Step I: White Paper Concept Approval Request

Project Title: Genome sequencing of clinical strains of *Entamoeba histolytica*

White Paper Submission Date (MM/DD/YY): 07/02/09

Investigator Contact:

Name:	William A. Petri, Jr. and Ibne Karim M. Ali
Position:	Professor and Research Associate, respectively
Institution:	University of Virginia
Address:	Division of Infectious Diseases and International Health Room 1709A Carter Harrison Research Building MR6 345 Crispell Drive Box 801340, Charlottesville
State:	Virginia
ZIP Code:	22908-1340
Telephone:	434-924-5167
Fax:	434-924-0075
E-Mail:	wap3g@virginia.edu and ika3t@virginia.edu

Project Objective:

- *(Please limit to a 1000 words)*
- *Strong and clear justification for the sequencing or genotyping study for the specific organism(s) proposed.*

The protozoan parasite *Entamoeba histolytica* causes invasive intestinal and extraintestinal infections in about 50 million people world-wide resulting in a death toll of up to 100,000 people annually (Anynomous, 1997; Haque et al, 2003a; Petri et al, 2000; Walsh, 1986) and it still remains a significant cause of human death in developing countries such as Bangladesh and Vietnam (Blessmann et al, 2002; Haque et al, 2003b; Stanley, 2003). However, four out of five *E. histolytica* infections remain asymptomatic (Blessmann et al, 2002; Gathiram and Jackson, 1987; Haque et al, 2003b). What determines the outcome of an *E. histolytica* infection is largely unknown. Infected persons can remain asymptomatically infected for months before developing any sign and symptom of disease. *E. histolytica* strains with variable virulence phenotype have been reported (Bracha and Mirelman, 1984). A recent tRNA gene-linked multi-locus genotyping study by our group using different clinical samples from asymptomatic, diarrhea/dysentery or amebic liver abscess (ALA) patients from Bangladesh showed that there exists a nonrandom distribution of parasite genotypes among these clinical groups, although the total number of genotypes detected in this study was very high - 85 from 111 unrelated samples (Ali et al, 2007). In addition, this study also identified one particular genotype (genotype #66) that was found to be associated with the occurrence of diarrhea/dysentery of the infected individuals. On the contrary, we also showed that almost all the liver abscess strains studied were different from strains from other clinical groups and also from each other (Ayeh-Kumi et al, 2001; Ali et al, 2007). Nevertheless, a sequencing of representative samples from these clinical groups revealed that some specific SNPs were confined in the liver abscess strains, and were never detected in any other sample groups (Ali et al, unpublished data). Since the loci used in the genotyping are non-coding repetitive DNA sequences (Ali et al, 2007), they are most likely not directly responsible for the observed differences in the outcome of infection. So, we think that a genome-wide sequencing of representative strains from 3 clinical groups (asymptomatic, diarrhea/dysentery and amebic liver abscess) may identify genomic regions that are responsible for variable outcome of infection.

E. histolytica is an asexual parasite (Tibayrenc et al, 1990). The existing evidences such as years of cultivation of *E. histolytica* strains in the laboratory or passage of ameba strains through the mouse intestine or liver do not show any changes in their genotypes (based on available markers used in genotyping), which points to *E. histolytica* being a clonal organism. Therefore, explaining the observed link between the parasite genotype and the outcome of infection presents us with some theoretical problems as newer genotypes are being evolved by time even in restricted geographic areas. So, the question arises "*what is the basis for generation of novel genotypes?*". One possibility would be that a novel genotype is evolved during the excystation of the parasite in a new host. However, what is the mechanism involved in this process is unknown at present, but the genome sequencing of strain in their cyst-stage and corresponding trophozoite-stage might give us some clues on whether excystation is accompanied by a change in genomic make-up that accounts for large number of genotypes exist in real world.

It has been shown recently by our group using the paired-samples from 3 countries (Bangladesh, Italy and USA) that parasite genotypes in the intestine and aspirated pus samples of amebic liver abscess (ALA) patients are different in the same patient (Ali et al, 2008). This suggests that either the initial intestinal infection was with more than one strain (or genotype) of *E. histolytica*, but only one of these strains (which has to be a minor population) had the ability to migrate and cause liver abscess in the infected patient or a DNA recombination event is taking place during the migration of ameba from intestine to the liver. A comparison of genomic sequences between intestinal and liver abscess strains from the same ALA patient may provide vital information as to what is actually happening. And if it is indeed a DNA recombination event, then this comparison might provide clues on how it is helping render the parasite capable of migrating to the liver site.

