ISSN 0026-8933, Molecular Biology, 2017, Vol. 51, No. 2, pp. 184–193. © Pleiades Publishing, Inc., 2017.
Original Russian Text © G.N. Chelomina, 2017, published in Molekulyarnaya Biologiya, 2017, Vol. 51, No. 2, pp. 215–226.

REVIEWS

UDC 575.11;576.89

# Genomics and Transcriptomics of the Chinese Liver Fluke *Clonorchis sinensis* (Opisthorchiidae, Trematoda)

**G. N. Chelomina**[a, b]

[a] *Institute of Biology and Soil Science, Far East Branch, Russian Academy of Sciences, Vladivostok, 690022 Russia*
[b] *School of Natural Sciences, Far Eastern Federal University, Vladivostok, 690012 Russia*
*e-mail: chelomina@ibss.dvo.ru*

Received April 5, 2016; in final form, April 19, 2016

**Abstract**—The review summarizes the results of first genomic and transcriptomic investigations of the liver fluke *Clonorchis sinensis* (Opisthorchiidae, Trematoda). The studies mark the dawn of the genomic era for opisthorchiids, which cause severe hepatobiliary diseases in humans and animals. Their results aided in understanding the molecular mechanisms of adaptation to parasitism, parasite survival in mammalian biliary tracts, and genome dynamics in the individual development and the development of parasite–host relationships. Special attention is paid to the achievements in studying the codon usage bias and the roles of mobile genetic elements (MGEs) and small interfering RNAs (siRNAs). Interspecific comparisons at the genomic and transcriptomic levels revealed molecular differences, which may contribute to understanding the specialized niches and physiological needs of the respective species. The studies in *C. sinensis* provide a basis for further basic and applied research in liver flukes and, in particular, the development of efficient means to prevent, diagnose, and treat clonorchiasis.

*Keywords*: genomics, transcriptomics, trematode, *Clonorchis sinensis*, liver fluke, retrotransposons, codon usage, microRNA, mobile elements

**DOI:** 10.1134/S0026893317020078

## INTRODUCTION

What is known as "omics" (genomics, transcriptomics, proteomics, etc.) is a priority field of modern molecular biology especially in applying to dangerous human parasites; this requires the use of new technologies and modern bioinformatics methods. Data from such studies are of general biological importance, providing a means to better understand the molecular basis of the advent and evolution of life and being of applied significance as well. The data shed further light on the biology of parasites and the specifics of their molecular evolution. Genomic and transcriptomic data are essential for understanding the molecular mechanisms of the fine regulation of parasite relationships with intermediate and definitive hosts, studying the molecular basis of the pathogenesis in parasitic invasions, and developing new means to diagnose, treat, and prevent helminthiases [1, 2]. Platyhelminthes species are currently used to study expression and splicing of genes involved in vital physiological processes, to design new and more effective anthelmintic drugs, and to understand the mechanisms of tissue regeneration [3–6].

Liver flukes of the family Opisthorchiidae are classified by the WHO as the most dangerous (group 1) biological carcinogens and by the International Agency for Research in Cancer as risk factors for of hepatobiliary cancer in humans [7, 8]. However, the genomes and transcriptomes of these carcinogenic liver flukes remained unstudied until recently. The review summarizes the data from new genomic and transcriptomic studies of the liver fluke *Clonorchis sinensis*, which mark the dawn of the genomic era for opisthorchiids. The data are compared with those available for other species, primarily liver and blood trematodes. Omics studies are of high theoretical and applied significance, providing for rapid progress in modern trematode research. Their results are promising for further developments in new and existing research fields.

## GENOMIC STUDIES

### *Organization of the Liver Fluke Clonorchis sinensis Genome*

Whole-genome sequencing was performed for ten species from various families of the flatworms Platyhelminthes, including the blood flukes *Schistosoma japonicum* (Genome Sequences and Functional Analysis Consortium, 2009) and *S. mansoni* [9]; the freshwater planarian *Schmidtea mediterranea* Benazzi et al., 1975 [10]; the fox tapeworm *Echinococcus multilocularis* Leuckart, 1863; the hydatid tapeworm *E. granulosus* Batsch, 1786 [11]; the pork tapeworm *Taenia solium*

Linnaeus, 1758; the rodent tapeworm *Hymenolepis microstoma* Dujardin, 1845 [12]; the monogenean ectoparasite *Gyrodactylus salaris* Malmberg, 1957 [13]; the liver flukes *C. sinensis* [14] and *Opisthorchis viverrini* [15].

First data on genome sequencing in *C. sinensis* were published only a few years ago [14]. The total *C. sinensis* genome is 516 Mb in size [14], smaller than the genome of its phylogenetically close relative *O. viverrini* (634.5 Mb) [15]. Of all trematodes, the common liver fluke *Fasciola hepatica* has the largest genome (1.3 Gb), but this is not a result of an increase in chromosome number, genome duplication, or repeat expansion. A major part of noncoding DNA in the *F. hepatica* genome is presumably involved in regulating gene expression [16]. The blood fluke genomes are far smaller, approximately 380–400 Mb [17], and cestodes have the smallest genomes (only 115–151 Mb) among all flatworms [11, 12]. The biological specifics of various species of Platyhelminthes make it possible to assume than an increase in genome size is associated with a more complex life cycle. For instance, a three-host life cycle is characteristic of the family Opisthorchiidae, while a two-host cycle is known for the family *Shistosoma* (i.e., with two or one intermediate host, respectively). In the case of nematodes, the genome size is possibly associated with the effective sizes of parasite and host populations [18].

The GC content of *C. sinensis* genomic DNA is 43.85% on average, higher than in blood trematodes of the genus *Schistosoma* (36% on average) and similar to that in *O. viverrini* (43.7%) [15]. The proportion of repeats in the *C. sinensis* genome (32.25%) [19] is comparable with the proportions reported for *O. viverrini* (30.6%) [15] and *F. hepatica* (32%) [16] and is lower than in the *Schistosoma* species (40.1–47.5%) [15, 19]. Therefore, repeats are AT rich in liver and blood flukes.

In total, 16268 protein-coding genes were found in the liver fluke genome [14]. The set includes genes involved in glycolysis, the tricarboxylic acid cycle, and fatty acid metabolism, but lacks the key genes of fatty acid biosynthesis. This circumstance is probably related to the parasitic lifestyle because liver trematodes take lipids from the host bile. There are genes for various proteases, kinases, phosphatases, tegument proteins, and excretory/secretory products, as well as host binding proteins and receptors, which play an important role in parasite life. Some of them are recommended as targets for designing vaccines, drugs, and diagnostic preparations. Genes for proteins involved in carcinogenesis are of particular interest. The set includes the genes for granulin, thioredoxin peroxidase, fatty acid-binding proteins, and phospholipase A2, which all belong to a group of cholangiocarcinoma-associated proteins. Key sex determination genes were not identified in *C. sinensis*, which is a hermaphrodite. At total of 53 genes related to sex determination, sexual differentiation, and reproduction were identified as a result of genome sequencing.

