



Taxon-Specific Proteins of the Pathogenic *Entamoeba* Species *E. histolytica* and *E. nuttalli*

Constantin König¹, Barbara Honecker¹, Ian W. Wilson², Gareth D. Weedall³, Neil Hall^{4,5}, Thomas Roeder^{6,7}, Nahla Galal Metwally¹ and Iris Bruchhaus^{1,8*}

¹ Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany, ² Institute of Infection, Veterinary & Ecological Sciences, University of Liverpool, Liverpool, United Kingdom, ³ School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, United Kingdom, ⁴ Earlham Institute, Norwich, United Kingdom, ⁵ School of Biological Sciences, University of East Anglia, Norwich, United Kingdom, ⁶ Zoology, Department of Molecular Physiology, Kiel University, Kiel, Germany, ⁷ Airway Research Center North (ARC/N), German Center for Lung Research (DZL), Kiel, Germany, ⁸ Department of Biology, University of Hamburg, Hamburg, Germany

OPEN ACCESS

Edited by:

Adrian Hehl,
University of Zurich, Switzerland

Reviewed by:

Shannon Moonah,
University of Virginia, United States
Elisa Azuara-Liceaga,
Universidad Autónoma de la Ciudad
de México, Mexico

*Correspondence:

Iris Bruchhaus
bruchhaus@bnitm.de

Specialty section:

This article was submitted to
Parasite and Host,
a section of the journal
Frontiers in Cellular
and Infection Microbiology

Received: 14 December 2020

Accepted: 03 March 2021

Published: 19 March 2021

Citation:

König C, Honecker B, Wilson IW, Weedall GD, Hall N, Roeder T, Metwally NG and Bruchhaus I (2021) Taxon-Specific Proteins of the Pathogenic *Entamoeba* Species *E. histolytica* and *E. nuttalli*. *Front. Cell. Infect. Microbiol.* 11:641472. doi: 10.3389/fcimb.2021.641472

The human protozoan parasite *Entamoeba histolytica* can live in the human intestine for months or years without generating any symptoms in the host. For unknown reasons, amoebae can suddenly destroy the intestinal mucosa and become invasive. This can lead to amoebic colitis or extraintestinal amoebiasis whereby the amoebae spread to other organs via the blood vessels, most commonly the liver where abscesses develop. *Entamoeba nuttalli* is the closest genetic relative of *E. histolytica* and is found in wild macaques. Another close relative is *E. dispar*, which asymptotically infects the human intestine. Although all three species are closely related, only *E. histolytica* and *E. nuttalli* are able to penetrate their host's intestinal epithelium. Lineage-specific genes and gene families may hold the key to understanding differences in virulence among species. Here we discuss those genes found in *E. histolytica* that have relatives in only one or neither of its sister species, with particular focus on the peptidase, AIG, Ariel, and BspA families.

Keywords: *Entamoeba*, peptidases, virulence, AIG, Ariel, BspA

INTRODUCTION

The intestinal protozoan *Entamoeba histolytica* is an important human parasite. Recent data clearly indicate that the life-threatening amoebic liver abscess (ALA) continues to be a common clinical complication of amoebiasis infection in Asian, African and Latin American countries with estimated 26700 death in 2016 (Collaborators, 2018; Shirley et al., 2019). *E. histolytica* can become invasive and cause amoebic colitis or amoebic liver abscess (ALA) formation. The life cycle of this parasite consists of infectious cysts that survive outside the host and vegetative trophozoites that proliferate in the human gut. In general, trophozoites persist asymptotically for months or years in the human intestine. However, in 10% of cases, the trophozoites become, under as yet unknown circumstances, invasive and induce extraintestinal amoebiasis. Invasion into the intestinal mucosa can lead to induction of amoebic colitis, whereas dissemination to the liver can result in ALA formation (Blessmann et al., 2003). A related species, *E. dispar*, is microscopically indistinguishable from *E. histolytica* and occurs only as a harmless commensal in the human

intestine. A key question in amoebic research is the elucidation of the mechanisms of *E. histolytica* invasion and tissue destruction. For decades, attempts have been made to identify the virulence factors of *E. histolytica* by comparative studies of both organisms at both biological and molecular level. These studies, and those that compared pathogenic and non-pathogenic *E. histolytica* isolates, have led to the identification of a number of virulence factors.

Three protein families (Gal/GalNAc lectins, cysteine peptidases and amoebapores) are of particular interest. The first step in the invasion process is to overcome the protective mucus barrier. Here, cysteine peptidases play an important role, along with a number of other molecules (Lidell et al., 2006; Thibaux et al., 2013). Subsequently, the adhesion of amoebae to target epithelial cells via the galactose/N-acetyl-D-galactosamine lectin (Gal/GalNAc) (Tannich et al., 1991; Petri and Schnaar, 1995). The Gal/GalNAc lectin is a 260 kDa heterodimer consisting of a disulfide-linked 170 kDa heavy chain (Hgl) and a GPI-anchored 35 kDa light chain (Lgl) (Petri et al., 2002). After establishing contact, amoebae can secrete amoebapores. These mediate a contact-dependent lysis of the target cells (Leippe, 1997). *E. histolytica* has three amoebapores (A, B and C), which are all characterized by a pore-forming activity. They are also capable of killing gram-positive bacteria by destroying bacterial cytoplasmic membranes. Thus, the functionality of these molecules is two-fold: they confer cytolytic activity to amoebae as well as an intracellular antimicrobial effect against phagocytosed bacteria (Leippe et al., 1994). Amoebae whose amoebapore expression is inhibited have a reduced antimicrobial activity and are non-pathogenic, i.e. they are not, or are only to a small extent, able to form ALAs in hamsters (Bracha et al., 1999; Bracha et al., 2003). Important for mucus degradation, invasion, as well as for the process of tissue degradation, are the cysteine peptidases (CPs) of *E. histolytica*. In the genome of *E. histolytica* 35 genes coding for CPs of the C1 papain superfamily can be identified. However, only four CPs (EhCP-A1, -A2, -A5, -A7) can be detected at the protein level in the trophozoite stage (Tillack et al., 2007; Irmer et al., 2009). These have previously been located in lysosomal-like vesicles, and some of them were found to be membrane-associated (Jacobs et al., 1998). The importance of CPs, particularly EhCP-A5, in ALA formation is evident in infections of laboratory animals, where overexpression of CPs leads to an increase in ALA size (Tillack et al., 2006). Conversely, reduced CP activity leads to a decrease in ALA formation ability in *E. histolytica* (Li et al., 1995; Stanley et al., 1995). In addition, an increase in the expression of some *ehcp* genes during ALA formation has previously been described, while non-pathogenic amoebae can be converted to pathogenic amoebae, simply by overexpression of some of specific *ehcp* genes (Matthiesen et al., 2013).

Homologs of the Gal/GalNAc lectins, cysteine peptidases and amoebapores are also present in non-pathogenic *E. dispar* (Nickel et al., 1999; Tillack et al., 2007; Weedall et al., 2011). Nevertheless, it has been shown that the CP activity of *E. dispar* is about 10-1000 times lower than that of *E. histolytica* and that

genes corresponding to *ehcp-a1* and *ehcp-a5* are absent or degenerated in *E. dispar* (Bruchhaus et al., 1996; Willhoeft et al., 1999b; Que and Reed, 2000). A lower amount of amoebapores and therefore reduced pore-forming activity was found in *E. dispar* in comparison to *E. histolytica* (Nickel et al., 1999). However, it is not yet clear whether these differences alone determine pathogenicity, whether additional genetic differences are involved, or whether pathogenicity is a result of complex interactions between various proteins.

In addition to the cysteine peptidases and the amoebapores, the AIG, BspA and Ariel families are always mentioned in connection with the virulence of *E. histolytica*. For members of the AIG family, a differential expression between pathogenic and non-pathogenic *E. histolytica* isolates could be shown (Biller et al., 2010). BspA-like molecules have also been described in *T. vaginalis* and are thought to play an important role in pathogenesis (Noel et al., 2010). The Ariel proteins are an *E. histolytica*-specific protein family that has not yet been found in any other *Entamoeba* species so far. Therefore, it is postulated that these molecules may have a role in virulence. However, the functions of the members of all three protein families are largely unclear.

