African Journal of Biotechnology Vol. 8 (8), pp. 1363-1369, 20 April, 2009
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2009 Academic Journals

Review

Genomics approaches in the understanding of *Entamoeba histolytica* virulence and gene expression regulation

César López-Camarillo[1]*, Nancy Guillen[2,3], Christian Weber[2,3], Esther Orozco[4], and Laurence A. Marchat[5]

[1]Universidad Autónoma de la Ciudad de México, Posgrado en Ciencias Genómicas, México D.F., México.
[2]Institut Pasteur, Unité Biologie Cellulaire du Parasitisme, Paris, France.
[3]INSERM U786, Paris, France.
[4]CINVESTAV-IPN, Departamento de Infectómica y Patogénesis Molecular, México DF, México.
[5]Instituto Politécnico Nacional, Escuela Nacional de Medicina y Homeopatía, Programa Institucional de Biomedicina Molecular, México D.F., México.

Accepted 5 February, 2009

Entamoeba histolytica **is the intestinal protozoan parasite responsible for amebic colitis and liver abscesses, which cause mortality in many developing countries. The sequencing of the parasite genome provides new insights into the cellular workings and genome evolution of this major human pathogen. Here, we reviewed recent advances in the efforts to understand virulence and gene expression regulation in *E. histolytica* by using genomic approaches based on microarray technology and bioinformatic analysis of genome sequence.**

Key words: *Entamoeba histolytica*, amoebiasis, genome sequence, cDNA microarrays, genomic expression profiling, virulence, gene families.

INTRODUCTION

Entamoeba histolytica, is the protozoan parasite causative of human amoebiasis, which has a world-wide distribution with a higher prevalence in developing countries, affecting more than 50 million people each year. The parasite causes intestinal dysentery and hepatic abscesses that result in approximately 70,000 - 100,000 deaths a year, making it a leading cause of parasitic death in humans. The World Health Organization places *E. histolytica* second after *Plasmodium falciparum* as causing most deaths annually among protist parasites (WHO, 1997). *E. histolytica* infection, which is spread through contaminated food and water, begin with the ingestion of the cyst, a non-dividing form that is able to survive in the environment due to a protective cell wall. Cyst undergoes excystation in the intestine to produce the proliferative trophozoite forms, which colonize the colon. In some cases, trophozoites can invade other organs to provoke abscesses that can be fatal (Haque et al., 2003). The dramatic virulence variability of trophozoites *in vivo* could be related to great genome plasticity and changes in gene expression program. Some virulence factors have been identified as key factors in pathogenesis, however most molecular mechanisms that are relevant for infection establishment are still unknown. Therefore, elucidation of molecular mechanisms modulating virulence and gene expression regulation is of major interest for researchers.

THE GENOME OF *E. HISTOLYTICA*

The *E. histolytica* genome sequence draft has been recently published (Loftus et al., 2005). The current *E. histolytica* genome assembly is 23.7 million base pairs in size, which predicts 9,938 genes that have an average size of 1.17 kb and are carried on 14 - 17 chromosomes.

*Corresponding author. E-mail: genomicas@yahoo.com.mx. Tel.: +55-5488-6661 ext 5353; Fax: +55 5559 0187

The length variation of chromosomes is thought to be caused by expansion and contraction of subtelomeric regions, which might be formed by tRNAs arranged in tandem arrays. The overall A+T content of the genome is around 75.3%. In addition, repeat sequences in *E. histolytica* account for 6% of the genome. They consist of long interspersed repetitive elements, LINEs, and short interspersed repetitive elements, SINEs: the LINEs (non-long terminal repeat elements) encode their own retro-transposition machinery, whereas SINEs (short interspersed elements) borrow this machinery (Bakre et al., 2005). Most genes only comprise a single exon; however as many as 25% may be potentially spliced and 6% contains two or more introns. *E. histolytica* genes are short, mainly due to the loss of introns. The 5´ and 3´ untranslated regions (UTRs) are usually short. The average length of *E. histolytica* predicted proteins is ~389 amino acids (Clark et al., 2007). The gene density is around 2.4 kb/gene. A deep analysis of genome sequence revealed a variety of metabolic adaptations, including reduction or elimination of most mitochondrial metabolic pathways and the use of oxidative stress enzymes generally associated with anaerobic prokaryotes (Loftus et al., 2005). Evidence for lateral gene transfer of bacterial genes into the *E. histolytica* genome was observed, and evidence that these genes have putative functions in central metabolism has been provided. The parasite genome also encodes a large number of novel receptor kinases and contains expansions of a variety of gene families associated with virulence, including cysteine and metallo-proteinases. Analysis of the *E. histolytica* genome provides new insights into the workings and genome evolution of this major human pathogen.