Interestingly, the *SOX6* and *DMRT1* genes were found, and their counterparts are known to act as sex-determining genes in vertebrates. Apart from protein-coding genes, there are genes for ribosomal (7), transport (235), small nuclear (169), small nucleolar (509), and micro (858) RNAs [14]. The genome of *O. viverrini*, another liver fluke, was predicted to include the same number (16379) of protein-coding genes, of which 3015 are unique to the species [15]. Similar numbers were observed for ribosomal (6), transport (189), and small nuclear (229) RNAs, while the microRNA (miRNA) genes are far fewer (178) than in *C. sinensis* [15]. The difference possibly reflects the fact that the two genomes have still not been studied comprehensively.

The number of gene families is similar among different trematode species (6910 in *C. sinensis*, 8898 in *S. japonicum*, and 7317 in *S. mansoni*) and is comparable with that in *Drosophila melanogaster* (7640) or *Homo sapiens* (8841) [14]. The mean gene size (11.548 kb), mean number of exons per gene (5.9), mean exon length (2.007 kb), and gene density (the portion of protein-coding sequences is 4.14%) estimated for the 16258 *C. sinensis* protein-coding genes are similar to those in *O. viverrini* [16], *S. mansoni*, and *S. japonicum* [14]. At the same time, the liver fluke genomes have longer introns (up to 2.8–3.5 kb) as compared with the *Schistosome* genomes; the intron size is 1.2- to 1.5-fold greater in the former [15]. The mean sizes of introns and exons tend to increase with the increasing genome size. Their estimates for the *F. hepatica* genome, which is the largest in liver trematodes, are 3.7 kb (from 33 bp to 17.5 kb) and 303 bp (from 36 bp to 1.369 kb), respectively [16].

A comparative analysis of the trematode genomes [14, 15] showed that structural genome variation is high in Opisthorchiidae. There is limited genome synteny not only between Opisthorchiidae and *Schistosoma*, but even between the closely related liver fluke species *C. sinensis* and *O. viverrini*. The genome synteny levels correlate with the similarity of amino acid composition in predicted proteins. Proteins of the two species have a greater similarity with each other (80.3%) than with *Schistosoma* (less than 60%) and especially tapeworm (less than 50%) proteins [15]. According to a functional classification of genes and protein domains [14], approximately 60% of all protein domains are common for *C. sinensis* and other taxa, being probably characteristic of all multicellular organisms. The majority (71%) of the domains that were not found in *C. sinensis* are also absent from the *Schistosoma* species, and only a minor portion of the domains found in *C. sinensis* were not detected in *Schistosoma*, suggesting a greater loss of domains for *C. sinensis* [14].

Genome comparisons in parasitic flatworms [13, 15] revealed genome features common for the ectoparasite and endoparasite lineages. The features include, for instance, a substantial reduction of the core gene set (including homeodomain-containing genes) in bilateral animals and loss of *piwi* and *vasa*, which code for regula-

tory proteins and are thought to be of importance for animal development. It was concluded that functional fatty acid biosynthesis pathways were generally lost and that peroxisomes were lacking, although they are characteristic of all eukaryotes other than parasitic protozoa. Additional evidence was obtained for the facts that opisthochriids are adapted to live in bile ducts and secrete proteins to modulate host cell proliferation. In total, 5160 conserved orthologs were shared among the flatworm species under study, and 2037 orthologs found in the opisthorchiids have diverged with respect to the blood flukes and tapeworms [15].

### *Codon Usage*

The frequencies of synonymous codons are known to vary in all forms of life, the phenomenon being known as codon usage bias. Such differences are now considered as an important factor of gene expression and cell functions, such as transcription factor binding, protein folding, ubiquitin modification, mRNA splicing, a regulation of tissue-specific gene products, etc. [20–26]. A mutation hypothesis and a natural selection hypothesis were advanced to explain codon usage bias. The former assumes that mutations are neutral and is often used to explain the interspecific variation in codon usage. The natural selection hypothesis assumes that synonymous substitutions are capable of affecting the viability of organisms and is usually used to explain codon usage bias in genomes or genes [25].

Codon usage biases may be due to various factors, including the gene expression level, RNA stability, adaptation to growth conditions, etc. [24]. The nucleotide composition is the strongest determinant of codon usage biases in interspecific comparisons. For instance, in the case of the AT-rich *S. mansoni* genome, an analysis of principal components reveals a main trend in codon usage, which is highly correlated with GC content in the third codon position [27]. In general, preferential usage of a particular codon may reflect a balance between mutational bias and natural selection for optimized translation; in other words, optimized codons should theoretically provide for faster rates and higher accuracy of translation [24, 28].

Several regularities were observed for codon usage. For instance, it is clear that the gene expression level correlates positively with the extent of variation in codon usage and negatively with the synonymous substitution rate between divergent species [25, 29]. Genes with maximum expression in the periods of rapid growth and intense protein synthesis show greater codon usage bias when compared with those of slower growth periods [29]. In addition, codon usage bias strongly affects heterologous gene expression, and the use of certain codons may increase the expression level by more than three orders of magnitude [25]. Therefore, codon optimization is often performed to design high-expression constructs suitable for gene therapy or gene vaccines. However, it should be noted that codon optimization can affect protein conformation and function, increase immunogenicity, and thus reduce the treatment efficacy [26].

A total of 12515 codons from 38 coding sequences were examined in a study of codon usage in *C. sinensis* [27]. Deviations in the total GC content and GC content in the third codon position were not detected. Theoretically, the species may therefore have rather different codon usage strategies. At the same time, codon usage bias was found to vary among certain amino acids in *C. sinensis*. Amino acids that are encode by two synonymous codons with A+U in the first and second positions and those encoded by four synonymous codons with G+C in the first and second positions are a main source of codon usage bias. Differences in combination of U and A in the first and second positions results in the preferential use of G or C in the third position, while a combination of G and C in the first and second positions is capable of limiting the use of G in the third position [27].

Several methods are designed to estimate the extent of synonymous codon usage bias. The effective number of codons (Nc) was used in the case of *C. sinensis* [28, 30]. The parameter Nc can theoretically vary from 20 in the case where a single codon is used for every amino acid to 61 in the case where all codons are used with equal frequencies. In *C. sinensis*, similar Nc values (51–53) were obtained only with complete coding sequences (CDSs) of more than 400 codons. A trend to a greater biases in short CDSs compared with longer genes was observed in several invertebrates [31].

A comparative analysis of the genomes and the amino acid composition of translated protein-coding domains in flatworms made it possible to assume that the genomes of the opisthorchiids *C. sinensis* and *O. viverrini* code for arginine, alanine, glycine, proline, and valine amino acid residues more frequently, and asparagine, isoleucine, serine, and tyrosine less frequently than the genomes of blood flukes [15]. It is promising to further study codon usage biases associated with mutational pressure or translational selection in liver flukes [27].

### *Retrotransposons*

Mobile genetic elements (MGEs) were found in the genomes of various species and are most diverse in invertebrates [32]. All MGEs are divided into two main classes by the mode of their transposition: retrotransposons, whose transposition involves a RNA intermediate (copy and paste), and DNA transposons, whose DNA is transposed directly. Each of the classes is divided into subclasses, superfamilies, and families by integration mechanism and structural features. Retrotransposons (LTR, non-LTR, and tyrosine recombinase ones) were found in all eukaryotes from fungi to mammals [33, 34]; these MGEs seem to be associated with sexual reproduction [35]. DNA transposons (self-replicating, cut-and-paste, and rolling-circle ones) are the most ancient MGE group and are found in all pro-

karyotes and eukaryotes [33, 34]. It is of interest that MGEs of the first class were observed in both flatworms and round worms, while MGEs of the other class are typical of parasitic nematodes, but not trematodes and tapeworms [36]. More than 30 different MGEs are currently known for mammalian helminths [34].

