Published in final edited form as:

Curr Trop Med Rep. 2016 December ; 3(4): 158–163. doi:10.1007/s40475-016-0088-9.

The *E. histolytica* Genome Structure and Virulence

Carol A. Gilchrist*

Department of Medicine, School of Medicine, University of Virginia, Charlottesville, Virginia,
United States of America

Abstract

The outcome of an *Entamoeba histolytica* infection is variable and the contribution of genetic diversity within *E. histolytica* to human disease is not fully understood. The information provided by the whole genome sequence of the *E. histolytica* reference laboratory strain (HM-1:IMSS) and thirteen additional laboratory strains have been made publically available. In this review theories on the source of the unexpected level of structural diversity found in *E. histolytica* will be discussed.

Keywords

single nucleotide polymorphisms; genome; cyclin-2; structural diversity

INTRODUCTION

Globally an estimated one million children die before their fifth birthday from diarrhea [1,2]. There is a need to not only develop better strategies to treat diarrheal disease in low and middle income countries but also to prevent its occurrence.

Entamoeba histolytica the causative agent of amebic colitis has been identified as one of the causes of severe diarrhea in young infants [3]. It has long been known that an *E. histolytica* infection does not necessarily result in symptomatic disease. The consequences of an infection are variable; it can cause diarrhea, amebic liver abscesses, latent disease or asymptomatic carriage [4]. Earlier investigations have shown that host genetics, immunity and bacterial milieu can influence the outcome of an infection [5–9]. *E. histolytica* strains however also exist with genetically predetermined differences in their potential to cause human disease [10–14]. For an extensive review of the potential virulence factors discovered in the sequenced genome of *Entamoeba* see Clark et al and Amebiasis: Biology and Pathogenesis of *Entamoeba* by Tomoyoshi Nozaki (Editor), Alok Bhattacharya (Editor)

Corresponding author: Carol A Gilchrist, Dept. of Medicine/Division of Infectious Diseases, Carter Harrison Building, MR-6, 345 Crispell Drive, Rm 1707A, PO Box 801365, Charlottesville VA 22908, USA, Tel: 434 924 5167/434 982 5473, cg2p@virginia.edu.

Compliance with Ethics Guidelines

Conflict of Interest

Carol A. Gilchrist declares grant funding from the Bill and Melinda Gates Foundation and from the NIAID.

Human and Animal Rights and Informed Consent

This article does not contain any studies with human or animal subjects performed by any of the authors.

[15,16]. Identification of the causal mutations and the mechanisms that increase parasite virulence in *E. histolytica* strains are however still to be discovered.

***E. histolytica* structural variation**

Extrapolating from the available whole genome sequence data of the *Entamoeba histolytica* reference laboratory strain (HM-1:IMSS) and the additional sequenced laboratory strains, the level of nucleotide variation in the *E. histolytica* is relatively modest [17–19]. Analysis of both the genomic and genotyping data suggest however that genetic reassortment events are common in this species [10,12,20–24]. Understanding the factors that can trigger genetic reassortment may provide important insights into parasite population dynamics and parasite phenotypes. It is feasible, for instance, that the association of different alleles made possible due to the high levels of recombination may allow the transient expression of polygenic traits which result in the increased virulence of *E. histolytica* strains.

The structural diversity in *E. histolytica* occurs despite the asexual reproduction of the vegetative trophozoite form. A fusion event between different trophozoite strains that co-infected the same host could be responsible for some of the observed diversity [12]. Thereafter homologous recombination may potentially be promoted by the highly variable ploidy levels of the *E. histolytica* trophozoite [25–27].

Possible Role of Repetitive Elements in promoting recombination in *Entamoeba*

In addition to LINE and SINE repeats Lorenzi et al identified novel repeat sequences in *Entamoeba* that were designated as Entamoeba Repeat Elements (ERE) 1 and 2 with the latter (ERE2) being specific to *E. histolytica*. Repeat sequences were estimated to account for 11.2% of the *E. histolytica* genome [28]. In examining their sequenced genomes Weedall et al observed substantial copy number variation in *E. histolytica* strains and repeat elements at the border of the duplicated DNA [18]. It is therefore feasible to speculate that the abundant repeat elements may play a role in promoting the expansion of the different paralogous gene families that occur in *E. histolytica* isolates [18,29]. This expansion in gene families provide an additional means of increasing gene diversity by permitting recombination between the paralogous gene copies as is postulated to have been increased the diversity in the gene families encoding the intermediate and light subunits of the Galactose and N-acetyl-D-galactosamine inhibitable lectin [30].

***E. histolytica* Single Nucleotide Polymorphisms (SNPs)**

The available whole genome sequencing data tends to support the findings of Ghosh *et al* that many of the same polymorphic alleles exist in both Asia and the America's [31]. Careful examination of the SNPs within the coding DNA of the sequenced strains suggest that over half of all SNPs could be found in at least one additional sequenced genome and geographical distance diminished but did not eliminate this effect. In this review SNPs found in more than one *E. histolytica* sequenced genome will be referred to as polymorphic SNPs as opposed to 'private' SNPs that have only been found in one strain.

Restricting the initial analysis of a strain to polymorphic SNPs provides a considerable advantage when examining the data from high throughput next generation sequencing as this

can introduce sequencing errors [32]. Illumina sequencing chemistry is usually considered to have an error rate of ~1% but the rapidly evolving Oxford Nanopore Technologies MinION system has a considerably higher rate of error (10–40%) [33,34]. Techniques exist to diminish the impact of these sequencing errors and a commonly used quality control method is to restrict confidence in a SNP until this region of DNA has been repeatedly sequenced. In Illumina sequencing 30–40X sequence depth or coverage is the most commonly used metric to ensure good quality SNP calls [34,35]. Most sequencing software is conservative and will substitute the reference allele in cases of uncertainty or low confidence in a predicted SNP [36]. In metazoan organisms this conservative approach has proven valuable however in *E. histolytica* where the default nucleotide in the HM-1:IMSS reference strain could be a SNP only found in this strain this system applied to low coverage data can prove highly misleading and should be avoided. The intergenic regions of *E. histolytica* often have regions of low sequence complexity. This lack of complexity can result in the misalignment of reads and sequencing artefacts which can increase sequence error rates [37,38]. While improvements in sequence read length may lessen this risk in future studies currently a high sequencing standard is required for analysis of *E. histolytica* SNPs in non-coding DNA [33,34].

The SNPs identified within the current sequenced *E. histolytica* genomes can be split into three groups; private (as described above; found only in one sequenced genome); rare (the variant was observed infrequently) and polymorphic (found at high frequency). Multilocus sequence typing (MLST), a PCR-based fingerprinting method, involves amplifying and sequencing several polymorphic DNA regions [39]. Gilchrist *et al* restricted their selection of loci for use in *E. histolytica* MLST typing to those where the minor allele (occurring in less than 50% of the sequenced genomes) was relatively common (occurring in >35%) [10]. To decrease errors arising from read misalignment and value to their analysis of their studies of the *E. histolytica* genome Gilchrist *et al* used non-synonymous SNPs within the coding DNA.

Analysis of the MLST results found that even in a restricted geographical area *E. histolytica* strains varied and vertical transmission of common SNP combinations was not observed [10]. This result was consistent with previous studies using the tRNA linked short tandem repeat arrays of *E. histolytica* which indicated that a high rate of recombination could exist within *E. histolytica* genomes [11,12,14].

This rate of recombination would suggest that genetic markers that can be associated with virulence may be genetically linked to a causal mutation. The MLST study identified one allele (SNPs 2725^{C/T} & 2730^{A/G}) that was significantly associated with the propensity of a *E. histolytica* to cause amebic liver abscesses in adults [10]. Interestingly it was not associated with diarrheal disease in children occurring there mainly in asymptomatic cases. The SNPs linked with invasive liver amebiasis (SNPs 2725^{C/T} & 2730^{A/G}) could be located within the sequenced genome within the cyclin-2 gene (AmoebaDB ID EHI_080100; XM_001914351) [10,19,40]. Currently however the sequence of the *E. histolytica* genome is not completely closed and EHI_080100 (cyclin-2) is present on a short region (5k) of contiguous DNA (DS571720) that could not be assembled into a larger DNA segment or sequence scaffold [19,40,41]. Additional studies are required to assess the strength of the

association between the cyclin-2 allele and SNPs in its neighboring DNA with the virulence phenotype. They are also needed to place the assembled DS571720 DNA within a larger genomic context and identify if any neighboring SNPs are also associated with parasite virulence. No sequences equivalent to the DNA within the *E. histolytica* DS571720 contiguous sequences have been identified in the non-pathogenic *E. dispar* or other sequenced *Entamoeba* species.

