

# The omic approach to parasitic trematode research—a review of techniques and developments within the past 5 years

Orçun Haçarız<sup>1</sup> · Gearóid P. Sayers<sup>2</sup>

Received: 15 April 2016 / Accepted: 19 April 2016 / Published online: 28 April 2016  
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** The evolution of technologies to explore parasite biology at a detailed level has made significant advances in recent years, particularly with the development of omic-based strategies. Whilst extensive efforts have been made in the past to develop therapeutic and prophylactic control strategies for trematode parasites, only the therapeutic anthelmintic approach can be regarded as usable in clinical practice. Currently, there is no commercialised prophylactic strategy (such as vaccination) for protection of the definitive host against any trematode parasite. Since 2010 in particular, the integration of omic technologies, including liquid chromatography-mass spectrometry (LC-MS) and next-generation sequencing (NGS), has been increasingly reported in trematode-related studies. Both LC-MS and NGS facilitate a better understanding of the biology of trematodes and provide a promising route to identifying clinically important biological characteristics of parasitic trematodes. In this review, we focus on the application, advantages, and disadvantages of omic technologies (LC-MS and NGS) in trematode research within the past 5 years and explore the use and translation of the omic-based research results into practical tools to deal with infection.

**Keywords** Infection · Liquid chromatography-mass spectrometry · Next-generation sequencing · Omic · Trematode

## Introduction

The therapeutic and prophylactic research efforts over the past 50 years, to address parasitic infection in human and veterinary medicine, have had profoundly different results. The therapeutic-based anthelmintics have been remarkably successful against a broad range of parasites and, although blighted by more recent reports of anthelmintic resistance (Brockwell et al. 2013; Ortiz et al. 2013; Elliott et al. 2015), stand in contrast to the failure of prophylactic research efforts to develop a viable product to effectively prevent trematode infection in both animals and humans.

Trematode parasites are eukaryotic organisms with a large protein repertoire ( $n > 10,000$  predicted for various species) (Young et al. 2012, 2014; Cwiklinski et al. 2015; Haçarız et al. 2015). Diseases caused by these parasites are responsible for substantial production losses in the livestock industry (e.g. fasciolosis) and/or for health problems in humans as over a million cases per year are reported for trematode infections worldwide (such as schistosomiasis, clonorchiasis and fasciolosis) (Toet et al. 2014; Cwiklinski et al. 2015; Walz et al. 2015; Qian et al. 2015). The previous approaches to develop vaccines against trematode parasites, largely adopted from virus and prokaryotic models, have failed to deliver an effective vaccine to prevent infection. Additionally, whilst recombinant antigens with various adjuvant combinations have been tested in experimental and field trials, none have showed full protection or been commercialised (Golden et al. 2010; Hotez et al. 2010; Toet et al. 2014; McNeilly and Nisbet 2014; van der Ree and Mutapi 2015; You and McManus 2015).

✉ Orçun Haçarız  
orcun.hacariz@tubitak.gov.tr; orcunhacariz@gmail.com

<sup>1</sup> TÜBİTAK Marmara Research Center, Genetic Engineering and Biotechnology Institute, Gebze, Kocaeli, Turkey

<sup>2</sup> Department of Science, Technology, Engineering and Mathematics, Institute of Technology Tralee, Tralee, Co. Kerry, Ireland

Currently, genuine progress into novel anti-trematode drug and vaccine development appears static, with no literary publications citing end stage production of commercial products, in human or veterinary medicine.

The application of omic technologies is widely regarded as a possible candidate tool to overcome this stasis (Toet et al. 2014; van der Ree and Mutapi 2015; Molina-Hernández et al. 2015). Adopting a whole proteome analysis approach to trematode parasitism is most useful to help pinpoint the essential parasite proteins relating to virulence, to investigate important parasite protein interactions (proteins that are correlated in expression or additive in their effect) and to identify the parasite's unique immuno-modulatory mechanisms to prolong infection (Robinson et al. 2009, 2013; Hotez et al. 2010; Haçarız et al. 2012, 2014; van der Ree and Mutapi 2015; You and McManus 2015; Khan et al. 2015).