TRANSCRIPTIONAL PROFILING OF *E. HISTOLYTICA*

The availability of *E. histolytica* genome sequence provides opportunities for the study of genomic expression profiles in parasites subjected to diverse stimuli and conditions. Microarrays made from genomic DNA fragments or synthetic oligonucleotides based on annotated genes and genomic data, have been successfully applied in *E. histolytica* to obtain transcriptional profiling in distinct conditions including virulence, stress and DNA damage (Debnath et al., 2004, 2007; Gilchrist et al., 2006; MacFarlane and Singh, 2006; Weber et al., 2006; Ehrenkaufer et al., 2007a, b; Ali et al., 2007; Davis et al., 2007, Vicente-Joao et al., 2008; Santi-Rocca et al., 2008; Weber et al., 2008). In addition, two recent papers on *E. histolytica* have reviewed the investigations focused in genomic expression profiles during cyst-trophozoites conversion and the role of epigenetic control on gene expression (Singh and Ehrenkaufer, 2009; Baxt and Singh, 2009), thus we recommend referring to these papers to review more specific data in these *E. histolytica* topics. Here, we discuss the recent genomic studies which have not been widely reviewed or from our own data on microarray and bioinformatic analysis of *E. histolytica*.

Genomic expression profiling during chemotaxis of trophozoites towards the pro-immflamatory TNF cytokine

During host invasion *E. histolytica* trophozoites migration is essential for colonization. Although the invasive process is central to pathogenesis and is known to be driven by the motility of the parasites, poorly is understood at the cellular and molecular level. The tumour necrosis factor (TNF) produced by host cells has a migratory effect on trophozoites. During chemotaxis, the cell senses the chemical gradient in the extracellular environment via chemoreceptors. The binding of the attractant to its receptors activates signaling cascades leading to cytoskeleton changes, cell polarization and oriented movement.

Blazquez et al. (2008) analyzed the genomic response associated to cellular changes leading to directional motility in response to TNF. They used microarray analysis to screen for genes that were potentially important during chemotaxis towards TNF. The microarray studies were performed using RNA isolated from amoeba that have chemotaxed towards TNF (close to the TNF source) compared to RNA isolated from amoeba that have not chemotaxed in the presence of a TNF gradient (far from the source). Intriguingly, the overall gene expression variation was lower than 2-fold but significant. Gene expression profiling showed that 263 genes were modulated; 139 were upregulated and 124 down regulated. Microarray screening also showed the upregulation of several genes encoding for proteins involved in actin cytoskeleton dynamics during chemotaxis towards TNF. These included proteins participating in microfilament dynamics, such as nucleation (Arp2/3, Formins), F-actin capping (Cap34α, Cap34β), microfilament network formation (ABP-120 (filamin), cortexellin), actin bundling array (α-actinin), as well as proteins necessary for microfilament functions, such as myosin IB and II, and coronin. The authors also observed the modulation of a putative α-actinin 2, which is one of the potential partners of the Gal/GalNAc lectin. Interestingly, Gal/GalNAc lectin genes hgl and lgl, were also upregulated during chemotaxis. In addition, trophozoites inhibited for Gal/GalNAc adhesion activities were impeded for chemotaxis towards TNF. Additionally, microarray data showed the upregulation of genes involved in signal transduction, which are functionally related to actin dynamics, including the adenylyl cyclase-associated protein (CAP) and an SH3 domain protein (Blazquez et al., 2008). Moreover, three amoebic virulence-related genes were upregulated; these encode ARIEL, the light subunit 1 (Lgl1) and the heavy subunit (Hgl) of the Gal/GalNAc lectin. Taken altogether, these results gave insights on *E. histolytica* transcription

changes during migration towards TNF and suggested how cytoskeleton dynamics could be regulated during chemotaxis.

Genomic expression profiling in response to DNA damage

To establish successful host colonization, trophozoites must resist oxidative stress by free radicals and drug treatments, which could cause DNA damage and cell death. Some anti-parasitic drugs produce genotoxic damage and parasites have developed molecular strategies to overcome DNA injure (Bhattacharyya et al., 2004). However, molecular mechanisms that safeguard genome integrity are unknown in this parasite.