Comparison between the genomes of *Entamoeba* Species

The *Entamoeba* species that are known to be clinically important and are commonly infecting the human population are *E. histolytica* and *E. moshkovskii*. Due to the fact that they are a confounding factor for the common microscopic method of diagnostics in Bangladesh *E. dispar* and *E. bangladeshi* have also been the subject of additional study [42–46]. At the time of writing the genome of *E. bangladeshi* has not been sequenced however preliminary evidence suggests that all four species have a high degree of homology [7,18,29,47]. Despite this similarity it has long been known that significant differences exist between the species and it has been hypothesized that these may result in the disparity observed in their potential virulence [18,20,48,49]. In addition the genomes of *E. histolytica* the genomes of *E. dispar* and *E. invadens* have been published and the genomes of both *E. moshkovskii* and *E. nuttelli* made available for study [47,50–52].

Interestingly there is no evidence of interspecies gene transfer between members of the *Entamoeba* genus. This is despite the data discussed above suggesting the existence of intraspecies genetic recombination in *E. histolytica* and the apparent lack of species specific immunity. Not only can different genotypes of *Entamoeba histolytica* infect the same host but co-infection of different *Entamoeba* species are relatively common in humans and other primates [12,53–56]. *E. histolytica* genes have been identified with a strong homology to those of prokaryotes which suggest that the mechanisms to integrate laterally transmitted genes exist in *E. histolytica* and even permit the usurping the role of even essential vertically transmitted paralogues [57–59].

Either mechanisms exist that prevents fusion occurring between different members of the *Entamoeba* genus or an *Entamoeba* proofreading function exists that efficiently targets for destruction somewhat similar sequences which reduces the lateral transfer of genetic material from co-infecting *Entamoeba* strains.

CONCLUSIONS

Molecular sequencing methods are evolving at a rapid rate and provide a powerful high-resolution method to study parasite evolution [60]. *E. histolytica* isolates in particular provide a unique opportunity to study the effect of SNPs on parasite virulence in a number of genetic contexts. Much remains to be learned however about the mechanism by which the high natural rate of recombination is induced in this parasite.

Acknowledgments

This work was supported by the National Institute of Allergy and Infectious Diseases (NIAID) grants AI103536-01 and AI043596. The funders had no role in study design, data collection and analysis, or decision to submit for publication

References

Papers of particular interest, published recently, have been highlighted as:

• Of importance

•• Of major importance

1. Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* [Internet]. 2010; 375:1969–87. [cited 2011 Jul 18]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20466419>.
2. Fischer Walker CL, Aryee MJ, Boschi-Pinto C, Black RE. Estimating Diarrhea Mortality among Young Children in Low and Middle Income Countries. *PLoS One* [Internet]. 2012; 7:e29151. [cited 2012 Jan 27]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3250411&tool=pmcentrez&rendertype=abstract>.
- *3. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* [Internet]. 2013; 382:209–22. [cited 2013 Aug 9]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23680352> The GEMS study is one of the largest studies in low and middle income countries of childhood enteric infections.
4. Haque R, Huston CD, Hughes M, Houpt E, Petri WAJ. Amebiasis. *N Engl J Med* [Internet]. 2003; 348:1565–73. [cited 2012 Jan 27]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12700377>.
5. Duggal P, Haque R, Roy S, Mondal D, Sack RB, Farr BM, et al. Influence of human leukocyte antigen class II alleles on susceptibility to *Entamoeba histolytica*. *J Infect Dis* [Internet]. 2004; 189:520–6. [cited 2012 Jan 29]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14745711>.
6. Duggal P, Guo X, Haque R, Peterson KM, Ricklefs S, Mondal D, et al. A mutation in the leptin receptor is associated with *Entamoeba histolytica* infection in children. *J Clin Invest* [Internet]. 2011; 121:1191–8. [cited 2012 Mar 27]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3049405&tool=pmcentrez&rendertype=abstract>.
- *7. Gilchrist CA, Petri SE, Schneider BN, Reichman DJ, Jiang N, Begum S, et al. Role of the gut microbiota of children in diarrhea due to the protozoan parasite *Entamoeba histolytica*. *J Infect Dis* [Internet]. 2016; 213:1579–85. [cited 2016 Jan 12]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26712950>.
- *8. Burgess SL, Buonomo E, Carey M, Cowardin C, Naylor C, Noor Z, et al. Bone marrow dendritic cells from mice with an altered microbiota provide interleukin 17A-dependent protection against *Entamoeba histolytica* colitis. *MBio* [Internet]. 2014; 5:e01817. [cited 2015 Jan 21]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4222101&tool=pmcentrez&rendertype=abstract>.
9. Korpe PS, Liu Y, Siddique A, Kabir M, Ralston K, Ma JZ, et al. Breast milk parasite-specific antibodies and protection from amebiasis and cryptosporidiosis in Bangladeshi infants: a prospective cohort study. *Clin Infect Dis* [Internet]. 2013; 56:988–92. [cited 2013 May 2]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3588117&tool=pmcentrez&rendertype=abstract>.
- **10. Gilchrist CA, Ali IKM, Kabir M, Alam F, Scherbakova S, Ferlanti E, et al. A Multilocus Sequence Typing System (MLST) Reveals a High Level of Diversity and a Genetic Component to *Entamoeba histolytica* Virulence. *BMC Microbiol* [Internet]. 2012; 12:151. [cited 2012 Aug 30]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22839995> The work described in this paper used polymorphic SNPs to describe the diversity of *E. histolytica* genotypes and their link to virulence.

- **11. Ali IKM, Maqsood R, Alam F, Kabir M, Siddique A, Petri WAJ. Evidence for a link between locus R-R sequence type and outcome of infection with *Entamoeba histolytica*. Clin Microbiol Infect [Internet]. 2012; 18:E235–7. [cited 2012 May 1]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22448930>The work described in this paper used tRNA gene-linked short tandem repeats to describe the diversity of *E. histolytica* genotypes and their link to virulence.
12. Ali IKM, Mondal U, Roy S, Haque R, Petri WAJ, Clark CG. Evidence for a link between parasite genotype and outcome of infection with *Entamoeba histolytica*. J Clin Microbiol [Internet]. 2007; 45:285–9. [cited 2012 Jan 29]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1829016&tool=pmcentrez&rendertype=abstract>.
13. Jaiswal V, Ghoshal U, Mittal B, Dhole TN, Ghoshal UC. Association between allelic variation due to short tandem repeats in tRNA gene of *Entamoeba histolytica* and clinical phenotypes of amoebiasis. Acta Trop [Internet]. 2014; 133:1–7. [cited 2016 Jun 6]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24495629>.
- **14. Das K, Mukherjee AK, Chowdhury P, Sehgal R, Bhattacharya MK, Hashimoto T, et al. Multilocus sequence typing system (MLST) reveals a significant association of *Entamoeba histolytica* genetic patterns with disease outcome. Parasitol Int [Internet]. 2014; 63:308–14. [cited 2015 Apr 8]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24321887>.
- **15. Nozaki T., Bhattacharya A., editors. Amebiasis: Biology and Pathogenesis of *Entamoeba*. Springer; 2014.
16. Clark CG, Alsmark UCM, Tazreiter M, Saito-Nakano Y, Ali V, Marion S, et al. Structure and content of the *Entamoeba histolytica* genome. Adv Parasitol [Internet]. 2007; 65:51–190. [cited 2011 Aug 6]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18063096>.
17. Loftus B, Anderson I, Davies R, Alsmark UCM, Samuelson J, Amedeo P, et al. The genome of the protist parasite *Entamoeba histolytica*. Nature [Internet]. 2005; 433:865–8. [cited 2011 Oct 10]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15729342>.
- *18. Weedall GD, Hall N. Evolutionary genomics of *Entamoeba*. Res Microbiol [Internet]. 2011; 162:637–45. [cited 2012 Nov 10]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3268252&tool=pmcentrez&rendertype=abstract>.
19. Lorenzi HA, Puiu D, Miller JR, Brinkac LM, Amedeo P, Hall N, et al. New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. PLoS Negl Trop Dis [Internet]. 2010; 4:e716. [cited 2012 Feb 6]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2886108&tool=pmcentrez&rendertype=abstract>.
- *20. Weedall GD, Clark CG, Koldkjaer P, Kay S, Bruchhaus I, Tannich E, et al. Genomic diversity of the human intestinal parasite *Entamoeba histolytica*. Genome Biol [Internet]. 2012; 13:R38. [cited 2012 Oct 5]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3446291&tool=pmcentrez&rendertype=abstract>This paper describes the resequencing 10 of the *E. histolytica* strains in laboratory culture.
- *21. Zermeno V, Ximénez C, Morán P, Valadez A, Valenzuela O, Rascón E, et al. Worldwide genealogy of *Entamoeba histolytica*: an overview to understand haplotype distribution and infection outcome. Infect Genet Evol [Internet]. 2013; 17:243–52. [cited 2016 Jun 6]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23624203>This paper describes the world wide diversity of the *E. histolytica* tRNA gene-linked short tandem repeats.
22. Haghighi A, Kobayashi S, Takeuchi T, Thammapalerd N, Nozaki T. Geographic diversity among genotypes of *Entamoeba histolytica* field isolates. J Clin Microbiol [Internet]. 2003; 41:3748–56. [cited 2012 Jan 30]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=179867&tool=pmcentrez&rendertype=abstract>.
23. Haghighi A, Kobayashi S, Takeuchi T, Masuda G, Nozaki T. Remarkable genetic polymorphism among *Entamoeba histolytica* isolates from a limited geographic area. J Clin Microbiol [Internet]. 2002; 40:4081–90. [cited 2012 Jan 30]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=139687&tool=pmcentrez&rendertype=abstract>.
24. Escueta-de Cadiz A, Kobayashi S, Takeuchi T, Tachibana H, Nozaki T. Identification of an avirulent *Entamoeba histolytica* strain with unique tRNA-linked short tandem repeat markers. Parasitol Int [Internet]. 2010; 59:75–81. [cited 2012 Jan 30]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19895903>.