As members of the repetitive DNA fraction, MGEs were initially thought to be junk or egoistic DNA or genome parasites. Then it became clear that, apart from exerting a negative effect, these genome elements positively contribute to evolution of the host genomes and can be renamed as "potentially useful domesticated elements" [34]. Genomic domestication may lead to substantial functional evolution of MGEs [34, 37–39]. MGEs were observed to coevolve with the host genome and to play an important role in shaping and maintaining the host chromosomes, modifying gene expression, and separating the genome into chromosome domains with epigenetic marks [34, 36, 40–43]. Presumably, transcriptional activity of retrotransposons is specifically controlled by the host and plays a certain role in organ regeneration in animals [41]. Understanding the mechanisms of regeneration in animals, including flatworms, can greatly contribute to human regenerative medicine [3, 6]. It is of both theoretical and applied interest that retroelements can perform important functions during embryo development and, in particular, affect the pluripotent status of cells [41]. Many retrotransposons are species or genus specific, providing a means to design sensitive and specific molecular diagnostic tests, which are especially important in the case of parasitic species [44]. Retrotransposons can generate specific genetic patterns associated with clinical manifestations of diseases, thus being of interest as the subjects of epidemiological and medical studies [36]. MGEs are main candidate vehicles that mediate horizontal gene transfer (HGT) between reproductively isolated species. HGT events are currently considered to be a main driving force in evolution of the eukaryotic genome [45, 46], and the parasite–host system provides an optimal model to study this phenomenon [47–51].

Moving in the genome and integrating into new sites, MGEs cause potentially harmful genetic variation from simple sequence polymorphism to dramatic changes in chromosome structure and integrity and inactivation of neighbor genes [34]. MGE excision and insertion induces mutagenesis, which may occur at a rate one order of magnitude higher than the spontaneous mutation rate [52]. Mutations due to retrotransposon insertion are more stable than mutations caused by DNA transposons [53].

Organisms developed special protective mechanisms to suppress uncontrollable MGE proliferation and transposition, including epigenetic processes, such as transcriptional and posttranscriptional gene silencing. However, various physiological stress factors, for instance, changes in temperature or host migration to new environments, can distort epigenetic silencing and activate MGEs, leading to their proliferation [51]. Based on the data on their great potential in creating genetic variation, retrotransposons are recognized as important evolutionary factors that contribute to the phenotypic variation, ecological adaptation, a remodeling of the host genome, and, in the long term, speciation [34, 36, 40, 54–58]. Understanding the MGE structure and activity will shed further light on the evolutionary history and phylogenetic relationships of between MGEs and the host genome and will facilitate the development of transgene technologies useful for studying the structure, function, and expression regulation of genes in parasitic species [36].

Various MGEs were found (mostly in silico) in species of the phylum Platyhelminthes and account for a substantial portion of their genomes [34, 36, 41, 42]. The *C. sinensis* genome harbors 691 families of repetitive sequences, including long terminal repeat (LTR) and non-LTR retrotransposons [14]. In the *O. viverrini* genome, 61.8% of all repeats are retrotransposons, including LTR retrotransposons and LINEs, which account for larger portions (13.1 and 5.85%, respectively) of genomic DNA than in *C. sinensis* [15]. In *Schistosoma* species, MGEs with copy numbers ranging from 20 to 50 thousand account for 40–50% of the genome [42]. Some sequences are capable of vertical transmission (to offspring), while some others lack this capability. The pattern of their presence or absence greatly varies in the course of the parasite life cycle, i.e., some sequences are found at the adult (marita) stage, while others occur in miracidia or cercariae, etc. [44, 50]. Virtually all flatworm MGEs are transcribed, sometimes, to a high level. Approximately 14% of all transcripts code for sequences with putative reverse transcriptase homology in *S. mansoni* cercariae [44].

Because retrotransposons induce variation via heterogeneous integration followed by sequence divergence, these polymorphic regions are possible to detect with arbitrary PCR primers as random amplified polymorphic DNA (RAPD) [59]. A RAPD analysis made it possible to isolate and characterize a new liver fluke retrotransposon, which was termed *C. sinensis* retrotransposon 1 (*CsRn1*). The full-length LTR retrotransposon *Gulliver* was described in the *S. japonicum* genome, first in Platyhelminthes. However, it was found that its sequence is damaged, although it is expressed at the transcriptional level. *CsRn1* is the first intact LTR retrotransposon that preserves its mobility and occurs at a high copy number, more than 100 per haploid genome. The element has a continuous open reading frame (ORF), which codes for 1304 amino acid residues and has a similarity to the Pol ORF of Ty3/gypsy-like LTR retrotransposons. It is noteworthy that Gag encoded by *CsRn1* has a unique cysteine-histidine (Cys-His) motif, CHCC, in place of the conventional CCHC motif. Owing to its high copy number and mobility, *CsRn1* may play an important role in evolution of the *C. sinensis* genome, including its remodeling [56, 60].

The phylogenetic relationships of retrotransposons are still poorly understood. In total, 29 retrotransposons were isolated in Digenea and grouped in one non-LTR and three LTR families. A phylogenetic analysis showed that *C. sinensis CsRn1* and *Paragonimus westermani PwRn1* (whose structure was studied comprehensively) are new LTR retrotransposons of the Ty3/gypsy superfamily (Metaviridae according to the International Committee on Taxonomy of Viruses), whose members lack *env*. The elements form a separate, highly conserved clade, which is thought to evolve only in the genomes of multicellular organisms [59, 60].

A phylogenetic analysis of the LTR sequences clustered multiple *C. sinensis CsRn1* copies into four subgroups that differ in the time of integration into the host genome during evolution of the species. It is noteworthy that *CsRn1* insertions occur predominantly in heterochromatic or gene-depleted chromosome regions. Several explanations were advanced for this selectivity. For instance, the selectivity was associated with an important role that heterochromatin plays in the chromosome structure and genome integrity, with tolerance of host defense mechanisms (due to a low density of functional genes in the regions involved), or the induction of phenotypic variation mediated by the effect that *CsRn1* exerts on expression of heterochromatin-adjacent genes. A likely consecutive transfer of elements from different *CsRn1* subgroups into the *C. sinensis* genome is supported by data on differential sequence divergence and heterogeneous integration patterns. The lowest divergence was observed for members of a group that spread recently, while its members are highly polymorphic among individual trematode genomes [59, 60].

### *Small Interfering RNAs*

MicroRNAs (miRNAs) were recently described as a new class of small noncoding RNAs that regulate protein expression by recognizing specific mRNAs. Since their first description, more than 28 thousand miRNAs of more than 200 species have been accumulated in the miRBase database (http://www.mirbase.org/). Small interfering RNAs are involved in key biological processes, such as development, differentiation, apoptosis, and proliferation [61, 62], and play an important role in the gene expression regulation, epigenetic modification, and the regulation of heterochromatin activity [63]. Although miRNAs and proteins bound with them are among the most abundant ribonucleoprotein complexes of the cell, some of them are expressed only in few cells or only in specific environmental conditions. Special computational approaches were developed to identify such miRNAs in silico [64, 65].