Using 70-mer oligonucleotides microarray carrying information of 3,000 ORFs, we identified genes that may contribute to adaptation and survival of trophozoites in response to DNA damage induced by UV-C irradiation (Weber et al., 2008). We found that 11.6 (350 ORFs) and 17.2% (522 ORFs) genes were modulated at 5 min and 3 h after UV-C irradiation, respectively. A subset of 180 genes was induced and 170 genes were repressed at 5 min after UV treatment, whereas at 3 hr, 230 ORFs were upregulated and 292 ORFs were downregulated. Most genes were < 2-fold modulated at both times, evidencing a weak transcriptional activation after DNA damage. We also identified 14 genes over expressed at both times after UV-C irradiation, including four putative kinases containing SH3 domains, a heat shock transcription factor and a DEAD/DExH box RNA helicase among others, which could have a role in the response to DNA damage and the maintenance of cellular processes involved in DNA integrity. In addition, 9 genes including 4 cytoskeleton predicted proteins were repressed at both times. Many regulated genes were involved in potential DNA damage-response pathways, including cell cycle, signal transduction, and oxidative stress response. However, the majority of predicted genes were involved in unexpected pathways, such as RNA processing, cell structure, protein synthesis and degradation, energy metabolism, vesicle trafficking, and a great number of hypothetical genes. Interestingly, at 5 min after UV-C irradiation, classical DNA repair genes including putative EhMre11, EhRad50, EhRad23, EhDdb1, and EhRad54 also presented a slight modulation. Our data confirmed that UV-C irradiation has subtle effects on gene transcription, despite its ability to efficiently produce DNA damage (Lopez-Casamichana et al., 2008). Gene expression profiling also showed that 18 genes were > 2-fold induced at 3 h after UV-C irradiation, including three ORFs codifying for iron-sulfur clusters-containing proteins, potentially involved in oxidative stress protection.

Transcriptional profiling showed that 26 genes were up to 2-fold repressed at 3 h after UV-C treatment. Interestingly, 12 downregulated genes correspond to cytoskeleton and associated proteins, including actins, actin binding proteins, coactosin, and myosin II, suggesting that the dynamics of amoebic cytoskeleton was impaired (Weber et al., 2008). These data revealed new directions for studying the response to DNA damage, and they will contribute to further elucidation of mechanisms regulating genome integrity in this early branch protozoan.

Influences of trichostatin A on gene expression in *E. histolytica*

Histone modifications regulate chromatin structure and influence gene expression associated with diverse biological functions including maintenance of genome architecture, cellular differentiation, cancer and pathogen virulence. This mechanism is regulated by the opposing activities of histone acetyltransferases and histone deacetylases (HDACs). In order to identify the genome-wide effects of histone acetylation in regulating *E. histolytica* gene expression, a genome expression profiling of parasites treated with trichostatin A (TSA), which is an inhibitor of HDACs, was developed (Isakov et al., 2008). TSA challenge enhanced the cytopathic and hemolytic activities of the parasite and its resistance to oxidative stress. Microarray analysis of parasites treated with 50 nM of TSA showed that 15% of the up-regulated genes were related to virulence, compared to 3% of the down-regulated genes. In contrast, 21% of the down-regulated genes were related to cellular organization and biogenesis, compared to 6% of the up-regulated genes. Over expressed genes up to 2-fold include the 20 kDa antigen-related protein, the Gal/GalNAc lectin small subunits lgl-2 and lgl-3, two iron-sulfur flavoproteins, a metal-dependent hydrolases, 5 hypothetical proteins and a gene encoding Jacob, a protein involved in cysts formation. However, the phenotype observed did not correlate with the expression of other genes related to cyst formation. Intriguingly, the pattern of gene expression was surprisingly different from that previously described (Ehrenkaufer et al., 2007b) after treatment with 150nM TSA.

Ehrenkaufer et al. (2007b) identified the genome-wide effects of histone acetylation in regulating gene expression using a whole-genome expression profiling of parasites treated with small chain fatty acids (SCFA) and TSA. Despite significant changes in histone acetylation patterns, exposure of parasites to SCFA resulted in minimal transcriptional changes (11 out of 9,435 regulated genes). In contrast, exposure to TSA, a more specific inhibitor of histone deacetylases, significantly affected transcription of 163 genes (122 upregulated genes and 41 downregulated genes). However, the set of genes regulated by TSA substantially overlapped with genes regulated during parasite development (Ehrenkaufer et al., 2007a), since 73 out of the 122 upregulated genes were upregulated in *E. histolytica* cysts and 15 out of the 41 downregulated genes were

downregulated in *E. histolytica* cysts.