The integration of omic approaches into molecular studies of trematode parasites has been ongoing since the beginning of this century. Proteins with different molecular functions for various trematode parasites have been identified, and candidate molecules have been proposed as drug and vaccine targets (Wilson et al. 2011; Dalton et al. 2013; van der Ree and Mutapi 2015).

This review focuses on the progress in the omic approach to address trematode parasitism within the past 5 years, and aims to complement and build upon previously published review studies prior to 2010 (Toledo et al. 2011). Targeted at researchers from various fields (e.g. biology, medicine, veterinary medicine, parasitology), the text addresses the tools, applied approaches, advantages and disadvantages of omic studies in trematode research, and importantly how to progress beyond candidate protein nomination to complete investigative viability studies.

## Frequently used omic technologies in trematode research

Two high-throughput omic approaches, liquid chromatography-mass spectrometry (LC-MS) to identify proteins at peptide level, and next-generation sequencing (NGS) to identify genes at nucleotide level, are key tools used to characterise the signature profile of each trematode parasite. Sample preparation and correct experimental design are central to a meaningful outcome from omic studies, which is valid for both proteomics and genomics/transcriptomics research. A workflow integrating LC-MS with NGS is illustrated in Fig. 1.

### LC-MS

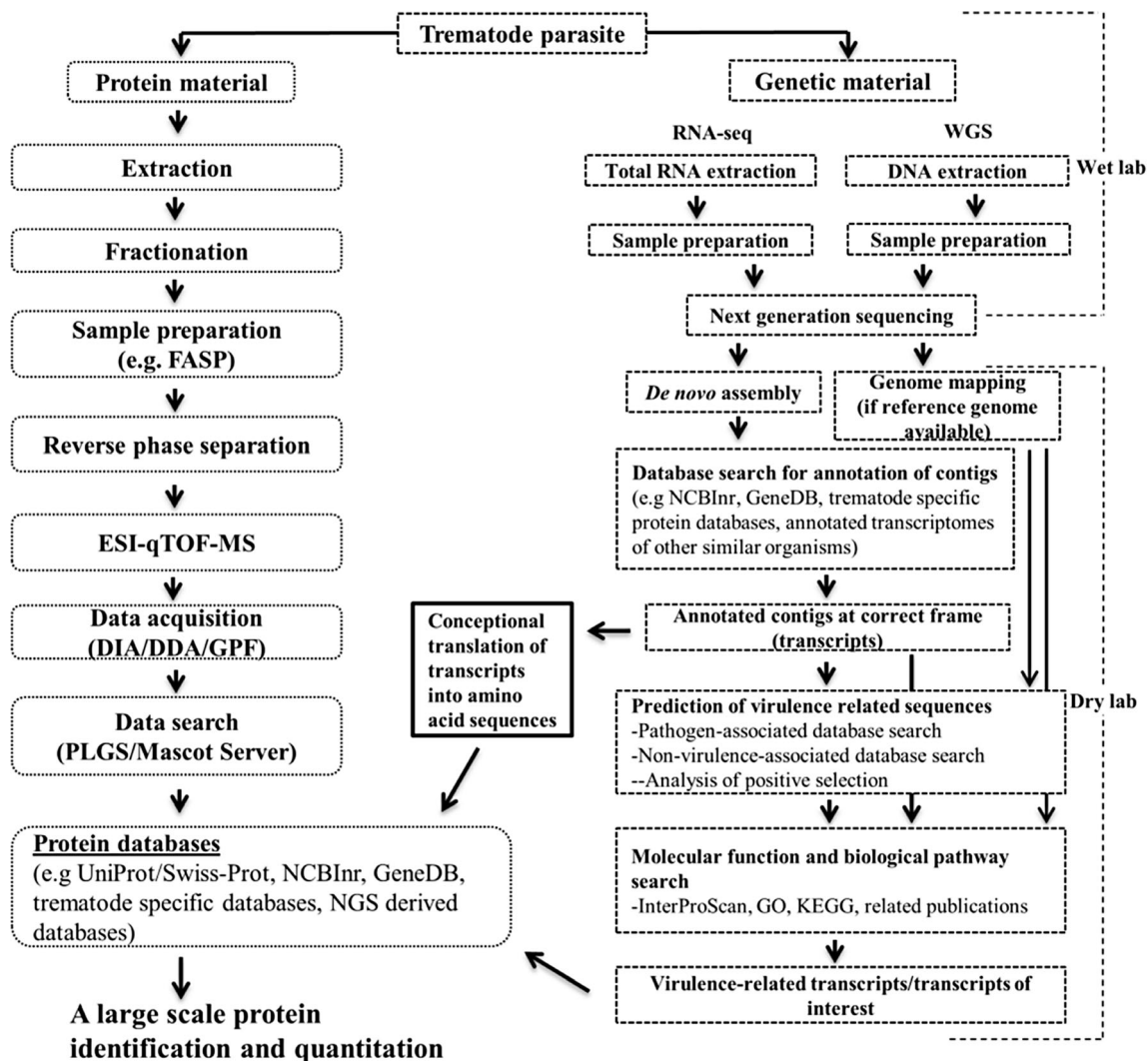
Earlier proteomic studies commenced by analysing the purified proteins of interest (which are mainly excised from