E. histolytica strains can be maintained in laboratory either xenically (i.e., in presence of bacteria) or axenically (in the absence of bacteria). Many investigations such as parasite virulence or gene expression analysis are being carried out using the axenic strains of *E. histolytica*, and as a result, we do not know how accurately this mimics the parasite in their actual host environment (intestine) where they encounter host microbiome. Also, it has been demonstrated for certain *E. histolytica* strains that they display increased virulence in *in vitro* or mouse model experiments if they are cultured in xenic condition. A genome sequencing of representative parasites before and after axenization may provide information on whether axenization may result in a change in the genome sequences predictive of increased virulence.

We, therefore, propose:

1. To investigate whether there exists a sequence difference that correlates with the clinical outcome by sequencing genomes of *E. histolytica* strains isolated from 3 different clinical populations such as asymptomatic (strain never causes disease such as Rahman from the ATCC or other such strains from ICDDR,B) versus diarrhea/dysenteric strains (such as DS4-868, 2592100 from our lab and two other such strains are available from ICDDR,B, and any strain with “genotype-66” from ICDDR,B, if available) versus liver abscess strains (such as HM-3:IMSS from the ATCC or any other strain available from ICDDR,B).

2. To investigate whether axenization changes the genome by sequencing genome of *E. histolytica* strains before and after axenization (4 axenized strains are available at our lab, and their xenic counter-parts are available from ICDDR,B, as well as axenic and xenic strains of SAW 891 are available at the ATCC and Stanford University, respectively).

3. To investigate whether excystation changes the genome by sequencing genome of *E. histolytica* strains before and after excystation (i.e. cyst-genome versus trophozoite-genome, respectively). These samples may come from ICDDR,B.

4. To investigate whether migration of trophozoite from intestinal site to the extra-intestinal site (such as liver) changes the genome by sequencing genome of *E. histolytica* strains isolated from feces (representative of intestinal site) and corresponding strain from liver abscess pus of ALA patients. DNA samples of these strains may come from ICDDR,B.

Reference:

1. (1997) WHO/PAHO/UNESCO report. A consultation with experts on amoebiasis. Mexico City, Mexico 28-29 January, 1997. Epidemiological bulletin 18: 13-14.
2. Haque R, Huston CD, Hughes M, Houpt E, Petri WA, Jr. (2003) Amebiasis. The New England journal of medicine 348: 1565-1573.
3. Petri WA, Jr., Haque R, Lyerly D, Vines RR (2000) Estimating the impact of amebiasis on health. Parasitology today (Personal ed) 16: 320-321.
4. Walsh JA (1986) Problems in recognition and diagnosis of amebiasis: estimation of the global magnitude of morbidity and mortality. Reviews of infectious diseases 8: 228-238.
5. Blessmann J, Van Linh P, Nu PA, Thi HD, Muller-Myhsok B, et al. (2002) Epidemiology of amebiasis in a region of high incidence of amebic liver abscess in central Vietnam. The American Journal of Tropical Medicine and Hygiene 66: 578-583.
6. Gathiram V, Jackson TF (1987) A longitudinal study of asymptomatic carriers of pathogenic zymodemes of *Entamoeba histolytica*. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde 72: 669-672.
7. Haque R, Mondal D, Duggal P, Kabir M, Roy S, et al. (2006) *Entamoeba histolytica* infection in children and protection from subsequent amebiasis. Infection and immunity 74: 904-909.
8. Haque R, Mondal D, Kirkpatrick BD, Akther S, Farr BM, et al. (2003) Epidemiologic and clinical characteristics of acute diarrhea with emphasis on *Entamoeba histolytica* infections in preschool children in an urban slum of Dhaka, Bangladesh. The American Journal of Tropical Medicine and Hygiene 69: 398-405.
9. Stanley SL, Jr. (2003) Amoebiasis. Lancet 361: 1025-1034.
10. Bracha R, Mirelman D (1984) Virulence of *Entamoeba histolytica* trophozoites. Effects of bacteria, microaerobic conditions, and metronidazole. J Exp Med 160: 353-368.
11. Ali IK, Mondal U, Roy S, Haque R, Petri WA, Jr., et al. (2007) Evidence for a link between parasite genotype and outcome of infection with *Entamoeba histolytica*. Journal of clinical microbiology 45: 285-289.
12. Ayeh-Kumi PF, Ali IK, Lockhart LA, Gilchrist CA, Petri WA, Haque R (2001) *Entamoeba histolytica*: genetic diversity of clinical isolates from Bangladesh as demonstrated by polymorphisms in the serine-rich gene. Experimental Parasitology 99:80-88.

13. Tibayrenc M, Kjellberg F, Ayala FJ (1990) A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. Proceedings of National Academy of Science, USA 87:2414-2418.
14. Ali IK, Solaymani-Mohammadi S, Akhter J, Roy S, et al. (2008) Tissue Invasion by *Entamoeba histolytica*: Evidence of Genetic Selection and/or DNA Reorganization Events in Organ Tropism. PLoS Neglected Tropical Diseases 2:e219.

