Published in final edited form as:

Trends Parasitol. 2005 October ; 21(10): 451–453.

The *Entamoeba histolytica* genome: something old, something new, something borrowed and sex too?

Samuel L. Stanley Jr

Departments of Medicine and Molecular Microbiology, Midwest Regional Center of Excellence for Biodefense and Emerging Infectious Diseases Research, Washington University School of Medicine, Saint Louis, MO 63110, USA

An ancient and potent pathogen

Entamoeba histolytica is an intestinal protozoan parasite of humans that causes amebic colitis and amebic liver abscess: diseases associated with significant levels of morbidity and mortality worldwide [1]. The organism has a simple life cycle, existing as either the motile trophozoite or the infectious, hardy cyst form. Trophozoites of *E. histolytica* reside within the anaerobic confines of the human colon, lack mitochondria, derive energy from fermentation and reproduce by binary fission. *E. histolytica* trophozoites can be potent pathogens, possessing an armamentarium that includes a galactose–N-acetyl-galactosamine-binding lectin that mediates adherence to host cells, pore-forming proteins (amebapores) that can lyse bacteria or eukaryotic cells, cysteine proteinases that can cleave extracellular-matrix proteins and facilitate invasion into the colonic mucosa, and impressive phagocytic capabilities [1]. During the past three decades, the tools of molecular biology have greatly increased the understanding of *E. histolytica* pathogenesis. The recent publication of the *E. histolytica* genome by Loftus *et al.* [2] provides remarkable new insights into the biology of *E. histolytica*, helps understanding of the requirements for intestinal parasitism and sets the stage for a new, far more detailed understanding of this important pathogen.

Repetition and redundancy

The *E. histolytica* strain HM1:IMSS, which is available from the American Type Culture Collection [ATCC (<http://www.lgcpromochem.com/atcc/>)] and is used in laboratories worldwide, served as the DNA source for the genome project. Ultimately, 12.5-fold coverage of the nearly 24-Mb genome was achieved, with identification of 9938 predicted genes, each averaging 1.7 kb in size, comprising ~49% of the genome [2]. Originally, introns were thought to be rare in *E. histolytica* genes, but they are contained within 25% of the putative genes of this parasite. A theme of the *E. histolytica* genome is repetition and redundancy. Almost 10% of the sequence reads consisted of tandem arrays of one of 25 different types of repeating tRNA unit, containing between one and five tRNA types per unit. Because all but four of the tRNAs required for translation were found exclusively within these arrays, they must be functional [2]. Redundancy also characterizes many of the genes encoding suspected or proven *E. histolytica* virulence factors. New cysteine proteinase genes were identified (there are now at least 20), some of which contain putative transmembrane domains – which are lacking in the cathepsins of higher eukaryotes [2,3]. Three new amebapore genes were identified (bringing the total to six) and 30 homologs of a gene encoding the intermediate subunit of an *E. histolytica* surface lectin were identified. These findings raise the obvious questions of how

Corresponding author: Stanley, S.L. (sstanley@id.wustl.edu)..

The recent publication of the protozoan parasite *Entamoeba histolytica* genome provides new insights into eukaryotic evolution, the role of lateral gene transfer in amebic biology and the adaptations required for eukaryotes that reside within the human intestine.

many copies of these redundant genes are expressed within *E. histolytica*, whether any exhibit stage-specific (trophozoite versus cyst), or host- or tissue-specific (e.g. expressed in trophozoites residing in the human colon but not in cultured ameba) expression and whether there are functional differences between members of these large gene families.

The issue of gene expression has been partially explored for the cysteine proteinase genes, in which it was observed that trophozoites in culture seem to express only eight of the 20 identified cysteine proteinase genes, with only six of the eight being expressed at significant levels [3]. It remains to be determined whether similar ratios exist for the other large gene families identified by the genome project. The repetitive nature of the genome and its high AT content, which made the cloning of large insert libraries difficult, led to the most important limitation of the study: the failure to generate a map of the genome. It is anticipated that this will be achieved in the future and, when completed, it will undoubtedly provide important insights into the organization of the multigene families, and clues as to how their expression is regulated.