Data on miRNAs of parasitic worms were obtained in a study of 17 nematode, 11 trematode, and 8 cestode species [66]. In particular, data on the genomics, transcriptomics, and bioinformatics of *Schistosoma* (*S. mansoni*, *S. japonicum*, and *S. haematobium*) suggest a miRNA-mediated silencing system for at least the two first species. The expression patterns of genes involved in RNA interference greatly vary among different developmental stages of parasites [67, 68]. In parasitic nematodes, miRNAs are probably involved in ecological adaptation and behavioral regulation [64]. In the pig nematode *Ascaris suum*, miRNA expression starts extremely early, immediately after fertilization and before pronuclear fusion and the first cleavage division of the zygote, that is, at the developmental stage that was long believed to be a transcriptional quiescence stage [69]. Some miRNAs are conserved among a broad range of species, while others are specific to helminths or individual species.

RNA interference is of special interest in studying the coevolutionary interactions in parasite–host and host–pathogen systems [63]. The most important problem is understanding the role that miRNAs play in the growth and development of parasites and evolution of their capability to regulate invasion of host mammals [70]. Parasites intensely release exosomes and other extracellular structures, which contain specific proteins, RNAs, and, sometimes, genomic DNA. Fusion of extracellular structures of the parasite and its host provides a channel for parasite–host interactions [71]. Extracellular vesicles are involved in the spreading of the pathogen and contribute to the regulation of the host immune response. In the case of *O. viverrini,* extracellular vesicles induce pro-inflammatory and carcinogenic phenotypes in human cholangiocytes [66].

MiRNAs hold a great potential as therapeutic and diagnostic targets for controlling human and animal parasitic diseases [66, 70, 72]. For instance, certain changes in miRNA profile were strongly associated with particular tumor subtypes or disease stages in intrahepatic cholangiocarcinoma [66]. Great attention is attracted by a group of circulating extracellular miRNAs identified recently. These miRNAs are passively released from necrotic cells in liver diseases or actively secreted into vesicular structures (e.g., exosomes) and are stable in biological fluids, such as the urine and blood serum or plasma. The serum levels of miR-192 and miR-21 in patients may be used as biomarkers of cholangiocarcinoma induced by the liver fluke *O. viverrini* [71]. In addition, miR-122 is one of the best studied and most interesting miRNAs as diagnosis and treatment of liver disorders [73–77].

Six new and 62512 conserved miRNAs of 284 families were identified and cloned in the first study of liver fluke miRNAs [78]. This miRNA diversity possibly helps *C. sinensis* to easily extend its host range in certain circumstances. On the other hand, the abundance of conserved miRNAs indicates that they are still evolving and that the formation or elimination of particular miRNA families is a common event [78, 79].

The majority of *C. sinensis* miRNAs are 20–23 nt in size, 21- and 22-nt miRNAs are the most common, and uracil is the predominant nucleotide. It is note-

worthy that the so-called seed regions were not detected in positions 1 and 9, while they are common in human, animal, and plant miRNAs. A classification of conserved miRNAs showed that *C. sinensis* miRNAs (all still innovative and conserved families) belong to three branches (vertebrate, insect, and nematode) on a phylogenetic tree. Two main strategies of using miRNAs were noted for *C. sinensis*. One is preserving the great diversity of miRNA families known for various animals. The other is maintaining strongly conserved regions and regions with high innovative activity (mostly in the central and terminal miRNA regions), which ideally suits the parasite in its complex life cycle and host changes [78]. To complete its life cycle, a parasite has to continuously adapt to intermediate and definitive hosts (which may change themselves or change their response to infestation) and to continuously changing environmental conditions. Both highly conserved regions and regions with high innovative activity are consequently necessary for parasite miRNAs, which contribute to the high adaptive potential of parasites and allow them to extend their host range.

Further studies using deep sequencing technologies identified 33 new and 18 conserved miRNAs in *C. sinensis*, 43 new and 17 conserved miRNAs in *O. felineus*, and 20 new and 16 conserved miRNAs in *O. viverrini* [80]. In silico methods made it possible to predict 178 conserved miRNA genes. Expression levels of conserved miRNAs vary among different species and developmental stages (*O. felineus*). An analysis of the genomic organization of miRNA genes in three opisthorchiid species (*C. sinensis*, *O. felineus*, and *O. viverrini*) supported the existence of two gene clusters (miR-71/miR-2 and let-7/miR-100/mir-125) and one intronic miRNA (the miR-190 gene is an intron of the talin gene) [80]. It is of interest that the copy number of the first cluster varies among different flatworm groups and that certain miRNAs of the miR-71 and miR-2 families display sex-associated expression (in *Schistosoma*) or are involved in regeneration (in planarians). The second cluster, which is conserved among almost all Deuterostomia taxa, was found to be disintegrated in flatworms, including opisthorchiids, with a complete loss of miR-100 [80].

In recent experiments, RNA interference was used to silence the enolase gene in the liver fluke to demonstrate the biological significance of enolase [81]. The approach is thought to be useful for further identification of functional genes in the liver fluke. A study in *F. hepatica* showed that a knockdown of the cathepsin L and B genes via RNA interference exerts a preventive effect against newly excysted juveniles and that a vaccine with cathepsin L1 partly protects against *F. hepatica* infestation and subsequent disease development [82, 83]. Use of miRNAs to study the mechanisms that regulate expression of functionally important genes will greatly contribute to understanding the biological basis of antigen variation and evasion of the host immune response in parasites [84] and developing new-generation antiparasitic drugs [65, 70, 72].

## TRANSCRIPTOME STUDIES

Approximately 2500000 ESTs are available from the dbEST database for human parasites, and 444049 of the ESTs belong to trematodes, including especially dangerous species, such as *F. hepatica*, *S. mansoni*, *S. japonicum*, *Paragonimus westermani*, *O. viverrini*, *O. felineus*, and *C. sinensis* (http:/www.ncbi.nlm.nih.gov/dbEST/dbEST_smary.html). A *C. sinensis* EST database was published [85] (http://pathod.cdc.go.kr/clonorestdb/) and is considered as a resource to study the functional genomics of parasites.

A series of transcriptome studies were performed in *C. sinensis* [86–92]. In total, 52745 ESTs were derived from *C. sinensis*, adults, metacercariae, and eggs, sharing 12830 EST sequences. The 12830 ESTs were further categorized by putative molecular function and relevant biological processes. The molecular function category included proteins binding with ATP, zinc ions, and protein molecules (48–49%); protein possessing enzymatic catalytic activity, including protein kinases and oxidoreductases (39%); and proteins related to transport activity (5%). The biological process category was mostly associated with proteins involved in cell processes, such as protein phosphorylation, translation, and transcription (35–36%); proteins involved in metabolism, including proteolysis, oxidation, and carbohydrate metabolism (33–34%); and proteins regulating biological processes, such as transcription (8%) and signal transmission (8%). Among particular categories, the most numerous in the molecular function group were proteins binding with nucleotides, nucleic acids, and ions and proteins possessing hydrolase and transferase activities. In the biological process group, the most numerous were proteins involved in polypeptide and nucleic acid metabolisms and proteins involved in transport and cell localization [89–91].

A transcriptome analysis of eggs, metacercaria, and adults showed that the majority of *C. sinensis* genes are differentially expressed at particular stages [88, 91] and that their expression correlates with biological and physical features of the particular developmental stage. Between-stage differences were observed in *C. sinensis*, *O. viverrini*, and *O. felineus* for homeodomain-coding genes and genes of G-proteins associated with receptors and neuroactive signaling systems of species [91, 93, 94]. High-level expression was observed for genes involved in energy metabolism, motility, and reproduction in adults; genes involved in minimal metabolism and definitive host adaptation in metacercariae; and embryonic genes in eggs [91, 93, 94]. In total, 30 genes showed the most intense expression in the *C. sinensis* genome. Metacercariae incubated in a bilious medium expressed 16 genes, which are involved in energy metabolism and modulation of regulatory signals for cell growth and proliferation and

the development of newly excysted juveniles [91]. In the *O. felineus* genome, 11114 are expressed at both of the developmental stages; 648 genes, only in adults; and 903, only in metacercariae [93, 94]. Substantial changes in transcriptome profile (especially in the protease and tubulin gene families) occur in *F. hepatica* as well when the parasite migrates from the stomach through the abdominal cavity and the liver to mature in bile ducts. Differential expression during infestation was observed for antioxidant system and detoxification genes, potentially accounting for the stage-specific effects of various anthelmintic drugs [16]. Genomic and transcriptomic studies in *C. sinensis* showed that a total of 9459 genes are expressed in all tissues examined. Of these, 272, 51, 648, and 81 are expressed exclusively in the oral sucker, the ovary, the testis, and muscles, respectively [92].

An interspecific transcriptome analysis yielded additional evidence for parasite adaptation to life in bile ducts; revealed certain specifics in metabolism and gene networks of parasites; and showed several molecular differences, which are possibly related to the specifics of the niche and physiological needs of particular species [89–91, 93–96]. For instance, the liver fluke transcriptomes lack polyamine synthesis enzymes, including catalase, methylthioribose 1-phosphate isomerase, and ornithine decarboxylase. Genes involved in methionine release and peroxisome biogenesis are lost (*O. felineus* has 17 orthologous groups, while there are 39 groups in soil nematode and 71 groups in human). In contrast, increased gene copy numbers were observed in comparisons with free-living species for several gene families, such as the MD-2 lipid-binding protein, calmodulin, and cathepsin families. The most intense transcription was observed for the genes that are necessary for parasitic life, that is, anaerobic respiration, reproduction, detoxification, outer surface integrity, and nutrition (e.g., the genes for myoglobin, vitellin, glutathione S-transferase, and cysteine proteases). Transcripts of cholangiocarcinoma-related genes and genes involved in neurotransmission were found in all opisthorchiids. The finding makes it possible to assume that opisthorchiids have serotoninergic, glutaminergic, and cholinergic neuronal networks (which agrees with the data on high conservation of the trematode nervous system) and that types of their ion channels are consistent with their parasitic lifestyle and positions of the phylogenetic tree of living organisms.

Transcriptome comparisons were performed for *S. japonicus*, *S. mansoni*, *F. hepatica*, *O. felineus*, *O. viverrini*, and *C. sinensis* and showed that the liver flukes have the highest predicted protein sequence homology with each other, reflecting their close phylogenetic relationships and biological similarity [89, 90, 94]. In total, 12665, 13634, and 14269 ORF-containing transcripts were found in *O. felineus*, *O. viverrini*, and *C. sinensis*, respectively. Of the 456 core eukaryotic genes, 445 were identified, and missing genes are absent from *S. mansoni* as well (and, probably, other trematode species) [94]. Interspecific comparisons of a larger scale can isolate the fraction of genes that are highly conserved among model eukaryotic organisms. The fraction includes the genes for actin, tubulin, translation elongation factor 1, valosin-containing proteins, glycogen phosphorylase, and heat shock proteins. Transcriptome comparisons between free-living (*Caenorhabditis elegans* and *Schmidtea mediterranea*) and parasitic worms showed that *C. sinensis* is closer to the parasitic species, consistently with its lifestyle [89, 90, 94]. At the same time, predicted *C. sinensis* proteins have higher homology (20%) with proteins of mammals (*Mus musculus* and *Homo sapiens*) than with proteins of the soil nematode *C. elegans* (15%). This molecular similarity probably allows the parasite to regulate the host response at the biochemical and immune levels [89, 90].

Transcriptome information on main regulators of the development, bile chemotaxis, and physico-metabolic pathways of *C. sinensis* is of importance for identifying new drug targets and diagnostic antigens. A set of candidate targets includes male sterility proteins, cathepsin L-like cysteine protease A, protein P5 precursor, *Cryptosporodium* mucin, TGF-ß1 receptor, and some other proteins [91].

To conclude, genomic and transcriptomic studies in the trematode *C. sinensis* provide new basic data that are essential for understanding the biology and molecular evolution of helminths pathogenic or carcinogenic to humans, thus being of social, economic, and medical importance. The studies open new opportunities for designing highly sensitive diagnostic tests and new-generation drugs to treat parasitic infestations and associated pathologies.

## REFERENCES

1. Hong S.-T., Fang Y. 2012. *C. sinensis* and clonorchiasis, an update. *Parasitol. Int.* **61**, 17–24.
2. Qian M.B., Chen Y.-D., Yan F. 2013. Time to tackle clonorchiasis in China. *Infect. Dis. Pov.* **2**, 4–7.
3. Boyle J.P., Yoshino T.P. 2003. Gene manipulation in parasitic helminths. *Int. J. Parasitol.* **33**, 1259–1268.
4. Kalinna B.H., Brindley P.J. 2007. Manipulating the manipulators: Advances in parasitic helminth transgenesis and RNAi. *Trends Parasitol.* **23**, 197–204.
5. Beckmann S., Grevelding C.G. 2012. Paving the way for transgenic schistosomes. *Parasitology*. **139**, 651–668.
6. Moguel B., Bobes R.J., Carrero J.C., Laclette J.P. 2015. Transfection of platyhelminthes. *BioMed Res. Int.* **2015**, ID 206161. http://dx.doi.org/. doi 10.1155/2015/206161
7. Bouvard V., Baan R., Straif K., Grosse Y., Secretan B., Ghissassi F.E.I., Benbrahim-Talaa L., Guha N., Freeman C., Galichey L., Cogliano V. 2009. A review of human carcinogenus – part B: Biological agents. *Lancet Oncology*. **10**, 321–322.
8. Qian M.B., Chen Y.D., Liang S., Yang G.J., Zhou X.N. 2012. The global epidemiology of clonorchiasis and its relation with cholangiocarcinoma. *Infect. Dis. Pov.* **1**, 4–15.

9. Berriman M., Haas B.J., LoVerde P.T., Wilson R.A., Dillon G.P., Cerqueira G.C., Mashiyama S.T., Al-Lazikani B., Andrade L.F., Ashton P.D., Aslett M.A., Bartholomeu D.C., Blandin G., Caffrey C.R., Coghlan A., et al. 2009. The genome of the blood fluke *Schistosoma mansoni. Nature*. **460**, 352–358. doi 10.1038/nature08160
10. Robb S.M.C., Ross E., Sánchez Alvarado A. 2008. SmedGD: The *Schmidtea mediterranea* genome database. *Nucleic Acids Res.* **36**, D599–D606.
11. Zheng H.J., Zhang W., Zhanf L., Zhang Z., Li J., Lu G., Zhu Y., Wang Y., Huang Y., Liu J., Kang H., Chen J., Wang L., Chen A., Yu S., et al. 2013. The genome of the hydatid tapeworm *Echinococcus granulosus. Nat. Genet.* **45**, 1168–1175.
12. Tsai I.J., Zaroweiecki M., Holroyd N., Garciarrubo A., Sanchez-Flores A., Brooks K.L., Tracey A., Bobes R.J., Fragoso G., Sciutto E., Aslett M., Beasley H., Bennett H.M., Cai J., Camicia F., et al. 2013. The genomes of four tapeworm species reveal adaptations to parasitism. *Nature*. **496**, 57–63.
13. Hahn C., FrommB., Bachmann L. 2014. Comparative genomics of flatworms (Platyhelminthes) reveals shared genomic features of ecto- and endoparastic neodermata. *Genome Biol. Evol.* **6**, 1105–1117. doi 10.1093/gbe/evu078
14. Wang X., Chen W., Huang Y., Sun J., Men J., Liu H., Luo F., Guo L., Lv X., Deng C., Zhou C., Fan Y., Li X., Huang L., Hu Y., Liang C., Hu X., Xu J., Yu X. 2011. The draft genome of the carcinogenic human liver fluke *Clonorchis sinensis. Genome Biol.* doi 10.1186/gb-2011-12-10-r107
15. Young N.D., Nagarajan N., Lin S.J., Korhonen P.K., Jex A.R., Hall R.S., Safavi-Hemami H., Kaewkong W., Bertrand D., Gao S., Seet Q., Wongkham S., Teh B.T., Wongkham C., Intapan P.M., et al. 2014. The *Opisthorchis viverrini* genome provides insights into life in the bile duct. *Nature Comm*. **5**, 4378–4388.
16. Cwiklinski K., Dalton J.P., Dufresne P.J., Course J.L., Williams D.J.L., Hodgkinson J., Paterson S. 2015. The *Fasciola hepatica* genome: Gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biol.* **16**, 71–82. doi 10.1186/s13059-015-0632-2
17. Young N.D., Jex A.R., Li B., Liu S., Yang L., Xiong Z., Li Y., Cantacessi C., Hall R.S., Xu X., Chen F., Wu X., Zerlotini A., Oliveira G., Hofmann A., Zhang G., et al. 2012. Whole-genome sequence of *Schistosoma haematobium. Nature Genet.* **44**, 221–228.
18. Lynch M., Conery J.S. 2003. The origins of genome complexity. *Science*. **302**, 1401–1404.
19. Huang Y, Chen W, Wang X, Liu H, Chen Y, Guo L., Luo F., Sun J., Mao Q., Liang P., Xie Z., Zhou C., Tian Y., Lv X., Huang L., et al. 2013. The carcinogenic liver fluke, *Clonorchis sinensis*: New assembly, reannotation and analysis of the genome and characterization of tissue transcriptomes. *PLOS ONE.* **8**, e54732. doi 10.1371/journal.pone.0054732
20. Duret L. 2002. Evolution of synonymous codon usage in metazoans. *Curr. Opin. Genet. Dev*. **12**, 640–649.
21. Plotkin J.B., Robins H., Levine A.J. 2004. Tissue-specific codon usage and the expression of human genes. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12588–12591.
22. Jiang P., Sun X., Lu Z. 2007. Analysis of synonymous codon usage in *Aeropyrum pernix* K1 and other *Crenarchaeota* microorganisms. *J. Genet. Genom*. **34**, 275–284.
23. Hershberg R., Petrov D.A. 2008. Selection on codon bias. *Annu. Rev. Genet*. **42**, 287–299.
24. Fox J.M., Erill I. 2010. Relative codon adaptation: A generic codon bias index for prediction of gene expression. *DNA Res.* **17**, 185–196.
25. Plotkin J.B., Kudla G. 2011. Synonymous but not the same: The causes and consequences of codon bias. *Nat. Rev. Genet*. **12**, 32–42. doi 10.1038/nrg2899
26. Mauro V.P., Chappell S.A. 2014. A critical analysis of codon optimization in human therapeutics. *Trends Mol. Med*. **20**, 604–613. doi 10.1016/j.molmed.2014.09.003
27. Tang Y., Cho P.Y., Kim T.I., Hong S.-J. 2007. *Clonorchis sinensis*: Codon usage in nuclear genes. *Exp. Parasitol.* **15**, 187–191.
28. Banerjee T., Gupta S.K., Ghosh T.C. 2005. Towards a resolution on the inherent methodological weakness of the "effective number of codons used by a gene." *Biochem. Biophys. Res. Comm*. **330**, 1015–1018.
29. Vicario S., Mason C.E., White K.P., Powell J.R. 2008. Developmental stage and level of codon usage bias in *Drosophila*. *Mol. Biol. Evol*. **25**, 2269–2277. doi 10.1093/molbev/msn189
30. Fuglsang A. 2005. On the methodological weakness of 'the effective number of codons,' *Biochim. Biophys. Res. Comm*. **327**, 1–3.
31. Powell J.R., Moriyama E.N. 1997. Evolution of codon usage bias in *Drosophila. Proc. Natl. Acad. Sci. U. S. A.* **94**, 7784–7790.
32. Holligan D., Zhang X., Jiang N., Pritham E.J., Wessler S.R. 2006. The transposable element landscape of the model legume *Lotus japonicus. Genetics*. **174**, 2215–2228.
33. Bao W., Kapitonov V.V., Jurka J. 2010. Ginger DNA transposons in eukaryotes and their evolutionary relationships with long terminal repeat retrotransposons. *Mobile DNA*. **1**, 3–13.
34. Thomas M.C., Macias F., Alonso C., López M.C. 2010. The biology and evolution of transposable elements in parasites. *Trends Parasitol.* **26**, 350–362.
35. Arkhipova I., Meselson M. 2000. Transposable elements in sexual and ancient asexual taxa. *Proc. Natl. Acad. Sci. U. S. A*. **97**, 14473–14477.
36. Brindley P.J., Laha T., McManus D.P., Loukas A. 2003. Mobile genetic elements colonizing the genome of metazoan parasites. *Trends Parasitol.* **19**, 79–87.
37. Cooper M.D., Alder M.N. 2006. The evolution of adaptive immune systems. *Cell*. **124**, 815–822.
38. Liu D., Bischerour J., Siddique A., Buisine N., Bigot Y., Chalmers R. 2007. The human STMAR protein preserves most of the activities of the ancestral Hs-mar1 transposase. *Mol. Cell. Biol*. **27**, 1125–1132.
39. Kaneko-Ishino T., Ishino F. 2012. The role of genes domesticated from LTR retrotransposons and retroviruses in mammals. *Front. Microbiol*. **3**, 2–11.
40. Flutre T., Permal E., Quesneville H. 2012. Transposable element annotation in completely sequenced eukaryote genomes. In: *Plant Transposable Elements*. Eds. Grandbastien M.-A., Casacuberta J.M. Heidelberg: Springer. *Topics Curr. Genetics*. **24**, pp. 17–40.

41. Mashanov V.S., Zueva O.R., Garcia-Arrarás J.E. 2012. Retrotransposons in animal regeneration. *Mobile Genet. Elements*. **2**, 244–246.
42. Wijayawardena B.K., DeWoody J.A., Minchella D.J. 2015. The genomic proliferation of transposable elements in colonizing populations: *Schistosoma mansoni* in the new world. *Genetica*. **143**, 287–298.
43. Feschotte C., Pritham E.J. 2007. DNA transposons and the evolution of eukaryotic genomes. *Ann. Rev. Genet*. **41**, 331–368.
44. Johnston D.A. 2006. Genomes and genomics of parasitic flatworms. In: *Parasitic Flatworms: Molecular Biology, Biochemistry, Immunology and Physiology.* Eds. Maule A.G., Marks N.J. Wallingford: CAB International, pp. 37–80.
45. Keeling P.J., Palmer J.D. 2008. Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* **9**, 605–618.
46. Gladyshev E.A., Meselson M., Arkhipova I.R. 2008. Massive horizontal gene transfer in bdelloid rotifers. *Science*. **320**, 1210–1213.
47. Houck M.A., Clark J.B., Peterson K.R., Kidwell M.G. 1991. Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis. Science*. **6**, 1125–1129.
48. Loreto E.L., Carareto C.M., Capy P. 2008. Revisiting horizontal transfer of transposable elements in *Drosophila*. *Heredity*. **100**, 545–554.
49. Gilbert C., Schaack S., Pace J.K., Brindley P.J., Feschotte C. 2010. A role for host–parasite interactions in the horizontal transfer of transposons across phyla. *Science*. **464**, 1347–1350.
50. Schaack S., Gilbert C., Feschotte C. 2010. Promiscuous DNA: Horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.* **25**, 537–546.
51. Wijayawardena B.K., Minchella D.J., DeWoody J.A. 2013. Hosts, parasites, and horizontal gene transfer. *Trends Parasitol.* **29**, 329–338.
52. Koga A., Iida A., Hori H., Shimada A., Shima A. 2006. Vertebrate DNA transposon as a natural mutator: The medaka fish Tol2 element contributes to genetic variation without recognizable traces. *Mol. Biol. Evol*. **23**, 1414–1419.
53. Georgiev G.P. 1984. Mobile genetic elements in animal cells and their biological significance. *Eur. J. Biochem*. **145**, 203–220.
54. McDonald J.F. 1990. Macroevolution and retroviral elements. *Bioscience*. **40**, 183–191.
55. Long A.D., Lyman R.F., Morgan A.H., Langley C.H., MacKay T.F. 2000. Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the *achaete–scute* complex are associated with variation in bristle number in *Drosophila melanogaster. Genetics*. **154**, 1255–1269.
56. Bae Y.A., Ahn J.-S., Kim S.-H., Rhyu M.-G., Kong Y., Cho S.-Y. 2008. PwRn1, a novel Ty3/gypsy-like retrotransposon of *Paragonimus westermani*: Molecular characters and its differentially preserved mobile potential according to host chromosomal polyploidy. *BMC Genomics.* doi 10.1186/1471-2164-9-482
57. Böhne A., Brunet F., Galiana-Arnoux D., Schultheis C., Volff J.-N. 2008. Transposable elements as drivers of genomic and biological diversity in vertebrates. *Chromosome Res.* **16**, 203–215.
58. Venancio T.M., Wilson R.A., Verjovski-Almeida S., DeMarco R. 2010. Burst of transposition from non-long terminal repeat retrotransposon families of the RTE clade in *Schistosoma mansoni. Int. J. Parasitol.* **40**, 743–749.
59. Bae Y.A., Moon S.Y., Kong Y., Cho S.Y., Rhyu M.G. 2001. CsRn1, a novel active retrotransposon in a parasitic trematode, *Clonorchis sinensis*, discloses a new phylogenetic clade of Ty3/gypsy-like LTR retrotransposons. *Mol. Biol. Evol.* **18**, 1474–1483.
60. Bae Y.A., Kong, Y. 2003. Evolutionary course of CsRn1 Long-terminal-repeat retrotransposon and its heterogeneous integrations into the genome of the liver fluke, *Clonorchis sinensis. Korean J. Parasitol.* **41**, 209–219.
61. Bartel D.P. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell*. **116**, 281–297.
62. Harfe B.D. 2005. MicroRNAs in vertebrate development. *Curr. Opin. Genet. Dev.* **15**, 410–415.
63. Obbard D.J., Gordon K.H., Buck A.H., Jiggins F.M. 2009. The evolution of RNAi as a defence against viruses and transposable elements. *Phil. Trans. R. Soc. Lond. B. Biol. Sci.* **364**, 99–115.
64. Huang Q.X., Cheng X.Y., Mao Z.C., Wang Y.S., Zhao L.L., Yan X., Ferris V.R., Xu R.M., Xie B.Y. 2010. MicroRNA discovery and analysis of pinewood nematode *Bursaphelenchus xylophilus* by deep sequencing. *PLoS ONE.* **5**, e13271.
65. Devaney E., Winter A.D., Britton C. 2010. MicroRNAs: A role in drug resistance in parasitic nematodes. *Trends Parasitol.* **26**, 428–433.
66. Cai P., Gobert G.N., McManus D.P. 2015. MicroRNAs in parasitic helminthiases: Current status and future perspectives. *Trends Parasitol.* 32 (1), 71–86. doi org10.1016/j.pt.2015.09.003
67. Gomes M.S., Cabral F.J., Jannotti-Passos L.K., Carvalho O., Rodrigues V., Baba E.H., Sá RG. 2009. Preliminary analysis of miRNA pathway in *Schistosoma mansoni. Parasitol. Int.* **58**, 61–68.
68. Xue X., Sun J., Zhang Q., Wang Z., Huang Y., Pan W. 2008. Identification and characterization of novel microRNAs from *Schistosoma japonicum*. *PLoS ONE.* **3**, e4034.
69. Wang J., Czech B., Crunk A., Wallace A., Mitreva M., Hannon G.J., Davis R.E. 2011. Deep small RNA sequencing from the nematode *Ascaris* reveals conservation, functional diversification, and novel developmental profiles. *Genome Res.* **21**, 1462–1477.
70. Manzano-Román R., Siles-Lukas M. 2012. MicroRNAs in parasitic diseases: Potential for diagnosis and targeting. *Mol. Biochem. Parasitol.* **186**, 81–86.
71. Silakit R., Loilome W., Yongvanit P., Thongchot S., Sithithaworn P., Boonmars T., Koonmee S., Titapun A., Khuntikeo N., Chamadol N., Techasen A., Namwat N. 2015. Urinary microRNA-192 and microRNA-21 as potential indicators for liver fluke-associated cholangiocarcinoma risk group. *Parasitol. Int.* doi.org/10.1016/j.parint.2015.10.001
72. Britton C., Winter A.D., Gillan V., Devaney E. 2014. microRNAs of parasitic helminths: Identification, characterization and potential as drug targets. *Int. J. Parasitol.: Drugs Drug Resist.* **4**, 85–94.
73. Lanford R.E., Hildebrandt-Eriksen E.S., Petri A., Persson R., Lindow M., Munk M.E., Kauppinen S.,

Ørum H. 2010. Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science*. **327**, 198–201.

