Phase I Grand Challenges Explorations Financial Report

Organization Name: Universidad Nacional Autónoma de México
Project Title: Software to identify and quantify pathogenic helminth eggs
Report Type: (Interim 01)
Reporting Period *From* **(01/01/2012)**
                        To **(05/30/2012)**

Total Expenses by Type	PHASE I PROJECT: ACTUAL EXPENSES (in US dollars)
Personnel	**$0.00**
Equipment	**$10,382.25**
Travel	**$0.00**
Consultants	**$0.00**
Supplies	**$1,847.54**
Subcontracts	**$8,461.54**
Project Expenses Subtotal	**$20,691.33**
Other sources of project support	**$100 gift (a genera of HE of Brazil) $ 5,000 (lab work)**
Total Project Cost	**$ 25,791.33**

Software to identify and quantify pathogenic helminth eggs

The analytical technique used to identify and quantify helminth eggs, besides its time consuming nature, involves the critical step of identification of the eggs. This is the major source of uncertainty in quantifying results, because microbiologists are not capable of performing it without a long training programme. This project is focused on creating software that will automatically perform helminth egg identification and quantification, reducing the consumption of materials and expenditure on highly trained personnel while obtaining reliable results. During this first four months of the project, a microscopic image processing workstation has been put together.

The first step of this process has now been completed, with the creation of a digital database of several helminth eggs images, including 23 genera and 36 representative species from different water and sludge samples. From the compilation, six species have been preliminarily selected to develop the automatic identification software based on their medical importance and worldwide ubiquity: *Ascaris lumbricoides*, *Hymenolepis nana*, *Hymenolepis diminuta*, *Trichuris trichiura*, *Toxocara canis* and *Taenia saginata.*

The study of the particular characteristics of each species (size, shape, texture, internal structure, etc.) has been performed using very detailed and clean images, but also pictures where the target structures are surrounded by an artifact saturated environment, which led to include multiple difficulty levels in an identification process (Fig. 1).
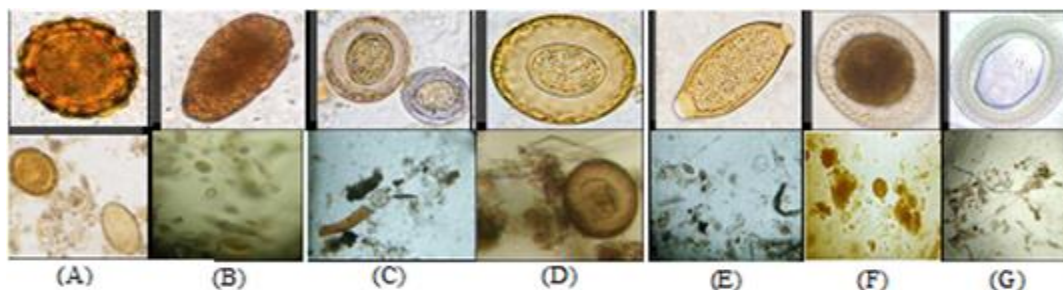

Figure 1. Clean and dirty images of the six species selected to develop the automatic identification software: (A) *Ascaris lumbricoides* fertile, (B) *Ascaris lumbricoides* infertile, (C) *Hymenolepis nana*, (D) *Hymenolepis diminuta*, (E) *Trichuris trichiura*, (F) *Toxocara canis* and (G) *Taenia saginata.*

The development of the software has been the second step of the project. Given the complexity of analyzing biological images, to obtain the most suitable system to identify helminth eggs, a comparative study of the available recognition protocols and image processing techniques was performed. This allowed the development of a specific image processing protocol for the helminth eggs. A brief description of the process is as follows:

The captured images are filtered using a smoothing Gaussian filter; then, the contrast is adjusted and edge detection is applied to identify objects. The holes in the objects in this image are filled, and the object labeling is performed eliminating the objects too small to be considered helminth eggs (Fig 2).

Figure 2. Preliminary image processing protocol for helminth eggs.

The area, perimeter, eccentricity, major axis length, minor axis length and other descriptors are calculated from each labeled object in the final images. These characteristics are expected to be useful to classify each object in the image to distinguish it from others. One class should be assigned to each helminth species and another one should be assigned to objects that appear in the images but are not helminth eggs like bubbles, pollen particles, etc. The system is currently in an experimental stage, and is being evaluated for different environmental samples and expanding its capabilities before reaching its final configuration and functionality.