Something new

One area of *E. histolytica* biology that has been understudied is how the parasite recognizes and responds to its environment. Although signaling in *E. histolytica* through G-proteins and mitogen-activated protein kinases has been described [4,5], the genome project has delineated an entirely new repertoire of potential signal-transduction pathways [2,6]. All of the major families contained within the eukaryotic protein kinase superfamily are represented within the *E. histolytica* genome, including tyrosine kinases with Src homology (SH)2 domains, tyrosine-kinase-like protein kinases and putative receptor serine/threonine kinases. The serine/threonine receptor kinases fall into three families, two of which are characterized by extracellular domains with CXXC or CXC repeats [2,6]. Interestingly, CXXC repeats are found in the highly abundant variant surface antigens of the intestinal protozoan *Giardia lamblia*, where they have been linked to the conferral of parasite resistance to host intestinal proteases [7]. More than 100 protein phosphatases were identified in the *E. histolytica* genome; among them was a small family of genes encoding proteins in which a serine/threonine phosphatase domain is linked to a region containing leucine-rich repeat motifs. These motifs mediate primarily protein-protein interactions and, although some proteins with leucine-rich repeats are known to contain kinase domains, leucine-rich repeat proteins with phosphatase activity have been described only recently [8,9]. There is extensive diversity in the leucine-rich repeats, with at least seven subfamilies identified [9]. The leucine-rich repeats linked to the *E. histolytica* phosphatases are primarily of the 'typical' or 'classic' subtype but the genome also contains a large gene family (>100 members) that encodes proteins with leucine-rich repeats of the *Treponema pallidum* subtype (TpLRR), similar to those seen in the BspA surface protein of the bacteria *Bacteroides forsythus* [9,10]. The BspA protein has been implicated in the binding of *B. forsythus* to host extracellular-matrix proteins, suggesting a potential function for the newly discovered *E. histolytica* BspA-like proteins.

Something borrowed

E. histolytica lacks mitochondria, probably through secondary loss, and the sequencing of the *E. histolytica* genome did not reveal evidence of a mitochondrial genome. Instead, the parasite derives energy from glycolysis and fermentation. More than two decades ago, Reeves defined the fermentation pathway of *E. histolytica* [11]. During the ensuing years, as specific enzymes in this pathway were identified by molecular tools, it became clear that many of the pathway components resembled prokaryotic enzymes rather than those found within higher eukaryotes [12,13]. This finding led to the concept that lateral gene transfer from prokaryotes provided the enzymes necessary for fermentation and enabled ameba to dispense with mitochondria [13]. Similar lateral gene transfer events, although apparently from different prokaryotic

donors, seem to have occurred in other amitochondriate protozoans, including *G. lamblia* and the genitourinary pathogen *Trichomonas vaginalis*. Following the completion of the genome project, the true importance of lateral gene transfer in shaping *E. histolytica* biology has become clear. Using phylogenetic screens, Loftus *et al.* [2] identified 96 *E. histolytica* genes that seem to represent cases of recent prokaryotic-to-eukaryote gene transfer. Most of these genes encode proteins involved in carbohydrate and protein metabolism, and many seem to expand the range of available energy substrates. How lateral gene transfer occurred, especially in an organism that has been so resistant to attempts to introduce homologous recombination, remains a fascinating and open question.

With the attention on lateral gene transfer, it is important to emphasize that one of the most crucial results of the genome project was the generation of an *E. histolytica* ‘metabolome’, whereby the identified genes were used to predict all of the parasite metabolic pathways [2]. This will be a valuable tool for researchers, and the gaps in these pathways might point to novel activities and/or vulnerabilities that could be exploited for new anti-amebic therapies.

Sex too?