All together, both works emphasize the role of epigenetic mechanisms in the control of *E. histolytica* virulence.

Transcriptomic analysis of *E. histolytica* in response to oxidative and nitrosative induced cellular stress

During tissue invasion, *E. histolytica* is confronted to cytotoxic reactive oxygen species (ROS) and reactive nitrogen species (RNS) that are released by activated phagocytes (Zhang et al., 2003). In order to survive and cause disease, parasites have developed common defense strategies against oxidative and nitrosative agents including detoxification enzymes and repair systems. In order to determine the molecular mechanisms by which *E. histolytica* reacts to oxidative and nitrosative stress, the transcriptional changes of pathogenic trophozoites were analyzed after a 60 min exposure to H_2O_2 (1 mM) or an NO donor (dipropylenetriamine-NONOate, 200 mM), using whole-genome DNA microarrays constituted by short oligonucleotides representing 9,435 of the 9,938 annotated amebic genes (Vicente-Joao et al., 2008).

Results showed that genes encoding reactive oxygen and nitrogen species detoxification enzymes had high transcriptional levels under basal conditions and upon exposure to both stresses. There was a significant modulation of gene expression by H_2O_2 (286 regulated genes) and dipropylenetriamine-NONOate (1,036 regulated genes) with a significant overlap of genes modulated under both conditions (164 genes). A number of regulated genes were potentially involved in signaling and repair/metabolic pathways. Interestingly, the majority of genes encode unknown proteins. In contrast, trophozoites of the non-pathogenic Rahman strain had fewer transcriptional changes in response to H_2O_2 and the overall fold-changes for the regulated genes were significantly lower in comparison with the pathogenic strain. These data suggested that one important difference between virulent and non-virulent trophozoites is their ability to deal with the stresses encountered during host invasion.

Transcriptional response to overexpression of a high mobility-group box protein induced during intestinal infection

A previous report using a microarray based genome-wide analysis of gene expression modulated by intestinal colonization allowed the identification of an upregulated gene that encoded a putative high-mobility-group box (HMGB) protein, EhHMGB1 (Gilchrist et al., 2006). Recently, Abhyankar et al. (2008) demonstrated that EhHMGB1 encodes a functional DNA bending protein that enhances the p53 DNA binding activity *in vitro*. Using microarray technology, they determined the role of EhHMGB1 in gene expression regulation of parasite. To define the target genes regulated by this putative transcription factor, EhHMGB1 protein was overexpressed in HM1: IMSS trophozoites. Results evidenced the modulation of 33 transcripts involved in a variety of cellular functions. Of these, 20 were also modulated, although in an opposite direction, at day one or day 29 in the mouse model of intestinal amoebiasis (Gilchrist et al., 2006). Notably, 4 transcripts with known roles in virulence, including two encoding Gal/GalNAc lectin light chains Lgl3 and Lgl4, one of the cysteine proteinases (EhCP-A7) and a potential enterotoxic peptide, were modulated in response to EhHMGB1 overexpression, suggesting a role of this putative transcription factor in regulating the virulence properties of parasite.

Genomic expression profiling of trophozoites in response to metronidazole drug treatment

Metronidazole has been used as an effective drug to treat invasive amoebiasis since more than 40 years. Although the activation of metronidazole has been well characterized, little is known on how the parasites modulated gene expression in response to chemotherapy stress. Tazreiter et al. (2008) used a 213 gene-focused low density microarray and quantitative RT-PCR assays to evaluate changes in transcripts levels in response to metronidazole challenge. Results showed that after 1 h of 50 µM metronidazole exposure, trophozoites slightly increased the levels of transcripts coding for superoxide dismutase, peroxiredoxin, 2 ferredoxins, thioredoxin reductase, long chain fatty CoA ligase, and the galactose/N-acetylgalactosamine specific lectin light and heavy chains. At prolonged times after metronidazole treatment (3 and 5 h) there were no significant further rise of these transcript levels. Only two genes encoding one of the ferredoxins and the lectin light chain were kept > 2-fold upregulated, evidencing a weak transcriptional response to metronidazole challenge. In contrast, a decrease of mRNAs encoding actin, 70 kDa and 101 kDa heat shock proteins was observed.Other repressed genes include a short dehydrogenase, a fatty acid elongase, diaphanous protein and a WD repeat protein. These data evidenced that parasite transcription seems to be modestly modulated when trophozoites were confronted with metronidazole stress.

