Viewpoints

Coming out of the Shell: Building the Molecular Infrastructure for Research on Parasite-Harbouring Snails

Cinzia Cantacessi[1]*, Sattrachai Prasopdee[1,2], Javier Sotillo[1], Jason Mulvenna[1], Smarn Tesana[2], Alex Loukas[1]

1 Centre for Biodiscovery and Molecular Development of Therapeutics, James Cook University, Smithfield, Queensland, Australia, 2 Department of Parasitology, Khon Kaen University, Khon Kaen, Thailand

In Thailand and Laos alone, approximately 10 million people are infected with the liver fluke *Opisthorchis viverrini* [1]. Chronic infection with this parasite is considered the leading cause of cholangiocarcinoma (CCA, or bile-duct cancer) in large areas of Southeast Asia [2]. In these regions, CCA caused by *O. viverrini* is typically diagnosed 30–40 years after infection, with death occurring within 3–6 months post diagnosis [3]. *O. viverrini* is characterised by a three-host life cycle, with prosobranch snails of the genus *Bithynia* and cyprinid fishes acting as first and second intermediate hosts, respectively, while piscivorous mammals, including dogs, cats, and humans, act as definitive hosts [2]. Over the last two decades, much attention has been paid to studies on the epidemiology, developmental biology, and diagnosis of *O. viverrini* [4], while recent biotechnological advances are contributing large-scale explorations of the fundamental molecular biology of this liver fluke, with a view toward identifying key molecules essential for its development, reproduction, and survival, as well as dissecting the molecular pathways leading to the development of CCA [5–8]. These advances provide a solid foundation for the development of novel strategies to fight this devastating disease. However, long-term control of *O. viverrini*–induced cancer strictly relies on the development of integrated approaches, targeting the parasite as well as its intermediate hosts.

Recently, an article in *PLOS Neglected Tropical Diseases* by Adema and colleagues [9] served to highlight the substantial gap in knowledge of aspects of the fundamental molecular biology of molluscs harbouring parasites, as well as the extraordinary opportunities that modern research toolkits, including microarray platforms, RNA interference, and high-throughput sequencing, offer for investigations of snail-parasite interactions [9]. Indeed, despite the massive expansion in the demand for and access to low-cost, high-throughput sequencing, large-scale genomic analyses of snails are limited to the draft genome sequence of the pulmonate snail intermediate host of

Schistosoma blood flukes (*Biomphalaria glabrata* Genome Initiative at http://biology.unm.edu/biomphalaria-genome/index.html). Until now, there has been no genomic or transcriptomic information available for the prosobranch snail intermediate hosts of carcinogenic liver flukes, such as *O. viverrini* and *Clonorchis sinensis*. In order to provide the research community with a solid resource for molecular studies of these organisms, we generated the first reference transcriptome of *Bithynia siamensis goniomphalos* (Gastropoda, Bithyniidae), the intermediate host of *O. viverrini* in areas of northeast Thailand, Lao PDR, Cambodia, and South Vietnam, where the incidence of CCA is highest [cf. 10] (cf. Figure 1). A cDNA library from adult snails [11] was constructed, sequenced using RNA-seq (Illumina), and annotated using an established bioinformatic workflow [12]. Briefly, snails were collected from a natural body of water in Muang District, Khon Kaen Province, northeast Thailand (Figure 1); the taxonomic identity of the specimens was confirmed based on characteristic morphological features of the shells [13]; RNA was extracted from whole adult, parasite-free *B. siamensis goniomphalos* (n = 5), reverse-transcribed to cDNA, adaptor-ligated, and paired-end sequenced on a Genome Analyzer II (Illumina). Almost 50 million high-quality (Phred score >28) reads were generated; the assembly, produced using the Trinity software [14], yielded 167,029 contigs >200 bp in length, with a GC content of 44.4% (Table 1). Using sequence

homology–based searches [12], approximately 40% of the assembled contigs could be annotated (cf. Table 1). A total of 32,026 contigs could be annotated with Gene Ontology terms (*via* Blast2GO; [15]), according to the categories "biological process," "cellular component," and "molecular function." Approximately 77,000 non-overlapping protein sequences (of which ~15,000 were full-length) could be inferred from the transcriptome of *B. siamensis goniomphalos via* BLASTx alignments with protein sequences available in public databases. Of the *B. siamensis goniomphalos* transcripts encoding proteins that could be mapped to orthologues in the Clusters of Orthologous Groups (COGs) database [16] (Table 1), the largest set was assigned to "general function prediction only" (16%), followed by "translation, ribosomal structure and biogenesis" (10%), "replication, recombination and repair" (7%), and "transcription" (7%) (Table 1). Approximately 54,000 protein-coding transcripts had orthologues in one of the 29 known biological pathways in the KEGG database (Table 1), including "regulation of actin cytoskeleton" (4%), "focal adhesion" (4%), and "spliceosome" (3.9%) (Table 1). Among the major protein classes included in the *B. siamensis goniomphalos* transcriptome were peptidases (n = 1,944; 4%), kinases (n = 4,757; 9%), phosphatases (n = 2,151; 4%), GTPases (n = 2,533; 5%), receptors (n = 6,243; 11%), transcription factors (n = 1,478; 3%), and channels and transporters (n = 1,287; 2%).

O. viverrini prevalence (%)

0	5	15	25	>30

0.1   10   20   30

CCA prevalence (%)

1.45   7.61   11.47   16.2   44.31

6.26   9.59   13.67   20.66   61.39

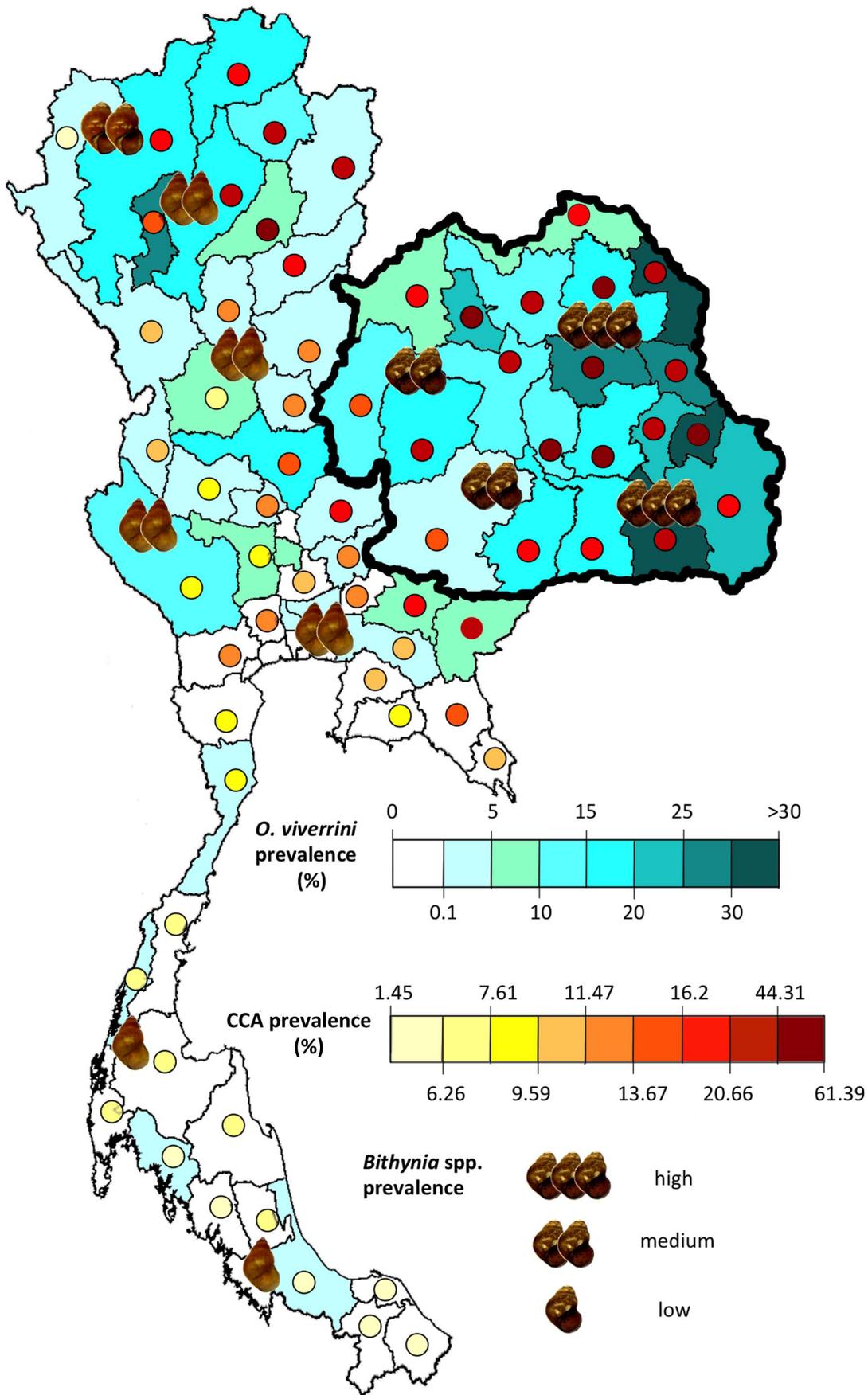***Bithynia* spp. prevalence**

high

medium

Figure 1. Distribution and prevalence of *Opisthorchis viverrini*, cholangiocarcinoma, and *Bithynia* spp. snails in Thailand. At present, three species of *Bithynia* act as intermediate hosts for *O. viverrini*, and their distribution differs all over Thailand. *B. siamensis goniomphalos* is located in northeast Thailand (thick border), where the prevalence of *O. viverrini*–associated cholangiocarcinoma (CCA) is highest; *B. funiculata* specifically distributes in north Thailand; and *B. siamensis siamensis* is mainly found in central and north Thailand, and rarely in the south [2,18–20].
doi:10.1371/journal.pntd.0002284.g001

Table 1. Summary of the RNA-seq data for *Bythinia siamensis goniomphalos* prior to and following assembly, and bioinformatics annotation and analyses.

Raw reads (paired-end)	49,952,248
Total contigs (average length ± SD)	167,029 (710±677)
GC content (%)	44.4
Number of BLASTn hits* (%)	47,220 (46)
Number of BLASTx hits* (%)	77,714 (77)
Containing an open reading frame (%)	76,963
Returning a COG result (%)	54,513 (71)
KEGG pathway result (%)	54,291 (70.5)
Gene Ontology results (%)	32,026 (41.6)

*e-value cut-off: 1e-05.
doi:10.1371/journal.pntd.0002284.t001

Annotation information linked to each transcript characterised in the present study, including top BLASTx hit, GO classification, KEGG pathway mapping, and COG orthologues, is available from Table S1. In the absence of a reference genome for *Bithynia* spp., the annotation of the sequence data analysed herein was based on comparison with data available in public databases. The relatively small proportion of *B. siamensis goniomphalos* annotated protein sequences (~60%) is likely to reflect the paucity of genomic sequence information available for prosobranch molluscan species in these databases [17]; supported by the availability of this dataset, as well as of the draft genome sequence of *B. glabrata*, future sequencing efforts will provide the depth of coverage required for the determination of the genome of *Bithynia* spp., which, in turn, will pave the way for comparative genomic studies of prosobranch and pulmonate snails harbouring parasite infections. This knowledge will be pivotal to improve our understanding of the biology of snail-borne parasites, and their "choice" of distinct snail species as their intermediate hosts.

Both raw and assembled sequence data generated in the present study are freely available from the Sequence Read Archive (SRA) and the Transcriptome Shotgun Assembly Sequence Database at NCBI (http://www.ncbi.nlm.nih.gov) under accession numbers SRR768418 and GAGS0000000, respectively. To our knowledge, this sequence data represents the first large-scale transcriptomic resource for a prosobranch mollusc intermediate host of a human platyhelminth and represents a major contribution to future fundamental explorations of the developmental biology of *O. viverrini* in *Bithynia* spp. snails, as well as the molecular interactions occurring at the snail-parasite interface. We suggest that the extensive sequence data generated in this study will be of great value in future studies aimed at exploring the changes in gene transcription occurring in *Bithynia* spp. upon infection with *O. viverrini* eggs and at different time points following infection. Besides yielding a general picture of the modified biology of trematode-infected snails, this data will set a basis for the identification of key genes, gene products, biological pathways, and/or bacterial and viral symbionts [cf. 9] involved in the cascade of molecular events leading to the development of the parasite through the stages of miracidium, sporocyst, redia, and cercaria. In turn, this advance will ultimately result in the development of novel targeted strategies to control snail-borne NTDs.

New molecular technologies have a tremendous potential to aid our efforts in controlling snail-borne diseases. However, despite the unanimous acknowledgment that an increased understanding of the fundamental molecular biology of parasite-harbouring snails will provide us with a range of new tools to aid current efforts aimed at controlling snail-borne infections, the progressive but steady decline of malacology expertise and funding for snail vector–related research [cf. 9] poses a serious obstacle to the application of such technologies to snail-oriented NTD control. With this first step toward the establishment of a reference database for genetic research of prosobranch snail vectors of parasitic helminths, we hope to stimulate integrated, interdisciplinary research across malacology, helminthology, genomics, and bioinformatics, in the bid to fight snail-borne infections.

Supporting Information

Table S1 The annotated transcriptome of *Bithynia siamensis goniomphalos*. Summary of annotation data linked to individual assembled transcripts. Abbreviations: reads per kilobase per million reads (RPKM); non-redundant protein database (Nr); non-redundant nucleotide database (Nt); SwissProt database (SwissProt); Clusters of Orthologous Groups of Proteins (COG); Kyoto Encyclopedia of Genes and Genomes (KEGG); Gene Ontology (GO).
(ZIP)

References

1. Sithithaworn P, Haswell-Elkins M (2003) Epidemiology of *Opisthorchis viverrini*. Acta Trop 88: 187–194.
2. Sripa B, Brindley PJ, Mulvenna J, Laha T, Smout MJ, et al. (2012) The tumorigenic liver fluke *Opisthorchis viverrini* – multiple pathways to cancer. Trends Parasitol 28: 395–407.
3. Sripa B, Kaewkes S, Sithithaworn P, Mairiang E, Laha T, et al. (2007) Liver fluke induces cholangiocarcinoma. PLoS Med 4: e201. doi:10.1371/journal.pmed.0040201.
4. Andrews RH, Sithithaworn P, Petney TN (2008) *Opisthorchis viverrini*: an underestimated parasite in world health. Trends Parasitol 24: 497–501.
5. Smout MJ, Laha T, Mulvenna J, Sripa B, Suttiprapa S, et al. (2009) A granulin-like growth factor secreted by the carcinogenic liver fluke, *Opisthorchis viverrini*, promotes proliferation of host cells. PLoS Pathog 5: e1000611. doi:10.1371/journal.ppat.1000611.
6. Mulvenna J, Sripa B, Brindley PJ, Gorman J, Jones MK, et al. (2010) The secreted and surface proteomes of the adult stage of the carcinogenic human liver fluke *Opisthorchis viverrini*. Proteomics 10: 1063–1078.
7. Young ND, Campbell BE, Hall RS, Jex AR, Cantacessi C, et al. (2010) Unlocking the tran-

scriptomes of two carcinogenic parasites, *Clonorchis sinensis* and *Opisthorchis viverrini*. PLoS Negl Trop Dis 4: e719. doi:10.1371/journal.pntd.0000719.

8. Jex AR, Young ND, Sripa J, Hall RS, Scheerlinck J-P, et al. (2012) Molecular changes in *Opisthorchis viverrini* (Southeast Asian liver fluke) during the transition from the juvenile to the adult stage. PLoS Negl Trop Dis 6: e1916. doi:10.1371/journal.pntd.0001916.

9. Adema CM, Bayne CJ, Bridger JM, Knight M, Loker ES, et al. (2012) Will all scientists working on snails and the diseases they transmit please stand up? PLoS Negl Trop Dis 6: e1835. doi:10.1371/journal.pntd.0001835.

10. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53–59.

11. Tesana S, Thapsripair P, Thammasiri C, Prasopdee S, Suwannatrai A, et al. (2012) Effects of Bayluscidae on *Bithyinia siamensis goniomphalos*, the first intermediate host of the human liver fluke,

Opisthorchis viverrini, in laboratory and field trials. Parasitol Int 61: 52–55.

12. Cantacessi C, Jex AR, Hall RS, Young ND, Campbell BE, et al. (2010) A practical, bioinformatics workflow system for large datasets generated by next-generation sequencing. Nucleic Acids Res 38: e171.

13. Chitramvong YP (1992) The Bithyniidae (Gastropoda: Prosobranchia) of Thailand: comparative external morphology. Malacol Rev 25: 21–38.

14. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29: 644–652.

15. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH et al. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36: 3420–3435.

16. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, et al. (2003) The COG database:

an updated version includes eukaryotes. BMC Bioinformatics 4: 41.

17. Sadamoto H, Takahashi H, Okada T, Kenmoku H, Toyota M, et al. (2012) *De novo* sequencing and transcriptome analysis of the central nervous system of mollusc *Lymnaea stagnalis* by deep RNA sequencing. PLoS ONE 7: e42546. doi:10.1371/journal.pone.0042546.

18. Sripa B, Bethony JM, Sithithaworn P, Kaewkes S, Mairiang E, et al. (2011) Opisthorchiasis and *Opisthorchis*-associated cholangiocarcinoma in Thailand and Laos. Acta Trop 120: S158–S168.

19. Suwannatrai A, Suwannatrai K, Haruay S, Piratae S, Thammasiri C, et al. (2011) Effect of soil surface salt on the density and distribution of the snail *Bithynia siamensis goniomphalos* in northeast Thailand. Geospat Health 5:183–190.

20. Petney T, Sithithaworn P, Andrews R, Kiatsopit N, Tesana S, et al. (2012) The ecology of the *Bithynia* first intermediate hosts of *Opisthorchis viverrini*. Parasitol Int 61: 38–45.