the reported cases are likely male homosexuals or bisexuals. Several groups previously reported a very high incidence of amebiasis among male homosexuals based on stool examination and serological tests (2–5). We recently observed sporadic cases of amebiasis transmitted through heterosexual intercourse (1) with an example of female commercial sex workers. Mentally handicapped persons are also severely affected by the disease (6–9). Mass infections of institutionalized mentally handicapped persons were often discovered during an onset of outbreaks as previously reported (6–9).

Recent Survey of Amebiasis in Institutions in Japan

Our recent survey to examine 484 individuals from six institutions [Institution B, C, D (10), E (11), and two other institutions (unpublished)] by a combination of microscopy, antigen capture ELISA, PCR, and serological tests showed that institutions were severely affected by amebiasis. Microscopic demonstration and antigen capture ELISA showed 9.7 and 12.3% overall positive, respectively. Serological tests, e.g., gel diffusion precipitin test and ELISA using whole parasite lysate as antigen showed 4.8 and 31.2% seropositive (unpublished). The positive rates of stool examination and serology varied significantly among institutions, suggesting that the intensity of infection varies among institutions or, alternatively, parasite strains spreading in each institution vary in virulent competence (see below). Importantly, these cases are often unreported or under-reported for several social reasons. In fact, the cases reported to the Ministry of Health, Labour, and Welfare included no cases of mentally handicapped persons. Thus, the number of amebiasis cases in Japan is largely underestimated. A very high incidence of *E. histolytica* infection in male homosexuals and mentally handicapped persons has not been reported in other developed countries (12–15) except for sporadic cases (16,17) and may be unique to Japan.

Address reprint requests to: Tomoyoshi Nozaki, Department of Parasitology, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan; E-mail: nozaki@med.gunma-u.ac.jp

0188-4409/06 $–see front matter. Copyright © 2006 IMSS. Published by Elsevier Inc.
doi: 10.1016/j.arcmed.2005.09.010

Molecular Basis of the Diversity of *E. histolytica* Isolates in Japan

Genetic diversity among approximately 40 clinical strains isolated from either male homosexuals or mentally handicapped persons in Japan was determined by high-resolution genotyping based on nucleotide sequencing of two protein-coding (SREHP and chitinase) and non-coding regions (locus 1–2 and 5–6) (10,11). Remarkable differences existed in the degree of heterogeneity of genotypes between the two groups. The strains isolated from male homosexuals are extremely heterogeneous; all the isolates derived from male homosexuals showed distinct genotypes. In contrast, isotypes from mentally handicapped persons were less heterogeneous. Isolates obtained from an institute showed an identical genotype. Moreover, one genotype was detected at three institutions at different times (see below). Because the intensity of infection is low prior to mass infection, probably due to previous mass treatment, mass infection was likely caused by a single source. One genotype was isolated from three institutions [Institutions B, A (10) and E (11)] located in three prefectures (Kanagawa, Shizuoka, and Yamagata) in 1994, 2001, and 2002, respectively. The emergence of the same genotype coincided with the movement of a single infected individual. This case is a good example to demonstrate that molecular fingerprinting is indeed a very reliable tool to determine a source of infection and a route of transmission. In addition, this case also raised serious concern on the effective treatment of amebiasis.

The genotypes of Japanese isolates were distinct from four representative reference strains used worldwide (HM-1:IMSS cl6, SAW755, SAW1627, and SAW1453) (10). In addition, none of 34 isolates from Thailand, Bangladesh, Cambodia, and Indonesia showed genotypes identical to Japanese isolates (11). Thus, the origin of Japanese strains is not understood and should be investigated in future studies.

Diversity of Clinical Manifestations and *In Vivo* Virulence among Japanese Strains

From a clinical point of view, heterogeneity of virulence attributable to genetic polymorphisms of the parasite likely exists. For instance, when we compared parasitological and serological results between Institutions D and E, which showed a similar rate of infection, notable differences in the serological marker for tissue invasion were observed. While parasitological stool examination gave a similar level of positive rate (28–30% positive) in these institutions, the seropositiveness evaluated by gel diffusion precipitin test significantly differed (0 or 16% in Institution D or E, respectively). The mean value of ELISA titer of the infected individuals was also significantly different between the two institutions (optical density at 405 nm of 0.13 or 0.50, respectively). The premise that this is not due to different rates of infection was also supported by the fact that the positive serology rate by ELISA was comparable between the two institutions (54–67%). These data strongly argue for the presence of genetic polymorphisms leading to distinct clinical manifestations.

Experimental animal infection using five isolates categorized into three representative genotypes from mentally handicapped persons also supported this premise. Hamsters were challenged with a direct inoculation of 5×10^4 trophozoites of KU13 (Institution A), KU19 (B), KU26 (C), KU27 (D), and KU33 (E), cultivated monoxenically with *Crithidia fasciculata* (18) to the liver, and abscess formation was evaluated a week later. All strains except for KU27 developed liver abscesses, while KU27 failed to cause abscess even in repeated attempts using a 4-times higher number of amebas (unpublished). These data agreed with the clinical manifestations in the patients infected with these strains and were consistent with the premise that a spectrum of virulence exists among the strains. *In vitro* virulence is conveniently assessed with the parasite's capacity to destroy a monolayer of mammalian cells (19). KU27 was incapable of destroying the monolayer of HeLa and Chinese hamster ovary cells, similar to *E. dispar* trophozoites (unpublished).