74. Hossein S. 2012. Dual role of miR-122 in molecular pathogenesis of viral hepatitis. *Hepat. Mon*. **12**, 312–314.
75. Hsu S., Wang B., Kota J., Yu J., Costinean S., Kutay H., Yu L., Bai S., La Perle K., Chivukula R.R., Mao H., Wei M., Clark K.R., Mendell J.R., Caligiuri M.A., et al. 2012. Essential metabolic, anti-inflammatory, and anti-tumorigenic functions of miR-122 in liver. *J. Clin. Invest*. **122**, 2871–2883.
76. Wang S., Qiu L., Yan X., Jin W., Wang Y., Chen L., Wu E., Ye X., Gao G.F., Wang F., Chen Y., Duan Z., Meng S. 2012. Loss of microRNA 122 expression in patients with hepatitis B enhances hepatitis B virus replication through cyclin G(1)-modulated P53 activity. *Hepatology*. **55**, 730–341.
77. Wen J., Friedman J.R. 2012. miR-122 regulates hepatic lipid metabolism and tumor suppression. *J. Clin. Invest*. **122**, 2773–2776.
78. Xu M.-J., Liu Q., Nisbet A.J., Cai X.-Q., Yan C., Lin R.-Q., Yuan Z.-G., Song H.-Q., He X.-H., Zhu X.-Q. 2010. Identification and characterization of microRNAs in *Clonorchis sinensis* of human health significance. *BMC Genomics*. **11**, 521–530.
79. Wei L., Zhang D., Xiang F., Zhang Z. 2009. Differentially expressed miRNAs potentially involved in the regulation of defense mechanism to drought stress in maize seedlings. *Int. J. Plant Sciences*. **170**, 979–989.
80. Ovchinnikov V.Y., Afonnikov D.A., Vasiliev G.V., Kashina E.V., Sripa B., Mordvinov V.A., Katokhin A.V. 2015. Identification of microRNA genes in three opisthorchiids. *PLoS Negl. Trop. Dis*. **9**, e0003680.
81. Wang X., Chen W., Tian Y., Huang Y., Li X., Yu X. 2014. RNAi-mediated silencing of enolase confirms its biological importance in *Clonorchis sinensis*. *Parasitol. Res*. **113**, 1451–1458.
82. McGonigle L., Mousley A., Marks N.J., Brennan G.P., Dalton J.P., Spithill T.W., Day T.A., Maule A.G. 2008. The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduces gut penetration. *Int. J. Parasitol*. **38**, 149–155.
83. Zafra R., Pérez-Écija R.A., Buffoni L., Moreno P., Bautista M.J., Martínez-Moreno A., Mulcahy G., Dalton J.P., Pérez J. 2013. Early and late peritoneal and hepatic changes in goats immunized with recombinant cathepsin L1 and infected with *Fasciola hepatica*. *J. Comp. Pathol*. **148**, 373–384. doi 10.1016/j.jcpa.2012.08.007
84. Liu Q., Tuo W., Gao H., Zhu X.-Q. 2010. MicroRNAs of parasites: Current status and future perspectives. *Parasitol. Res*. **107**, 501–507.
85. Kim D.-W., Yoo W.G., Lee S., Lee M.-R., Kim Y.-J., Cho S.-H., Lee W.-J., Ju J.-W. 2014. ClonorESTdb: A comprehensive database for *Clonorchis sinensis* EST sequences. *BMC Res. Not*. **7**, 388. http://www.biomedcentral.com/1756-0500/7/388.
86. Lee J.S., Lee J., Park S.J., Yong T.S. 2003. Analysis of the genes expressed in *Clonorchis sinensis* adults using the expressed sequence tag approach. *Parasitol. Res*. **91**, 283–289.
87. Cho P.Y., Lee M.J., Kim T.I., Kang S.Y., Hong S.J. 2006. Expressed sequence tag analysis of adult *Clonorchis sinensis*, the Chinese liver fluke. *Parasitol. Res*. **99**, 602–608.
88. Cho P.Y., Kim T.I., Whang S.M., Hong S.J. 2008. Gene expression profile of *Clonorchis sinensis* metacercariae. *Parasitol. Res*. **102**, 277–282.
89. Young N.D., Campbell B.E., Hall R.S., Jex A.R., Cantacessi C., Laha T., Sohn W.-M., Sripa B., Loukas A., Brindley P.J., Gasser R.B. 2010. Unlocking the transcriptomes of two carcinogenic parasites, *Clonorchis sinensis* and *Opisthorchis viverrini*. *PLoS Negl. Trop. Dis*. **4**, e719.
90. Young N.D., Jex A.R., Cantacessi C., Campbell B.E., Laha T., Sohn W.-M., Sripa B., Loukas A., Brindley P.J., Gasser R.B. 2010. Progress on the transcriptomics of carcinogenic liver flukes of humans: Unique biological and biotechnological prospects. *Biotechnol. Adv*. **28**, 859–870.
91. Yoo W.G., Kim D.-W., Ju J.-W., Cho P.Y., Kim T.I., Cho S.-H., Choi S.-H., Park H.-S., Kim T.-S., Hong S.-J. 2011. Developmental transcriptomic features of the carcinogenic liver fluke, *Clonorchis sinensis*. *PLoS Negl. Trop. Dis*. **5**, e1208. doi 10.1371/journal.pntd.0001208
92. Huang Y., Chen W., Wang X., Liu H., Chen Y., Guo L., Luo F., Sun J., Mao Q., Liang P., Xie Z., Zhou C., Yanli Tian Y., Lv X., Huang L., et al. 2014. The carcinogenic liver fluke, *Clonorchis sinensis*: New assembly, reannotation and analysis of the genome and characterization of tissue transcriptomes. *PLOS ONE*. **8**, e54732. pone. 0054732. doi 10.1371/journal
93. Pomaznoy M., Tatkov S., Katokhin A., Afonnikov D., Babenko V., Furman D., Brusentsov I., Belavin P., Najakshin A., Guselnikov S., Vasiliev G., Sivkov A., Prokhortchouk E., Skryabin K., Mordvinov V. 2013. Adult *Opisthorchis felineus* major protein fractions deduced from transcripts: Comparison with liver flukes *Opisthorchis viverrini* and *Clonorchis sinensis*. *Exp. Parasitol*. **135**, 297–306, http://dx.doi.org/10.1016/j.exppara.2013.07.011.
94. Pomaznoy M.Y., Logacheva M.D., Young N.D., Penin A.A., Ershov N.I., Katokhin A.V., Mordvinov V.A. 2016. Whole transcriptome profiling of adult and infective stages of the trematode *Opisthorchis felineus*. *Parasitol. Int*. **65**, 12–19.
95. Young N.D., Hall R.S., Jex A.R., Cantacessi C., Gasser R.B. 2010. Elucidating the transcriptome of *Fasciola hepatica:* A key to fundamental and biotechnological discoveries for a neglected parasite. *Biotechnol. Adv*. **28**, 222–231.
96. Laha T., Pinlaor P., Mulvenna J., Sripa B., Sripa M., Smout M.J., Gasser R.B., Brindley P.J., Loukas A. 2015. Gene discovery for the carcinogenic human liver fluke, *Opisthorchis viverrini*. *BMC Genomics*. **8**, 189–203. doi 10.1186/1471-2164-8-189

*Translated by T. Tkacheva*