Besides the characterization of individual proteins, one straightforward method for the identification of pathogenicity factors is a direct comparison of pathogenic and non-pathogenic *E. histolytica* isolates using comparative genomic, transcriptomic and proteomic approaches (Davis et al., 2006a; MacFarlane and Singh, 2006; Ehrenkaufer et al., 2007; Biller et al., 2010; Meyer et al., 2016; Nakada-Tsukui et al., 2018). Unfortunately, two isolates with very different genetic backgrounds were used in the majority of the studies cited (pathogenic isolate HM-1:IMSS and non-pathogenic isolate Rahman) (Davis et al., 2006a; MacFarlane and Singh, 2006; Davis et al., 2007; Ehrenkaufer et al., 2007). The non-pathogenic Rahman isolate has several serious functional defects (e.g. defective phagocytosis, reduced cytotoxicity, inability to grow in animals, and a truncated glycan chain of the proteophosphoglycan coating surface) (Davis et al., 2006a). The genomic differences between HM-1:IMSS and Rahman are small, however, DNA fragment duplications have been detected (Weedall et al., 2012). An alternative approach compared the transcriptomes of pathogenic and non-pathogenic clones derived from the isolate HM-1:IMSS were compared (Biller et al., 2009; Biller et al., 2010; Meyer et al., 2016). In total, approximately 90 genes are differentially expressed between the investigated non-pathogenic and pathogenic clones (Meyer et al., 2016). Based on transfectants, in which the identified genes were either overexpressed or silenced, it was possible to identify another pathogenicity factor, namely the hypothetical protein EHI_127670. When EHI_127670 is silenced in pathogenic amoebae their ability to form ALAs is reduced. On the other hand, overexpression of EHI_127670 in non-pathogenic amoebae leads to restoration of ALA formation ability (Meyer et al., 2016; Matthiesen et al., 2019). However, nothing is yet known about the function of the protein and it is not yet known with certainty that the RNA actually encodes a protein.

Using a transcriptome approach, Naiyer and colleagues were able to identify downstream regulatory motifs in very highly expressed *E. histolytica* genes that are very likely important for gene expression. These motifs were also detected in various genes encoding virulence factors, which are also highly expressed under axenic conditions. It can therefore be assumed that corresponding encoded proteins are important for optimal growth, but in addition also play a role in tissue invasion and virulence (Naiyer et al., 2019b). Comparative transcriptome analysis also makes it possible to better understand the biology of *E. histolytica* in general. With the help of this method, it is possible to identify molecules that play a role in phagocytosis, the stress response, the enzyme station and the excystation, among others [for review (Naiyer et al., 2019a)].

Besides *E. histolytica* and *E. dispar*, the genus *Entamoeba* includes many other species, some of which colonize the human intestine (*E. moshkovskii*, *E. bangladeshi*, *E. polecki*, *E. coli*, and *E. hartmanni*). However, in humans a severe course of disease with extraintestinal abscesses has so far only been described for *E. histolytica*. Pathogenic amoebae have also been described in reptiles (*E. invadens*) and macaques (*E. nuttalli*). *E. nuttalli* is the species most closely related to *E. histolytica*. Different species of wild macaques, as well as other non-human primates kept in captivity, have been identified as hosts for *E. nuttalli* (Tachibana et al., 2007; Tachibana et al., 2009; Levecke et al., 2010; Wei et al., 2018; Tanaka et al., 2019). In addition to the genomes of *E. histolytica* and *E. dispar*, the genome of *E. nuttalli* was recently sequenced and all genomes are available on AmoebaDB (Tanaka et al., 2019). The ability to compare the genomes of *E. histolytica*, *E. dispar* and *E. nuttalli* now opens up the possibility of identifying additional molecules responsible for the development of extraintestinal amoebiasis in human or non-human primates.

IDENTIFICATION OF HOMOLOGOUS PROTEINS UNIQUE TO PATHOGENIC *E. HISTOLYTICA* AND *E. NUTTALLI*

Recently, using an OrthoMCL approach, Wilson and colleagues studied the genetic diversity and gene family expansions for *E. histolytica*, *E. dispar*, *E. moshkovskii* and *E. invadens*. They identified 984 genes that were unique to *E. histolytica* (Table S8) (Wilson et al., 2019). Here, this list was used as a basis for a BlastP approach to analyze whether homologous proteins are found in *E. nuttalli*. Additionally, *E. dispar* was again included in this analysis. Pseudogenes were not included in the analysis and a sequence identity of $\geq 50\%$ was assumed to be a functionally homologous protein. In a BlastP analysis of all 927 translatable “unique” *E. histolytica* genes, 309 of them had homologs in *E. dispar* and *E. nuttalli* with an identity of $\geq 50\%$ (Figure 1, Table S7) (Wilson et al., 2019). Another 376 “unique” *E. histolytica* genes encode for proteins with homologs in *E. dispar* but not in *E. nuttalli* (Figure 1, Table S3). A further 67 putative proteins share $\geq 50\%$ identity only with *E. nuttalli* proteins (Figure 1, Table S6). Only 175 genes remain, whose encoded proteins do

not show homology with proteins of *E. dispar* or *E. nuttalli* (Figure 1, Table S1). That many of the proteins encoded by the “unique” *E. histolytica* genes were also found in *E. dispar* was very surprising. However, looking at the genomic organization on AmoebaDB, one finds that many of the “unique” *E. histolytica* genes are located at the edge of contigs and therefore do not contain the complete gene sequence (Wilson et al., 2019).

PROTEINS THAT COULD ONLY BE DETECTED IN *E. HISTOLYTICA*

Almost all 175 proteins found to be unique for *E. histolytica* are hypothetical proteins (Table S1). However, 18 genes encoding the asparagine-rich *E. histolytica* antigens (Ariel) could be identified (including two genes with an incomplete sequence) (Figure 1, Table S2) (Mai et al., 2000; Wilson et al., 2019). The fact that Ariel proteins are encoded by an *E. histolytica*-specific multicopy gene family was described by Willhoeft and colleagues 20 years ago (Willhoeft et al., 1999a). No homologous proteins can be detected in *E. dispar* or *E. invadens* (Wilson et al., 2019). However, one homologous protein (ENU1_012640) is found in *E. nuttalli*, with 92% identity to EHI_005260/EHI_188600. Except for the Ariel proteins EHI_057430 and EHI_172730, which did not contain a signal peptide, and EHI_131360 and EHI_185110 which did not contain a transmembrane domain, an N-terminal signal peptide and a C-terminal transmembrane domain can be predicted for all other Ariel proteins. Therefore, a surface localization can be postulated, however it is unknown whether the protein is a deterministic factor in the virulence of *E. histolytica*.

Two further genes unique to *E. histolytica* unique genes (EHI_107560 and EHI_157010) encode for proteins with approximately 55% sequence identity to aldo-keto reductases of plant chloroplasts and are annotated as alcohol dehydrogenases. However, both genes are located at the edges of contigs and are thus only partially represented in the genome assembly. EHI_107560 (187 aa) is located at the edge of contig DS571548, while EHI_157010 (158 aa) is located at the edge of contig DS571869 (AmoebaDB, release 48 beta, 20 Aug 2020). BlastP analysis of these proteins revealed two genes (EHI_029620, EHI_039190) in the genome of *E. histolytica* encoding the same protein sequence of 305 amino acids, termed aldolase reductase. If these proteins are used as a template for a search in the *Entamoeba* genomes, a protein with a sequence identity of 55% was also found in pathogenic *E. invadens* (EIN_497000), along with one with 73% sequence identity in *E. dispar* (EDI_260680). A phylogenetic analysis using the online tool Clustal Omega with the respective protein sequences as input (Sievers et al., 2011), showed the closest relationship to *Dictyostelium discoideum* and plant chloroplast-like aldo-keto reductases (Figure 2). In various Kinetoplastidae such as *Trypanosoma* and *Leishmania*, genes encoding chloroplast-like proteins were identified. It is assumed that the organelles were taken up by endosymbiosis before divergence. Later, the organelles were lost, but several genes

