

Research article

Open Access

Gene discovery for the carcinogenic human liver fluke, *Opisthorchis viverrini*

Thewarach Laha^{†1}, Porntip Pinlaor^{†2}, Jason Mulvenna^{†3}, Banchob Sripa², Manop Sripa², Michael J Smout³, Robin B Gasser⁴, Paul J Brindley⁵ and Alex Loukas^{*3}

Address: ¹Department of Parasitology, Khon Kaen University, Khon Kaen, Thailand, ²Department of Pathology, Khon Kaen University, Khon Kaen, Thailand, ³Division of Infectious Diseases and Immunology, Queensland Institute of Medical Research, Brisbane, Australia, ⁴Department of Veterinary Science, University of Melbourne, Melbourne, Australia and ⁵Department of Tropical Medicine, Tulane University Health Sciences Center, New Orleans, USA

Email: Thewarach Laha - thewa_la@kku.ac.th; Porntip Pinlaor - porawa@kku.ac.th; Jason Mulvenna - jason.mulvenna@qimr.edu.au; Banchob Sripa - banchob@kku.ac.th; Manop Sripa - thewa_la@kku.ac.th; Michael J Smout - michael.smout@qimr.edu.au; Robin B Gasser - robinbg@unimelb.edu.au; Paul J Brindley - paul.brindley@tulane.edu; Alex Loukas* - Alex.Loukas@qimr.edu.au

* Corresponding author †Equal contributors

Published: 22 June 2007

Received: 8 March 2007

BMC Genomics 2007, 8:189 doi:10.1186/1471-2164-8-189

Accepted: 22 June 2007

This article is available from: <http://www.biomedcentral.com/1471-2164/8/189>

© 2007 Laha et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Cholangiocarcinoma (CCA) – cancer of the bile ducts – is associated with chronic infection with the liver fluke, *Opisthorchis viverrini*. Despite being the only eukaryote that is designated as a 'class I carcinogen' by the International Agency for Research on Cancer, little is known about its genome.

Results: Approximately 5,000 randomly selected cDNAs from the adult stage of *O. viverrini* were characterized and accounted for 1,932 contigs, representing ~14% of the entire transcriptome, and, presently, the largest sequence dataset for any species of liver fluke. Twenty percent of contigs were assigned GO classifications. Abundantly represented protein families included those involved in physiological functions that are essential to parasitism, such as anaerobic respiration, reproduction, detoxification, surface maintenance and feeding. GO assignments were well conserved in relation to other parasitic flukes, however, some categories were over-represented in *O. viverrini*, such as structural and motor proteins. An assessment of evolutionary relationships showed that *O. viverrini* was more similar to other parasitic (*Clonorchis sinensis* and *Schistosoma japonicum*) than to free-living (*Schmidtea mediterranea*) flatworms, and 105 sequences had close homologues in both parasitic species but not in *S. mediterranea*. A total of 164 *O. viverrini* contigs contained ORFs with signal sequences, many of which were platyhelminth-specific. Examples of convergent evolution between host and parasite secreted/membrane proteins were identified as were homologues of vaccine antigens from other helminths. Finally, ORFs representing secreted proteins with known roles in tumorigenesis were identified, and these might play roles in the pathogenesis of *O. viverrini*-induced CCA.

Conclusion: This gene discovery effort for *O. viverrini* should expedite molecular studies of cholangiocarcinogenesis and accelerate research focused on developing new interventions, drugs and vaccines, to control *O. viverrini* and related flukes.

Background

Throughout East Asia, there is a strikingly high prevalence of cholangiocarcinoma (CCA – cancer of the bile ducts) in regions where the human liver fluke is endemic. No stronger link occurs between a human malignancy and infection with a eukaryotic parasite than that between CCA and infection with the liver fluke, *Opisthorchis viverrini* (Digenea) [1]. Indeed, the International Agency for Research on Cancer (IARC) recognizes *O. viverrini* as a 'category I carcinogen' [2,3]. CCA is highly prevalent in Northeast Thailand, areas where uncooked cyprinoid fish are a dietary staple. Due to poor sanitation practices and inadequate sewerage infrastructure, *O. viverrini*-infected people pass the trematode's eggs in their feces into natural bodies of fresh water. Aquatic snails, which represent the first intermediate hosts of *O. viverrini*, ingest the eggs from which the miracidia undergo asexual reproduction before a population of the free swimming larval stage, called a cercaria, is shed from the infected snails. The cercaria then locates a cyprinoid fish, encysts in the fins, skin and musculature of the fish, and becomes a metacercaria. The metacercarial stage is infective to humans and other fish-eating mammals. Infection is acquired when people ingest raw or undercooked fish. The young adult worm escapes from the metacercarial cyst in the upper small intestine and then migrates through the ampulla of Vater into the biliary tree, where it develops to sexual maturity over four to six weeks, thus completing the life cycle. The adult worms, which are hermaphrodites, can live for many years in the liver, even decades, shedding as many as 200 eggs per day which pass out via bile into the chyme and feces [4].

In Thailand, ~6 million people are infected with *O. viverrini*. Despite widespread chemotherapy with the compound, praziquantel, the prevalence of *O. viverrini* in some endemic areas approaches 70% (reviewed in [1]). Moreover, in Thailand, liver cancer is the most prevalent of the malignant/fatal neoplasias, and the prevalence of CCA in regions in which the parasite is endemic is unprecedented [5].

While sexual reproduction takes place in the mature adults of *O. viverrini* within the bile ducts, asexual reproduction in the snail leads to a massive increase in the number of infectious cercarial stages exiting and swimming off to locate then infect the fish host. The adult fluke is a diploid organism which reproduces by meiosis; self fertilization of the male and female organs occurs, but it is believed that cross-fertilization between adjacent adult worms is the normal pattern. Although the genome size of *O. viverrini* has not yet been reported, it is known to have six pairs of chromosomes, i.e. $2n = 12$ [6], distinct from the closely related liver fluke, *Clonorchis sinensis*, which possesses $2n = 56$ chromosomes [7].

Despite its public health importance, only a small number of *O. viverrini* sequences (mostly ribosomal genes) have been available in public databases prior to the present study. Characterization of the genes expressed in this organism should provide a foundation for elucidating the immunopathogenesis of CCA, particularly the molecular mechanisms by which infection with this parasite induces cancer. Indeed, secreted proteins of *O. viverrini* induce hyper-proliferation of cells (or hyperplasia) *in vitro* [8], implying that carcinogenesis may not be just a consequence of chronic inflammation, but that the parasite actively secretes gene products which initiate neoplasia.

Here, we undertake gene discovery for *O. viverrini* after the construction of a cDNA library and characterization of ~5,000 expressed sequence tags (ESTs) from this carcinogenic parasite. A similar dataset exists for *C. sinensis* [9], which, despite its widespread prevalence [10], is not recognized as a 'class I carcinogen' [3]. Therefore, we compared the available transcriptomic dataset from *O. viverrini* with those from *C. sinensis*, and from several other flatworms, both free-living and parasitic in humans.

Results and Discussion

Features of the dataset

Of 5,159 randomly selected ESTs, a total of 4,241 yielded acceptably high quality sequences. These in turn were clustered into contigs, establishing a catalogue of 1,932 non-redundant OvAEs. This apparently represents the largest dataset thus far for any of the liver flukes. Table 1 summarizes the key features of the dataset. Of note is that the identities for 1,070 (55%) of these OvAEs could not yet be established, as they did not share sequence homology (BLASTx/tBLASTx) with any other predicted or known molecules in public databases, including dbEST which contains 2,678 ESTs from the related liver fluke, *C. sinensis* [9]. The average insert size of these novel OvAEs was 550 nt; 47 of these 1,070 OvAEs had insert sizes of less than 150 nt. These ESTs may in fact be *O. viverrini*-specific or even digenetic fluke-specific genes. A similar situation currently pertains to the human blood fluke where a large percentage of known transcripts, and indeed proteins, are assumed to be *Schistosoma*- or indeed phylum Platyhelminthes-specific [11,12]. If the *O. viverrini* genome has 14,000 protein-coding genes (like the blood fluke *S. mansoni*) [13], and if each of the 1,932 *O. viverrini* contigs represented a protein coding gene, these newly discovered genes from the adult stage of *O. viverrini* are predicted to represent ~14% of the entire transcriptome of this liver fluke. EST sequences described herein have been deposited in dbEST under accession numbers [EL618683–EL620614](#).

Table 1: Features of the *Opisthorchis viverrini* EST catalogue.

Feature	Number
Initial Sequences	5159
Usable sequences ^a	4241
Contigs	1995 (1632 singletons; 363 clusters)
Contigs after clean-up ^b	1932
Contigs identical to known proteins ^c	68
Contigs similar to other proteins ^d	794
Contigs with gene ontology assignments	383
Novel contigs	1070
Novel contigs with signal sequences	75 (29 signal peptides; 46 signal anchors)
Average insert size	548 bp (ESTs); 660 bp (contigs)
Percentage of recombinant clones	95%
Number of ribosomal seqs	1184 ESTs; 136 clusters

^aUsable sequences were determined using seqclean – sequences that were removed were either non-recombinant, of low complexity and/or quality and those of < 100 nt in length.

^bclean-up refers to removal of sequences from contaminating sources; eg. *Mycoplasma*

^cidentity determined by ≥ 95% identity over ≥ 50 amino acids.

^dbased on BLASTx and tBLASTx searchers of GenBank nr and dbEST respectively.

Abundantly expressed transcripts

After manual filtering of 136 ribosomal sequences, the 10 most abundantly represented mRNAs encoded proteins with known or unknown functions, including one contig that did not have homologues in any public databases (Table 2). Abundant contigs encoded proteins involved in a range of physiological functions which are considered essential to parasitism, such as anaerobic respiration (myoglobin) [14], reproduction (vitelline precursors and egg shell proteins) and detoxification of xenobiotic compounds (glutathione-S-transferase). Other abundantly expressed OvAEs encoded proteins of likely key relevance to the host-parasite relationship, and included proteases (papain-like and legumain-like enzymes), saposin-like proteins and dynein light chains. Homologues of some of the most abundantly represented OvAEs were also highly represented in *C. sinensis* ESTs (cysteine proteases, myoglobin, vitelline B precursor), whereas others were uniquely over-expressed in each species. In particular, structural molecules, including tubulin and actin-binding proteins, were among the 10 most abundant clones from *C. sinensis* [9], but were not highly represented in the dataset from *O. viverrini*. An in-depth comparison of the *O. viverrini* and *C. sinensis* datasets is presented below.