Genomic expression profiling in response to heath shock

Microarray assays revealed that when exposed to stress, *E. histolytica* trophozoites exhibit a specific heat shock response, associated with a dramatic overall reduction in gene transcription (Weber et al., 2006). Of the 1,131 unique genes probed by the microarray, almost 471 (42%) were significantly repressed during heat treatment (42°C for 4 h). 135 (29%) of the 471 downregulated

genes potentially encode proteins of unknown functions. Interestingly, some lectin genes, which represent key factors participating in virulence (Petri et al., 2002) were identified as differentially regulated. A differential allelic expression of the galactose/N-acetylgalactosamine (Gal/GalNAc) lectin) was observed under heat shock conditions. Data also showed the expression of additional virulence factors, such as cysteine proteinases and the so-called 20-kDa antigen. The analysis of heat shock modulated genes by categories showed that gene downregulation is rather indiscriminate and may correspond to a general response for adaptation to new environmental conditions. Data reported by Weber et al. (2006) provide insights to understand how parasites survive to stress conditions.

BIOINFORMATICS-BASED STUDIES

Taking advantage of the recent completion of the *E. histoly*tica genome, we and others have extensively screened databases to define potential genes encoding proteins involved in virulence and survival under stress conditions, as well as conserved sequences potentially participating in gene expression regulation. These studies showed the usefulness of computer-based methods to define novel regulatory sequences and genes with potential roles in biological processes, which act as guides for experimental investigation.

Large scale *in silico* prediction of 3´ UTR motifs involved in gene expression regulation

In most eukaryotic cells, the formation of the poly(A) tail at pre-mRNA 3´ end is an integral part of gene expression regulation, affecting mRNA stability, translatebility and nuclear-to-cytoplasmic export. Multiprotein complexes bind specific sequences in pre-mRNA 3´ UTR and act in a coordinated way to perform pre-mRNA cleavage and poly(A) tail synthesis (Proudfoot, 2004). In human cells and yeast, the cleavage and polyadenylation site (poly(A) site) is flanked by the polyadenylation signal (A($^{A}/_{U}$)UAAA) and the U/GU-rich element, as well as additional U-rich and G/C-rich elements (Hu et al., 2005). Through an extensive computational analysis of a large *E. histolytica* EST and genomic sequences dataset, we evidenced an AU-rich domain corresponding to the consensus UA($^{A}/_{U}$)UU polyadenylation signal previously reported in amoeba (Brucchaus et al., 1993) or variants, the poly(A) site that is denoted by a U residue and flanked by U-rich tracts, and a novel A-rich element in 3´ UTR of parasite genes (Zamorano et al., 2008). This predicted array was validated through the analysis of genes with experimentally characterized poly(A) site. The molecular organization of pre-mRNA 3´ UTR cis-regulatory elements appears to be roughly conserved through an evolutionary scale; particularly, the polyadenylation signal is the conserved A($^{A}/_{U}$)UAAA hexanucleotide in plants, animals and yeast, but it appears to be specie-specific in the few protozoan parasites studied. Remarkably, the novel A-rich region that we identified downstream the poly(A) site seems to be unique for the primitive eukaryote *E. histolytica* transcripts (Zamorano et al., 2008).

Bioinformatic prediction of microRNAs in *E. histolytica* genome

MicroRNAs (miRNAs) are endogenous and abundant short double stranded RNA hairpins, typically 20 - 24 nucleotides in length, that downregulate gene expression at the posttranscriptional level by binding to their target mRNA. These noncoding RNAs control the expression of several genes, regulating many important biochemical pathways, although their targets and mechanism of action are not well known. miRNAs have been identified in many vertebrates including humans, and also in flies, nematodes, plants, and viruses. De S et al. (2008) developed a computational method to identify the putative miRNAs in the *E. histolytica* genome. They identified 17 putative candidate miRNAs (Eh-mir1-17). Some Eh-miRNAs target two or more different genes, which are involved in diverse cellular processes. Predicted proteins include two BspA-like leucine rich proteins, a DEAD-box RNA helicase, small-GTP binding protein Rab 8, protein kinase, an UDP n-acetylglucosamine transporter, a Rap/Ran GTPase activating protein, an acid sphingomyelinase-like phosphodiesterase, a Rho guanine nucleotide exchange factor, a RNA pseudouridylate synthase, a putative splicing factor Prp8, and 19 hypothetical proteins. This bioinformatic based work will help future attempts on experimental validation of the putative Eh-*miRNA* candidates in this parasite.