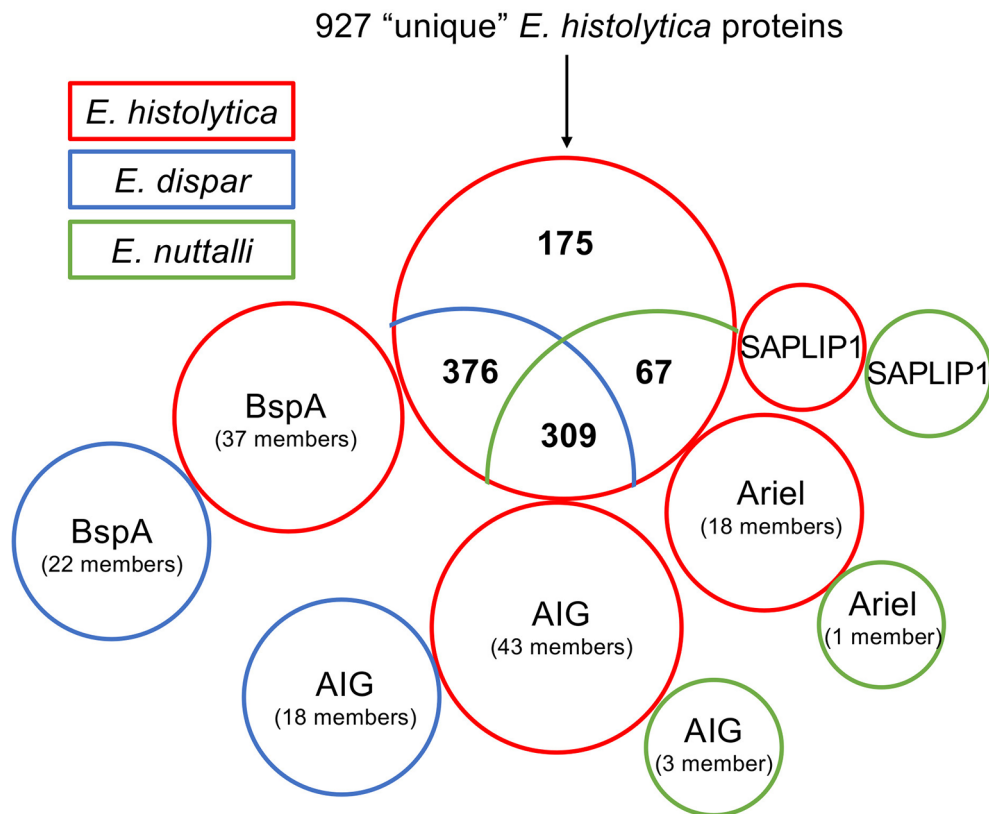


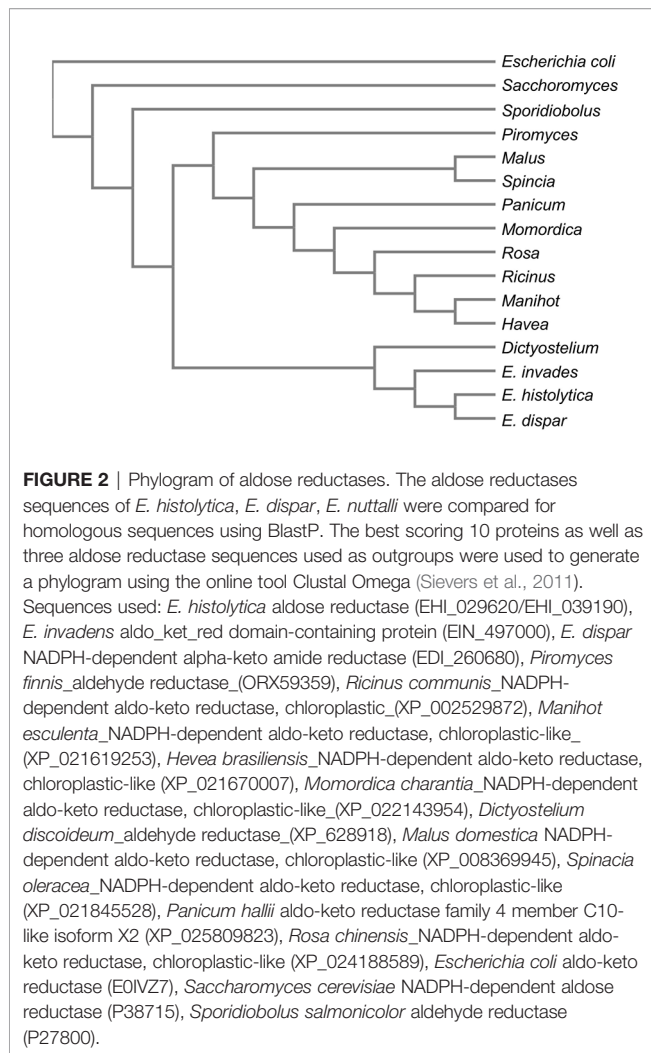
FIGURE 1 | Comparative blastp analysis of 927 proteins identified as unique for *E. histolytica* (Wilson et al., 2019). Shown is the comparison of *E. histolytica* with *E. nuttalli* and *E. dispar* as a venn diagram. The three protein families Ariel, AIG and BspA as well SAPLIP1 are shown separately. The proteins found in *E. histolytica* are boxed in red, those of *E. dispar* in blue and those of *E. nuttalli* in green.

from these organelles were integrated into the genome (Hannaert et al., 2003). Horizontal gene transfer has also been described for *Entamoeba*. However, almost all identified genes could be traced back to bacteria (mainly Bacteroidetes) so far (Romero et al., 2016). Comparison of the amino acid sequences revealed that *E. dispar* EDI_260680, contains a deletion of about 50 aa. Comparison of the sequence with those of plant aldo-keto reductases shows that this area contains amino acids that are important for the enzymes' activity (Sengupta et al., 2015). In plants aldo-keto reductases play a role in stress response and detoxification of toxic aldehydes among other things (Sengupta et al., 2015). Since nothing is known about the function of the protein in *E. histolytica*, no statement can be made as to whether the deletion present in *E. dispar* has an influence on its activity.

PROTEINS OCCURRING IN *E. HISTOLYTICA* AND *E. DISPAR*, BUT NOT IN *E. NUTTALLI*

In the gene set of *E. histolytica* "unique" genes underlying this work, 37 genes encoding members of the BspA family (leucine-

rich repeat protein) were identified (Wilson et al., 2019). BlastP analysis of the proteins encoded by *bspa* showed that for all but two (EHI_098720/EHI_173850: 49% identity), homologous proteins (in total 22 members) with an identity of $\geq 50\%$ could be detected in *E. dispar* (Figure 1, Tables S3 and S4). However, no homologous proteins could be found in *E. nuttalli*. The length of *bspa*-encoded proteins listed here ranges from 101 to 447 aa. It is striking that almost all *bspas* are found on very short contigs and then often on the edge of the contigs. Therefore, it is not clear whether all of these are really full-length proteins since, with a few exceptions, the BspA proteins characterized so far have an average length of about 550 amino acids (Davis et al., 2006b). For some members of the BspA family a surface localization was described, although neither signal sequences nor transmembrane domains were detected (Davis et al., 2006b, Silvestre et al., 2015). However, the phenomenon of a surface localization without the detection of protein domains that would allow membrane anchoring has been described for a number of other proteins (Biller et al., 2014). It has been shown that BspA-like proteins are involved in adhesion to extracellular membranes, epithelial cell invasion and fibronectin and fibrinogen binding (Mengaud et al., 1996; Sharma et al., 1998; Inagaki et al., 2006). A BspA-like gene family could also be



identified in *Trichomonas vaginalis* (Hirt et al., 2002; Noel et al., 2010). The BspA proteins of *T. vaginalis* are believed to play various and important roles in the pathobiology of this parasite by contributing to invasion and long-term infections of the urogenital tract (Noel et al., 2010). However, the exact functions of the proteins in *T. vaginalis* as well as in *E. histolytica* and *E. dispar* are not yet deciphered. As these are large gene families, it can be assumed that they play an important role in the life cycle of the amoebae. It is therefore all the more surprising that homologous proteins are not detectable in the closest relative of *E. histolytica*, *E. nuttalli*.