Gene ontology assignments of ESTs from *O. viverrini* and related flukes

Three hundred and eighty three (383) of the total 1,932 OvAEs (19.8%) could be assigned GO classifications (Figure 1). The most abundant groups represented under the molecular function category were linked to binding (34.8%), catalytic activity (27.9%) and structural molecule activity (13.9%). Other sequences of interest identi-

fied in this category were inferred to relate to caspase activity (0.2%) and transporter activity (5.3%). The most abundant groups represented under the biological process category corresponded to physiological processes (41.2%), cellular processes (39.7%) and unknown biological processes (15.6%).

We then undertook a comparative assessment of GO assignments of sequences from *O. viverrini* and two other trematode parasites of humans in Asia – the liver fluke, *C. sinensis* (2,679 contigs) and the blood fluke, *S. japonicum* (107,427 contigs). In general, the percentages of ESTs allocated to each GO category among these three flukes was similar (Figure 2). However, some categories were over- or under-represented in one species. For example, contigs encoding structural proteins were ~four times more abundantly represented in the two liver flukes than in *S. japonicum*, whereas contigs encoding motor proteins were ~three times more abundant in *O. viverrini* than they were in *C. sinensis* or *S. japonicum*. Sample sizes were too small to determine whether these differences were statistically significant. Both structural and motor proteins are important components of fluke teguments [15], playing roles in surface maintenance and turnover in schistosomes [16,17] and liver flukes [18]. Therefore, these molecular differences might reflect the specialised niches and physiological requirements of each parasite. From just 1932 OvAEs, 15 different contigs had ORFs encoding components of the dynein complex of motor proteins, a category of motion-related, and surface and gut-localized EF-hand motif-containing antigens which, at least in schistosomes, represent potent allergens and targets of protective immunological responses [17,19].

Table 2: The 10 most abundant contigs^a from the *Opisthorchis viverrini* EST dataset.

Contig	ESTs/contig	Closest homologue in GenBank nr (accession no.)	%identities (no. of aa)	Score (Bits)	Closest homologue in dbEST (accession no.)	%identities (no. of aa)	Score (Bits)
OvAEI587	100	vitelline B precursor, <i>O. viverrini</i> (AAL23712)	99% (230)	493	<i>C. sinensis</i> cDNA clone CSAD-01-D02 (AT007557)	92% (225)	524
OvAEI588	77	17 kDa myoglobin, <i>Clonorchis sinensis</i> (AAM18464)	81% (149)	244	<i>C. sinensis</i> cDNA clone CSAD-29-A12 (AT009373)	77% (188)	344
OvAEI585	77	hypothetical protein, <i>C. sinensis</i> (AAM55183)	84% (90)	156	<i>C. sinensis</i> cDNA clone CS30 (AT006763)	84% (100)	199
OvAEI593	41	egg protein, <i>C. sinensis</i> (AAN64160)	89% (253)	389	<i>C. sinensis</i> cDNA clone CSAD-20-B05 (AT008604)	82% (237)	477
OvAEI584	37	hypothetical protein, <i>Macaca fascicularis</i> (BAE73006)	67% (59)	82	SJA_AAF_D11.T3 SJA S. japonicum (CX857852)	85% (94)	183
OvAEI602	21	histone H1, <i>Schistosoma japonicum</i> (AAP06509)	74% (70)	112	<i>C. sinensis</i> cDNA clone CSAD-25-H03 (AT009091)	85% (177)	302
OvAEI607	17	egg protein, <i>C. sinensis</i> (AAN64160)	59% (252)	288	<i>C. sinensis</i> cDNA clone CSAD-01-B01 (AT007532)	82% (252)	493
OvAEI595	16	retrotransposon gag region, <i>Monascus pilosus</i> (ABC24965)	31% (57) ^b	33	NA	NA	NA
OvAEI608	15	translationally controlled tumor protein, <i>C. sinensis</i> (AAX84199)	98% (169)	306	<i>C. sinensis</i> cDNA clone CSAD-24-E07 (AT008979)	92% (122)	228
OvAEI601	11	glutathione-S-transferase, <i>O. viverrini</i> (AAL23713)	98%(213)	429	<i>C. sinensis</i> cDNA clone CSAD-32-E05 (AT009695)	86% (232)	484

^aNon-ribosomal sequences only were used in this analysis.

^bSequence identity was low but diagnostic motifs of gag were detected over 57 amino acids.

Evolutionary relationships between *O. viverrini* and other platyhelminths

To assess the evolutionary relationships between *O. viverrini* and other platyhelminths (both parasitic and free-living), we used SimiTri [20] to plot the relative similarities of predicted polypeptide sequences (Figure 3). OvAEs shared most sequence identity with sequences from *C. sinensis* (2,679 publicly available ESTs), and *S. japonicum* (107, 427 ESTs). OvAEs were less similar to sequences from the free-living turbellarian platyhelminth, *Schmidtea mediterranea* (171,472 ESTs) [21], which was not altogether surprising given the phylogenetic distance between parasitic and free-living members of the phylum Platyhelminthes [22,23]. The bulk of this phylum (including those species analysed here) represents a monophyletic group based on 18S rDNA sequences [22] and morphological characters [24], and is often referred to as the Rhabditophora [22]. Therefore, the members of the Rhabditophora are considered to be more closely related to each other than to other turbellarian clades, such as the Polycladida [22]. A total of 105 OvAEs had homologues in the ESTs from the two parasitic flukes but not in the free-living *Schmidtea* ([see additional file 1]; selected examples are shown in the table in Figure 3), suggesting that at least some of these are parasitism-specific genes. Thirty-three (33) of the conserved parasitic fluke genes were novel and did not have homologues of known function. Predicted proteins of known function included homologues of legumain, fatty acid binding proteins, myoglobin and potential anti-inflammatory proteins such as Ly6/UPAR domain-containing proteins. Of the

parasitic fluke-specific genes, 38 encoded ORFs with N-terminal signal sequences; 14 of these OvAEs had homologues in just *S. japonicum* and 24 had homologues in both *S. japonicum* and *C. sinensis*.

Secreted and membrane proteins

We conducted an analysis of ORFs containing an N-terminal signal peptide or signal anchor. A total of 164 OvAEs contained ORFs with signal sequences. The dataset was divided into three categories – sequences that were (i) novel; (ii) platyhelminth-specific; (iii) conserved across multiple phyla. Novel sequences constituted 55.4% of the total, but only 5.2% of them encoded proteins with a signal sequence. Conserved sequences constituted 36.3% of the total, and 10.7% of these encoded proteins with signal sequences. Finally, the sequences inferred to be platyhelminth-specific accounted for 8.3% of the total dataset, but 20.6% of these encoded proteins with a signal sequence (Figure 4). It should be noted, however, that not all of the OvAEs contained full-length nt sequences, and therefore the true percentage of sequences with signal peptides cannot be definitively inferred in the absence of full genome coverage.

Sequences encoding novel secreted and/or membrane proteins (without orthologues/paralogues in other organisms or phyla) may be of particular interest for the development of vaccines and drugs, because the absence of host homologues enhances the prospect for therapeutic margins of safety. OvAEs encoding secreted/membrane proteins involved in many aspects of parasitism were