In silico identification of EhDEAD/DExH-box RNA helicases family

RNA helicases are evolutionary conserved enzymes that catalyze ATP-dependent-unwinding of double stranded-RNA structures to perform numerous genetic processes, such as transcription, splicing, ribosome biogenesis, pre-mRNA processing, mRNA export, degradation and translation. In addition, they act as RNPases and remodel RNA-proteins interactions (Linder, 2006). By a deep survey of the *E. histolytica* genome, we found that this parasite has a large gene family corresponding to 20 EhDEAD and 13 EhDExH-box RNA helicases with the characteristic DEAD and DExH motifs in the conserved helicase domain, respectively (Marchat et al., 2008). Phylogenetic studies and sequence analyses suggested that EhDEAD/DExH-box RNA helicases family has been generated by gene or internal region duplications, mutation events, intron formations, and motif deletions.

Moreover, EhDexh1 and EhDexh10 genes appear to result from gene fusion events of two independent RNA helicase genes, to generate multidomain proteins that could have enhanced RNA binding and unwinding activities. Functional characterization of the recombinant EhDEAD1 protein, which is homologous to yeast DED1 and human DDX3X RNA helicases, both involved in translation and cell cycle regulation (Chuang et al., 1997; Yedavali et al., 2004), showed that it is a conserved DEAD-box RNA helicase with ATPase and ATP-dependent RNA unwinding activities. Moreover, it seems to participate in S to G2/M phase transitions during cycle progression (López-Camarillo et al., 2008). Interestingly, RT-PCR assays, microarray and proteomic data showed that several EhDead genes are differentially expressed in relation to distinct culture conditions including virulence and in response to heat shock and DNA damage (Marchat et al., 2008). These data showed the power of bioinformatics analyses to identify gene families and provide novel information on the evolution of EhDEAD/EhDExH-box RNA helicases and their potential relevance for RNA metabolism in *E. histolytica*.

Bioinformatic prediction of protease families

Parasite proteases, especially cysteine protease, constitute major pathogenicity factors in *E. histolytica*, being involved in cytopathic activity and amoebic liver abscesses formation in laboratory animals. Homology searches based on the conservation of active site regions revealed that *E. histolytica* genome contains a considerable number of gene coding for proteolytic enzymes (Tillack et al., 2007). The larger family includes 50 gene encoding potential cysteine peptidases that belong to C1 papain superfamily and are phylogenetically clustered into 3 distinct clades (Clark et al., 2007). *E. histolytica* also possesses 22 metallo peptidases from at least 11 different families, 10 serine peptidases belonging to 3 different families, and 4 aspartic peptidases from only one family. Furthermore, mRNA expression profile of various proteolytic enzymes was assessed in different conditions through microarray and quantitative real time PCR assays. These data confirmed the existence of a large family of proteases that are relevant for *E. histolytica* survival and cell damage in the host.

CONCLUSION

DNA microarray and bioinformatic analyses have been useful to begin the dissection of the genome of *E. histolytica* and provide a functional context to the genes identified in the genome sequencing effort. Future directions will include further analyses of the parasite's transcriptome in invasive hepatic disease, as well as additional characterization of the developmental conversion to the cyst form. These data may be useful in the development of novel diagnostic and therapeutic options. Additionally, genetic approaches can now be applied to definitively assign a role to novel genes in amoebic biology and pathogenesis. The computational and functional data reviewed here provide new information on the evolution of protein families and their potential relevance for virulence and gene expression regulation in *E. histolytica*.

REFERENCES

Abhyankar MM, Hochreiter AE, Hershey J, Evans C, Zhang Y, Crasta O, Sobral BW, Mann BJ, Petri WA Jr, Gilchrist CA (2008). Characterization of an Entamoeba histolytica high-mobility-group box protein induced during intestinal infection. Eukaryot Cell. 7: 1565-1572.

Ali IK, Ehrenkaufer GM, Hackney JA, Singh U (2007). Growth of the protozoan parasite Entamoeba histolytica in 5-azacytidine has limited effects on parasite gene expression. BMC Genomics. 8: 7-11.

Bakre AA, Rawal K, Ramaswamy R, Bhattacharya A, Bhattacharya S (2005). The LINEs and SINEs of Entamoeba histolytica: comparative analysis and genomic distribution. Exp Parasitol. 110: 207-213.