A fascinating recent discovery from mining the genome project is that *E. histolytica* possesses an essentially complete complement of the genes known to be required for meiosis [14]. *E. histolytica* trophozoites reproduce by binary fission and, although the existence of genetic recombination has been proposed on at least one occasion [15], there have been no confirmed sightings of amebic sex. It remains unknown whether *E. histolytica* is capable of meiosis but these findings (and the presence of meiotic genes in *Giardia*) are consistent with the idea that meiosis must be an early component of eukaryotic evolution and that, if they are not conserved for sex, these genes must have a role in other parasite functions.

A genome has been called a Rosetta Stone, a blueprint, the combination to a safe, a roadmap and, perhaps most accurately, a list of parts. Whatever the metaphor, the completion of the *E. histolytica* genome is having the greatest impact on the amebiasis field since Diamond reported the successful axenic cultivation of *E. histolytica* trophozoites in 1961 [2,16]. It is already changing the way in which the *E. histolytica* research community does science, and is providing a powerful new resource for investigators who are interested in parasitic protozoa, eukaryotic evolution and microbial pathogenesis.

Acknowledgements

Research in the Stanley laboratory is supported by NIH grants AI30084, AI51621 and U54AI57160. S.L.S. Jr is a Burroughs Wellcome Scholar in molecular parasitology.

References

1. Stanley SL Jr. Amoebiasis . Lancet 2003;361:1025–1034. [PubMed: 12660071]
2. Loftus B, et al. The genome of the protist parasite . *Entamoeba histolytica* Nature 2005;433:865–868.
3. Bruchhaus I, et al. 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;2:501–509. [PubMed: 12796295]
4. Meza I. Extracellular matrix-induced signaling in *Entamoeba histolytica*: its role in invasiveness. Parasitol Today 2000;16:23–28. [PubMed: 10637584]
5. Cruz-Vera J, et al. Collagen-induced STAT family members activation in *Entamoeba histolytica* trophozoites. FEMS Microbiol Lett 2003;229:203–209. [PubMed: 14680700]
6. Beck DL, 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;4:722–732. [PubMed: 15821132]

7. Nash TE. Surface antigenic variation in *Giardia lamblia*. *Mol Microbiol* 2002;45:585–590. [PubMed: 12139606]
8. Gao T, et al. PHLPP: a phosphatase that directly dephosphorylates Akt, promotes apoptosis, and suppresses tumor growth . *Mol Cell* 2005;18:13–24. [PubMed: 15808505]
9. Kobe B, Kajava AV. The leucine-rich repeat as a protein recognition motif . *Curr Opin Struct Biol* 2001;11:725–732. [PubMed: 11751054]
10. Sharma A, et al. Cloning, expression, and sequencing of a cell surface antigen containing a leucine-rich repeat motif from *Bacteroides forsythus* ATCC 43037 . *Infect Immun* 1998;66:5703–5710. [PubMed: 9826345]
11. Reeves RE. Metabolism of *Entamoeba histolytica* Schaudinn, 1903 . *Adv Parasitol* 1984;23:105–142. [PubMed: 6085216]
12. Yang W, et al. *Entamoeba histolytica* has an alcohol dehydrogenase homologous to the multifunctional *adhE* gene product of *Escherichia coli*. *Mol Biochem Parasitol* 1994;64:253–260. [PubMed: 7935603]
13. Rosenthal B, et al. Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica*. *J Bacteriol* 1997;179:3736–3745. [PubMed: 9171424]
14. Ramesh MA, et al. A phylogenomic inventory of meiotic genes; evidence for sex in *Giardia* and an early eukaryotic origin of meiosis . *Curr Biol* 2005;15:185–191. [PubMed: 15668177]
15. Blanc D, et al. Experimental production of new zymodemes of *Entamoeba histolytica* supports the hypothesis of genetic exchange . *Trans R Soc Trop Med Hyg* 1989;83:787–790. [PubMed: 2617647]
16. Diamond LS. Axenic cultivation of *Entamoeba histolytica*. *Science* 1961;134:336–337. [PubMed: 13722605]