25. Parfrey LW, Lahr DJG, Katz LA. The dynamic nature of eukaryotic genomes. *Mol Biol Evol* [Internet]. 2008; 25:787–94. [cited 2016 Jun 6]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2933061&tool=pmcentrez&rendertype=abstract>.
26. Lohia A, Mukherjee C, Majumder S, Dastidar PG. Genome re-duplication and irregular segregation occur during the cell cycle of *Entamoeba histolytica*. *Biosci Rep* [Internet]. 2007; 27:373–84. [cited 2012 May 1]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17592766>.
27. Mukherjee C, Clark CG, Lohia A. *Entamoeba* shows reversible variation in ploidy under different growth conditions and between life cycle phases. *PLoS Negl Trop Dis* [Internet]. 2008; 2:e281. [cited 2012 May 1]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2500184&tool=pmcentrez&rendertype=abstract>.
28. Lorenzi H, Thiagarajan M, Haas B, Wortman J, Hall N, Caler E. Genome wide survey, discovery and evolution of repetitive elements in three *Entamoeba* species. *BMC Genomics* [Internet]. 2008; 9:595. [cited 2012 May 1]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19077187>.
29. Das K, Ganguly S. Evolutionary genomics and population structure of *Entamoeba histolytica*. *Comput Struct Biotechnol J* [Internet]. 2014; 12:26–33. [cited 2016 May 20]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4262060&tool=pmcentrez&rendertype=abstract>.
30. Weedall GD, Sherrington J, Paterson S, Hall N. Evidence of gene conversion in genes encoding the Gal/GalNac lectin complex of *Entamoeba*. *PLoS Negl Trop Dis* [Internet]. 2011; 5:e1209. [cited 2015 Nov 24]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3125142&tool=pmcentrez&rendertype=abstract>.
31. Ghosh, S., Frisardi, M., Ramirez-Avila, L., Descoteaux, S., Sturm-Ramirez, K., Newton-Sanchez, OA., et al. *J Clin Microbiol* [Internet]. Vol. 38. American Society for Microbiology; 2000. Molecular epidemiology of *Entamoeba* spp.: evidence of a bottleneck (Demographic sweep) and transcontinental spread of diploid parasites; p. 3815–21. [cited 2012 Feb 1] Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=87481&tool=pmcentrez&rendertype=abstract>
32. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* [Internet]. 2012; 30:434–9. [cited 2013 Feb 10]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22522955>.
33. Laver T, Harrison J, O'Neill PA, Moore K, Farbos A, Paszkiewicz K, et al. Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomol Detect Quantif* [Internet]. 2015; 3:1–8. [cited 2016 Jun 15]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26753127>.
34. Loman NJ, Quick J, Simpson JT. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods* [Internet]. 2015; 12:733–5. [cited 2016 Jun 15]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26076426>.
35. Lou, DI., Hussmann, JA., McBee, RM., Acevedo, A., Andino, R., Press, WH., et al. High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc Natl Acad Sci* [Internet]. 2013. [cited 2013 Nov 16]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24243955>
36. Altmann A, Weber P, Bader D, Preuss M, Binder EB, Müller-Myhsok B. A beginners guide to SNP calling from high-throughput DNA-sequencing data. *Hum Genet* [Internet]. 2012; 131:1541–54. [cited 2016 Jun 15]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22886560>.
37. Bhattacharya D, Haque R, Singh U. Coding and noncoding genomic regions of *Entamoeba histolytica* have significantly different rates of sequence polymorphisms: implications for epidemiological studies. *J Clin Microbiol* [Internet]. 2005; 43:4815–9. [cited 2012 Jan 30]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1234137&tool=pmcentrez&rendertype=abstract>.
38. Altschul SF, Boguski MS, Gish W, Wootton JC. Issues in searching molecular sequence databases. *Nat Genet* [Internet]. 1994; 6:119–29. [cited 2016 Jun 15]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8162065>.
39. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A*. 1998; 95:3140–5. [PubMed: 9501229]