These avirulent phenotypes of KU27 are associated with a specific genotype of locus 1–2 type C and SREHP type A, neither of which was found among isolates examined in our laboratory, except two other isolates from the same institution (KU28 and KU29). Whether or not this specific marker is associated with the avirulent phenotype is not known. There is also no causal connection between this particular SREHP type and a lack of virulence. There are no notable differences in clinical manifestations of amebiasis between Japan and other countries. Hepatic, pulmonary, and brain abscesses are seen in 5–20% of cases (almost exclusively in male). None of the four genetic markers was found to be associated with a tissue tropism.

Acknowledgment

This work was supported by a grant for research on emerging and re-emerging infectious diseases from the Ministry of Health, Labour, and Welfare of Japan.

References

1. Infectious Disease Surveillance Center. Amebic dysentery. Infectious Agents Surveillance Report 2003;24:79–80.
2. Takeuchi T, Kobayashi S, Asami K, Yamaguchi N. Correlation of positive syphilis serology with invasive amebiasis in Japan. Am J Trop Med Hyg 1987;36:321–324.
3. Nozaki T, Motta SR, Takeuchi T, Kobayashi S, Sargeaunt PG. Pathogenic zymodemes of *Entamoeba histolytica* in Japanese male homosexual population. Trans R Soc Trop Med Hyg 1989;83:525.

4. Takeuchi T, Okuzawa E, Nozaki T, Kobayashi S, Mizokami M, Minoshima N, Yamamoto M, Isomura S. High seropositivity of Japanese homosexual men for amebic infection. J Infect Dis 1989;159:808.
5. Ohnishi K, Kato Y, Imamura A, Fukayama M, Tsunoda T, Sakaue Y, Sakamoto M, Sagara H. Present characteristics of symptomatic *Entamoeba histolytica* infection in the big cities of Japan. Epidemiol Infect 2004;132:57–60.
6. Nagakura K, Tachibana H, Kaneda Y, Suzuki H, Sasaoka K, Kobayashi S, Takeuchi T. Amebiasis in institutions for the mentally retarded in Kanagawa Prefecture, Japan. Jpn J Med Sci Biol 1990; 43:123–131.
7. Abe N, Nishikawa Y, Yasukawa A, Haruki K. *Entamoeba histolytica* outbreaks in institutions for the mentally retarded. Jpn J Infect Dis 1999;52:135–136.
8. Kaneda Y, Nagakura K, Tachibana H, Tanaka T, Sasao M. *Entamoeba histolytica* infection in a rehabilitation center for mentally retarded persons in Japan. Scand J Infect Dis 1988;20:687.
9. Nagakura K, Tachibana H, Tanaka T, Kaneda Y, Tokunaga M, Sasao M, Takeuchi T. An outbreak of amebiasis in an institution for the mentally retarded in Japan. Jpn J Med Sci Biol 1989;42:63–76.
10. Haghighi A, Kobayashi S, Takeuchi T, Masuda G, Nozaki T. Remarkable genetic polymorphism among *Entamoeba histolytica* isolates from a limited geographic area. J Clin Microbiol 2002;40:4081–4090.
11. Haghighi A, Kobayashi S, Takeuchi T, Thammapalerd N, Nozaki T. Geographic diversity among genotypes of *Entamoeba histolytica* field isolates. J Clin Microbiol 2003;41:3748–3756.
12. Allason-Jones E, Mindel A, Sargeaunt P, Williams P. *Entamoeba histolytica* as a commensal intestinal parasite in homosexual men. N Engl J Med 1986;315:353–356.
13. Allason-Jones E, Mindel A, Sargeaunt P, Katz D. Outcome of untreated infection with *Entamoeba histolytica* in homosexual men with and without HIV antibody. BMJ 1988;297:654–657.
14. Gatti S, Cevini C, Atzori C, Muratori S, Zerboni R, Cusini M, Scaglia M. Non-pathogenic *Entamoeba histolytica* in Italian HIV-infected homosexuals. Zentralbl Bakteriol 1992;277:382–388.
15. Weinke T, Friedrich-Janicke B, Hopp P, Janitschke K. Prevalence and clinical importance of *Entamoeba histolytica* in two high-risk groups: travelers returning from the tropics and male homosexuals. J Infect Dis 1990;161:1029–1031.
16. Gatti S, Lopes R, Cevini C, Ijaoba B, Bruno A, Bernuzzi AM, de Lio P, Monco A, Scaglia M. Intestinal parasitic infections in an institution for the mentally retarded. Ann Trop Med Parasitol 2000;94:453–460.
17. Scaglia M, Gatti S, Bruno A, Cevini C, Marchi L, Sargeaunt PG. Autochthonous amoebiasis in institutionalized mentally-retarded patients: preliminary evaluation of isoenzyme patterns in three isolates. Ann Trop Med Parasitol 1991;85:509–513.
18. Clark CG, Diamond LS. Methods for cultivation of luminal parasitic protists of clinical importance. Clin Microbiol Rev 2002;15:329–341.
19. Dvorak JA, Kobayashi S, Nozaki T, Takeuchi T, Matsubara C. Induction of permeability changes and death of vertebrate cells is modulated by the virulence of *Entamoeba* spp. isolates. Parasitol Int 2003;52:169–173.