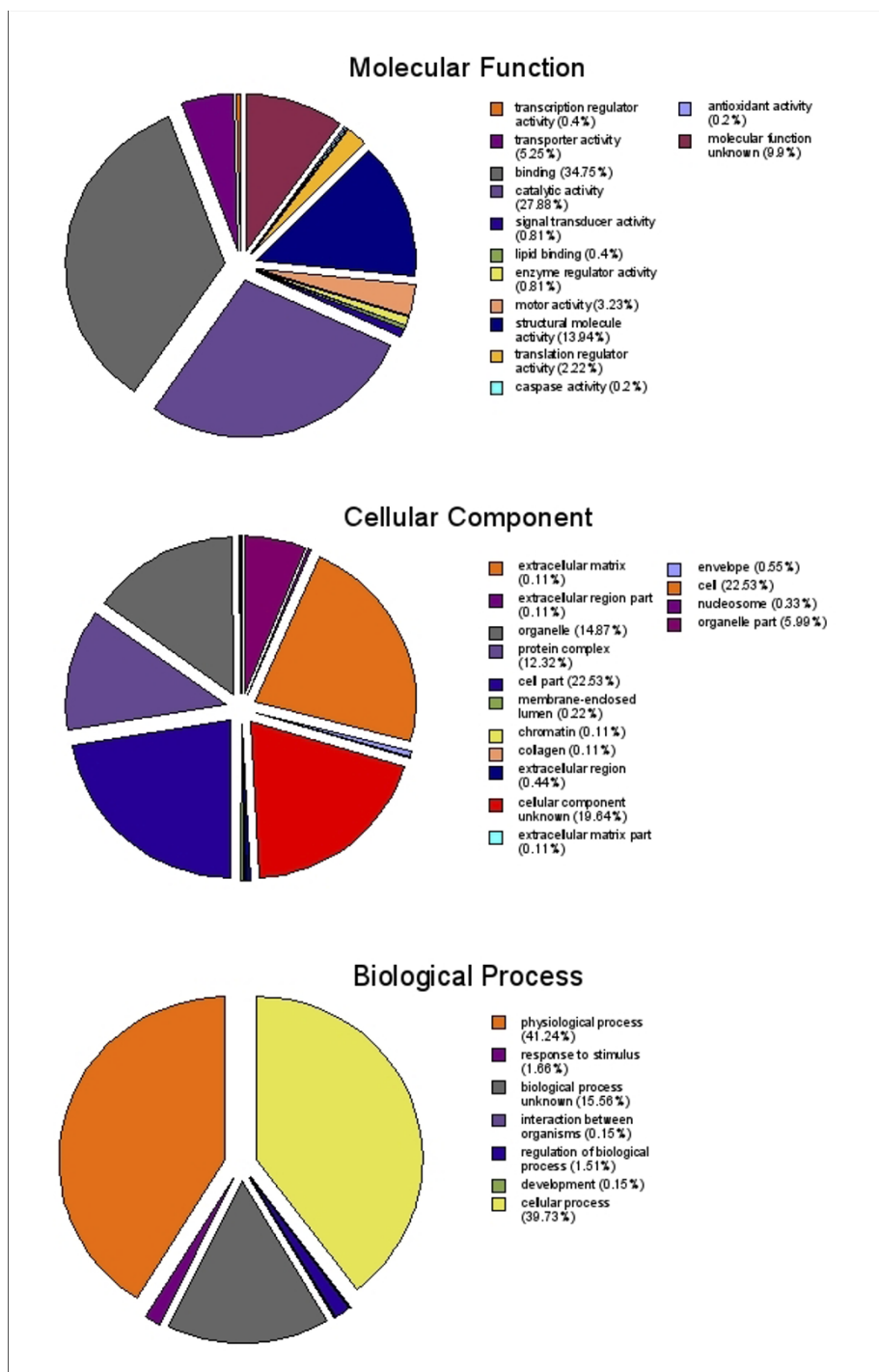


Figure 1
Summary of predicted gene product function and location using gene ontology terms. Gene ontology (GO) terms for annotated *Opisthorchis viverrini* assembled ESTs were extracted, if present, from the GO database and sorted into the immediate subcategories for molecular function, cellular component and biological process. The GO subcategory and percentage relative to the total number of extracted terms is indicated in the legend. Although cellular and physiological processes, structural proteins and catalytic activity were strongly represented other categories of interest include the caspases and transporter activity that may represent proteins important for a parasitic lifestyle. The large number of unknowns in each of the three categories highlights the lack of knowledge regarding many of the proteins found in these parasites.

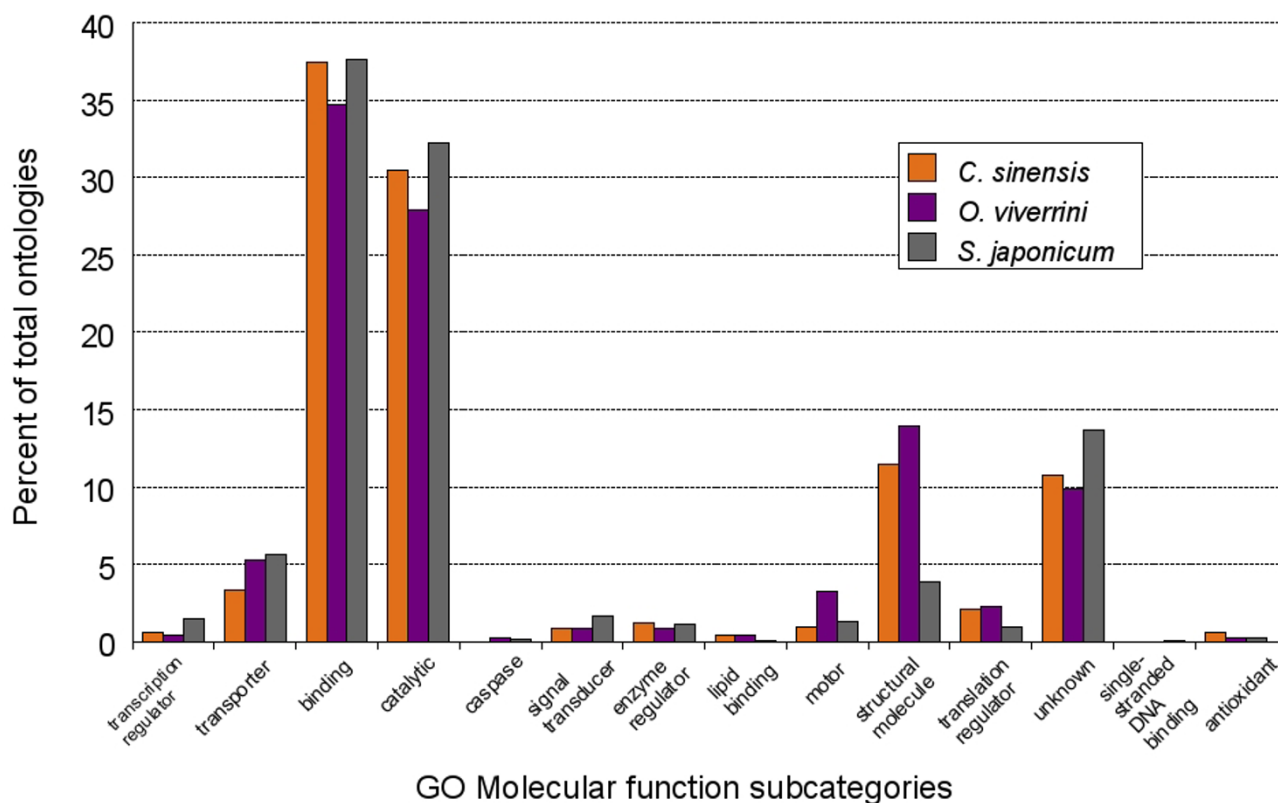


Figure 2
Comparison of the gene ontology molecular function terms for expressed sequence tags from *Opisthorchis viverrini*, *Clonorchis sinensis* and *Schistosoma japonicum*. Expressed sequence tags from *C. sinensis* and *S. japonicum* were downloaded from NCBI and subjected to the same analyses used for *O. viverrini* sequences. A comparison of the percentage of terms correlating to the molecular function subcategory for each organism shows a broad similarity, although in some cases, such as categories for structural or motor proteins, categories are over- or under-represented in certain species.

identified (Table 3), and some of these are discussed in the following section. Two of the OvAEs presented in Table 3, which are inferred to encode transforming growth factor β receptor (see section "Host-parasite cross-talk") and calumenin, were more similar in sequence to vertebrate proteins from both the non-redundant and dbEST databases than they were to platyhelminth sequences, suggesting that they have evolved independently to bind host ligands. These results are reminiscent of reports of the schistosome transcriptome where, for example, receptors for mammalian hormones, including insulin, fibroblast growth factor and cytokines, have been hypothesized to bind host molecules (reviewed in [25]).

Proteases

As with other parasitic helminth transcriptomes [11-13,25,26], proteins with catalytic activity were abundantly represented in the *O. viverrini* dataset (27.9% of contigs that were assigned GO molecular functions). Many of these enzymes encoded endo- and exo-proteases belonging to established families (MEROPS classification), but

which have not yet been described from liver flukes (Table 3). Of particular interest were members of the S1A serine protease family with sequence similarity to kallikrein and chymotrypsin, and, therefore, potentially involved in feeding or tissue migration [27]. Other proteases included homologues of enzymes that digest hemoglobin in blood-feeding helminths, including cathepsin D-like aspartic and cathepsin B-like cysteine proteases [28-30] as well as an asparaginyl endopeptidase, which is known to activate the gastrodermal cathepsin B enzyme, and probably other gut proteases in *S. mansoni* [31]. We also identified *O. viverrini* homologues of the cell death enzyme, caspase-2, and the neutral cysteine protease from the tegument of schistosomes, calpain.

Multiple membrane-spanning proteins

Predicted proteins with multiple membrane spanning domains were identified. Tetraspanins, an abundantly represented family of four-transmembrane proteins in the tegument of schistosomes [32,33], were identified from *O. viverrini* (Figure 5). These proteins are thought to stabi-