40. Aurrecochea C, Barreto A, Brestelli J, Brunk BP, Caler EV, Fischer S, et al. AmoebaDB and MicrosporidiaDB: functional genomic resources for Amoebozoa and Microsporidia species. *Nucleic Acids Res* [Internet]. 2011; 39:D612–9. [cited 2012 Jan 30]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3013638&tool=pmcentrez&rendertype=abstract>.
41. Branscomb E, Predki P. On the high value of low standards. *J Bacteriol* [Internet]. 2002; 184:6406–9. [cited 2016 Jun 8]. discussion 6409. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12426325>.
42. Ali IKM, Hossain MB, Roy S, Ayeh-Kumi PF, Petri, William AJ, Haque R, et al. Entamoeba moshkovskii infections in children, Bangladesh. *Emerg Infect Dis*. 2003; 9:580–4. [PubMed: 12737742]
43. Shimokawa C, Kabir M, Taniuchi M, Mondal D, Kobayashi S, Ali IKM, et al. Entamoeba moshkovskii is associated with diarrhea in infants and causes diarrhea and colitis in mice. *J Infect Dis* [Internet]. 2012; 206:744–51. [cited 2014 Jul 14]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3491739&tool=pmcentrez&rendertype=abstract>.
44. Royer TL, Gilchrist C, Kabir M, Arju T, Ralston KS, Haque R, et al. *Entamoeba bangladeshi* nov. sp., Bangladesh. *Emerg Infect Dis* [Internet]. 2012; 18:1543–5. [cited 2012 Nov 7]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3437720&tool=pmcentrez&rendertype=abstract>.
45. Diamond LS, Clark CG. A redescription of *Entamoeba histolytica* Schaudinn, 1903 (Emended Walker, 1911) separating it from *Entamoeba dispar* Brumpt, 1925. *J Eukaryot Microbiol* [Internet]. 2012; 40:340–4. [cited 2012 May 1]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8508172>.
46. Haque R, Faruque AS, Hahn P, Lyerly DM, Petri WJ. Entamoeba histolytica and Entamoeba dispar infection in children in Bangladesh. *J Infect Dis*. 175:734–6.
47. Wilson IW, Weedall GD, Hall N. Host-Parasite interactions in Entamoeba histolytica and Entamoeba dispar: what have we learned from their genomes? *Parasite Immunol* [Internet]. 2012; 34:90–9. [cited 2012 May 1]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21810102>.
48. Willhoeft U, Buss H, Tannich E. DNA sequences corresponding to the ariel gene family of Entamoeba histolytica are not present in E. dispar. *Parasitol Res* [Internet]. 1999; 85:787–9. [cited 2016 Jun 13]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10431751>.
49. Willhoeft U, Hamann L, Tannich E. 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. *Infect Immun* [Internet]. 1999; 67:5925–9. [cited 2016 Jun 13]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10531249>.
50. Tachibana H, Yanagi T, Pandey K, Cheng X-J, Kobayashi S, Sherchand JB, et al. An Entamoeba sp. strain isolated from rhesus monkey is virulent but genetically different from Entamoeba histolytica. *Mol Biochem Parasitol* [Internet]. 2007; 153:107–14. [cited 2016 Sep 15]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17403547>.
51. Ehrenkaufer GM, Weedall GD, Williams D, Lorenzi HA, Caler E, Hall N, et al. The genome and transcriptome of the enteric parasite Entamoeba invadens, a model for encystation. *Genome Biol* [Internet]. 2013; 14:R77. [cited 2016 Sep 15]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23889909>.
52. Tawari B, Ali IKM, Scott C, Quail MA, Berriman M, Hall N, et al. Patterns of evolution in the unique tRNA gene arrays of the genus Entamoeba. *Mol Biol Evol* [Internet]. 2008; 25:187–98. [cited 2016 Sep 15]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17974548>.
53. Haque R, Ali IK, Akther S, Petri WA. Comparison of PCR, isoenzyme analysis, and antigen detection for diagnosis of Entamoeba histolytica infection. *J Clin Microbiol* [Internet]. 1998; 36:449–52. [cited 2015 May 8]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=104557&tool=pmcentrez&rendertype=abstract>.
54. Regan CS, Yon L, Hossain M, Elsheikha HM. Prevalence of Entamoeba species in captive primates in zoological gardens in the UK. *PeerJ* [Internet]. 2014; 2:e492. [cited 2016 Jun 14]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25097822>.
55. Ngui R, Angal L, Fakhrurrazi SA, Lian YLA, Ling LY, Ibrahim J, et al. Differentiating Entamoeba histolytica, Entamoeba dispar and Entamoeba moshkovskii using nested polymerase chain reaction