Baxt LA, Singh U (2009). New insights into Entamoeba histolytica pathogenesis. Curr. Opin. Infect. Dis. 21: 489-494.

Bhattacharyya MK, Norris DE, Kumar N (2004). Molecular players of homologous recombination in protozoan parasites: implications for generating antigenic variation. Infect. Genet. Evol. 4: 91-98.

Blazquez S, Guigon G, Weber C, Syan S, Sismeiro O, Coppée JY, Labruyère E, Guillén N (2008). Chemotaxis of Entamoeba histolytica towards the pro-inflammatory cytokine TNF is based on PI3K signalling, cytoskeleton reorganization and the Galactose/N-acetylgalactosamine lectin activity. Cell Microbiol. 10: 1676-1686.

Brucchaus M, Leippe C, Lioutas C, Tannich E (1993). Unusual gene organization in the protozoan parasite *Entamoeba histolytica*. DNA Cell Biol. 12: 925-933.

Chuang R, Weaver P, Liu Z, Chang T (1997). Requirement of the DEAD-box protein Ded1p for messenger RNA translation. Science. 275: 1468-1471.

Clark CG, Alsmark UC, Tazreiter M, Saito-Nakano Y, Ali V, Marion S, Weber C, Mukherjee C, Bruchhaus I, Tannich E, Leippe M, Sicheritz-Ponten T, Foster PG, Samuelson J, Noël CJ, Hirt RP, Embley TM, Gilchrist CA, Mann BJ, Singh U, Ackers JP, Bhattacharya S, Bhattacharya A, Lohia A, Guillén N, Duchêne M, Nozaki T, Hall N (2007). Structure and content of the Entamoeba histolytica genome. Adv. Parasitol. 65(2): 51-190.

Davis PH, Schulze J, Stanley SL Jr (2007). Transcriptomic comparison of two Entamoeba histolytica strains with defined virulence phenotypes identifies new virulence factor candidates and key differences in the expression patterns of cysteine proteases, lectin light chains, and calmodulin. Mol. Biochem. Parasitol. 151: 118-128.

De S, Pal D, Ghosh SK (2006). Entamoeba histolytica: computational identification of putative microRNA candidates. Exp. Parasitol. 113: 239-243.

Debnath A, Das P, Sajid M, McKerrow JH (2004). Identification of genomic responses to collagen binding by trophozoites of *Entamoeba histolytica*. J. Infect. Dis. 190: 448-457.

Debnath A, Tashker JS, Sajid M, McKerrow JH (2007). Transcriptional and secretory responses of Entamoeba histolytica to mucins, epithelial cells and bacteria. Int. J. Parasitol. 37: 897-906.

Ehrenkaufer GM, Haque R, Hackney JA, Eichinger DJ, Singh U (2007a). Identification of developmentally regulated genes in *Entamoeba histolytica*: insights into mechanisms of stage conversion in a protozoan parasite. Cell Microbiol. 9: 1426-1434.

Ehrenkaufer GM, Eichinger DJ, Singh U (2007b). Trichostatin. A effects on gene expression in the protozoan parasite Entamoeba histolytica. BMC Genomics 8: 216.

Ehrenkaufer GM, Singh U (2008).Transcriptional regulatory networks in *Entamoeba histolytica*. Curr Drug Targets 9: 931-937.

Gilchrist CA, Houpt E, Trapaidze N, Fei Z, Crasta O, Asgharpour A, Evans C, Martino-Catt S, Baba DJ, Stroup S, Hamano S, Ehrenkaufer G, Okada M, Singh U, Nozaki T, Mann BJ, Petri WA Jr. (2006). Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome. Mol. Biochem. Parasitol. 147: 163-176.

Haque R, Huston CD, Hughes M, Houpt E, Petri WA Jr (2003). Amebiasis. N. Engl. J. Med. 348: 1565-1573.

Hu J, Lutz CS, Wilusz J, Tian B (2005). Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. RNA. 11: 1485-1493.

Isakov E, Siman-Tov R, Weber C, Guillen N, Ankri S (2008). Trichostatin. A regulates peroxiredoxin expression and virulence of the parasite *Entamoeba histolytica*. Mol. Biochem. Parasitol. 158: 82-94.

Linder P (2006). Dead-box proteins: a family affair-active and passive players in RNP-remodeling. Nucleic Acids Res. 34: 4168-4180.

López-Camarillo C, García-Hernández ML, Marchat LA, Luna-Arias JP, Hernandez de la Cruz O, Mendoza L, Orozco E (2008). Entamoeba histolytica EhDEAD1 protein is a conserved DEAD-box RNA helicase with ATPase and ATP-dependent RNA unwinding activities. Gene. 14: 19-31.

Loftus B, Anderson I, Davies R, Alsmark UC, Samuelson J, Amedeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, 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, Quail MA, Rabbinowitsch E, Norbertczak H, Price C, Wang Z, Guillén N, Gilchrist C, Stroup SE, Bhattacharya S, Lohia A, Foster PG, Sicheritz-Ponten T, Weber C, Singh U, Mukherjee C, El-Sayed NM, Petri WA Jr, Clark CG, Embley TM, Barrell B, Fraser CM, Hall N (2005). The genome of the protist parasite Entamoeba histolytica. Nature, 433: 865-868.

López-Casamichana M, Orozco E, Marchat LA, López-Camarillo C (2008). Transcriptional profile of the homologous recombination machinery and characterization of the EhRAD51 recombinase in response to DNA damage in *Entamoeba histolytica*. BMC Mol. Biol. 9: 35.

MacFarlane RC, Singh U (2006). Identification of differentially expressed genes in virulent and nonvirulent Entamoeba species: potential implications for amebic pathogenesis. Infect. Immun. 74: 340-351.

Marchat LA, Orozco E, Guillen N, Weber C, López-Camarillo C (2008). Putative DEAD and DExH box RNA helicases families in *Entamoeba histolytica*. Gene. 424: 1-10.

Petri WA Jr, Haque R, Mann BJ (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.

Proudfoot N (2004).New perspectives on connecting messenger RNA 3' end formation to transcription. Curr. Opin. Cell Biol. 16: 272-278.

Santi-Rocca J,Weber C,Guigon G,Sismeiro O,Coppee JY,Guillen N (2008).The lysine- and glutamic acid-rich protein KERP1 plays a role in Entamoeba histolytica liver abscess pathogenesis. Cell Microbiol. 10: 202-217.

Singh U, Ehrenkaufer GM (2009). Recent insights into Entamoeba development: identification of transcriptional networks associated with stage conversion. Int. J. Parasitol. 39: 41-47.

Tillack M, Biller L, Irmer H, Freitas M, Gomes MA, Tannich E, Bruchhaus I (2007). The Entamoeba histolytica genome: primary structure and expression of proteolytic enzymes. BMC Genomics. 8: 170.

Tazreiter M, Leitsch D, Hatzenbichler E, Mair-Scorpio GE, Steinborn R, Schreiber M, Duchêne M (2008). Entamoeba histolytica: response of the parasite to metronidazole challenge on the levels of mRNA and protein expression. Exp. Parasitol. 120: 403-410.

Vicente-Joao B, Ehrenkaufer GM, Saraiva LM, Teixeira M, Singh U (2008). Entamoeba histolytica modulates a complex repertoire of novel genes in response to oxidative and nitrosative stress: implications for amebic pathogenesis. Cell Microbiol. 11: 51-69.

Weber C, Guigon G, Bouchier C, Frangeul L, Moreira S, Sismeiro O, Gouyette C, Mirelman D, Coppee JY, Guillen N (2006). Stress by heat shock induces massive down regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in Entamoeba histolytica. Eukaryot. Cell, 5: 871-875.

Weber C, Marchat LA, Guillen N, López-Camarillo C (2008). Effects of DNA damage induced by UV irradiation on gene expression in the protozoan parasite Entamoeba histolytica. Molec. Biochem. Parasitol. Published on-line 24 december 2008. doi: 10.1016/; j.molbiopara.2008.12. 005

WHO (1997). Entamoeba taxonomy. Bull. World Health Organ. 75: 291-294

Yedavali V, Neuveut C, Chi Y, Kleiman L, Jeang K (2004). Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. Cell; 119: 381-392.

Zamorano A, López-Camarillo C, Orozco E, Weber C, Guillen N, Marchat LA (2008). In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for Entamoeba histolytica mRNA polyadenylation. Comput. Biol. Chem. 32: 256-263.

Zhang Z, Mahajan S, Zhang X, Stanley SL Jr (2003). Tumor necrosis factor alpha is a key mediator of gut inflammation seen in amebic colitis in human intestine in the SCID mouse-human intestinal xenograft model of disease. Infect Immun. 71: 5355-5359.