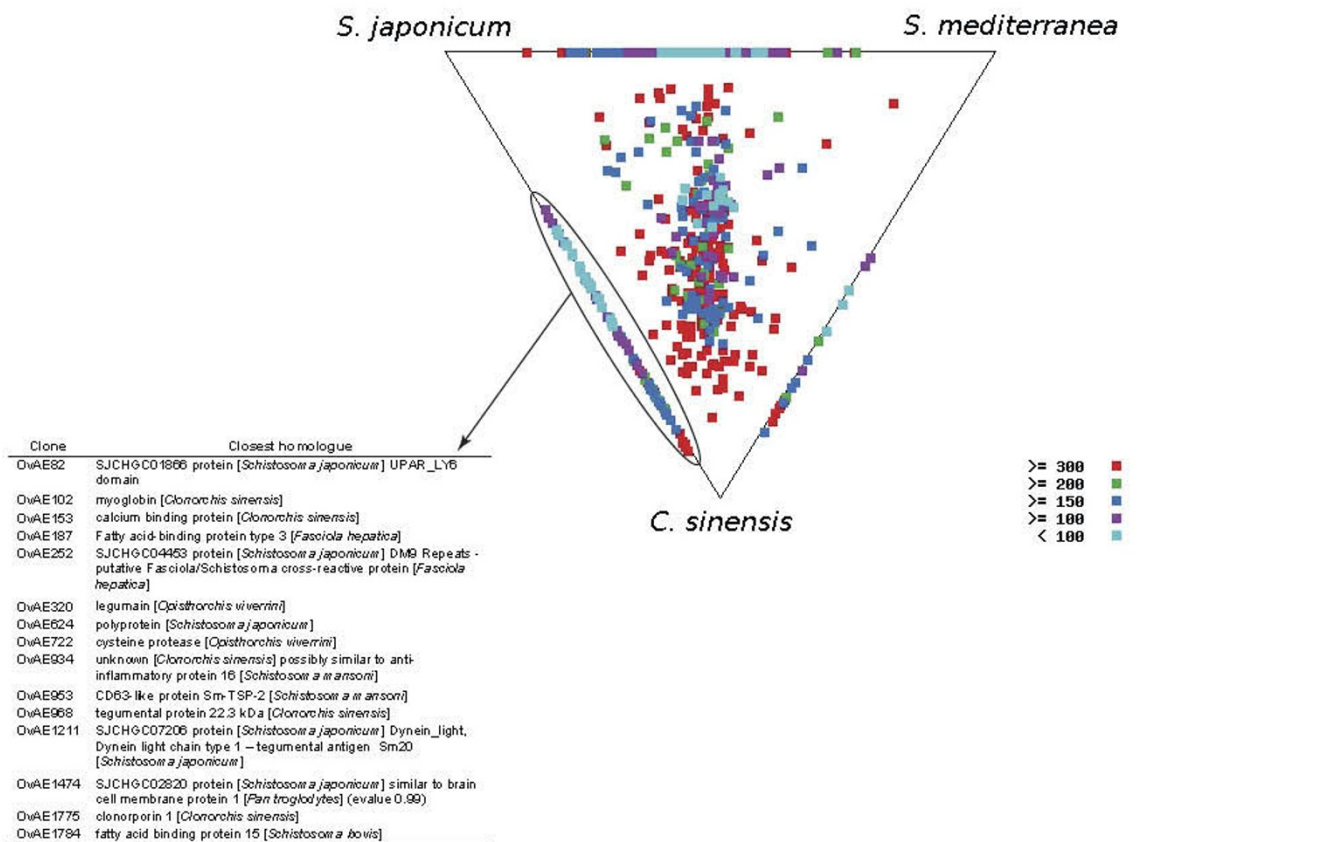


Figure 3
Evolutionary relationships between *Opisthorchis viverrini* and related platyhelminths based on similarities of protein coding genes using SimiTri. Similarity of *O. viverrini* ORFs (1,932 ESTs) to those from the liver fluke *Clonorchis sinensis* (2,679 ESTs), the blood fluke *Schistosoma japonicum* (107,427 ESTs) and the free-living turbellarian *Schmidtea mediterranea* (171,472 ESTs). SimiTri [20] was used to plot 1,932 *O. viverrini* contigs against related species database entries (A). Each spot represents a unique contig and its sequence similarity to each of the three selected databases as determined by tBLASTx scores. Sequences showing similarity to only one database are not shown. Sequences showing sequence similarity to only two databases appear on the lines joining the two databases. Spots are coloured by their highest tBLASTx score to each of the databases. *O. viverrini* sequences with homologues in the parasitic flukes only (not in *Schmidtea*) are highlighted in the dotted region and the identities of selected examples are shown in the table (B). The entire list (105) of these putative parasite-specific proteins is shown in Table S1.

lize the cell membrane by forming a network of interactions, called the tetraspanin web, with other membrane-bound and -associated proteins, particularly on the surface of cells of the immune system [34]. A homologue of the six-transmembrane domain family of water channel proteins, aquaporin [35], was identified. Seven transmembrane proteins are common drug targets [36], and at least three distinct members of this family were identified, including receptors for dystroglycan and lamin b, and a protein with homology to a the DC-STAMP receptor from the surface of dendritic cells (Table 3).

Host-parasite "cross talk"

Parasitic helminths receive host-derived signals for growth and reproduction *via* surface receptors for host ligands, [37-39]. Convergent evolution of extracellular parasite proteins to promote their interactions with host tissues is well documented [40,41], and we identified *O. viverrini* ORFs encoding membrane and secreted proteins, some of which were clearly more similar to vertebrate than to invertebrate homologues (Table 3). Transforming growth factor-beta (TGF-β) regulates cell growth and differentiation and is acquired on the cell surface by specific

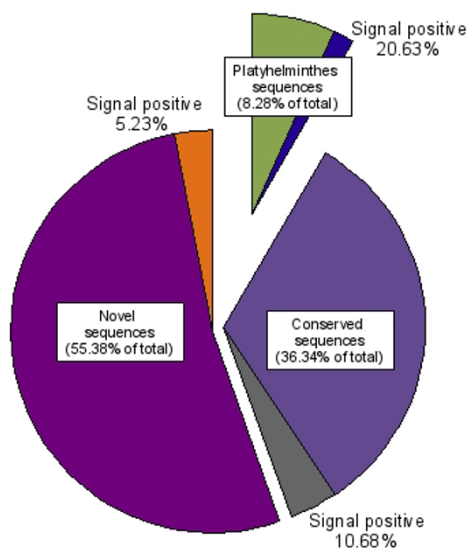


Figure 4
Distribution of *Opisthorchis viverrini* assembled ESTs (OvAEs) that contain predicted signal peptides or signal anchors. OvAEs that had BLAST hits greater than 1×10^{-5} were sorted into conserved (those matching entries for species other than platyhelminths), phylum Platyhelminthes-specific (only matching platyhelminth entries) and novel (no significant homology to any database entry). The sequences in each category were then analysed for the presence of a signal sequence using SignalP. The relative percentages of each category are indicated along with the sub-category of signal sequence positive contigs.

TGF- β receptors [42]. An ORF encoding a member of the TGF- β receptor type Ib family was identified in *O. viverrini*. The ORF included a 28 amino acid insertion absent from other type I TGF- β receptors, except for TR1 from the hydatid tapeworms of the genus *Echinococcus* (also members of the phylum Platyhelminthes) [43]. However, these two insertions did not share sequence identity (Figure 6A). Unlike many of the ESTs identified for which the closest homologues were from parasitic trematodes, the *O. viverrini* TGF- β receptor type I was divergent from SmRK-I of *S. mansoni* [44] and instead grouped more closely with proteins from *Echinococcus multilocularis* and from parasitic and free-living nematodes (Figure 6B). In pairwise sequence comparisons, however, the *O. viverrini* partial ORF was more similar to pig and macaque sequences (44% identity over 181 amino acids) than it was to *Echinococcus* TR1 (40% over 180 residues) or SmRK1 (40% over 182 residues). SmRK-I is known to bind to human TGF- β [40], suggesting that the *O. viverrini* receptor might also bind host growth factors for maturation and reproduction. Another OvAE encoding a protein which is potentially involved in the acquisition of host

signals (and subsequent signaling) for growth and development was a fibroblast growth factor (FGF) receptor substrate 2. Parasitic flatworms induce fibrosis (and FGF) [45], and the parasites might acquire and utilize the host FGF that they induce for development and reproduction. Indeed, schistosomes are dependent upon FGF and transferrin for growth and maturation *in vitro* [46]. Of the sequences presented in Table 3, another OvAE which shared greatest identity with vertebrate homologues, encoded for calumenin, an EF-hand calcium binding protein localized to the secretory pathway. Calumenin is an inhibitor of the gamma-carboxylation system [47] and is expressed in thrombin-activated thrombocytes. It has a modulating effect on the organization of the actin cytoskeleton and may be involved in the pathophysiology of thrombosis or in wound healing [48]. The predicted calumenin of *O. viverrini* was most similar to rat and frog orthologues/paralogues, suggesting that it might interact with actin on the surface of host cells which are damaged during parasite feeding and migration.

Molecules associated with cancer?

O. viverrini is the major cause of CCA in South-East Asia [1]. The molecular mechanisms underlying induction of *O. viverrini*-induced CCA are thought to be multi-factorial (reviewed in [49]), but recent evidence suggests that *O. viverrini* secretes mitogenic proteins into host tissues [1,8]. OvAEs encoding secreted proteins with prospective mitogenic activity were identified in the EST dataset. Of note, first, progranulin (*pgrn*) is a pluripotent secreted growth factor that mediates cell cycle progression, cell motility [50] and wound repair [51]. We identified an OvAE (OvAE1732) that shared sequence identity with *pgrn* (data not shown). Of particular importance is that *pgrn* has been implicated in regulating the proliferation of tumour cells, and its expression is up-regulated in more aggressive cancers (reviewed in [50]). The kallikrein-like serine proteases are another family of enzymes whose over-expression has been linked to cancer. The expression of some kallikreins in prostate cells leads to changes indicative of an epithelial to mesenchymal transition, an important process in cancer progression [52]. An OvAE (OvAE1918) with sequence identity to kallikrein-like secreted proteases is present in the new *O. viverrini* gene catalogue. Phospholipase A2 (PLA-2) regulates the provision of arachidonic acid to both cyclooxygenase- and lipoxygenase-derived eicosanoids (reviewed by [53]), and the upregulation of cyclooxygenase-2 is thought to be an important feature of cholangiocarcinogenesis in both humans and experimental rodent models [49,54,55]. We identified an OvAE (OvAE1644) that encodes a secreted PLA-2 which shared greatest sequence identity with PLA-2 from venom of *Heloderma* (Gila monster) and an EST from *C. sinensis* (Table 3). Parasites utilize secreted serine proteases [56] and PLA-2s [57] to invade host tissues, and

Table 3: Selected *Opisthorchis viverrini* contigs that encode families of secreted/membrane proteins that potentially interact with or are exposed to host tissues. Genera of the closest homologues from BLAST × (nr) searches are shown. Where the closest homologue was from a vertebrate (bold font), a tBLASTx search against dbEST was conducted.

Predicted function	Examples/comments and genera of closest orthologues/paralogues	Contigs	%identities
TGF-β receptor	bone morphogenic protein receptor type I (<i>Sus nr/ Macaca est</i>)	OvAE22	44
Seven transmembrane receptor	DC-STAMP (<i>Strongylocentrotus</i>); laminin receptor (<i>Bos nr/ Clonorchis est</i>)	OvAE92, OvAE1722	51
Tetraspanin	stabilize cell membranes – expressed in the tegument of schistosomes (<i>Schistosoma</i>)	OvAE953	34
C1 family papain-like cysteine protease	cathepsin L (<i>Paragonimus</i>), cathepsin B (<i>Fasciola; Clonorchis</i>)	OvAE1795, OvAE813, OvAE1171, OvAE532, OvAE1070, OvAE1613, OvAE1711, OvAE615, OvAE398	> 80
C13 family asparaginyl endopeptidase	legumain (<i>O. viverrini</i>)	OvAE1624, OvAE1824	94
S1 family serine protease	HtrA-like (<i>Macaca nr/Schistosoma est</i>) and kallikrein-like (<i>Schistosoma</i>) peptidases	OvAE622, OvAE1918	47–53
A1 family aspartic protease	cathepsin D-like; digestive enzyme in helminths (<i>Clonorchis</i>)	OvAE1300	80
M41 family metalloprotease	mitochondrial membrane proteinase (<i>Schistosoma</i>)	OvAE65	91
Granulin	mitogen associated with cancer (<i>Bos nr/Clonorchis est</i>)	OvAE1732	45
Aquaporin	water channel protein (<i>Schistosoma</i>)	OvAE6	48
Tyrosinase	critical for <i>S. mansoni</i> egg shell production (<i>Schistosoma</i>)	OvAE1900, OvAE1854	63
Phospholipase A2	similar to vertebrate venom proteins; (<i>Heloderma nr/ Clonorchis est</i>)	OvAE1644	55
Thioredoxin peroxidase	immunomodulatory in fasciolosis (<i>Schistosoma</i>)	OvAE54	74
EF-hand secreted Ca ²⁺ -binding protein	calumenin (<i>Rattus nr/Xenopus est</i>)	OvAE61	47
Saposin-like protein	pore forming; similar to fluke cytolysins (<i>Clonorchis</i>)	OvAE1692	64
Pathogenesis related protein	similar to helminth venom allergen homologues (<i>Schistosoma</i>)	OvAE534, OvAE1862	38
Glutathione-S-transferase	detoxification of heme and free radicals (<i>Clonorchis</i>)	OvAE1057, OvAE1892, OvAE1601, OvAE1729	86
Synaptobrevin	neurotransmission/vesicular docking – vesicle associated (<i>Schistosoma</i>)	OvAE1001	73
Innexin	integral membrane protein forming gap junctions (<i>Schistosoma</i>)	OvAE631	78
Fibroblast growth factor (FGF) receptor substrate 2	host FGF is essential for growth of schistosomes (<i>Schistosoma</i>)	OvAE1563	32
Ly6c	Immune cell differentiation antigen	OvAE82	26

homologues of these proteins (and granulin) are potentially secreted by *O. viverrini* into host tissues where they might promote cell proliferation, mutagenesis and ultimately carcinogenesis. Ongoing studies in our laboratories are now focused on the physiological roles of these putative carcinogens in the host-parasite relationship and in cholangiocarcinogenesis induced by *O. viverrini* infection.

Potential vaccines

Digenean flukes develop through a series of morphologically and developmentally discrete stages within their mammalian hosts, and each stage can be expected to display a characteristic transcriptome, confounding efforts to develop new control measures. Adult parasitic flukes are bound by an outer epithelial tegument, a structure that is widely regarded as the most vulnerable target for vaccines and drugs [32]. Homologues and orthologues of vaccine

antigens identified in the tegument (and other structures from larval stages) of other flatworms were identified in the *O. viverrini* dataset (Table 4). Of particular note were the membrane spanning proteins, including an orthologue of the protective tetraspanin from *S. mansoni*, Sm-TSP-2 [32,33] (Figure 5) and the 22.6 kDa family of antigens from the schistosome tegument [58]. Homologues of gut proteases used by blood-feeding helminths to digest their blood-meal were identified from *O. viverrini*, including cathepsin D-like aspartic proteases [59,60], 11 distinct papain-like cysteine proteases [61–63] and the neutral protease, calpain, which associates with the inner tegument of schistosomes [64]. Other potential immunogens include lipid-binding proteins which are efficacious vaccines in the rabbit model against the western liver fluke, *Fasciola hepatica*, including saposin-like proteins [65] and fatty acid-binding proteins [66,67].

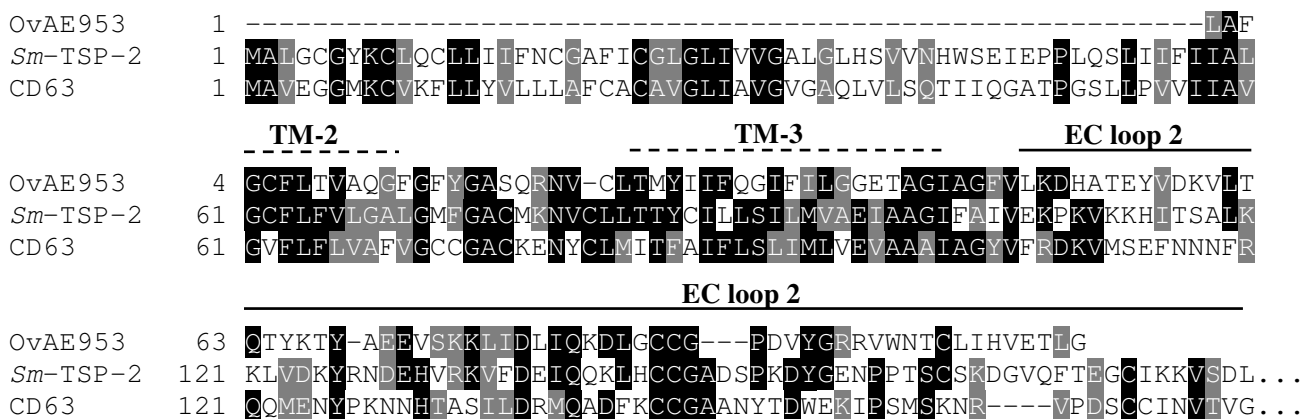


Figure 5
An *Opisthorchis viverrini* homologue of *Sm-TSP-2*, a vaccine antigen expressed in the tegument of *Schistosoma mansoni*. Multiple sequence alignment comparing the ORF of OvAE953 with *Sm-TSP-2* from *S. mansoni* (GenBank AF521091) and human CD63 (NM_001780). Both *Sm-TSP-2* and CD63 sequences shown here are truncated at the C-terminus (fourth transmembrane domain and C-terminal tail are not shown) for comparative purposes because OvAE953 is a partial sequence. Black boxes denote identical residues shared by two or more of the sequences. Grey boxes denote conservative substitutions. Dashed lines denote the predicted transmembrane (TM) domains of *Sm-TSP-2*; the solid line represents the extracellular (EC) loop 2 region of *Sm-TSP-2* [33].

Conclusion

This report provides the first description of gene discovery for the liver fluke *O. viverrini*. Infection with *O. viverrini* is an important tropical health issue, but even more important and enigmatic is that chronic *O. viverrini* infection leads to the development of CCA. Indeed, there is no stronger link between a human parasite and cancer than that between *O. viverrini* and CCA [68]. The new gene catalogue for *O. viverrini* represents the largest EST dataset in the public domain for any species of liver fluke, and provides a platform for explorations into the molecular basis of host-helminth parasite interactions. We [1] and others [8] are interested in the molecules secreted into host tissues by *O. viverrini* that induce hyper-proliferation of biliary cells which can subsequently undergo malignant transformation. Given the number of *O. viverrini* ESTs sequenced herein, it is possible that mRNAs corresponding to these parasite mitogens are already present in the current dataset. Proteomic analysis of proteins secreted by adult *O. viverrini* maintained *in vitro* also is underway in our laboratories, and linking peptide sequences to corresponding mRNAs can be expected to be facilitated by this gene discovery program [12]. Finally, this gene discovery information for *O. viverrini* should expedite molecular studies of cholangiocarcinogenesis and accelerate research focused on developing new interventions, drugs and vaccines, to control *O. viverrini* and related flukes.

Methods

Parasite material

Adult *O. viverrini* were collected from experimentally infected hamsters (*Mesocricetus auratus*) maintained at the animal facility of the Khon Kaen University Faculty of Medicine. Protocols approved by the Khon Kaen University Animal Ethics Committee were used for all animal research conducted in this study. Briefly, metacercariae of *O. viverrini* were collected from naturally infected cyprinoid fish by pepsin digestion. Metacercariae (100 per hamster) were administered intragastrically to hamsters. Hamsters were euthanized 6 weeks after inoculation, and adult worms were flushed with saline from the bile ducts [69]. Worms were washed extensively with sterile phosphate-buffered saline (pH 7.2), after which they were snap frozen and stored in liquid nitrogen or employed immediately as the source of fluke RNA.

Construction and mass excision of cDNA library

Total RNA from adult *O. viverrini* was extracted using Trizol (Invitrogen), following the manufacturer's instructions. Ten µg of *O. viverrini* total RNA was used as a template for the synthesis of double-stranded cDNA using the SMART cDNA kit (BD Bioscience), after which the cDNA modified with adapters was cloned into the *Sfi* I site of the pTriplEx2 plasmid (BD Bioscience) and packaged into λ arms. The titer and percentage of recombinant phages in the library were determined using the protocols