Demonstration of the relevant scientific community's size and depth of interest in the proposed sequencing or genotyping data for this organism or group of organisms.

There are several laboratories in the world working on amebiasis, and quite a few of them are interested in sequencing or genotyping data for *Entamoeba histolytica* strains. The laboratory of Dr. William A. Petri, Jr. at the University of Virginia supports the genotyping of *E. histolytica* strains from clinical specimens. Dr. Petri's lab has a collaboration with a number of other renowned labs working on amebiasis world-wide including Bangladesh, Japan, India, Nepal, Turkey etc. In Bangladesh, there is an excellent field study going on of amebiasis at the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) led by Dr. Rashidul Haque, in collaboration with the University of Virginia, and clinical specimens for genome sequencing will be available from there. In addition, 4 axenic clinical strains for sequencing are available from Dr. Petri's lab.

Utility of the new sequencing or genotyping information

Successful sequencing of the proposed strains will help the scientific community to better understand the genomic basis and differences between asymptomatic and disease-causing *E. histolytica* strains.

Certain SNPs may be identified that will show association with a particular clinical sample group, which will ultimately help discover a new genotyping system with a very high predictive value for clinical outcome.

Genomic sequences of xenic versus corresponding axenic strains will help us understand if axenization causes change in genome sequences. If this is so, it would be interesting to look into sequences of known virulence factors, such as cysteine proteases, amebapores, lectins etc to learn why axenic strains are generally less virulent.

Genomic sequences of cyst-stage versus corresponding excysted trophozoite will help us understand if excystation causes change in genome sequences, and it may explain how a new genotype of *E. histolytica* evolves in actual world.

 Status of other projects on the same organism.

I do not know the details of sequencing project at the Liverpool University in the UK. I recently learnt from Graham Clark that they have sequenced *E. histolytica* Rahman strain and in the process of sequencing another strain “2592100”.

 Arrangements for deposit of resulting reagents, resources, and datasets in NIAID approved repositories.

The resulting genome sequences will be deposited to the NIAID approved repositories.

 List availability of other funding sources for the project.

Not known.

Nature, Availability & Source of Reagents/Samples:

Indicate availability of laboratory strains and clinical isolates. State sample types to be sent to the GSC, their sources and their current readiness. If samples need to be prospectively prepared indicate method of preparation and expected date of delivery.

1. Four axenic strains are available from our laboratory (see Table 1). The xenic cultures of these strains are available at ICDDR,B. They are preserved at liquid nitrogen. They will be ready in 3 months. In addition, two more axenic diarrhea/dysentery strains are available from ICDDR,B. It would take around 4-6 months to isolate high quality genomic DNA from these strains to be sent to the GSC for sequencing.
2. *E. histolytica* HM-3:IMSS strain was isolated directly from the aspirated pus sample of an adult ALA patient from Mexico, and this strain is available at the ATCC. I anticipate that it would take 3-6 months to get high quality genomic DNA from this strain for sequencing.
3. Avirulent (asymptomatic) strain Rahman is available from the ATCC. I anticipate that it would take 3-6 months to get high quality genomic DNA from this strain for sequencing.
4. Axenic *E. histolytica* SAW 891 strain is available from the ATCC. The xenically maintained SAW 891 strain will be available from the Stanford University. I anticipate that it would take 4-6 months to get high quality genomic DNA from this strain for sequencing.
5. The laboratory of Dr. Rashidul Haque will investigate to find out an *E. histolytica* cyst-passer individual at their field site in Mirpur, Bangladesh. I anticipate that it would take 6-12 months before we can get DNA isolated directly from the cyst of this *E. histolytica* strain, as well as DNA purified from the corresponding excysted trophozoite. Alternatively, we could try and get some cyst-specific DNA from the

Stanford University for the strain SAW 891 (which I know that it encysts, although inefficiently, at Gretchen's hand in there).

6. There is a possibility that we would get paired-strains' DNA (i.e. DNA purified from the intestinal strain as well as corresponding LA pus strain from the same ALA patient) from ICDDR,B. However, the isolation of strain from the patient's LA pus material has been unsuccessful so far.

Collaborator Role:

List all potential project collaborators and their role in the project

Dr. Rashidul Haque, International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B), will provide xenic and any liver abscess strains of *E. histolytica* and will also provide DNA samples isolated from the purified *E. histolytica* cysts.

Dr. Upinder Singh, Stanford University, will provide xenic *E. histolytica* strain SAW891 or its DNA, as well as cyst DNA from this strain, if necessary.

NIAID's Genomic Sequencing Center Reagent, Data & Software Release Policy:

Accept

http://www3.niaid.nih.gov/research/resources/mscs/data.htm

Accept Decline

Describe arrangements for deposit of resulting reagents, resources, and datasets in NIAID approved repositories.

Investigator Signature:

Investigator Name: Date: