A proposal for sequencing the etiological agents of the Food-Borne Trematodiases (FBTs)
by Makedonka Mitreva and Paul Brindley on behalf of the International Food-Borne Trematodiases Community

Executive Summary

The Food-borne trematodiases (FBTs) represent a major group of the neglected tropical diseases (NTDs) – more than 40 million people are infected with one or more of the FBTs, and 750 million (>10% of the world's population) others worldwide remain at risk on contracting FBTs. Over 100 species of food-borne trematodes are known to infect humans, although only several are responsible for much of the FBT disease burden. As with the NTDs at large, most FBTs affect the poorest people in rural areas of the endemic countries.

Human FBTs targeted for genome sequencing.

			WGS Coverage		Transcriptome		Clinical	Illumina[e]
	Site of adult fluke	Species	454/Roche[a]	Illumina[b] (# reads)	454/Roche[c] (# reads)	Illumina[d] (# reads)	isolates (#)	(coverage per strain)
Tier 1	Liver	***Opisthorchis viverrini***	40X	200M	1M	50 M	8	25X
		Fasciola hepatica	40X	200M	1M	50 M	4	25X
		Clonorchis sinensis	40X	200M	1M	50 M	-	-
	Lung	***Paragonimus westermani***	40X	200M	1M	50 M	8	25X
	Intestines	***Haplorchis taichui***	40X	200M	1M	50 M	-	-
		Fasciolopsis buski	40X	200M	1M	50 M	-	-
Tier 2	Liver	***Opisthorchis felineus***	25X	100M	1M	25 M	-	-
		Fasciola gigantica	25X	100M	1M	50 M	4	25X
	Lung	***Paragonimus spp. (3 species)***	25X	100M	1M	25 M	-	-
	Intestines	***Metagonimus yokagawi***	25X	100M	1M	25 M	-	-
		Echinostoma spp (2 species)	25X	100M	1M	25 M	-	-

[a] 454/Roche, combination of fragments, 3KB PE, 8KB PE and 20 KB PE reads; [b] WGS Illumina, 200bp PE and 3kb PE (200M reads=4lns); [c] 1M reads = 1 run; [d] 50M reads = 1 Illumina ln; [e] 25X coverage (2 Illumina lns per clinical isolate).

This document proposes sequencing of 14 FBT genomes (expected size ~ 400 Mb per genome). The Tier 1 flukes include the most important species in terms of public health significance (numbers of people infected, morbidity and mortality) and cosmopolitan distribution. The Tier 2 flukes, while having more localized distribution, will provide key information on the association of infection with bile duct cancer (*O. felineus)*, enigmatic lung parasites (*Paragonimusspp.)* and a large number of human intestinal trematodes (*Metagonimus yokagawi* and two species of *Echinostoma)*, which while usually not as pathogenic as liver and lung flukes, are certainly much more common infections. We propose transcriptome sequencing for each species in support of genome annotation and functional comparative studies as a high priority.

To better understand diversity among populations of key organisms, we propose to sequence eight separate clinical isolates from each of three groups. Sequencing multiple clinical isolates of *Paragonimus* will inform the long standing debate on whether polyploid flukes are more pathogenic than diploid specimens. Sequencing multiple isolates of *O. viverrini* will resolve an existing controversy about the presence of a species complex for *O. viverrini* in the Mekong River basin in Thailand, Laos and Cambodia and address the prediction that certain cryptic species are even more carcinogenic than others. Finally, we will sequence multiple Fasciola isolates to test the long-held speculation that hybrids exist between *Fasciola gigantica* and *F. hepatica*, with increased pathogenic, and even carcinogenic capacity, but unable to produce viable progeny (non-fertile). Sequencing clinical isolates will direct address this hypothesis. Generating polymorphism information, including catalogs of SNPs, by sequencing multiple isolates of each of these organisms will enable functional genomics studies and future laboratory, molecular epidemiological and field investigations.

Genomic information from these major, representative FBT pathogens will provide the research community with lists of genes and other novel information important for: i) Development of new intervention targets (drugs, vaccine, diagnostics); ii) Genetic and biochemical leads in relation to drug sensitivity, especially for praziquantel; iii) Clarification of evolution of the Class Trematoda and Phylum Platyhelminthes, at large, and many other important evolutionary questions including; iv) Development of dioecism in the Platyhelminthes; v) Carcinogenic nature of *Clonorchis* and *Opisthorchis*; vi) Host range, host organ preferences, host finding behavior; vii) Other fundamental aspects of the FBT host-parasite relationship.

This white paper has strong support from worldwide FBT research community. Furthermore, it has the commitment and interest of parasitologists and the NTDs communities at large, geneticists, evolutionary and computational biologists whose contributions to this project will aid in data analysis and hasten the pace of discovery.

1. Introduction and Background

Food-borne trematodiases (FBTs) are one major group of the neglected tropical diseases (NTDs) worldwide with more than 40 million people infected and 750 million (>10% of the world's population) at risk (Hotez et al. 2008; Keiser and Utzinger 2009). Over 100 species of food-borne trematodes are known to infect humans. Most trematodiases affect the poorest people in rural areas of the endemic countries. Many factors contribute to the high prevalence of these infections, including lack of education, poor recognition of the trematode infections because of their vague clinical presentation, poverty, malnutrition, a lack of food inspection and poor sanitation (Graczyk and Fried 2007; Sripa et al. 2010).

The life-cycle of FBTs is complex and diverse, with species-specific characteristics. The common feature is that aquatic snails act as intermediate hosts; Figure 1 depicts typical life cycles of five different food-borne trematodes, including intestinal, liver, and lung flukes.

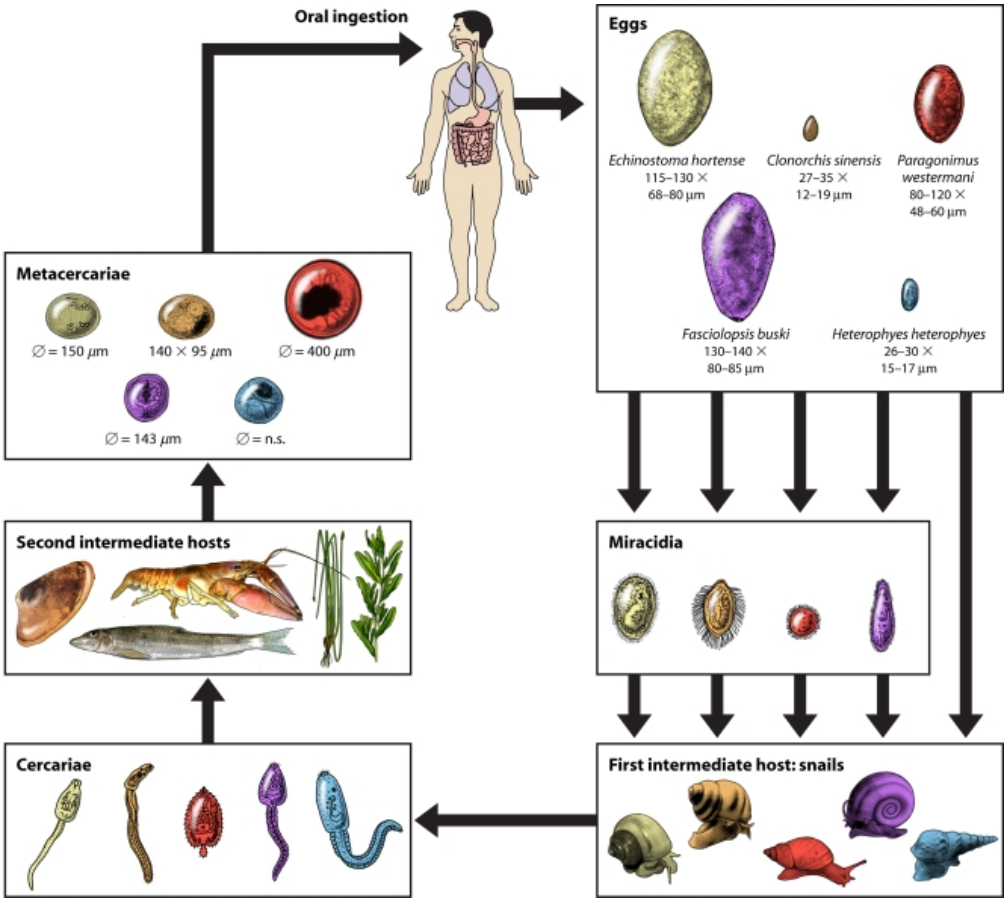


Figure 1. Life cycles of five different food-borne trematodes including intestinal flukes (*Echinostoma hortense*, *Fasciolopsis buski*, and *Heterophyes heterophyes*), a liver fluke (*Clonorchis sinensis*), and a lung fluke (*Paragonimus westermani*) (Keiser and Utzinger, 2009).

Opisthorchiasis caused by *Opisthorchis viverrini* and *O. felineus* and clonorchiasis, caused by *Clonorchis sinensis*, are closely related parasitoses which people contract by ingestion of metacercariae in flesh of raw or undercooked freshwater fishes. Fascioliasis is caused by *Fasciola* species (*F. hepatica* and *F. gigantica*), where infection arises from ingestion of metacercariae on water plants such as watercress (Lun et al. 2005; Sripa et al. 2007). Paragonimiasis is caused by several species of *Paragonimus*, the lung flukes, which use freshwater crabs and other crustaceans as intermediate hosts. Other FBTs cause intestinal infections including *Fasciolopsis buski*, the echinostomes and the many 'minute intestinal flukes' of the family Heterophyidae. These infections usually occur focally and are still endemic in many parts of the world, particularly Southeast Asia (WHO 2002; Graczyk and Fried 2007) (Table 1).

All FBTs can be treated with anthelmintic drugs, praziquantel for most, triclabendazole for fascioliasis. However, people typically become re-infected because it is difficult to convince them to change age-old culinary habits.

Table 1. Food-borne trematodes of medical importance and public health significance (WHO, 2002).

Family	Genus	Species	Source of human	Location in human body
Opisthorchidae	*Clonorchis*	*C. sinensis*	Fish	Liver and biliary system
	Opisthorchis	*O. viverrini*	Fish	Biliary system
		O. felineus	Fish	Biliary
Fasciolidae	*Fasciola*	*F. hepatica*	Plants	Liver and biliary system
		F. gigantica	Plants	Liver and biliary system
	Fasciolopsis	*Fas. buski*	Plants	Small intestine
Troglotrematidae	*Paragonimus*	Several species	Crabs and crayfish	Pleural cavity and lung; occasional
Heterophyidae	*Haplorchis*	*H. taichui*	Fish	Mucosa of small intestine
	Metagonimus	*M. yokogawai*	Fish	Mucosa of small intestine

Vaccines would be a valuable adjunct to chemotherapy but no vaccines are available. In addition since several of the FBTs lead to cancer, future vaccine targeting specific FBTs including *Opisthorchis* and/or *Clonorchis* infections would also represent anti-cancer vaccines (Sripa et al. 2010).

1. Significance of Diseases

1.1. OPISTHORCHIASIS AND CLONORCHIASIS

The human liver flukes, *O. viverrini*, *O. felineus* and *C. sinensis* remain important public health problems in many parts of the world. *Clonorchis sinensis* is widespread in China, Korea and Vietnam, while *O. viverrini* is endemic in Southeast Asia, including Thailand, Lao People's Democratic Republic (Lao PDR), Cambodia and central Vietnam. Human infection follows the consumption of raw or undercooked cyprinoid (freshwater) fish haboring infective metacercariae. Recent reports suggested that about 35 million people are infected with *C. sinensis* globally (Keiser and Utzinger 2009); with up to 15 million human infections in China alone and another 8-10 million individuals infected with *O. viverrini* in Thailand and Lao PDR (Sayasone et al. 2007; Andrews et al. 2008). More than 600 million people, mainly in Asia, are at risk of infection with these two liver flukes (Keiser and Utzinger 2009).

The infections are associated with hepatobiliary diseases including hepatomegaly, cholangitis, fibrosis of the periportal system, cholecystitis, gallstones and are major aetiological agents of bile duct cancer, cholangiocarcinoma (CCA). *O. viverrini* and *C. sinensis* are classified as Group 1 carcinogens – metazoan parasites that are carcinogenic to humans – by the the International Agency for Research on Cancer, World Health Organization (WHO) (Bouvard 2009). Therefore, not only do these liver flukes cause pathogenic helminth infections, they also are carcinogenic in humans in similar fashion to several other more well known biological carcinogens, in particular hepatitis viruses, human papilloma virus and *Helicobacter pylori*.

Liver fluke-induced CCA appears to result from chronic inflammation in the vicinity of adult flukes within the bile ducts. Several mechanisms by which *O. viverrini or C. sinensis* infection may enhance cholangiocarcinogenesis have been proposed. It is almost certain that a combination of mechanical damage caused by the feeding and movement of the worms, parasite secretions, and immunopathology in response to the flukes culminates in CCA after chronic infection with *O. viverrini* or C. sinensis.

The primary pathologic change, epithelial desquamation, may be due to mechanical irritation by the fluke or its metabolic products. Immunopathologic processes contribute to the long-standing hepatobiliary damage. During liver fluke infection, inflammation, periductal fibrosis, and proliferative responses, including epithelial hyperplasia, goblet cell metaplasia, and adenomatous hyperplasia, may represent predisposing lesions that enhance susceptibility of DNA to carcinogens. *N*-nitroso compounds and their precursors occur at low levels in fermented food such as preserved mud fish paste, *pla-ra*, a condiment that is a ubiquitous component of the cuisine of northeastern Thailand and Laos. Indeed, it has been hypothesized that *N*-nitroso compounds (e.g., nitrosamine) are a primary carcinogen leading to CCA in humans in this region (Sripa et al. 2007).

The liver fluke endemic area of Khon Kaen, Northeast Thailand has reported the highest incidence of liver cancer in the world (Sripa et al. 2007).

Table 2. Top 10 diseases – ranked using DALYs - in the population of Thailand (population ~70 million) stratified by sex for the year 2004 (Sripa and Pairojkul 2008). Liver and bile duct cancer, highlighted in bold, primarily results from chronic infection with *Opisthorchis viverrini*.

	Male				Female	
Rank	Disease	DALY ('000)	%	%	DALY ('000)	Disease
1	HIV/AIDS	645	11.3	7.4	313	Stroke
2	Traffic accidents	584	10.2	6.9	291	HIV/AIDS
3	Stroke	332	5.8	6.4	271	Diabetes
4	Alcohol dependence/harmful use	332	5.8	4.5	191	Depression
5	**Liver and bile duct cancer**	**280**	**4.9**	3.4	142	Ischemic heart disease
6	COPD	187	3.3	3	125	Traffic accidents
7	Ischemic heart disease	184	3.2	**3**	**124**	**Liver and bile duct cancer**
8	Diabetes	175	3.1	2.8	118	Osteoarthritis
9	Cirrhosis	144	2.5	2.7	115	COPD
10	Depression	137	2.4	2.6	111	Cataracts

In regard to socioeconomic impact, it was estimated 20 years ago that the total direct cost of *O. viverrini* infection to the work force (between the age of 15 and 60 years) in Northeast Thailand was US$ 80 million per annum. More recently, it has been reported that liver and bile duct cancer, the end-stage consequence of liver fluke disease, ranks number five in Thai males among all diseases with highest number of disability-adjusted life years (DALYs) (Table 2)(Sripa and Pairojkul 2008).

1.2. FASCIOLIASIS

There are an estimated 2.4-17 million people worldwide, excluding Asia, infected with one or both species of the liver fluke *Fasciola*, namely *F. hepatica* and *F. gigantica*, often causing serious acute and chronic morbidity (Mas-Coma et al. 2009). The estimated at-risk population is 80 million (Keiser and Utzinger 2009). In the past, *Fasciola* infections were limited to specific and typical geographical areas, but more recently, this liver fluke has spread throughout the world, especially where sheep and cattle are farmed intensively.

Human cases are increasingly reported from Europe, the Americas and Oceania (where only *F. hepatica* is transmitted), and from Africa and Asia (where the two species overlap). These parasites commonly infect domestic ruminants, particularly sheep, goats and cattle. Humans usually become infected by eating aquatic plants grown in water that is contaminated with feces from animals harboring *Fasciola*. With regard to Southeast Asia, particularly in Vietnam, *Fasciola* infections are known to be increasing (Mas-Coma et al. 2009; Sripa et al. 2010). The parasites cause considerable mortality in sheep and cattle, and human morbidity, which is dependent on the number of worms and stage of infection. The acute phase occurs during migration of the immature flukes through the liver. Severe pathology results from ingestion of liver tissues by the flukes and destruction of parenchymal tissue, haemorrhage, parasite death and inflammatory responses largely mediated by eosinophils. Repair mechanisms can lead to extensive periportal fibrosis. Radiologic imaging in the acute phase may present with multiple, small, clustered, necrotic cavities or abscesses in the peripheral parts of the liver, showing 'tunnels and caves' sign, reflecting parasite migration in the liver parenchyma (Lim et al., 2008). The chronic phase, during which parasites are present in the bile ducts, tends to be less severe. Tissue change, including bile duct proliferation, dilatation and fibrosis, is largely caused by mechanical obstruction of the ducts, inflammatory responses and the activity of proline, which the fluke excretes in large quantities. In this phase, the flukes are demonstrated in the intra- and extrahepatic bile ducts and the gallbladder as small intraluminal flat objects, sometimes moving spontaneously. Bile ducts are dilated. Anemia may result from blood loss through bile duct lesions. Death can result, especially in juvenile sheep, cattle (and humans), as the result of hemorrhage into the bile ducts (Sripa et al. 2010).

1.3. PARAGONIMIASIS

Paragonimus spp., the lung flukes, represents one of the most injurious of the food-borne helminths. These flukes cause paragonimiasis in people and other crab-eating mammals in Asia, parts of West Africa, and South and Central America. About 20 million people are infected with lung flukes and an estimated 293 million people are at-risk (Keiser and Utzinger 2009).

There are about 15 species of *Paragonimus* known to infect humans. *P. westermani* infection is the most common, while *P. heterotremus* is the etiologic agent of human paragonimiasis in China, Lao PDR, Vietnam and

Thailand (Blair et al. 1999; Sripa et al. 2010). Species of *Paragonimus* are reported to infect humans outside Asia, including *P. africanus* and *P. kellicotti* in North America. *Paragonimus miyazakii* may be synonymous with *P. skrjabini*.

Pathogenesis ensues because of the migration of the lung form from the gut to the lungs and indeed through not infrequent ectopic migrations to aberrant sites including the brain and subcutaneous sites at the extremities, and toxin and other mediators released by the migratory parasites (Lv et al. 2010). The presence of the flukes in the lung causes hemorrhage, inflammatory reaction with leukocytic infiltration and necrosis of lung parenchyma that gradually proceeds to the development of fibrotic encapsulation except for an opening from the evolving lesion to the respiratory tract to allow the fluke eggs to exit to the outside environment. There are signs and symptoms that allow characterization of acute and chronic stages of paragonimiasis. In pulmonary paragonimiasis, for example, the most noticeable clinical symptom of an infected individual is a chronic cough with gelatinous, rusty brown, pneumonia-like, bloodstreaked sputum. Hemoptysis is commonly induced by heavy work. Pneumothorax, empyema from secondary bacterial infection and pleural effusion might also be presented. When symptoms include only a chronic cough, the disease may be misinterpreted as chronic bronchitis and bronchiectasis or bronchial asthma. Pulmonary paragonimiasis is frequently confused with pulmonary tuberculosis (Liu et al. 2008). In extra-pulmonary paragonimiasis, the symptoms of extra-pulmonary lesions vary depending on the location of the fluke, including cerebral (Lv et al. 2010) and abdominal paragonimiasis (reviewed in (Sripa et al. 2010).

1.4. CLINICAL ISOLATES

In addition to draft genome sequences, we anticipate that clinical isolates in the context of important medical issues can also be addressed by genome sequencing for the FBTs. First, clinical isolates of *Paragonimus* will inform the long standing debate on whether polyploid flukes are more pathogenic than diploid specimens (Blair et al. 1999). Second, there is controversy about the presence or not of a species complex for *O. viverrini* in the Mekong River basin in Thailand, Laos and Cambodia; e.g. see contrasting in views in (Saijuntha et al. 2007) and (Thaenkham et al. 2010). Clinical isolates from infective persons in informative sites in this region will shed light on this issue, including predictions that cryptic species are even more carcinogenic than others. Third, for *Fasciola gigantica* and *F. hepatica*, it has long been speculated that hybrid species occur in tropical regions including Vietnam and, moreover, that the hybrid fluke is all of more pathogenic, and even carcinogenic, in contrast to the two parental species (Nguyen et al. 2009) and references therein). The hybrids, while viable and pathogenic, cannot themselves produce viable progeny i.e are believed that might be sterile (non-fertile).

2. Rationale

Fundamental molecular biological investigations are needed to underpin the development of novel methods of control, treatment and diagnosis for the FBTs. However, to date, most molecular studies of trematodes have focused on human blood flukes, the schistosomes. Complementing the nuclear genome sequencing of schistosomes (Berriman et al. 2009; Liu et al. 2009) have been sustained efforts to establish *in vitro* systems for functional genomic analyses (Brindley and Pearce 2007; Kalinna and Brindley 2007; Ndegwa et al. 2007; Morales et al. 2008; Rinaldi et al. 2009). This is in stark contrast to the situation for the FBT flukes, for which the potential to explore transcriptomes and gene function is only now being realized (McGonigle et al. 2008; Rinaldi et al. 2008). Despite their major socioeconomic impact, no draft or complete nuclear genomic sequence is available for liver, lung or intestinal flukes, and transcriptomic data fro FBT pathogens have until recently been scant (Young et al. 2010a; Young et al. 2010b).

Genomic and transcriptomics information from the major, representative FBT pathogens will redress this problem and will provide the research community with genes and novel information important to address new interventions, questions of host-parasite relationships, host range, evolution and other fundamental concepts associated with this major grouping of NTDs. Some of these key questions are highlighted below.

A. Why are some food-borne trematodes carcinogenic?

Three helminth pathogens are categorized by the WHO as Group 1 carcinogens – known cancer causing agents. These are *Schistosoma haematobium*, the causative agent of uro-genital schistosomiasis and two major FBTS pathogens, *C. sinensis* and *O. viverrini* (Bouvard 2009). Why these two FBT worms cause cancer whereas other liver flukes including *Fasciola hepatica* do not is unexplained. However, it is highly likely that information revealed from the availability of a draft genome will inform rational scrutiny of this enigmatic phenomenon.

B. Host range, host organ preferences, host finding behavior, etc.

The physiological and behavioral basis of host range in FBT pathogens (and indeed in parasitic helminths at large) is not well understood. Yet issues such as the basis of host preference and, in addition, organ site preferences within suitable hosts are fundamental to understanding parasitism by the FBT flukes. Hypotheses on some aspects have been postulated. For example, Dalton and colleagues, for instances, have noted that helminth pathogens express papain-like cysteine peptidases, termed cathepsins, which have important roles in virulence, including host entry, tissue migration and the suppression of host immune responses. The liver fluke *F. hepatica*, expresses the largest cathepsin L cysteine protease family yet described. Recent phylogenetic, biochemical and structural studies indicate that this family contains five separate clades, which exhibit overlapping but distinct substrate specificities created by a process of gene duplication followed by subtle residue divergence within the protease active site. The developmentally regulated expression of these proteases correlates with the passage of the fluke through intestinal wall then liver capsule host tissues and its encounters with different host macromolecules. Other liver flukes including *C. sinensis* and *Opisthorchis* spp express more cathepsin Fs rather than cathepsin L cysteine proteases, but use a different migration route from the intestine to the liver, via the Ampulla of Vater and entry through the bile duct (Robinson et al. 2008; Pinlaor et al. 2009a). As noted above, information from draft FBT genomes will facilitate investigation of these fundamental problems.

C. Genetic and biochemical leads in relation to drug sensitivity, especially praziquantel

In general, infection with the FBT flukes can be treated successfully with the broad-spectrum, anthelmintic drug, praziquantel. PZQ is a generally safe medication, is given by mouth, is usually well tolerated, can also be used for children older than two years, and is inexpensive. Also, it is the drug of choice for treatment of all forms of human schistosomiasis, and is active against cestode (tapeworm) infections of people and livestock. Thus PZQ is broadly active against platyhelminth parasites. Problematically, however, the mode of action of PZQ is not understood. Also, curiously, PZQ it is only poorly active against *Fasciola* hepatica and *F. gigantica*. (Fortunately, other drugs including triclabendazole are available for treatment of human fascioliasis.) However, why *Fasciola* is not susceptible to PZQ remains to be established, yet this information would be valuable for future development of new drug(s) based on the structure of PZQ. Again, information from draft FBT genomes will facilitate investigation of these critical issues relating to the action of this key NTDs medication.

D. Development of new intervention targets - drugs, vaccine, diagnostics

The availability of genome and transcriptome data for the FBT pathogens will provide the basis for a comprehensive understanding of the molecular mechanisms involved in helminth nutrition and metabolism, host-dependent development and maturation, immune evasion and evolution. These data will also predict new potential vaccine candidates and drug targets. Treatment for the FBTs, as noted, depends almost exclusively on PZQ. Vaccines and new drugs are needed, certainly because drug resistance in human helminth parasites would present a major problem for current treatment and control strategies. Pharmacogenomics with the new helminth genomes represents a practicable route forward toward new drugs. For example, chemogenomics screening of the genome sequence of *S. mansoni* recently identified >20 parasite proteins for which potential drugs are available approved for other human ailments; indeed, in the case of schistosome thioredoxin glutathione reductase, one of these – auranofin, an anti-arthritis medication - was subsequently demonstrated to exhibit potent anti-schistosomal activity (reviewed in (Brindley et al. 2009)). With the new genome data, we can predict similar advances in development of new intervention targets for the FBTs.

In addition, given the extraordinary linkage between eukaryotic pathogens and cancer, represented by liver fluke induced cholangiocarcinoma (above), characterization of the predicted nature and action of proteins of *C. sinensis* and/or *O. viverrini* at the host-parasite interface, such as cathepsin F (Pinlaor et al. 2009b), can be expected to provide insights into this linkage, and indeed fundamental insights into carcinogenesis at large. Membrane and secreted proteins have potential as intervention targets including as vaccine candidates, given recent successes with chemotherapy targeting these kinds of proteins in schistosomes (e.g. (Abdulla et al. 2007)) and with vaccines targeting cathepsin L of *F. hepatica* (McManus and Dalton 2006) and tetraspanins of *S. mansoni* (Tran et al. 2006). Indeed, in view of the recent implementation of an acclaimed vaccination of adolescents against papilloma-virus infection to provide protection from cervical cancer (Schiller et al. 2008), there is the tantalizing prospect that vaccination to prevent liver fluke infection caused by *C. sinensis* and *O. viverrini* could provide protection against another infection-related cancer, liver fluke-induced cholangiocarcinoma.

E. Evolutionary considerations

Genome sequences would facilitate clarification of evolution of the Sub-Class Digenea, the Class Trematoda and Phylum Platyhelminthes, at large. The sequence information would assist with other fundamental evolutionary questions relationships among the Lophotrochozoan assemblage, and relationships among the Lophotrochozoa and other assemblages of the Animalia. They would assist with the evolution of dioecism in the Platyhelminthes (restricted to the schistosomes), to the evolution of secondary host range (e.g. fish for the opistorchiids, shrimp for paragonimids, vegetation for fasciolids), and to parasitism as an ecological strategy.

Moreover, the new sequence information will undoubtedly allow revision of our understanding of the human and other mammalian host-parasite relationships, intermediate host (snail, etc)-FBT pathogens, evolution, and so forth.

F. Functional genomics

Despite difficulties with investigation of helminth parasites, including the inability to maintain any life cycles completely in vitro, the absence of cell lines, complex and expensive animal models, etc., new insights into fundamental helminth biology are accumulating through the schistosome and related genome projects and the application of genome manipulation technologies including RNA interference and transgenesis. We can anticipate the availability of genome sequences for the FBT flukes will hasten functional genomics investigations for these NTDs, including studies to determine the role and importance of the newly discovered genes. The availability of the draft genome sequences of two species of human schistosomes certainly has enabled nascent functional genomics approaches for these parasites, as evidenced in recent reports by Krause-Peterson et al. (Krautz-Peterson et al. 2010), Tchoubrieva et al., (Tchoubrieva et al. 2010) and Yang et al. (Yang et al. 2010).

3. Phylogenetic considerations prior to the whole genome sequencing

Food-borne trematodes are classified into the phylum Platyhelminthes, class Trematoda, and subclass Digenea. To provide the basic taxonomy of FBTs, emphasizing species that are infective to humans and proposed for sequencing, the 18S sequences were downloaded from SILVA rRNA database (www.arb-silva.de).

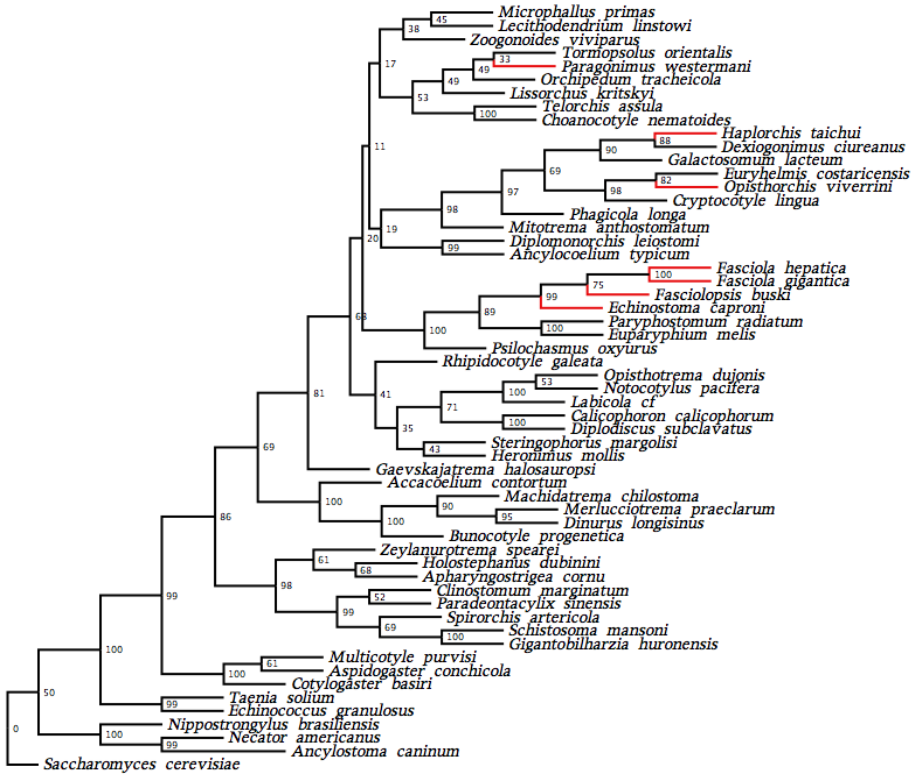


Figure 2. Molecular phylogeny of the class Trematoda based on ribosomal RNA sequence data. For Genera/groups proposed for genome sequencing are included (in red). The final data set of 57 sequences was aligned using MUSCLE. Sequence bootstrapping was done with SEQBOOT of Phylip. Phylogenetic trees was generated using DNADIST and NEIGHBOR of Phylip. Consensus tree together with the bootstrapping values were displayed by FigTree.

Redundant sequences were removed by choosing the longest one for each genus, resulting of 189 representative sequences (one per genus). This list is far from complete. For example, a total of 70 species (14 families and 36 genera) of intestinal flukes have been reported to have been isolated from humans (Chai, 2009). For better visualization, we implemented random sampling approach and excluded very closely related sequences, resulting in inclusion of a total of 51 species. Three nematodes, 2 cestodes and yeast were used as outgroups (Figure 2).

4. Selected species for genome sequencing

Several hundred labs, including groups in the U.S., Europe, Asia, Australia and endemic countries worldwide, study a variety of human parasitic food-borne trematodes and their close relatives. As an indication of the community and active labs, a simple PubMed keyword search produces:

Flatworm	34,490 references
Trematode	24,562
Trematode infections	26,684

However, reports to date have mainly focused on epidemiology, pathogenesis, diagnosis, drug resistance and systematics, rather than systematic genomic approaches that will enable better understanding of the pathogen at a molecular level. The parasitology and Neglected Tropical Diseases (NTDs) communities have available only very limited numbers of sequences originated from FBTs, although these pathogens can be considered as emerging food-borne pathogens and pose significant public health and economic problems. Other parasitic helminths, particularly the nematodes, cestodes and schistosomes have already benefited from genomic research (e.g. ref, nemaESts, Brugia, Schisto). By contrast, the FBTs/trematodes not only remain as a major group of NTDS, but certainly from the genomics, transcriptomics and proteomics perspectives they have truly been neglected.

4.1. Whole Genome Sequencing strategy and assembly

We propose sequencing of 14 species (Table 3), members of 5 families of the three major orders (Opisthorchiida, Plagiorchiida and Echinostomida; Figure 3). When we categorized the species by priority, 2 tiers were identified:

- Tier 1, first priority: these parasites are clearly of major human health importance;
- Tier 2, second priority: species whose genomes are worthy of sequencing, but for which lower coverage is the most cost effective yet still informative approach, since most of them have a close relative proposed in Tier 1.

We propose to sequence the genomes of 6 Tier 1 FBTs to 40X coverage (using Roche/454), and 8 genomes of species in Tier 2 to 25X coverage supplemented by deep coverage on the Illumina platform. Unfortunately, to our knowledge, genome sizes have yet to be determined for any of the FBTs. Based on the genomes of the sequenced blood flukes (*Schistosoma japonicum* and *S. mansoni*), we speculate that the genome size of FBTs is ~400MB. We also propose sequencing eight clinical isolated of each of the three main examples: polyploid *Paragonimus,* cryptic O. viverrini and hybrid Fasciola gigantica *and F. hepatica. By resequencing of these 24 isolates will* begin to explore genome-wide variants in clinical isolates.

Genetic polymorphism is a challenge for *de novo* assembly of population based genome sequencing using next generation technology. Our recent studies of the effect of polymorphism on de novo assembly showed that the polymorphism dramatically reduces assembly contiguity. To address this, we have developed an approach triggered by the increased number of eukaryotic parasite genomes coming through the production and assembly pipelines, and the need to simultaneously assemble and annotate them (http://www.genome.gov/10002154). The approach includes a basic backbone of 454/Roche fragment sequences obtained from a single individual whenever possible. This is the case with the tier 1 trematodes due to the recent achievements that allow WGS library production from as low as 100-500ng of DNA. The body sizes of the Tier 1 species are not diminutive; adult *F. hepatica* flukes are up to 10 cm in length, *O. viverrini* and *C. sinensis* are ~ 2 cm in length, and *Paragonimus westermani* is about the size and shape of a coffee bean; extraction of genomic DNA from a single individual of these flukes typically yields > one µg of gDNA. Hence, having the main backbone from single individual will minimize the effect that heterozygosity has on assembly accuracy and contiguity. For genome improvements this coverage will be supplemented by 8kb and

20kb 454/Roche pair-ends and deep coverage of PCR-free Illumina short insert size and 3kb paired ends for scaffolding.

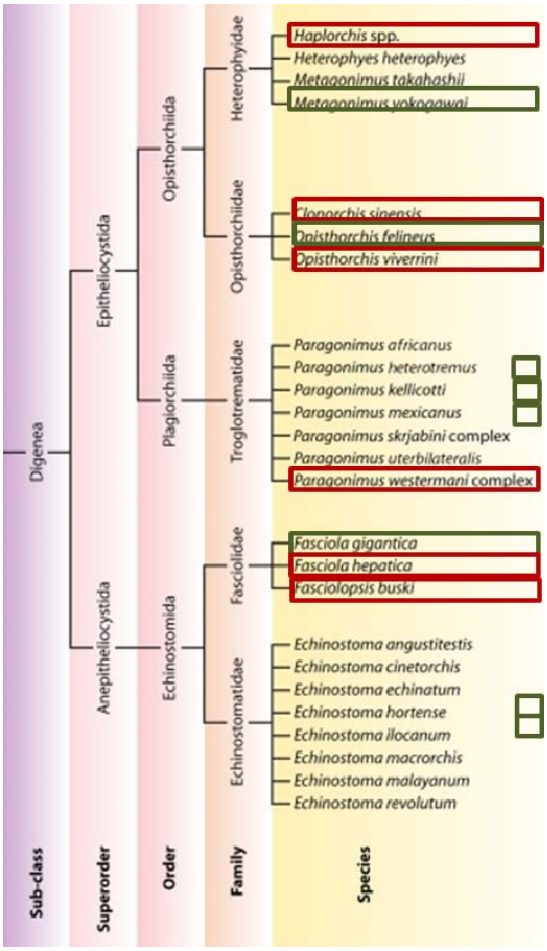


Figure 3. Order Digenea. Species are grouped into major taxonomic groups based on phylogenic analysis of data for the small subunit of ribosomal DNA (SSU) (modified from Keiser and Utzinger, 2009). Species proposed for sequencing are boxed in: in red (tier 1) and in green (tier 2). For Genera for which we have not specified down to a species level the corresponding boxes are just indicating the genus.

While for some of these PE libraries we will use DNA from populations of trematodes these reads are primarily used for scaffolding therefore the polymorphism is not an issue. In addition publicly available (e.g. Tsai et al, 2010) and in-house developed at the WUGC tools (Merger, unpublished) scaffolding tools will be used to close gaps within scaffold and to extend off the ends of scaffolds.

4.2. Automated annotation and variant detection

The genomes will undergo so-called Annotation Grade Improvements. For this standard, gene models (gene calls) and annotation of the genomic content should fully support the biology of the organism and the scientific questions being investigated. This standard is of finished quality in annotated gene regions. Repeat regions at this level will not be resolved, so errors in those regions are much more likely. The first step in gene prediction is masking repeated sequences. We use RECON (25) for automated, *de novo* identification and classification of repeat sequence families. Initial RECON results are checked to remove gene families and further classified before being used to mask the genome. In addition to RECON repeats, the genome is also masked for simple and low-complexity repeats using Repeatmasker. Local tandem and inverted repeats are annotated but not masked.

Furthermore, building on our past experience with manual annotation of ESTs using a multi-tiered approach, we will use the output from our automated gene prediction pipeline described above, and orthologous protein alignments to further validate start codons, length and conservation of final gene predictions in the genome. However, polyploidy is well known in some species, especially in the genera *Paragonimus* and *Fasciola*, and indeed it has been suggested that triploid individuals of *P. westermani* are more pathogenic than diploid specimens (Blair et al., 1999). Karyotypes have also been described for other

FBTs including *C. sinensis* and species of *Opisthorchis* (e.g. see Park et al., 2000; Laha et al., 2007). In any event, prior to sequencing, the genome sizes will be measured using flow-sorted nuclei stained with propidium iodide. *C. elegans* and *D. melanogaster* will be included as standards (Abubacker et al., 2008).

To facilitate better gene prediction and to add depth to the analysis of structure-function for species without substantial number of ESTs, up to 1 million cDNAs (this is one 454 run) will be generated for Tier 1 and for Tier 2 species. To increase the diversity of identified transcripts stage- or tissue-specific cDNA libraries will be constructed whenever possible. cDNA clusters, representing sequences from a single gene, will be produced by interconnecting groups of contigs followed by proper validation (e.g. evaluation of the number of reads supporting a connection) and filtering (e.g. translations of all possible paths). Final sub-graphs will be outputted individually as splicing graphs to record alternative splicing for each cDNA cluster. Coverage will be further enhanced by Illumina sequencing, where we will generate 1 lane per stage (50×10^6 reads of 100 nt each stage based on current capabilities). If HiSeq used the cDNAs will be indexed and pooled resulting in the same total number of cDNAs per species.

Table 3. Targeted human FBTs in regard to priority for genome sequencing.

			GenBank Division			
	Site of adult fluke	**Species**	**Nucleotide**	**EST**	**Proteins**	**Mitochondrion**
Tier 1	Liver	***Opisthorchis viverrini***	120	647,112	14778*	-
		Fasciola hepatica	368	-	264	yes
		Clonorchis sinensis	394	580,418	12368*	yes
	Lung	***Paragonimus westermani***	232	505	140	yes
	Intestines	***Haplorchis taichui***	9	-	1	-
		Fasciolopsis buski	11	-	1	-
Tier 2	Liver	***Opisthorchis felineus***	65	-	56	yes
		Fasciola gigantica	173	-	99	-
	Lung	***Paragonimus spp. (3 species)***	-	-	-	-
	Intestines	***Metagonimus yokagawi***	10	-	9	-
		Echinostoma spp (2 species)	86	358	53	-

* these are partial protein sequences obtained by cDNA contig translations of 454/Roche reads (Young et al, 2010a, 2010b)

These short reads will be mapped to predicted genes from the genome and to the 454-based cDNA contigs. Samples will be non-normalized to provide quantitative and qualitative data on gene expression among regions. The quantitative nature of the Illumina cDNA data will enable measuring the mRNA levels with unparalleled precision. The adults and metacercariae are in general most accessible - either from lab models, or at necropsy of naturally infected definitive or intermediate hosts (e.g. metacercariae [MCs] of *Opisthorchis viverrini* or *Clonorchis sinensis*). Egg stage will be included for species for which they are obtainable.

Clinical isolates: Using the Illumina WGS sequencing approach, we will sequence eight clinical isolated of each of the three main examples: polyploid *Paragonimus,* cryptic O. viverrini and hybrid Fasciola gigantica *and F. hepatica*. By resequencing of these 24 isolates will begin to explore genome-wide variants in clinical isolates. This is enabled in great part by our recent achievement of generating good quality WGS libraries from only 100ng DNA, while retaining the complexity of the DNA population. Our experience with cancer genomes taught us that deep sequencing of tumor and normal genomes (e.g. 30X) captures >99% of single nucleotide variants, allowing identification of nearly all somatic changes in coding and non-coding sequences. The Illumina GAII machine produces ~40Gb data per run (pair-end reads, 2x100bp) and the error rate is <1%. Hence, we propose ~25X coverage (2 Illumina lanes per isolate with the current capacity for 400MB genomes, resulting in only 6 Illumina runs for the 24 clinical isolates). The whole-genome re-sequencing will generate large sets of sequence reads, which will be subjected to quality control on a per-lane or per-region basis. Data sets with excessive levels of non-unique reads will be flagged using a combination of BWA and our existing de-duplication software pipeline and duplicate reads will be filtered prior to the next analysis step. In addition,

datasets comprising primarily reads with insufficient read length, base quality, mapping quality, and paired-end reads with excessive or atypical distribution of insert sizes will be flagged and discarded when necessary.

The processing pipeline will align clinical isolate reads to a reference sequence (Tier 1 species) and then predict variants, using the combination of mappers/aligners and our in-house tools. The consensus caller is run with a cutoff, such that it ignores zero-quality alignments when calling single-nucleotide variations. Because the aligner gives a zero alignment score to reads which fit equally well in multiple locations, this prevents many false positives. Complete sequencing of 5 clinical isolate s will allow us to calculate the background rate in individual genomes, as well as estimate the overall mutation rate in a given fluke type.

4.3. Data deposition and Dissemination

The data will be made available to the public by deposition of traces within 24 hours of data collection and of assemblies > 1kb. All the assemblies will be also added to the sequencing center's ftp site. Gene and protein sets will be deposited into GenBank. We will also work with the sequencing center for possible making the data available to the public on a web-based genome browser (GBrowse), to enable easier design and interpretation of experiments by the scientific community.

6. Community support for this proposal

Consortium members: Paul J. Brindley (Department of Microbiology, Immunology, and Tropical Medicine, George Washington University, Washington DC); Thewarach Laha (Department of Parasitology, Khon Kaen Unity, Faculty of Medicine, Thailand); David Blair (School of Marine and Tropical Biology, James Cook University, Townsville Australia); John P. Dalton (Institute of Parasitology, McGill University, Montreal Canada); Makedonka Mitreva (Department of Genetics, The Genome Center, Washington University, St. Louis, MO).

In addition to the authors of this document, investigators worldwide have endorsed, supported and provided input to this proposal (Table 4). For most of these parasites, DNA and/or RNA is already available in the labs of the consortium members as demonstrated by the recent published papers (e.g. Devi et al., 2010; Rinaldi et al, 2010; Young et al, 2010a, 2010b) and the examples Letters of support (see page 16 and 17).

Table 4. Investigators active in the research area of FBTs who support this proposal.

Investigator	**Affiliation**	**FBT expertise**
Banchob Sripa	Khon Kaen University, Khon Kaen, Thailand	*O. viverrini*
Gabriel Rinaldi	George Washington University, Washington DC	*Fasciola hepatica*
Jose F. Tort	Facultad de Medicina, Universidad de la República,Montevideo, Uruguay	*Fasciola hepatica*
Grace Mulchay	University College Dublin, Dublin, Ireland	*Fasciola hepatica*
Thanh Hoa Le	Institute of Biotechnology, Hanoi, Viet Nam	*Clonorchis* , *Paragonimus, Fasciola*
Hu Wei	National Institute of Parasitic Diseases, China CDC, Shanghai, China	Food borne trematodiases
Zhou Xiao Nung		
Ross H. Andrews	National University of Singapore, Singapore	*O. viverrini* , FBTs
Aaron Maule	Queens University, Belfast, Northern Ireland, UK	*Fasciola hepatica*
Robin B. Gasser	University of Melbourne, Australia	Food borne trematodiases
Terrence Spithill	Charles Sturt University, Australia	*Fasciola* spp
Yoon Kong	Sungkyunkwan University, Suwon, Kora	*Clonorchis sinensis*
Young-An Bae		
Eric S. Loker	University of New Mexico, Albuquerque, NM	*Echinostoma* spp
J. Mahanta	Indian Council of Medical Research, Dibrugarh, Assam, India	*Paragonimus* spp
Takeshi Agatsuma	Kochi Medical School, Nankoku City, Japan	*Paragonimus* spp
Jennifer Keiser	Swiss Tropical Institute, Basel, Switzerland	Food borne trematodiases

References

Abdulla MH, Lim KC, Sajid M, McKerrow JH, Caffrey CR. 2007. Schistosomiasis mansoni: novel chemotherapy using a cysteine protease inhibitor. *PLoS Med* **4**(1): e14.

Andrews RH, Sithithaworn P, Petney TN. 2008. Opisthorchis viverrini: an underestimated parasite in world health. *Trends Parasitol* **24**(11): 497-501.

Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC, Mashiyama ST, Al-Lazikani B, Andrade LF, Ashton PD et al. 2009. The genome of the blood fluke *Schistosoma mansoni*. *Nature* **460**(7253): 352-358.

Blair D, Xu ZB, Agatsuma T. 1999. Paragonimiasis and the genus Paragonimus. *Adv Parasitol* **42**: 113-222.

Bouvard V, Baan, R., Straif, K., Grosse, Y., Secretan, B., Ghissassi F., et al., . 2009. WHO International Agency for Research on Cancer Monograph Working Group. A review of human carcinogens--part B: biological agents. *Lancet Oncol* **10**: 321-322.

Brindley PJ, Mitreva M, Ghedin E, Lustigman S. 2009. Helminth genomics: The implications for human health. *PLoS Negl Trop Dis* **3**(10): e538.

Brindley PJ, Pearce EJ. 2007. Genetic manipulation of schistosomes. *Int J Parasitol* **37**(5): 465-473.

Graczyk TK, Fried B. 2007. Human waterborne trematode and protozoan infections. *Adv Parasitol* **64**: 111-160.

Hotez PJ, Brindley PJ, Bethony JM, King CH, Pearce EJ, Jacobson J. 2008. Helminth infections: the great neglected tropical diseases. *J Clin Invest* **118**(4): 1311 - 1321.

Kalinna BH, Brindley PJ. 2007. Manipulating the manipulators: advances in parasitic helminth transgenesis and RNAi. *Trends Parasitol* **23**(5): 197-204.

Keiser J, Utzinger J. 2009. Food-borne trematodiases. *Clin Microbiol Rev* **22**(3): 466-483.

Krautz-Peterson G, Bhardwaj R, Faghiri Z, Tararam CA, Skelly PJ. 2010. RNA interference in schistosomes: machinery and methodology. *Parasitology* **137**(3): 485-495.

Liu F, Zhou Y, Wang ZQ, Lu G, Zheng H, Brindley PJ, McManus DP, Blair D, Zhang QH, Zhong Y et al. 2009. The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature* **460**(7253): 345-351.

Liu Q, Wei F, Liu W, Yang S, Zhang X. 2008. Paragonimiasis: an important food-borne zoonosis in China. *Trends Parasitol* **24**(7): 318-323.

Lun ZR, Gasser RB, Lai DH, Li AX, Zhu XQ, Yu XB, Fang YY. 2005. Clonorchiasis: a key foodborne zoonosis in China. *Lancet Infect Dis* **5**(1): 31-41.

Lv S, Zhang Y, Steinmann P, Zhou XN, Utzinger J. 2010. Helminth infections of the central nervous system occurring in Southeast Asia and the Far East. *Adv Parasitol* **72**: 351-408.

Mas-Coma S, Valero MA, Bargues MD. 2009. Chapter 2. Fasciola, lymnaeids and human fascioliasis, with a global overview on disease transmission, epidemiology, evolutionary genetics, molecular epidemiology and control. *Adv Parasitol* **69**: 41-146.

McGonigle L, Mousley A, Marks NJ, Brennan GP, Dalton JP, Spithill TW, Day TA, Maule AG. 2008. The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduces gut penetration. *Int J Parasitol* **38**(2): 149-155.

McManus DP, Dalton JP. 2006. Vaccines against the zoonotic trematodes Schistosoma japonicum, Fasciola hepatica and Fasciola gigantica. *Parasitology* **133 Suppl**: S43-61.

Morales ME, Rinaldi G, Gobert GN, Kines KJ, Tort JF, Brindley PJ. 2008. RNA interference of *Schistosoma mansoni* cathepsin D, the apical enzyme of the hemoglobin proteolysis cascade. *Mol Biochem Parasitol* **157**(2): 160-168.

Ndegwa D, Krautz-Peterson G, Skelly PJ. 2007. Protocols for gene silencing in schistosomes. *Experimental Parasitology* **117**(3): 284-291.

Nguyen TG, Van De N, Vercruysse J, Dorny P, Le TH. 2009. Genotypic characterization and species identification of Fasciola spp. with implications regarding the isolates infecting goats in Vietnam. *Exp Parasitol* **123**(4): 354-361.

Pinlaor P, Kaewpitoon N, Laha T, Sripa B, Kaewkes S, Morales ME, Mann VH, Parriott SK, Suttiprapa S, Robinson MW et al. 2009a. Cathepsin F cysteine protease of the human liver fluke, Opisthorchis viverrini. *PLoS Negl Trop Dis* **3**(3): e398.

Pinlaor S, Prakobwong S, Boonmars T, Wongkham C, Pinlaor P, Hiraku Y. 2009b. Effect of praziquantel treatment on the expression of matrix metalloproteinases in relation to tissue resorption during fibrosis in hamsters with acute and chronic Opisthorchis viverrini infection. *Acta Trop* **111**(2): 181-191.
Rinaldi G, Morales ME, Alrefaei YN, Cancela M, Castillo E, Dalton JP, Tort JF, Brindley PJ. 2009. RNA interference targeting leucine aminopeptidase blocks hatching of *Schistosoma mansoni* eggs. *Mol Biochem Parasitol* **167**(2): 118-126.
Rinaldi G, Morales ME, Cancela M, Castillo E, Brindley PJ, Tort JF. 2008. Development of functional genomic tools in Trematodes: RNA interference and luciferase reporter gene activity in *Fasciola hepatica*. *PLoS Neglect Trop Dis* **2**(7): e260.
Robinson MW, Dalton JP, Donnelly S. 2008. Helminth pathogen cathepsin proteases: it's a family affair. *Trends Biochem Sci* **33**(12): 601-608.
Saijuntha W, Sithithaworn P, Wongkham S, Laha T, Pipitgool V, Tesana S, Chilton NB, Petney TN, Andrews RH. 2007. Evidence of a species complex within the food-borne trematode Opisthorchis viverrini and possible co-evolution with their first intermediate hosts. *Int J Parasitol* **37**(6): 695-703.
Sayasone S, Odermatt P, Phoumindr N, Vongsaravane X, Sensombath V, Phetsouvanh R, Choulamany X, Strobel M. 2007. Epidemiology of Opisthorchis viverrini in a rural district of southern Lao PDR. *Trans R Soc Trop Med Hyg* **101**(1): 40-47.
Schiller JT, Castellsague X, Villa LL, Hildesheim A. 2008. An update of prophylactic human papillomavirus L1 virus-like particle vaccine clinical trial results. *Vaccine* **26 Suppl 10**: K53-61.
Sripa B, Kaewkes S, Intapan PM, Maleewong W, Brindley PJ. 2010. Food-borne trematodiases in Southeast Asia epidemiology, pathology, clinical manifestation and control. *Adv Parasitol* **72**: 305-350.
Sripa B, Kaewkes S, Sithithaworn P, Mairiang E, Laha T, Smout M, Pairojkul C, Bhudhisawasdi V, Tesana S, Thinkamrop B et al. 2007. Liver fluke induces cholangiocarcinoma. *PLoS Med* **4**(7): e201.
Sripa B, Pairojkul C. 2008. Cholangiocarcinoma: lessons from Thailand. *Curr Opin Gastroenterol* **24**(3): 349-356.
Tchoubrieva EB, Ong PC, Pike RN, Brindley PJ, Kalinna BH. 2010. Vector-based RNA interference of cathepsin B1 in Schistosoma mansoni. *Cell Mol Life Sci* **67**(21): 3739-3748.
Thaenkham U, Nuamtanong S, Sa-nguankiat S, Yoonuan T, Touch S, Manivong K, Vonghachack Y, Sato M, Waikagul J. 2010. Monophyly of Opisthorchis viverrini populations in the lower Mekong Basin, using mitochondrial DNA nad1 gene as the marker. *Parasitol Int* **59**(2): 242-247.
Tran MH, Pearson MS, Bethony JM, Smyth DJ, Jones MK, Duke M, Don TA, McManus DP, Correa-Oliveira R, Loukas A. 2006. Tetraspanins on the surface of Schistosoma mansoni are protective antigens against schistosomiasis. *Nat Med* **12**(7): 835-840.
WHO. 2002. Report of the Joint WHO/FAO Workshop on Food-borne Trematode Infections in Asia. Ha Noi, Vietnam, 26-28 November 2002.
Yang S, Brindley PJ, Zeng Q, Li Y, Zhou J, Liu Y, Liu B, Cai L, Zeng T, Wei Q et al. 2010. Transduction of Schistosoma japonicum schistosomules with vesicular stomatitis virus glycoprotein pseudotyped murine leukemia retrovirus and expression of reporter human telomerase reverse transcriptase in the transgenic schistosomes. *Mol Biochem Parasitol* **174**(2): 109-116.
Young ND, Campbell BE, Hall RS, Jex AR, Cantacessi C, Laha T, Sohn WM, Sripa B, Loukas A, Brindley PJ et al. 2010a. Unlocking the transcriptomes of two carcinogenic parasites, Clonorchis sinensis and Opisthorchis viverrini. *PLoS Negl Trop Dis* **4**(6): e719.
Young ND, Hall RS, Jex AR, Cantacessi C, Gasser RB. 2010b. Elucidating the transcriptome of Fasciola hepatica - a key to fundamental and biotechnological discoveries for a neglected parasite. *Biotechnol Adv* **28**(2): 222-231.

FACULTY OF MEDICINE
KHON KAEN UNIVERSITY

Our ref: K.K.U.0507.1 /

24 September 2010

Drs. Paul Brindley and Makedonka Mitreva
Department of Microbiology, Immunology & Tropical Medicine
School of Medicine, The George Washington University, Washington DC 20037
And Department of Genetics, The Genome Center, Washington University, St. Louis, MO 63108

Dear Paul and Makedonka,

In regard to our white paper, "The food-borne trematodiases", I write to confirm that we can supply individual adult worms of *Opisthorchis viverrini*. We obtain the metacercariae from naturally infected fish, and adult worms from experimentally infected hamsters. These procedures are routine in our laboratory, and infected fish are plentiful in this region of Northeastern Thailand.

Best wishes for success!

Yours sincerely

Dr. Thewarach Laha, thewa_la@kku.ac.th
Vice-Chair, Department of Parasitology,
Faculty of Medicine,
Khon Kaen University
Khon Kaen, 40002 Thailand

JAMES COOK UNIVERSITY
AUSTRALIA

Townsville Campus
Townsville Qld 4811 Australia
Telephone (07) 4781 5255
International +61 7 4781 5255
www.jcu.edu.au
CRICOS Provider Code 00117J

School of Marine and Tropical Biology
27 Sept 2010

Drs. Paul Brindley and Makedonka Mitreva
Department of Microbiology, Immunology & Tropical Medicine,
School of Medicine, The George Washington University, Washington DC 20037
And Department of Genetics, The Genome Center, Washington University, St. Louis, MO 63108

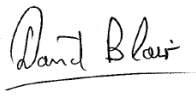
Re: Foodborne Trematodes White Paper

Dear Paul and Makedonka,

In regard to our white paper, "The food-borne trematodiases", I write to confirm that we can access adults of species of *Paragonimus* – the lung flukes – reasonably readily. I have studied flukes of this genus for many years, including undertaking research on their taxonomy and systematics, their mitogenomics, ploidy, biogeography and evolution.

Best wishes for success!

Yours sincerely



Prof. David Blair
School of Marine and Tropical Biology
James Cook University
Townsville,
Queensland 4811