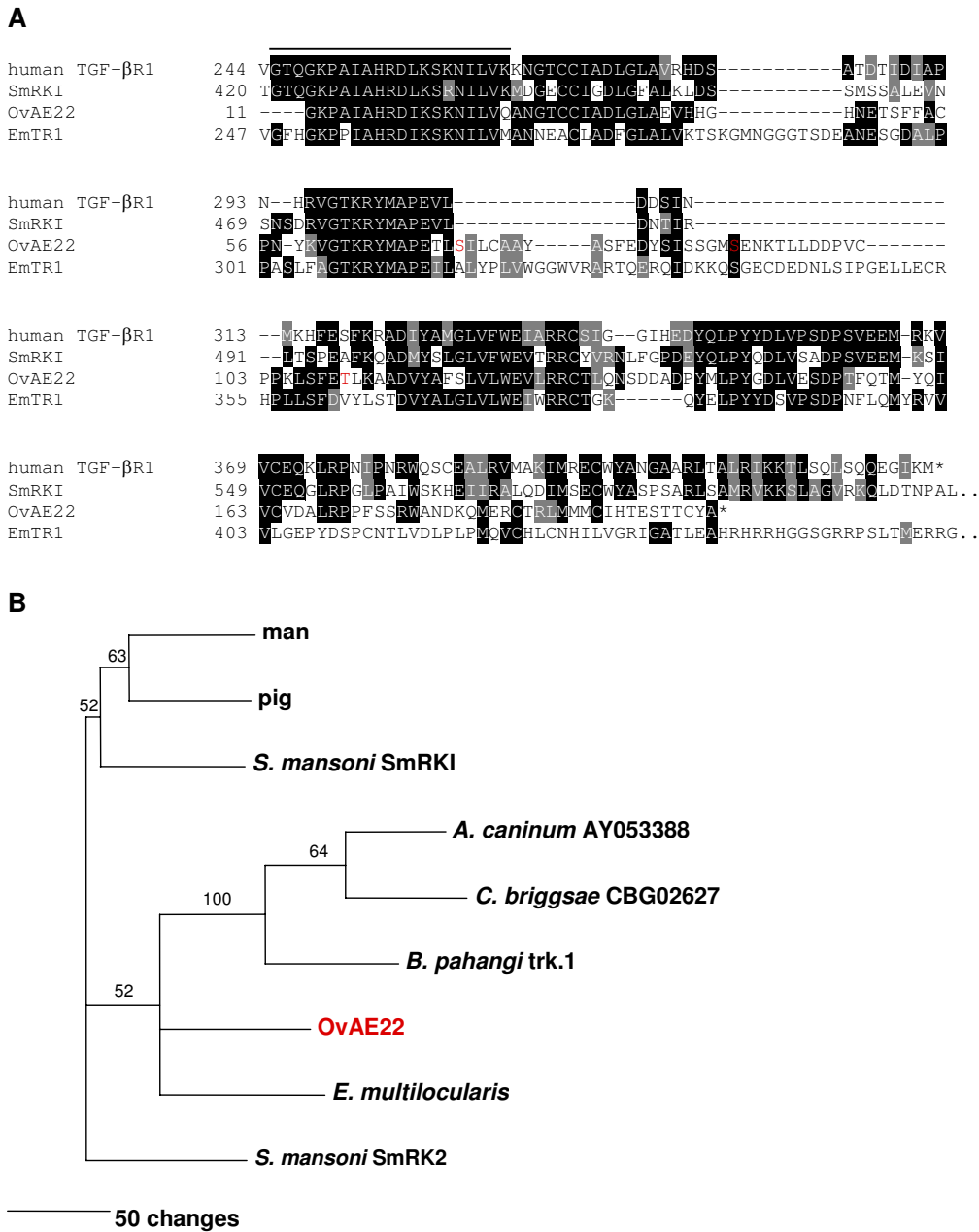


Figure 6
A TGF-β receptor type I from *Opisthorchis viverrini*. Multiple sequence alignment of the ORFs of OvAE22 with homologues from *Schistosoma mansoni* (SmRK-1 – GenBank AF031557), the hydatid tapeworm *Echinococcus multilocularis* (TRI – AJ841786) and human (TGF-β receptor type I – LI1695) (**A**). The overlined region denotes the putative serine-threonine kinase active site in SmRK-1 [44]. Residues highlighted in red font in OvAE22 are putative sites of serine/threonine phosphorylation. Both SmRK-1 and human TGF-β receptor type I sequences shown here are truncated at the N-terminus and SmRK-1 is truncated at the C-terminus for comparative purposes with the partial sequence from *O. viverrini*. Black boxes denote identical residues shared by two or more of the sequences. Grey boxes denote conservative substitutions. Neighbour joining phylogenetic tree showing the relationship between the ORF of OvAE22 and other members of the TGF-β receptor type I family (**B**). Numbers on branches denote bootstrap values from 100 samplings. The nominated outgroup was the type 2 receptor, SmRK-2. GenBank accession numbers not already provided above are as follows: pig bone morphogenic protein (BMP) receptor type I (AY065994); dog hookworm *Ancylostoma caninum* S/T kinase (AY053388); *Caenorhabditis briggsae* CBG02627 (CAAC01000012); filarial nematode *Brugia pahangi* trk-1 (AF013991); *S. mansoni* SmRK-2 (AY550912).

Table 4: *Opisthorchis viverrini* ESTs with sequence identity to mRNAs encoding proteins efficacious as vaccines against other flatworm and nematode parasites.

Predicted protein family	Vaccine antigen and helminth targeted	References
Aspartic protease	APR-1 for hookworm; cathepsin D for <i>Schistosoma japonicum</i>	[59, 60]
Glutathione-S-transferase	Ac-GST-1 for hookworm; bilvax for schistosomes	[79, 80]
Cysteine protease	TSBP for <i>Haemonchus contortus</i> ; AcCPI for hookworm; Cathepsins L1 and L2 for <i>Fasciola hepatica</i>	[61–63]
Tetraspanin	TSP-1, TSP-2 and Sm23 for <i>Schistosoma mansoni</i>	[33, 81]
Pathogenesis related protein calpain	ASP-2 for hookworm; ASP-1 for <i>Onchocerca volvulus</i>	[82, 83]
Fatty acid binding protein	Smp80 for schistosomes	[64]
Saposin-like protein	Sm14 for <i>S. mansoni</i> and <i>F. hepatica</i>	[66, 67]
14-3-3	FhSAP-2 for <i>F. hepatica</i>	[65]
22.6 (unknown function)	Sm14-3-3 for schistosomes	[84]
	Sm22.6 for <i>S. mansoni</i>	[58]

recommended by the manufacturer. *Escherichia coli* strain BM25.8 cells were transduced with recombinant phage, from which the excision of the pTriplEx phagemid library was accomplished.

Clone selection, sequencing and annotation

Five thousand clones were randomly selected from the phagemid library and grown overnight in Luria Bertani (LB) broth supplemented with ampicillin to a final concentration of 25 µg/ml. Overnight cultures were shipped at 4°C in LB broth/ampicillin to the University of Melbourne (Department of Veterinary Science). The sequencing was performed by AgGenomics Inc., Australia, using a 3730xl DNA analyzer (Applied Biosystems). The TempliPhi™ DNA Sequencing Template Amplification system (GE Healthcare) was used to sequence each clone using the 5'λ TriplEx2 sequencing primer.

Bioinformatic analyses

Edited sequences were condensed into contigs or singletons using TGICL [70] with the default parameters of 40 bp overlap, a minimum of 95% identity and a 30 bp maximum mismatched overhang. Sequences of less than 100 nt were discarded. Sequences were named using the same convention as that used for the human blood fluke, *Schistosoma mansoni* [13]; OvAE for *O. viverrini* Assembled EST. Sequences were compared with those available in the NCBI non-redundant protein and nucleotide databases using BLASTx and BLASTn searches, respectively in October 2006. The dbEST database was queried using BLASTn and tBLASTx searches. BLAST alignments with an *E*-value of $\leq 1.0 \times 10^{-5}$ were reported. OvAEs were functionally categorized by querying a local copy of the Gene Ontology (GO) database [71] (downloaded November, 2006) with an *E*-value cutoff of 1.0×10^{-5} . All ESTs from *C. sinensis* [9] and *Schistosoma japonicum* [11,12] were downloaded from NCBI [72], and the same methodology was used to derive ontology classifications for the *C. sinensis* ESTs. ORF predictions were performed using GENSCAN [73] using the

HumanIso parameter set. Signal sequence prediction was accomplished using SignalP 3.0 [74], incorporating both hidden Markov models and neural networks. Positive signal sequence predictions from either method and positive signal anchor predictions using Markov models were reported. Predictions of transmembrane domains were conducted using TMPred [75]. All multiple sequence alignments were carried out using ClustalW. Clan and family assignments of proteolytic enzymes were analyzed via the MEROPS protease database [76]. Putative phosphorylation sites were predicted using the NetPhos 2.0 server [77].

Phylogenetic trees

Multiple sequence alignments were assembled using ClustalW. Only regions which completely overlapped with partial ORFs of *O. viverrini* ESTs were used for tree construction. Alignments were imported into PAUP version 4.0 beta [78] to construct trees using the neighbour joining and maximum parsimony methods. Robustness was assessed by bootstrap analysis using 100 replicates. Clades with more than 50% support were denoted with bootstrap values on the branches.

Cross-taxon similarity analysis

OvAEs were compared with all entries for other organisms in the NCBI dbEST database using tBLASTx. The highest BLAST scores (above a cut-off value of 50) were used to generate SimiTri plots [20] using software developed in-house (J. Mulvenna, unpublished).

Authors' contributions

TL, PP and JM generated and analyzed the data and contributed to drafting of the ms. BS provided parasite material, helped conceive the study and edited the drafted ms. MS and MJS provided technical assistance and edited the drafted ms. RBG facilitated interactions with the sequencing unit and edited the drafted ms. PB helped conceive the study, supervised the research and helped draft the ms. AL

helped conceive the study, supervised the research, and took the lead on drafting the ms. All authors read and approved the final ms.

Additional material

Additional file 1

Opisthorchis viverrini sequences that had homologues in the parasitic flukes, Clonorchis sinensis and Schistosoma japonicum, but not in the free-living platyhelminth, Schmidtea mediterranea.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-8-189-S1.doc>]

Acknowledgements

We are grateful for grant support from the Sandler Foundation, NIH-NIAID (award number AI065871), Khon Kaen University Research Fund, Thailand-Tropical Diseases Research Program and the Australian Research Council (LP0667795 and DP0665230). JM was supported by a Peter Doherty Training Fellowship from the National Health and Medical Research Council of Australia (NHMRC). AL was supported by an R. Douglas Wright Career Development Award from NHMRC. We thank Ian Smith, Jacqui Batley, Nonglack Kewgrai, Bronwyn Campbell and David Blair for advice or support.

References

- Sripa B, Kaewkes S, Sithithaworn P, Mairiang E, Laha T, Smout M, Paironkul C, Bhudhisawasdi V, Tesana S, Thinkamrop B, Bethony JM, Loukas A, Brindley PJ: **Liver fluke induces cholangiocarcinoma.** *PLoS Med* 2007, **in press**.
- IARC: **Infection with liver flukes (Opisthorchis viverrini, Opisthorchis felinus and Clonorchis sinensis).** *IARC Monogr Eval Carcinog Risks Hum* 1994, **61**:121-175.
- IARC monographs on the evaluation of carcinogenic risks to humans.**
- Sithithaworn P, Tesana S, Pipitgool V, Kaewkes S, Paironkul C, Sripa B, Paupairoj A, Thaiklar K: **Relationship between faecal egg count and worm burden of Opisthorchis viverrini in human autopsy cases.** *Parasitology* 1991, **102 Pt 2**:277-281.
- Srivatanakul P: **Epidemiology of Liver Cancer in Thailand.** *Asian Pac J Cancer Prev* 2001, **2(2)**:117-121.
- Komalamisra C: **Chromosomes and C-banding of Opisthorchis viverrini.** *Southeast Asian J Trop Med Public Health* 1999, **30(3)**:576-579.
- Park GM, Im K, Huh S, Yong TS: **Chromosomes of the liver fluke, Clonorchis sinensis.** *Korean J Parasitol* 2000, **38(3)**:201-206.
- Thuwajit C, Thuwajit P, Kaewkes S, Sripa B, Uchida K, Miwa M, Wongkham S: **Increased cell proliferation of mouse fibroblast NIH-3T3 in vitro induced by excretory/secretory product(s) from Opisthorchis viverrini.** *Parasitology* 2004, **129(Pt 4)**:455-464.
- Cho PY, Lee MJ, Kim TI, Kang SY, Hong SJ: **Expressed sequence tag analysis of adult Clonorchis sinensis, the Chinese liver fluke.** *Parasitol Res* 2006, **99**:602-608.
- Lun ZR, Gasser RB, Lai DH, Li AX, Zhu XQ, Yu XB, Fang YY: **Clonorchiasis: a key foodborne zoonosis in China.** *Lancet Infect Dis* 2005, **5(1)**:31-41.
- Hu W, Yan Q, Shen DK, Liu F, Zhu ZD, Song HD, Xu XR, Wang ZJ, Rong YP, Zeng LC, Wu J, Zhang X, Wang JJ, Xu XN, Wang SY, Fu G, Zhang XL, Wang ZQ, Brindley PJ, McManus DP, Xue CL, Feng Z, Chen Z, Han ZG: **Evolutionary and biomedical implications of a Schistosoma japonicum complementary DNA resource.** *Nat Genet* 2003, **35(2)**:139-147.
- Liu F, Lu J, Hu W, Wang SY, Cui SJ, Chi M, Yan Q, Wang XR, Song HD, Xu XN, Wang JJ, Zhang XL, Zhang X, Wang ZQ, Xue CL, Brindley PJ, McManus DP, Yang PY, Feng Z, Chen Z, Han ZG: **New perspectives on host-parasite interplay by comparative transcriptomic and proteomic analyses of Schistosoma japonicum.** *PLoS Pathog* 2006, **2(4)**:e29.
- Verjovski-Almeida S, DeMarco R, Martins EA, Guimaraes PE, Ojopi EP, Paquola AC, Piazza JP, Nishiyama MY Jr., Kitajima JP, Adamson RE, Ashton PD, Bonaldo MF, Coulson PS, Dillon GP, Farias LP, Gregorio SP, Ho PL, Leite RA, Malaquias LC, Marques RC, Miyasato PA, Nascimento AL, Ohlweiler FP, Reis EM, Ribeiro MA, Sa RG, Stukart GC, Soares MB, Gargioni C, Kawano T, Rodrigues V, Madeira AM, Wilson RA, Menck CF, Setubal JC, Leite LC, Dias-Neto E: **Transcriptome analysis of the acelomate human parasite Schistosoma mansoni.** *Nat Genet* 2003, **35(2)**:148-157.
- Sim S, Park GM, Yong TS: **Cloning and characterization of Clonorchis sinensis myoglobin using immune sera against excretory-secretory antigens.** *Parasitol Res* 2003, **91(4)**:338-343.
- Jones MK, Gobert GN, Zhang L, Sunderland P, McManus DP: **The cytoskeleton and motor proteins of human schistosomes and their roles in surface maintenance and host-parasite interactions.** *Bioessays* 2004, **26(7)**:752-765.
- Matsumoto Y, Perry G, Levine RJ, Blanton R, Mahmoud AA, Aikawa M: **Paramyosin and actin in schistosomal teguments.** *Nature* 1988, **333(6168)**:76-78.
- Zhang LH, McManus DP, Sunderland P, Lu XM, Ye JJ, Loukas A, Jones MK: **The cellular distribution and stage-specific expression of two dynein light chains from the human blood fluke Schistosoma japonicum.** *Int J Biochem Cell Biol* 2005, **37(7)**:1511-1524.
- Tansatit T, Sahaphong S, Riengrojpitak S, Vivanant V, Sobhon P: **Immunolocalization of cytoskeletal components in the tegument of the 3-week-old juvenile and adult Fasciola gigantica.** *Vet Parasitol* 2006, **135(3-4)**:269-278.
- Fitzsimmons CM, McBeath R, Joseph S, Jones FM, Walter K, Hoffmann KF, Kariuki HC, Mwatha JK, Kimani G, Kabatereine NB, Vennervald BJ, Ouma JH, Dunne DW: **Factors affecting human IgE and IgG responses to allergen-like Schistosoma mansoni antigens: Molecular structure and patterns of in vivo exposure.** *Int Arch Allergy Immunol* 2007, **142(1)**:40-50.
- Parkinson J, Blaxter M: **SimiTri--visualizing similarity relationships for groups of sequences.** *Bioinformatics* 2003, **19(3)**:390-395.
- Zayas RM, Hernandez A, Habermann B, Wang Y, Stary JM, Newmark PA: **The planarian Schmidtea mediterranea as a model for epigenetic germ cell specification: analysis of ESTs from the hermaphroditic strain.** *Proc Natl Acad Sci U S A* 2005, **102(51)**:18491-18496.
- Carranza S, Baguna J, Riutort M: **Are the Platyhelminthes a monophyletic primitive group? An assessment using 18S rDNA sequences.** *Mol Biol Evol* 1997, **14(5)**:485-497.
- Le TH, Blair D, McManus DP: **Mitochondrial genomes of parasitic flatworms.** *Trends Parasitol* 2002, **18(5)**:206-213.
- Karling TG: **On the anatomy and affinities of the turbellarian orders.** In *Biology of the turbellaria* Edited by: Riser NW, Morse MP. New York, McGraw Hill; 1974:1-16.
- Hu W, Brindley PJ, McManus DP, Feng Z, Han ZG: **Schistosome transcriptomes: new insights into the parasite and schistosomiasis.** *Trends Mol Med* 2004, **10(5)**:217-225.
- Ranjit N, Jones MK, Stenzel DJ, Gasser RB, Loukas A: **A survey of the intestinal transcriptomes of the hookworms, Necator americanus and Ancylostoma caninum, using tissues isolated by laser microdissection microscopy.** *Int J Parasitol* 2006, **36(6)**:701-710.
- Paliouras M, Diamandis EP: **The kallikrein world: an update on the human tissue kallikreins.** *Biol Chem* 2006, **387(6)**:643-652.
- Brindley PJ, Kalinna BH, Wong JY, Bogtish BJ, King LT, Smyth DJ, Verity CK, Abbenante G, Brinkworth RI, Fairlie DP, Smythe ML, Milburn PJ, Bielefeldt-Ohmann H, Zheng Y, McManus DP: **Proteolysis of human hemoglobin by schistosome cathepsin D.** *Mol Biochem Parasitol* 2001, **112(1)**:103-112.
- Delcroix M, Sajid M, Caffrey CR, Lim KC, Dvorak J, Hsieh I, Bahgat M, Dissous C, McKerrow JH: **A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite.** *J Biol Chem* 2006, **281(51)**:39316-39329.
- Williamson AL, Lecchi P, Turk BE, Choe Y, Hotez PJ, McKerrow JH, Cantley LC, Sajid M, Craik CS, Loukas A: **A multi-enzyme cascade of hemoglobin proteolysis in the intestine of blood-feeding hookworms.** *J Biol Chem* 2004, **279(34)**:35950-35957.
- Sajid M, McKerrow JH, Hansell E, Mathieu MA, Lucas KD, Hsieh I, Greenbaum D, Bogoy M, Salter JP, Lim KC, Franklin C, Kim JH, Caf-

- frey CR: **Functional expression and characterization of Schistosoma mansoni cathepsin B and its trans-activation by an endogenous asparaginyl endopeptidase.** *Mol Biochem Parasitol* 2003, **131(1)**:65-75.
32. Loukas A, Tran MH, Pearson MS: **Schistosome membrane proteins as vaccines.** *Int J Parasitol* 2007, **37**:257-263.
 33. Tran MH, Pearson MS, Bethony JM, Smyth DJ, Jones MK, Duke M, Don TA, McManus DP, Correa-Oliveira R, Loukas A: **Tetraspanins on the surface of Schistosoma mansoni are protective antigens against schistosomiasis.** *Nat Med* 2006, **12(7)**:835-840.
 34. Levy S, Shoham T: **The tetraspanin web modulates immune-signalling complexes.** *Nat Rev Immunol* 2005, **5(2)**:136-148.
 35. Agre P: **The aquaporin water channels.** *Proc Am Thorac Soc* 2006, **3(1)**:5-13.
 36. Simmons DL: **What makes a good anti-inflammatory drug target?** *Drug Discov Today* 2006, **11(5-6)**:210-219.
 37. Amiri P, Locksley RM, Parslow TG, Sadick M, Rector E, Ritter D, McKerrow JH: **Tumour necrosis factor alpha restores granulomas and induces parasite egg-laying in schistosomium-infected SCID mice.** *Nature* 1992, **356(6370)**:604-607.
 38. Blank RB, Lamb EV, Tocheva AS, Crow ET, Lim KC, McKerrow JH, Davies SJ: **The common gamma chain cytokines interleukin (IL)-2 and IL-7 indirectly modulate blood fluke development via effects on CD4+ T cells.** *J Infect Dis* 2006, **194(11)**:1609-1616.
 39. Davies SJ, Grogan JL, Blank RB, Lim KC, Locksley RM, McKerrow JH: **Modulation of blood fluke development in the liver by hepatic CD4+ lymphocytes.** *Science* 2001, **294(5545)**:1358-1361.
 40. Beall MJ, Pearce EJ: **Human transforming growth factor-beta activates a receptor serine/threonine kinase from the intravascular parasite Schistosoma mansoni.** *J Biol Chem* 2001, **276(34)**:31613-31619.
 41. Osman A, Niles EG, Verjovski-Almeida S, LoVerde PT: **Schistosoma mansoni TGF-beta receptor II: role in host ligand-induced regulation of a schistosoma target gene.** *PLoS Pathog* 2006, **2(6)**:e54.
 42. Keah HH, Hearn MT: **A molecular recognition paradigm: promiscuity associated with the ligand-receptor interactions of the activin members of the TGF-beta superfamily.** *J Mol Recognit* 2005, **18(5)**:385-403.
 43. Zavala-Gongora R, Kroner A, Bernthaler P, Knaus P, Brehm K: **A member of the transforming growth factor-beta receptor family from Echinococcus multilocularis is activated by human bone morphogenetic protein 2.** *Mol Biochem Parasitol* 2006, **146(2)**:265-271.
 44. Davies SJ, Shoemaker CB, Pearce EJ: **A divergent member of the transforming growth factor beta receptor family from Schistosoma mansoni is expressed on the parasite surface membrane.** *J Biol Chem* 1998, **273(18)**:11234-11240.
 45. Wynn TA: **Fibrotic disease and the T(H)1/T(H)2 paradigm.** *Nat Rev Immunol* 2004, **4(8)**:583-594.
 46. Clemens LE, Basch PF: **Schistosoma mansoni: effect of transferin and growth factors on development of schistosomula in vitro.** *J Parasitol* 1989, **75(3)**:417-421.
 47. Wajih N, Sane DC, Hutson SM, Wallin R: **The inhibitory effect of calumenin on the vitamin K-dependent gamma-carboxylation system. Characterization of the system in normal and warfarin-resistant rats.** *J Biol Chem* 2004, **279(24)**:25276-25283.
 48. Ostergaard M, Hansen GA, Vorum H, Honore B: **Proteomic profiling of fibroblasts reveals a modulating effect of extracellular calumenin on the organization of the actin cytoskeleton.** *Proteomics* 2006, **6(12)**:3509-3519.
 49. Sirica AE: **Cholangiocarcinoma: molecular targeting strategies for chemoprevention and therapy.** *Hepatology* 2005, **41(1)**:5-15.
 50. He Z, Bateman A: **Progranulin (granulin-epithelin precursor, PC-cell-derived growth factor, acrogranin) mediates tissue repair and tumorigenesis.** *J Mol Med* 2003, **81(10)**:600-612.
 51. He Z, Ong CH, Halper J, Bateman A: **Progranulin is a mediator of the wound response.** *Nat Med* 2003, **9(2)**:225-229.
 52. Whitbread AK, Veveris-Lowe TL, Lawrence MG, Nicol DL, Clements JA: **The role of kallikrein-related peptidases in prostate cancer: potential involvement in an epithelial to mesenchymal transition.** *Biol Chem* 2006, **387(6)**:707-714.
 53. Dong Q, Patel M, Scott KF, Graham GG, Russell PJ, Sved P: **Oncogenic action of phospholipase A2 in prostate cancer.** *Cancer Lett* 2006, **240(1)**:9-16.
 54. Endo K, Yoon BI, Pairojkul C, Demetris AJ, Sirica AE: **ERBB-2 over-expression and cyclooxygenase-2 up-regulation in human cholangiocarcinoma and risk conditions.** *Hepatology* 2002, **36(2)**:439-450.
 55. Sirica AE, Lai GH, Endo K, Zhang Z, Yoon BI: **Cyclooxygenase-2 and ERBB-2 in cholangiocarcinoma: potential therapeutic targets.** *Semin Liver Dis* 2002, **22(3)**:303-313.
 56. Salter JP, Choe Y, Albrecht H, Franklin C, Lim KC, Craik CS, McKerrow JH: **Cercarial elastase is encoded by a functionally conserved gene family across multiple species of schistosomes.** *J Biol Chem* 2002, **277(27)**:24618-24624.
 57. Saffer LD, Schwartzman JD: **A soluble phospholipase of Toxoplasma gondii associated with host cell penetration.** *J Protozool* 1991, **38(5)**:454-460.
 58. Pacifico LG, Fonseca CT, Chiari L, Oliveira SC: **Immunization with Schistosoma mansoni 22.6 kDa antigen induces partial protection against experimental infection in a recombinant protein form but not as DNA vaccine.** *Immunobiology* 2006, **211(1-2)**:97-104.
 59. Loukas A, Bethony JM, Mendez S, Fujiwara RT, Goud GN, Ranjit N, Zhan B, Jones K, Bottazzi ME, Hotez PJ: **Vaccination with recombinant aspartic hemoglobinase reduces parasite load and blood loss after hookworm infection in dogs.** *PLoS Med* 2005, **2(10)**:e295.
 60. Verity CK, McManus DP, Brindley PJ: **Vaccine efficacy of recombinant cathepsin D aspartic protease from Schistosoma japonicum.** *Parasite Immunol* 2001, **23(3)**:153-162.
 61. Redmond DL, Knox DP: **Further protection studies using recombinant forms of Haemonchus contortus cysteine proteinases.** *Parasite Immunol* 2006, **28(5)**:213-219.
 62. Loukas A, Bethony JM, Williamson AL, Goud GN, Mendez S, Zhan B, Hawdon JM, Elena Bottazzi M, Brindley PJ, Hotez PJ: **Vaccination of dogs with a recombinant cysteine protease from the intestine of canine hookworms diminishes the fecundity and growth of worms.** *J Infect Dis* 2004, **189(10)**:1952-1961.
 63. Dalton JP, McGonigle S, Rolph TP, Andrews SJ: **Induction of protective immunity in cattle against infection with Fasciola hepatica by vaccination with cathepsin L proteinases and with hemoglobin.** *Infect Immun* 1996, **64(12)**:5066-5074.
 64. Hota-Mitchell S, Siddiqui AA, Dekaban GA, Smith J, Tognon C, Podesta RB: **Protection against Schistosoma mansoni infection with a recombinant baculovirus-expressed subunit of calpain.** *Vaccine* 1997, **15(15)**:1631-1640.
 65. Espino AM, Hillyer GV: **A novel Fasciola hepatica saposinlike recombinant protein with immunoprophylactic potential.** *J Parasitol* 2004, **90(4)**:876-879.
 66. Almeida MS, Torloni H, Lee-Ho P, Vilar MM, Thaumaturgo N, Simpson AJ, Tendler M: **Vaccination against Fasciola hepatica infection using a Schistosoma mansoni defined recombinant antigen, Sm14.** *Parasite Immunol* 2003, **25(3)**:135-137.
 67. Tendler M, Brito CA, Vilar MM, Serra-Freire N, Diogo CM, Almeida MS, Delbem AC, Da Silva JF, Savino VV, Garratt RC, Katz N, Simpson AS: **A Schistosoma mansoni fatty acid-binding protein, Sm14, is the potential basis of a dual-purpose anti-helminth vaccine.** *Proc Natl Acad Sci U S A* 1996, **93(1)**:269-273.
 68. Parkin DM: **The global health burden of infection-associated cancers in the year 2002.** *Int J Cancer* 2006, **118(12)**:3030-3044.
 69. Sithithaworn P, Pipitgool V, Srisawangwong T, Elkins DB, Haswell-Elkins MR: **Seasonal variation of Opisthorchis viverrini infection in cyprinoid fish in north-east Thailand: implications for parasite control and food safety.** *Bull World Health Organ* 1997, **75(2)**:125-131.
 70. Perteau G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, Quackenbush J: **TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets.** *Bioinformatics* 2003, **19(5)**:651-652.
 71. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25(1)**:25-29.

72. National Center for Biotechnology Information. .
73. Burge C, Karlin S: **Prediction of complete gene structures in human genomic DNA.** *J Mol Biol* 1997, **268(1)**:78-94.
74. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340(4)**:783-795.
75. Hofmann K, Stoffel W: **TMbase - A database of membrane spanning proteins segments.** *Biol Chem Hoppe-Seyler* 1993, **374**:166.
76. **MEROPS the peptidase database.** .
77. **NetPhos 2.0 server.** .
78. Swofford DL: **PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.** Sunderland, MA , Sinauer Associates; 1998.
79. Capron A, Riveau G, Capron M, Trottein F: **Schistosomes: the road from host-parasite interactions to vaccines in clinical trials.** *Trends Parasitol* 2005, **21(3)**:143-149.
80. Zhan B, Liu S, Perally S, Xue J, Fujiwara R, Brophy P, Xiao S, Liu Y, Feng J, Williamson A, Wang Y, Bueno LL, Mendez S, Goud G, Bethony JM, Hawdon JM, Loukas A, Jones K, Hotez PJ: **Biochemical characterization and vaccine potential of a heme-binding glutathione transferase from the adult hookworm *Ancylostoma caninum*.** *Infect Immun* 2005, **73(10)**:6903-6911.
81. Da'dara AA, Skelly PJ, Wang MM, Harn DA: **Immunization with plasmid DNA encoding the integral membrane protein, Sm23, elicits a protective immune response against schistosome infection in mice.** *Vaccine* 2001, **20(3-4)**:359-369.
82. Bethony J, Loukas A, Smout M, Brooker S, Mendez S, Plieskatt J, Goud G, Bottazzi ME, Zhan B, Wang Y, Williamson A, Lustigman S, Correa-Oliveira R, Xiao S, Hotez PJ: **Antibodies against a secreted protein from hookworm larvae reduce the intensity of hookworm infection in humans and vaccinated laboratory animals.** *Faseb J* 2005, **19(12)**:1743-1745.
83. MacDonald AJ, Tawe W, Leon O, Cao L, Liu J, Oksov Y, Abraham D, Lustigman S: **Ov-ASP-1, the *Onchocerca volvulus* homologue of the activation associated secreted protein family is immunostimulatory and can induce protective anti-larval immunity.** *Parasite Immunol* 2004, **26(1)**:53-62.
84. Schechtman D, Tarrab-Hazdai R, Arnon R: **The I4-3-3 protein as a vaccine candidate against schistosomiasis.** *Parasite Immunol* 2001, **23(4)**:213-217.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