- (PCR) in rural communities in Malaysia. *Parasit Vectors* [Internet]. 2012; 5:187. [cited 2016 Jun 14]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22947430>.
56. Lau YL, Anthony C, Fakhrurrazi SA, Ibrahim J, Ithoi I, Mahmud R. Real-time PCR assay in differentiating *Entamoeba histolytica*, *Entamoeba dispar*, and *Entamoeba moshkovskii* infections in Orang Asli settlements in Malaysia. *Parasit Vectors* [Internet]. 2013; 6:250. [cited 2014 Oct 13]. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3765902&tool=pmcentrez&rendertype=abstract>.
57. Nixon JEJ, Wang A, Field J, Morrison HG, McArthur AG, Sogin ML, et al. Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokaryotes to *Giardia lamblia* and *Entamoeba histolytica*. *Eukaryot Cell* [Internet]. 2002; 1:181–90. [cited 2012 Sep 11]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12455953>.
58. Rosenthal B, Mai Z, Caplivski D, Ghosh S, de la Vega H, Graf T, et al. Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica*. *J Bacteriol* [Internet]. 1997; 179:3736–45. [cited 2016 Jun 13]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9171424>.
59. Alsmark UC, Sicheritz-Ponten T, Foster PG, Hirt RP, Embley TM. Horizontal gene transfer in eukaryotic parasites: a case study of *Entamoeba histolytica* and *Trichomonas vaginalis*. *Methods Mol Biol* [Internet]. 2009; 532:489–500. [cited 2016 Jun 13]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19271203>.
60. Quick J, Loman NJ, Durauffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* [Internet]. 2016; 530:228–32. [cited 2016 Jun 15]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26840485>.