Wilson and colleagues found 17 members of the AIG family that were “unique” to *E. histolytica* (Wilson et al., 2019). However, by BlastP analysis 8 AIG proteins were identified in *E. dispar*, showing 49 – 77% homology to the 17 different *E. histolytica* AIG1 proteins. Eight of the *E. histolytica* AIGs are most similar (53 – 63% identity) to *E. dispar* EDI_185310. For only one AIG1, EHI_126560, homologs were identified in both *E. dispar* (EDI_274460, 77% identity) and *E. nuttalli* (ENU1_158210, 97% identity) (Tables S3, S5). In addition to

the 17 genes, 26 more genes encoding AIG proteins can be identified in the *E. histolytica* genome. For all these proteins, homologs with a sequence identity between 53 and 83% can be detected in *E. dispar*. However, seven *E. histolytica* AIGs each show the highest homology to only two *E. dispar* AIGs (EDI_036000 and EDI_243490). Furthermore, two more AIGs, ENU1_161270 and ENU1_207600, are found in *E. nuttalli*, which are homologous to AIGs from *E. histolytica* (EHI_180390, 95% identity, and EHI_191790, 98% identity, respectively) (Figure 1, Table S5).

AIGs belong to the GTPases, were originally isolated from *Arabidopsis thaliana* and are thought to confer resistance to bacterial infections. It is also believed that AIG proteins are involved in the development of *A. thaliana* and its response to environmental stimuli (Reuber and Ausubel, 1996). Orthologous proteins are also found in mammals and play a role in B-cell and T-cell development via interaction with proteins of the Bcl2 family (Nitta et al., 2006; Nitta and Takahama, 2007). However, very little is known about the function of AIG proteins in *E. histolytica*. In a comparative genome analysis of *E. histolytica* isolated from a patient presenting with diarrhea and an asymptomatic patient, it was shown that one *aig* gene from a tandem array of three *aig* genes was deleted by homologous recombination in the isolate from the asymptomatic patient. Overexpression of this *aig* (EHI_176590) resulted in increased formation of cell surface protrusions and increased adhesion to human erythrocytes. Furthermore, the gene EHI_176590 was detected in approximately 60% of stool samples from symptomatic patients, but only in 15% of stool samples from asymptomatic individuals infected with *E. histolytica*. It is therefore postulated that the AIG protein plays a central role in the virulence of *E. histolytica* by regulating host cell adhesion (Nakada-Tsukui et al., 2018). In addition, a quantitative real-time PCR approach showed that 18 of 34 investigated *aig* genes are increasingly expressed in pathogenic compared to non-pathogenic HM-1:IMSS cell lines (Biller et al., 2010). But as with the Ariel family and the BspA family, the function of the AIG family is not yet clear.

PROTEINS OCCURRING IN *E. HISTOLYTICA* AND *E. NUTTALLI*, BUT NOT IN *E. DISPAR*

In the set of “unique” *E. histolytica* genes on which this work is based, 67 had homologs in *E. nuttalli* but not in *E. dispar* (Figure 1, Table S6) (Wilson et al., 2019). As already described for the other comparisons, these are mostly hypothetical proteins. However, the list also contains the non-pathogenic pore-forming peptide EHI_169350. EHI_169350 (entbd24tf or SAPLIP1) belongs to a family of 15 saposin-like proteins (SAPLIPs) which were first identified by Bruhn and Leippe (Bruhn and Leippe, 2001; Winkelmann et al., 2006). Like amoebapore A, B and C the SAPLIPs are characterized by a conserved sequence motif consisting of six cysteine residues.

Interestingly, of the entire SAPLIP family, only EHI_169350 can be clearly assigned to the amoebapore subfamily, with the largest sequence identity of about 65% shared with amoebapore A (Bruhn and Leippe, 2001). Only two members of the SAPLIP family including EHI_169350 possess a typical signal peptide and only EHI_169350 resembles amoebapores in its net charge and dispersed charge distribution, with the greatest similarity being to amoebapore C. However, EHI_169350 lacks the typical C-terminal histidine residue of amoebapores, which is essential for oligomerization during channel formation (Andra and Leippe, 1994). EHI_169350 has an asparagine residue at position 90 instead. It can therefore be assumed that EHI_169350 has no pore-forming activity (Bruhn and Leippe, 2001). No statement can therefore be made about the function of EHI_169350 especially concerning its role in pathogenicity.

THE PEPTIDASES OF *E. HISTOLYTICA*, *E. DISPAR* AND *E. NUTTALLI*

As described above, CPs play an important role in the destruction and invasion of human tissue. This has been demonstrated in a variety of *in vitro* and *in vivo* studies (Bruchhhaus and J, 2015; Gilmartin et al., 2020). A new *in silico* analysis of the genome of *E. histolytica* (AmoebaDB, release 48 beta, 27 Aug 2020) revealed a total of 33 genes encoding CPs of clan CA, C1 (papain-like) family (Table S9). Twelve of the CPs could be assigned to the CP-A family, ten to the CP-B family and eleven to the CP-C family. In *E. dispar* there are only nine members of the CP-A family. Genes coding for EdCP-A1, -A5, -A7 are not present or are present as pseudogenes. The CP-B family consists of seven members; here the genes coding for EdCP-B1, -B8 and -B9 are missing. The EdCP-C family has

eleven members, like the EhCP-C family. In *E. nuttalli* there are no homologous proteins to EhCP-A1, -A6, -A7, and -A8; thus the EnCP-A family comprises eight members. The EnCP-B family consists of nine members (EnCP-B2 is missing here), and EnCP-C5 is missing in the CP-C family, so there are ten members in total (Figure 3, Table S9).

Only for four of the CP encoding genes (*ehcp-a1*, -a2, -a5 and -a7) could it be shown that they are highly expressed in *E. histolytica* under the standard axenic culture conditions (Bruchhhaus et al., 2003; Clark et al., 2007; Tillack et al., 2007). Interestingly, the ability to disrupt a cell monolayer (cytopathic activity) was dramatically increased for amoebae overexpressing *ehcp-a5*, whereas it showed only a moderate increase in transfectants overexpressing *ehcp-a1* or *ehcp-a2*. Furthermore, the overexpression of *ehcp-a1* and *ehcp-a2* in *E. histolytica* trophozoites did not increase ALA formation in gerbils, whereas overexpression of *ehcp-a5* resulted in significantly larger ALAs compared to controls. If the *ehcp5* is overexpressed in the HM-1:IMSS derived G3 isolate, in which the *amoebapore a* gene is silenced and therefore has only low virulence, this is sufficient to compensate for the reduction in virulence (Hellberg et al., 2001; Tillack et al., 2006). In previous studies EhCP-A5 was also believed to play an important role in intestinal invasion. EhCP-A5 has been shown to interact directly with the integrins on the surface of human colon epithelial cells and induce the secretion of pro-inflammatory cytokines (Hou et al., 2010). Thus, amoebae silenced for *ehcp-a5* expression do not trigger an inflammatory response of the host and do not induce the collagen remodeling required for invasion. Further investigations showed that CP-A5 can convert the pro-matrix metalloproteinase (MMP)-3 into its active form, which in turn activates pro-MMP-1 (Bansal et al., 2009; Thibeaux et al., 2012; Thibeaux et al., 2014). The

Cysteine peptidases			Cysteine peptidases			Aspartic peptidases			Metallo peptidases		
<i>E. dispar</i>	<i>E. histolytica</i>	<i>E. nuttalli</i>	<i>E. dispar</i>	<i>E. histolytica</i>	<i>E. nuttalli</i>	<i>E. dispar</i>	<i>E. histolytica</i>	<i>E. nuttalli</i>	<i>E. dispar</i>	<i>E. histolytica</i>	<i>E. nuttalli</i>
EdCP-A2	EhCP-A1	EnCP-A2	EdCALP1	EhCALP1	EnCALP1	EdAsP22-1	EhAsP22-1	EnAsP22-1	EdMP1-1	EhMP1-1	EnMP1-1
EdCP-A3	EhCP-A2	EnCP-A3	EdCALP2	EhCALP2	EnCALP2	EdAsP22-1	EhAsP22-1	EnAsP22-1	EdMP3-1	EhMP3-1	EnMP3-1
EdCP-A4	EhCP-A3	EnCP-A4	EdUBHY	EhUBHY	EnUBHY	EdAsP22-1	EhAsP22-1	EnAsP22-1	EdMP3-2	EhMP3-2	EnMP3-2
EdCP-A5	EhCP-A4	EnCP-A5	EdUOH	EhUOH	EnUOH	EdAsP22-1	EhAsP22-1	EnAsP22-1	EdMP8-1	EhMP8-1	EnMP8-1
EdCP-A6	EhCP-A5	EnCP-A6	EdUlp1-1	EhUlp1-1	EnUlp1-1				EdMP8-2	EhMP8-2	EnMP8-2
EdCP-A7	EhCP-A6	EnCP-A7	EdUlp1-2	EhUlp1-2	EnUlp1-2				EdMP48-1	EhMP48-1	EnMP48-1
EdCP-A8	EhCP-A7	EnCP-A8	EdUlp1-3	EhUlp1-3	EnUlp1-3				EdMP16-1	EhMP16-1	EnMP16-1
EdCP-A9	EhCP-A8	EnCP-A9	EdAUTO1	EhAUTO1	EnAUTO1				EdMP24-1	EhMP24-1	EnMP24-1
EdCP-A10	EhCP-A9	EnCP-A10	EdAUTO2	EhAUTO2	EnAUTO2				EdMP24-2	EhMP24-2	EnMP24-2
EdCP-A11	EhCP-A10	EnCP-A11	EdAUTO3	EhAUTO3	EnAUTO3				EdMP24-3	EhMP24-3	EnMP24-3
EdCP-A12	EhCP-A11	EnCP-A12	EdAUTO4	EhAUTO4	EnAUTO4				EdMP24-5	EhMP24-5	EnMP24-5
EdCP-B1	EhCP-A12	EnCP-B1	EdOTU	EhOTU	EnOTU				EdMP24-6	EhMP24-6	EnMP24-6
EdCP-B2	EhCP-B1	EnCP-B2							EdMP18-1	EhMP18-1	EnMP18-1
EdCP-B3	EhCP-B2	EnCP-B3							EdMP18-2	EhMP18-2	EnMP18-2
EdCP-B4	EhCP-B3	EnCP-B4							EdMP20-1	EhMP20-1	EnMP20-1
EdCP-B5	EhCP-B4	EnCP-B5							EdMP20-2	EhMP20-2	EnMP20-2
EdCP-B6	EhCP-B5	EnCP-B6							EdMP20-3	EhMP20-3	EnMP20-3
EdCP-B7	EhCP-B6	EnCP-B7							EdMP20-4	EhMP20-4	EnMP20-4
EdCP-B8	EhCP-B7	EnCP-B8							EdMP22-1	EhMP22-1	EnMP22-1
EdCP-B9	EhCP-B8	EnCP-B9							EdMP48-1	EhMP48-1	EnMP48-1
EdCP-B10	EhCP-B9	EnCP-B10									
EdCP-C1	EhCP-B10	EnCP-C1									
EdCP-C2	EhCP-C1	EnCP-C2									
EdCP-C3	EhCP-C2	EnCP-C3									
EdCP-C4	EhCP-C3	EnCP-C4									
EdCP-C5	EhCP-C4	EnCP-C5									
EdCP-C6	EhCP-C5	EnCP-C6									
EdCP-C7	EhCP-C6	EnCP-C7									
EdCP-C8	EhCP-C7	EnCP-C8									
EdCP-C9	EhCP-C8	EnCP-C9									
EdCP-C10	EhCP-C9	EnCP-C10									
EdCP-C11	EhCP-C10	EnCP-C11									
EdCP-C12	EhCP-C11	EnCP-C12									
EdCP-C13	EhCP-C12	EnCP-C13									

FIGURE 3 | Cysteine, asparagine, serine and metallopeptidases of *E. histolytica*, *E. dispar* and *E. nuttalli* (for more details see Table S9).

observation that CP-A5 occurs in the two pathogenic amoeba species *E. histolytica* and *E. nuttalli*, but not in the non-pathogenic species *E. dispar*, supports the significance of this peptidase as an important pathogenicity factor (**Figure 3, Table S9**). However, a CP-A5 homologue is not detectable in the genome of the reptile pathogen *E. invadens*.

In the genome of *E. histolytica* another twelve genes encoding CPs can be identified and assigned to five different families (C2, C19, C48, C54, C65; all clan CA) (**Figure 3, Table S9**). Except for the Ulp1 protease Ulp1-3 (C48 family) in *E. nuttalli*, homologs for all CPs in *E. dispar* and *E. nuttalli* could be found.

Homologs for the four members of the aspartic peptidase family (clan AD, family A22, A) of *E. histolytica* are found in both *E. dispar* and *E. nuttalli*. However, nothing is known about the function of these peptidases (**Figure 3, Table S9**).

Nine serine peptidases can be detected, which can be assigned to four families (clan SC, family S9, C; Clan SF, family S26, B; Clan SC, S28; clan ST, family S54) (**Figure 3, Table S9**). For the serine peptidases EhSP26-2, EhSP28-1 and EhSP28-2 no homologs can be detected in *E. nuttalli*. Functional analysis has only been performed for the S28 and rhomboid proteases. Both are found associated with the amoeba membrane and the rhomboid protease probably plays an important role in mobility and adhesion of amoebae to the host tissue (Barrios-Ceballos et al., 2005; Baxt et al., 2008; Rastew et al., 2015; Welter et al., 2020).

A total of 21 metallopeptidases belonging to eleven different families were identified in the genome of *E. histolytica* (**Figure 3, Table S9**). Only for 2 of them no homologous proteins can be identified in *E. dispar*. These are MP8-1 (cell surface protease gp63) and U48-1, a CAAX prenyl protease. In *E. nuttalli* the two peptidases can be detected, just like in *E. histolytica*. For EhMP8-1, Teixeira and colleagues showed that it is a functional metallopeptidase localized on the surface of *E. histolytica* trophozoites. By silencing *ehmp8-1* expression, the adherence of trophozoites to cells was increased, while the surface staining of several antigens, including the Gal/GalNAc lectin, remained unchanged. Amoebae which were silenced for *ehmp8-1* expression also showed decreased cytopathic activity and reduced mobility, but phagocytic activity was increased (Teixeira et al., 2012). In contrast to the EhMP8-1 there is unfortunately very little information about the EhMP8-2. The *ehmp8-2* gene is approximately 150 times more highly expressed in non-pathogenic amoebae than in pathogenic amoebae. Furthermore, it was shown that overexpression of *ehmp8-2* in pathogenic amoebae significantly reduced ALA formation in the mouse model. Thus, the presence of EhMP8-2 leads to a non-pathogenic phenotype of the amoebae (Meyer et al., 2016). Another peptidase, which is found only in *E. histolytica* and *E. nuttalli*, but not in *E. dispar*, is a CAAX prenyl protease of unknown function in the U48 family. The question arises for EhMP8-2 as well as for EhU48-1 whether the absence of these molecules has an influence on the virulence of amoebae.

CONCLUSION

Unfortunately, the majority of proteins found exclusively in pathogenic *E. histolytica*, or in *E. histolytica* and *E. nuttalli* but not non-pathogenic *E. dispar*, lack functional annotations. With Ariel, BspA and AIG, *E. histolytica* has three large protein families whose members are repeatedly discussed as virulence factors. The Ariel family, consisting of 18 members, is almost exclusively found in *E. histolytica*. Only one member was detected in *E. nuttalli*. In contrast, proteins of the BspA family are found in *E. histolytica* and *E. dispar*, and members of the AIG family in all three organisms. However, it is noticeable that *E. histolytica* contributes the most members to each of these protein families. One reason for this observation could be that, in analyses like these, incomplete or fragmented genome assemblies may contain gaps and genes may be missing in the annotated gene set rather because of this than because of the actual absence of these genes. For example, the lysine and glutamic acid-rich protein KERP1 is a virulence factor that is active in the development of amoebic liver abscesses (Seigneur et al., 2005; Santi-Rocca et al., 2008). Originally, it was assumed to be absent in the genome of *E. dispar*, but in fact it is present but only partially represented in the genome assembly (Weedall, 2020). Such a situation is probably a factor in defining genes as lineage-specific. It is probably also a wide-ranging confounding factor for *Entamoeba* genomes, which are particularly difficult to assemble for a number of reasons (Weedall and Hall, 2011), and is likely to have a greater impact on genome assemblies with low coverage such as *E. dispar*. Therefore, the correctness of the genome assembly should first be verified before further investigations on the influence of the *E. histolytica*-unique proteins on virulence are performed.

AUTHOR CONTRIBUTIONS

All authors (JK, BH, IW, GDW, NH, TR, NM, and IB) wrote, edited, and reviewed the drafts of the manuscript, and agree to be accountable for the content of the work. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft (BR1744/17-1), Joachim Herz Stiftung (Joachim Herz Graduate School (BH)) und Jürgen Manchot Stiftung (CK).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcimb.2021.641472/full#supplementary-material>

REFERENCES

- Andra, J., and Leippe, M. (1994). Pore-forming peptide of *Entamoeba histolytica*. Significance of positively charged amino acid residues for its mode of action. *FEBS Lett.* 354, 97–102. doi: 10.1016/0014-5793(94)01103-6
- Bansal, D., Ave, P., Kerneis, S., Frileux, P., Boche, O., Baglin, A. C., et al. (2009). An ex-vivo human intestinal model to study *Entamoeba histolytica* pathogenesis. *PLoS Negl. Trop. Dis.* 3, e551. doi: 10.1371/journal.pntd.0000551
- Barrios-Ceballos, M. P., Martinez-Gallardo, N. A., Anaya-Velazquez, F., Mirelman, D., and Padilla-Vaca, F. (2005). A novel protease from *Entamoeba histolytica* homologous to members of the family S28 of serine proteases. *Exp. Parasitol.* 110, 270–275. doi: 10.1016/j.exppara.2005.02.022
- Baxt, L. A., Baker, R. P., Singh, U., and Urban, S. (2008). An *Entamoeba histolytica* rhomboid protease with atypical specificity cleaves a surface lectin involved in phagocytosis and immune evasion. *Genes Dev.* 22, 1636–1646. doi: 10.1101/gad.1667708
- Biller, L., Schmidt, H., Krause, E., Gelhaus, C., Matthiesen, J., Handal, G., et al. (2009). Comparison of two genetically related *Entamoeba histolytica* cell lines derived from the same isolate with different pathogenic properties. *Proteomics*. 9, 4107–4120. doi: 10.1002/pmic.200900022
- Biller, L., Davis, P. H., Tillack, M., Matthiesen, J., Lotter, H., Stanley, S. L. Jr., et al. (2010). Differences in the transcriptome signatures of two genetically related *Entamoeba histolytica* cell lines derived from the same isolate with different pathogenic properties. *BMC Genomics* 11, 63. doi: 10.1186/1471-2164-11-63
- Biller, L., Matthiesen, J., Kuhne, V., Lotter, H., Handal, G., Nozaki, T., et al. (2014). The Cell Surface Proteome of *Entamoeba histolytica*. *Mol. Cell Proteomics*. 13, 132–144. doi: 10.1074/mcp.M113.031393
- Blessmann, J., Ali, I. K., Nu, P. A., Dinh, B. T., Viet, T. Q., Van, A. L., et al. (2003). Longitudinal study of intestinal *Entamoeba histolytica* infections in asymptomatic adult carriers. *J. Clin. Microbiol.* 41, 4745–4750. doi: 10.1128/JCM.41.10.4745-4750.2003
- Bracha, R., Nuchamowitz, Y., Leippe, M., and Mirelman, D. (1999). Antisense inhibition of amoebapore expression in *Entamoeba histolytica* causes a decrease in amoebic virulence. *Mol. Microbiol.* 34, 463–472. doi: 10.1046/j.1365-2958.1999.01607.x
- Bracha, R., Nuchamowitz, Y., and Mirelman, D. (2003). Transcriptional silencing of an amoebapore gene in *Entamoeba histolytica*: molecular analysis and effect on pathogenicity. *Eukaryot. Cell*. 2, 295–305. doi: 10.1128/EC.2.2.295-305.2003
- Bruchhaus, I., Jacobs, T., Leippe, M., and Tannich, E. (1996). *Entamoeba histolytica* and *Entamoeba dispar*: differences in numbers and expression of cysteine proteinase genes. *Mol. Microbiol.* 22, 255–263. doi: 10.1046/j.1365-2958.1996.00111.x
- Bruchhaus, I., Loftus, B. J., Hall, N., and Tannich, E. (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. doi: 10.1128/EC.2.3.501-509.2003
- Bruchhaus, I., and Matthiesen, J. (2015). “Cysteine Peptidases in Pathogenesis,” in *Amoebiasis - Biology and Pathogenesis of Entamoeba* Springer. Eds. T. Nozaki and A. Bhattacharya, ISBN 978-4-431-55199-7.
- Bruhn, H., and Leippe, M. (2001). Novel putative saposin-like proteins of *Entamoeba histolytica* different from amoebapores. *Biochim. Biophys. Acta* 1514, 14–20. doi: 10.1016/S0005-2736(01)00345-5
- Clark, C. G., Alsmark, U. C., Tazreiter, M., Saito-Nakano, Y., Ali, V., Marion, S., et al. (2007). Structure and content of the *Entamoeba histolytica* genome. *Adv. Parasitol.* 65, 51–190. doi: 10.1016/S0065-308X(07)65002-7
- Collaborators GBDDD (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect. Dis.* 18, 1211–1228. doi: 10.1016/S1473-3099(18)30362-1
- Davis, P. H., Zhang, X., Guo, J., Townsend, R. R., and Stanley, S. L. Jr. (2006a). Comparative proteomic analysis of two *Entamoeba histolytica* strains with different virulence phenotypes identifies peroxiredoxin as an important component of amoebic virulence. *Mol. Microbiol.* 61, 1523–1532. doi: 10.1111/j.1365-2958.2006.05344.x
- Davis, P. H., Zhang, Z., Chen, M., Zhang, X., Chakraborty, S., and Stanley, S. L. Jr. (2006b). Identification of a family of BspA like surface proteins of *Entamoeba histolytica* with novel leucine rich repeats. *Mol. Biochem. Parasitol.* 145, 111–116. doi: 10.1016/j.molbiopara.2005.08.017
- Davis, P. H., Schulze, J., and Stanley, S. L. 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. doi: 10.1016/j.molbiopara.2006.10.014
- Ehrenkauf, G. M., Haque, R., Hackney, J. A., Eichinger, D. J., and Singh, U. (2007). Identification of developmentally regulated genes in *Entamoeba histolytica*: insights into mechanisms of stage conversion in a protozoan parasite. *Cell Microbiol.* 9, 1426–1444. doi: 10.1111/j.1462-5822.2006.00882.x
- Gilmartin, A. A., Ralston, K. S., and Petri, W. A. (2020). Inhibition of Amebic Cysteine Proteases Blocks Amebic Trophocytosis but Not Phagocytosis. *J. Infect. Dis.* 221, 1734–1739. doi: 10.1093/infdis/jiz671
- Hannaert, V., Saavedra, E., Duffieux, F., Szikora, J. P., Rigden, D. J., Michels, P. A., et al. (2003). Plant-like traits associated with metabolism of Trypanosoma parasites. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1067–1071. doi: 10.1073/pnas.0335769100
- Hellberg, A., Nickel, R., Lotter, H., Tannich, E., and Bruchhaus, I. (2001). Overexpression of cysteine proteinase 2 in *Entamoeba histolytica* or *Entamoeba dispar* increases amoeba-induced monolayer destruction in vitro but does not augment amoebic liver abscess formation in gerbils. *Cell Microbiol.* 3, 13–20. doi: 10.1046/j.1462-5822.2001.00086.x
- Hirt, R. P., Harriman, N., Kajava, A. V., and Embley, T. M. (2002). A novel potential surface protein in *Trichomonas vaginalis* contains a leucine-rich repeat shared by micro-organisms from all three domains of life. *Mol. Biochem. Parasitol.* 125, 195–199. doi: 10.1016/S0166-6851(02)00211-6
- Hou, Y., Mortimer, L., and Chadee, K. (2010). *Entamoeba histolytica* cysteine proteinase 5 binds integrin on colonic cells and stimulates NFκB-mediated pro-inflammatory responses. *J. Biol. Chem.* 285, 35497–35504. doi: 10.1074/jbc.M109.066035
- Inagaki, S., Onishi, S., Kuramitsu, H. K., and Sharma, A. (2006). Porphyromonas gingivalis vesicles enhance attachment, and the leucine-rich repeat BspA protein is required for invasion of epithelial cells by “Tannerella forsythia”. *Infect. Immun.* 74, 5023–5028. doi: 10.1128/IAI.00062-06
- Irmer, H., Tillack, M., Biller, L., Handal, G., Leippe, M., Roeder, T., et al. (2009). Major cysteine peptidases of *Entamoeba histolytica* are required for aggregation and digestion of erythrocytes but are dispensable for phagocytosis and cytopathogenicity. *Mol. Microbiol.* 72, 658–667. doi: 10.1111/j.1365-2958.2009.06672.x
- Jacobs, T., Bruchhaus, I., Dandekar, T., Tannich, E., and Leippe, M. (1998). Isolation and molecular characterization of a surface-bound proteinase of *Entamoeba histolytica*. *Mol. Microbiol.* 27, 269–276. doi: 10.1046/j.1365-2958.1998.00662.x
- Leippe, M., Andra, J., Nickel, R., Tannich, E., and Muller-Eberhard, H. J. (1994). Amoebapores, a family of membranolytic peptides from cytoplasmic granules of *Entamoeba histolytica*: isolation, primary structure, and pore formation in bacterial cytoplasmic membranes. *Mol. Microbiol.* 14, 895–904. doi: 10.1111/j.1365-2958.1994.tb01325.x
- Leippe, M. (1997). Amoebapores. *Parasitol. Today* 13, 178–183. doi: 10.1016/S0169-4758(97)01038-7
- Leveck, B., Dreesen, L., Dorn, P., Verweij, J. J., Vercammen, F., Casaert, S., et al. (2010). Molecular identification of *Entamoeba* spp. in captive nonhuman primates. *J. Clin. Microbiol.* 48, 2988–2990. doi: 10.1128/JCM.00013-10
- Li, E., Yang, W. G., Zhang, T., and Stanley, S. L. Jr. (1995). Interaction of laminin with *Entamoeba histolytica* cysteine proteinases and its effect on amoebic pathogenesis. *Infect. Immun.* 63, 4150–4153. doi: 10.1128/IAI.63.10.4150-4153.1995
- Lidell, M. E., Moncada, D. M., Chadee, K., and Hansson, G. C. (2006). *Entamoeba histolytica* cysteine proteases cleave the MUC2 mucin in its C-terminal domain and dissolve the protective colonic mucus gel. *Proc. Natl. Acad. Sci. U. S. A.* 103, 9298–9303. doi: 10.1073/pnas.0600623103
- MacFarlane, R. C., and Singh, U. (2006). Identification of differentially expressed genes in virulent and nonvirulent *Entamoeba* species: potential implications for amoebic pathogenesis. *Infect. Immun.* 74, 340–351. doi: 10.1128/IAI.74.1.340-351.2006
- Mai, J., Waisman, D. M., and Sloane, B. F. (2000). Cell surface complex of cathepsin B/annexin II tetramer in malignant progression. *Biochim. Biophys. Acta* 1477, 215–230. doi: 10.1016/S0167-4838(99)00274-5
- Matthiesen, J., Bar, A. K., Bartels, A. K., Marien, D., Ofori, S., Biller, L., et al. (2013). Overexpression of specific cysteine peptidases confers pathogenicity to

- a nonpathogenic *Entamoeba histolytica* clone. *MBio*. 4, e00072–e00013. doi: 10.1128/mBio.00072-13
- Matthiesen, J., Lender, C., Haferkorn, A., Fehling, H., Meyer, M., Matthies, T., et al. (2019). Trigger-induced RNAi gene silencing to identify pathogenicity factors of *Entamoeba histolytica*. *FASEB J.* 33, 1658–1668. doi: 10.1096/fj.201801313R
- Mengaud, J., Lecuit, M., Lebrun, M., Nato, F., Mazie, J. C., and Cossart, P. (1996). Antibodies to the leucine-rich repeat region of internalin block entry of *Listeria monocytogenes* into cells expressing E-cadherin. *Infect. Immun.* 64, 5430–5433. doi: 10.1128/IAI.64.12.5430-5433.1996
- Meyer, M., Fehling, H., Matthiesen, J., Lorenzen, S., Schuldt, K., Bernin, H., et al. (2016). Overexpression of differentially expressed genes identified in non-pathogenic and pathogenic *Entamoeba histolytica* clones allow identification of new pathogenicity factors involved in amoebic liver abscess formation. *PLoS Pathog.* 12, e1005853. doi: 10.1371/journal.ppat.1005853
- Naiyer, S., Bhattacharya, A., and Bhattacharya, S. (2019a). Advances in *Entamoeba histolytica* Biology Through Transcriptomic Analysis. *Front. Microbiol.* 10, 1921. doi: 10.3389/fmicb.2019.01921
- Naiyer, S., Kaur, D., Ahamed, J., Singh, S. S., Singh, Y. P., Thakur, V., et al. (2019b). Transcriptomic analysis reveals novel downstream regulatory motifs and highly transcribed virulence factor genes of *Entamoeba histolytica*. *BMC Genomics* 20, 206. doi: 10.1186/s12864-019-5570-z
- Nakada-Tsukui, K., Sekizuka, T., Sato-Ebine, E., Escueta-de Cadiz, A., Ji, D. D., Tomii, K., et al. (2018). AIG1 affects in vitro and in vivo virulence in clinical isolates of *Entamoeba histolytica*. *PLoS Pathog.* 14, e1006882. doi: 10.1371/journal.ppat.1006882
- Nickel, R., Ott, C., Dandekar, T., and Leippe, M. (1999). Pore-forming peptides of *Entamoeba dispar*. Similarity and divergence to amoebapores in structure, expression and activity. *Eur. J. Biochem.* 265, 1002–1007. doi: 10.1046/j.1432-1327.1999.00807.x
- Nitta, T., and Takahama, Y. (2007). The lymphocyte guard-IANs: regulation of lymphocyte survival by IAN/GIMAP family proteins. *Trends Immunol.* 28, 58–65. doi: 10.1016/j.it.2006.12.002
- Nitta, T., Nasreen, M., Seike, T., Goji, A., Ohigashi, I., Miyazaki, T., et al. (2006). IAN family critically regulates survival and development of T lymphocytes. *PLoS Biol.* 4, e103. doi: 10.1371/journal.pbio.0040103
- Noel, C. J., Diaz, N., Sicheritz-Ponten, T., Safarikova, L., Tachezy, J., Tang, P., et al. (2010). *Trichomonas vaginalis* vast BspA-like gene family: evidence for functional diversity from structural organisation and transcriptomics. *BMC Genomics* 11, 99. doi: 10.1186/1471-2164-11-99
- Petri, W. A. Jr., and Schnaar, R. L. (1995). Purification and characterization of galactose- and N-acetylgalactosamine-specific adhesin lectin of *Entamoeba histolytica*. *Methods Enzymol.* 253, 98–104. doi: 10.1016/S0076-6879(95)53011-8
- Petri, W. A. Jr., Haque, R., and Mann, B. J. (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. doi: 10.1146/annurev.micro.56.012302.160959
- Que, X., and Reed, S. L. (2000). Cysteine proteinases and the pathogenesis of amebiasis. *Clin. Microbiol. Rev.* 13, 196–206. doi: 10.1128/CMR.13.2.196
- Rastew, E., Morf, L., and Singh, U. (2015). *Entamoeba histolytica* rhomboid protease 1 has a role in migration and motility as validated by two independent genetic approaches. *Exp. Parasitol.* 154, 33–42. doi: 10.1016/j.exppara.2015.04.004
- Reuber, T. L., and Ausubel, F. M. (1996). Isolation of Arabidopsis genes that differentiate between resistance responses mediated by the RPS2 and RPM1 disease resistance genes. *Plant Cell*. 8, 241–249. doi: 10.1105/tpc.8.2.241
- Romero, M., Cerritos, R., and Ximenez, C. (2016). Horizontal Gene Transfers from Bacteria to *Entamoeba* Complex: A Strategy for Dating Events along Species Divergence. *J. Parasitol. Res.* 2016, 3241027. doi: 10.1155/2016/3241027
- Santi-Rocca, J., Weber, C., Guigon, G., Sismeiro, O., Coppee, J. Y., and 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. doi: 10.1111/j.1462-5822.2007.01030.x
- Seigneur, M., Mounier, J., Prevost, M. C., and Guillen, N. (2005). A lysine- and glutamic acid-rich protein, KERP1, from *Entamoeba histolytica* binds to human enterocytes. *Cell Microbiol.* 7, 569–579. doi: 10.1111/j.1462-5822.2005.00487.x
- Sengupta, D., Naik, D., and Reddy, A. R. (2015). Plant aldo-keto reductases (AKRs) as multi-tasking soldiers involved in diverse plant metabolic processes and stress defense: A structure-function update. *J. Plant Physiol.* 179, 40–55. doi: 10.1016/j.jplph.2015.03.004
- Sharma, A., Sojar, H. T., Glurich, I., Honma, K., Kuramitsu, H. K., and Genco, R. J. (1998). Cloning, expression, and sequencing of a cell surface antigen containing a leucine-rich repeat motif from *Bacteroides forsythus* ATCC 43037. *Infect. Immun.* 66, 5703–5710. doi: 10.1128/IAI.66.12.5703-5710.1998
- Shirley, D. T., Watanabe, K., and Moonah, S. (2019). Significance of amebiasis: 10 reasons why neglecting amebiasis might come back to bite us in the gut. *PLoS Negl. Trop. Dis.* 13, e0007744. doi: 10.1371/journal.pntd.0007744
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. doi: 10.1038/msb.2011.75
- Silvestre, A., Plaze, A., Berthon, P., Thibaux, R., Guillen, N., and Labruyere, E. (2015). In *Entamoeba histolytica*, a BspA family protein is required for chemotaxis toward tumour necrosis factor. *Microb. Cell.* 2, 235–246. doi: 10.15698/mic2015.07.214
- Stanley, S. L. Jr., Zhang, T., Rubin, D., and Li, E. (1995). Role of the *Entamoeba histolytica* cysteine proteinase in amoebic liver abscess formation in severe combined immunodeficient mice. *Infect. Immun.* 63, 1587–1590. doi: 10.1128/IAI.63.4.1587-1590.1995
- Tachibana, H., Yanagi, T., Pandey, K., Cheng, X. J., Kobayashi, S., Sherchand, J. B., et al. (2007). An *Entamoeba* sp. strain isolated from rhesus monkey is virulent but genetically different from *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* 153, 107–114. doi: 10.1016/j.molbiopara.2007.02.006
- Tachibana, H., Yanagi, T., Akatsuka, A., Kobayashi, S., Kanbara, H., and Tsutsumi, V. (2009). Isolation and characterization of a potentially virulent species *Entamoeba nuttalli* from captive Japanese macaques. *Parasitology*. 136, 1169–1177. doi: 10.1017/S0031182009990576
- Tanaka, M., Makiuchi, T., Komiyama, T., Shiina, T., Osaki, K., and Tachibana, H. (2019). Whole genome sequencing of *Entamoeba nuttalli* reveals mammalian host-related molecular signatures and a novel octapeptide-repeat surface protein. *PLoS Negl. Trop. Dis.* 13, e0007923. doi: 10.1371/journal.pntd.0007923
- Tannich, E., Ebert, F., and Horstmann, R. D. (1991). Primary structure of the 170-kDa surface lectin of pathogenic *Entamoeba histolytica*. *Proc. Natl. Acad. Sci. U. S. A.* 88, 1849–1853. doi: 10.1073/pnas.88.5.1849
- Teixeira, J. E., Sateriale, A., Bessoff, K. E., and Huston, C. D. (2012). Control of *Entamoeba histolytica* adherence involves metalloproteinase 1, an M8 family surface metalloproteinase with homology to leishmanolysin. *Infect. Immun.* 80, 2165–2176. doi: 10.1128/IAI.06389-11
- Thibaux, R., Dufour, A., Roux, P., Bernier, M., Baglin, A. C., Frileux, P., et al. (2012). Newly visualized fibrillar collagen scaffolds dictate *Entamoeba histolytica* invasion route in the human colon. *Cell Microbiol.* 14, 609–621. doi: 10.1111/j.1462-5822.2012.01752.x
- Thibaux, R., Weber, C., Hon, C. C., Dillies, M. A., Ave, P., Coppee, J. Y., et al. (2013). Identification of the virulence landscape essential for *Entamoeba histolytica* invasion of the human colon. *PLoS Pathog.* 9, e1003824. doi: 10.1371/journal.ppat.1003824
- Thibaux, R., Ave, P., Bernier, M., Morcelet, M., Frileux, P., Guillen, N., et al. (2014). The parasite *Entamoeba histolytica* exploits the activities of human matrix metalloproteinases to invade colonic tissue. *Nat. Commun.* 5, 5142. doi: 10.1038/ncomms6142
- Tillack, M., Nowak, N., Lotter, H., Bracha, R., Mirelman, D., Tannich, E., et al. (2006). Increased expression of the major cysteine proteinases by stable episomal transfection underlines the important role of EhCP5 for the pathogenicity of *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* 149, 58–64. doi: 10.1016/j.molbiopara.2006.04.009
- Tillack, M., Biller, L., Irmer, H., Freitas, M., Gomes, M. A., Tannich, E., et al. (2007). The *Entamoeba histolytica* genome: primary structure and expression of proteolytic enzymes. *BMC Genomics* 8, 170. doi: 10.1186/1471-2164-8-170
- Weedall, G. D., and Hall, N. (2011). Evolutionary genomics of *Entamoeba*. *Res. Microbiol.* 162, 637–645. doi: 10.1016/j.resmic.2011.01.007
- Weedall, G. D., Sherrington, J., Paterson, S., and Hall, N. (2011). Evidence of gene conversion in genes encoding the Gal/GalNAc lectin complex of *Entamoeba*. *PLoS Negl. Trop. Dis.* 5, e1209. doi: 10.1371/journal.pntd.0001209

- Weedall, G. D., Clark, C. G., Koldkjaer, P., Kay, S., Bruchhaus, I., Tannich, E., et al. (2012). Genomic diversity of the human intestinal parasite *Entamoeba histolytica*. *Genome Biol.* 13, R38. doi: 10.1186/gb-2012-13-5-r38
- Weedall, G. D. (2020). The *Entamoeba* lysine and glutamic acid rich protein (KERP1) virulence factor gene is present in the genomes of *Entamoeba nuttalli*, *Entamoeba dispar* and *Entamoeba moshkovskii*. *Mol. Biochem. Parasitol.* 238, 111293. doi: 10.1016/j.molbiopara.2020.111293
- Wei, M., Feng, M., Guan, Y., Guo, C., Zhou, H., Fu, Y., et al. (2018). Correlation of genetic diversity between hosts and parasites in *Entamoeba nuttalli* isolates from Tibetan and rhesus macaques in China. *Biosci. Trends.* 12, 375–381. doi: 10.5582/bst.2018.01157
- Welter, B. H., Walters, H. A., and Temesvari, L. A. (2020). Reduced expression of a rhomboid protease, EhROM1, correlates with changes in the submembrane distribution and size of the Gal/GalNAc lectin subunits in the human protozoan parasite, *Entamoeba histolytica*. *PLoS One* 15, e0219870. doi: 10.1371/journal.pone.0219870
- Willhoeft, U., Buss, H., and Tannich, E. (1999a). DNA sequences corresponding to the ariel gene family of *Entamoeba histolytica* are not present in *E. dispar*. *Parasitol. Res.* 85, 787–789. doi: 10.1007/s004360050633
- Willhoeft, U., Hamann, L., and Tannich, E. (1999b). A DNA sequence corresponding to the gene encoding cysteine proteinase 5 in *Entamoeba histolytica* is present and positionally conserved but highly degenerated in *Entamoeba dispar* [In Process Citation]. *Infect. Immun.* 67, 5925–5929. doi: 10.1128/IAI.67.11.5925-5929.1999
- Wilson, I. W., Weedall, G. D., Lorenzi, H., Howcroft, T., Hon, C. C., Deloger, M., et al. (2019). Genetic Diversity and Gene Family Expansions in Members of the Genus *Entamoeba*. *Genome Biol. Evol.* 11, 688–705. doi: 10.1093/gbe/evz009
- Winkelmann, J., Leippe, M., and Bruhn, H. (2006). A novel saposin-like protein of *Entamoeba histolytica* with membrane-fusogenic activity. *Mol. Biochem. Parasitol.* 147, 85–94. doi: 10.1016/j.molbiopara.2006.01.010

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 König, Honecker, Wilson, Weedall, Hall, Roeder, Metwally and Bruchhaus. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.