polyacrylamide gels after two-dimensional electrophoresis) through peptide mass fingerprint using matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF), with an appropriate database search to reveal its identify (Thiede et al. 2005). This method, which is efficient for the identification of targeted proteins (such as antigenic proteins), has been used in trematode research (Acosta et al. 2008; Tian et al. 2013), but the isolation of hydrophobic proteins from the gels (after the electrophoresis application) has been reported to be challenging because of the protein aggregation (Fischer et al. 2006).

In the recent past (no more than 10 years), researchers in the trematodology field aimed to elucidate the whole proteome of trematode organisms in order to enhance our biological understanding of the parasites, and develop effective strategies for the disease control. This whole proteome approach was favoured, as the one-by-one molecule-based strategy (such as the use of an antigen) proved difficult in developing new efficient prevention strategy (such as vaccination) in many experimental trials. As the gel electrophoresis-based separation of proteins can be difficult to perform, difficult to reproduce, time consuming and economically inefficient, particularly for a large number of identification targets, the use of separation techniques (such as liquid chromatography) with mass spectrometry took hold amongst the research community.

The LC-MS approach, which is increasingly used in many fields, has been particularly applied to analyse samples containing thousands of proteins and, therefore, can be used to identify large proteomes of trematode parasites. The principle of LC-MS is the chromatographic separation (such as reverse phase separation-based on molecule hydrophobicity) of digested peptides and deduction of amino acid sequences from mass-to-charge ratios ( $m/z$ ) determined by mass spectrometry (de Hoffmann 1996; Chen and Pramanik 2009). The commonly used version of LC-MS, called LC-MS/MS, previously described in detail (de Hoffmann 1996; Chen and Pramanik 2009; Grebe and Singh 2011; Rauh 2012), typically includes technical stages such as peptide ionisation (by electrospray ionisation), peptide fragmentation (by collision energy), and mass detection of the resulting products (by time-of-flight-based detection), yielding spectral data of peptide and amino acid sequences. Finally, raw data obtained by this approach is searched against a protein database using *in silico* tools (such as Mascot (Matrix Science; <http://www.matrixscience.com/>), ProteinLynx Global Server (Waters; <http://www.waters.com/>)) to identify proteins in the biological samples.

In addition to protein identification, LC-MS/MS is also useful for protein quantitation using labelled (e.g. isotopes) or label-free methods. In the label-free methods, which is now regarded as the more practical and effective choice (Sandin et al. 2015), the relative abundance of proteins is possible using internal standard with known concentration



**Fig. 1** Workflow diagram describing LC-MS and NGS (RNA-Seq and WGS), as adopted (Haçarız et al. 2012, 2014, 2015; Haçarız and Baykal 2014). A protein database derived from NGS provides parasite specific database, leading to a significant increase in protein identification. *ESI* electrospray ionisation, *qTOF* quadrupole time-of-flight, *MS* mass

(such as the use of enolase peptides of *Saccharomyces cerevisiae*) (Haçarız et al. 2012; Haçarız and Baykal 2014) or using a more advanced approach such as peptide ion abundance (Haçarız et al. 2014). Other than protein detection, LC-MS (and also other technologies such as proton nuclear magnetic resonance spectroscopy (NMR)) can also be helpful to detect metabolites which are small molecules derived from enzymatic reactions in biological systems (Becker et al. 2012).

Overall, LC-MS technology is mainly useful for (a) the identification and quantitation of proteins (including isoforms and splice variants) at peptide level, including protein-protein interactions and ancillary comparative and correlative analysis to PCR-based RNA expression levels (Haçarız et al. 2012, 2014; Schaefer et al. 2012; Larance and Lamond 2015), (b) assessment of post-translation modifications, as previously described (Gupta et al. 2007), (c) understanding the micro

spectrometry, *DDA* data-dependent acquisition, *DIA* data-independent acquisition, *GFP* gas phase fractionation, *PLGS* ProteinLynx Global Server, *WGS* whole genome sequencing, *FASP* filter-aided sample preparation, *GO* gene ontology, *KEGG* Kyoto Encyclopaedia of Genes and Genomes

RNA (miRNA) biology, influencing protein expression levels (Huang et al. 2013a) and (d) detection of metabolites which indicate metabolic pathways (Becker et al. 2012).

### Important issues for LC-MS

#### Protein identification

For proteomic studies involving trematodes, careful isolation, recovery and maintenance of the parasite is important for the deduction of meaningful proteomic and genomic data. Procedures to achieve this for trematodes are outlined by Haçarız et al. (2011, 2012, 2014, 2015); however, variations may need to be applied depending on the experimental purpose.

The sample could be prepared using homogenisation (such as beads and homogeniser) and/or detergents (such as nonidet P40, for enriching proteins of surface layers in particular) (Wilson et al. 2011; Haçarız et al. 2011, 2014). Excretory/secretory (ES) proteins could be obtained in a cell culture environment and concentrated with methanol/chloroform precipitation (Robinson et al. 2009). However, enrichment protocols do not guarantee pure isolation of the experimental fractions. A small concentration of surface proteins can be found in ES fractions due to shedding of the tegument in vitro (Morphew et al. 2007) and a quantity of ES proteins can potentially remain in surface fractions (Wilson et al. 2011; Morphew et al. 2013). In addition, the proteome fractions can be influenced by the duration of incubation, as the in vitro conditions can cause damage to the parasites (Wilson et al. 2011). In the case of trematodes, the addition of protease inhibitor cocktails to prevent protein sample degradation is important, as parasitic trematodes release various proteases (Haçarız et al. 2014). Following this, an efficient protein extraction technique, such as filter-aided sample preparation (FASP), is required (Wiśniewski et al. 2009).

In LC-MS-based proteomic studies, proteomic grade trypsin is generally used to digest proteins into tryptic peptides, but using other endoproteases (e.g. Glu C, and Asp N) cleaving proteins at different amino acid points could be useful to identify other proteins that have not been identified with the trypsin protocol (Wiśniewski et al. 2009; Swaney et al. 2010). One of the major obstacles of LC-MS/MS technology is that peptides with high relative concentrations may mask the identification of other peptides expressed at lower concentrations (Gingras et al. 2005). This ‘masking effect’ is due to the measurement of protein concentration in starting material, which mostly relates to abundant proteins, whilst the concentrations of some proteins that are relatively low become undetectable for LC-MS (Liu et al. 2006). This problem could be solved by fractionation of proteins/peptides based on their chemical properties such as isoelectric point and molecular weight, prior to the LC-MS/MS analysis. The integration of different data acquisition approaches (e.g. data-dependent acquisition, data-independent acquisition, gas phase fractionation) could be important to identify more proteins (Haçarız et al. 2014). In addition, the quality and amount of data depends on the performance of the device (e.g. the resolution capacity of mass spectrometry) used.

Protein identification is highly dependent on the size of the database provided, as LC-MS/MS requires a database to identify proteins. To date, LC-MS/MS has been applied to a number of trematode parasites, and a number of proteins have been identified by comparing the raw proteomic data with protein data of non-specific organisms and other datasets derived from EST

resources. However, in many cases, findings have been limited because of the absence of well-established protein databases due to the lack of transcriptome/genome data for trematode parasites. Since 2010, NGS-based studies have provided a number of genomes/transcriptomes of various trematodes and this trend is likely to continue for others where genomes/transcriptomes are not yet deciphered. The integration of current advanced LC-MS/MS systems with protein databases derived from NGS data (for both genomic and transcriptomic components) would allow researchers to visualise the whole proteome of the trematode parasite. Technical parameters such as sequence coverage, the number of peptides identified, parent and fragment ion error tolerances, and false positive rate (<4 % is accepted in general) are important in the proteomic analysis (Gingras et al. 2005; Wiśniewski et al. 2009; Haçarız et al. 2014). The integration of two main protein identification approaches, called *de novo* peptide sequencing and database searching, is suggested to increase the identification coverage and accuracy (Wang and Wilson 2013). As LC-MS/MS has been widely used for whole profiling omic studies in trematode research, other useful mass spectrometry techniques (e.g. MALDI-TOF) will not be reviewed here.

#### *Metabolite identification*

In metabolome-based studies, sample preparation is an important issue and an extraction procedure for metabolome profiling of helminths has been proposed (Saric et al. 2012). After spectral data is obtained by the sample run, the general experimental procedure of the metabolome analysis involves spectral processing (including baseline correction, noise filtering, peak detection and alignment, normalisation and deconvolution), data analysis [including various methods such as uni-variant, multi-variant (supervised/unsupervised) and multi-way] and metabolite identification (see for technical details; Alonso et al. 2015). An organism-specific metabolome database is an important tool to increase the number of identifications (Jewison et al. 2012). Currently, fully annotated metabolome databases are limited for many infectious organisms (Preidis and Hotez 2015), but a comprehensive metabolome database is available for other comprehensively studied organisms such as human (Wishart et al. 2013).

Metabolome analyses have been mainly focused on the study of prokaryotic organisms, yielding useful results. Metabolome analysis of *Staphylococcus aureus* led to discovery of the target (pyruvate dehydrogenase) of an antibacterial product (triphenylbismuth dichloride) (Birkenstock et al. 2012) in multiresistant bacterial pathogens, and fatty acid

metabolism has been found to be altered in rifampicin-resistance *Mycobacterium tuberculosis* (du Preez and Loots 2012).

## NGS

The other important tool in high throughput sequencing is NGS, which is based on the fragmentation of genetic material with various techniques such as heat enzymatic cleavage or sonication, followed by attachment and amplification of the labelled gDNA or complementary DNA (cDNA) fragments on an appropriate surface (e.g. chip, flow cell), and deduction of sequences of the attached fragments (~100–400 bases) using optical signals (such as in 454 sequencing, Roche; HiSeq 2000/2500, Illumina) or pH change (Ion Torrent, Life Technologies) (Cantacessi et al. 2012a; Mardis 2013). The base information of each sequence is called ‘read or sequence read’. These technologies enable extensive parallel sequencing, which results in large volumes of data (at gigabase levels for trematodes).

In general, NGS of whole cDNA (reverse-transcribed from mRNA) and gDNA are referred to as RNA-seq (or whole transcriptome sequencing) and whole genome sequencing (WGS), respectively (Ng and Kirkness 2010; Smith et al. 2013). Millions of reads can be assembled to get contiguous sequences (called ‘contigs’) using various software applications (called de novo assembly) (such as Velvet, Oases) or can be mapped to a reference genome sequence, if available in silico (such as Bowtie, TopHat) (Zerbino and Birney 2008; Wang et al. 2009; Trapnell et al. 2009, 2012; Ng and Kirkness 2010; Langmead and Salzberg 2012; Schulz et al. 2012; Smith et al. 2013; Kim et al. 2013). With WGS, intronic and exonic DNA regions of trematodes can be sequenced, as previously described (Smith et al. 2013; Ng and Kirkness 2010; Young et al. 2014). RNA-seq data can be aligned to reference trematode genomes in order to identify splice junctions between exons (using the appropriate software such as TopHat, Bowtie), and gene expression levels can be analysed based on read abundance using the reference trematode genome (e.g. TopHat, Cufflinks) (Langmead and Salzberg 2012; Trapnell et al. 2012; Kim et al. 2013; Cwiklinski et al. 2015), or even in the absence of the reference genome (Li and Dewey 2011). Some genome assemblies of trematodes (to use as reference genome for NGS applications) are publicly available (e.g. WormBase, <http://www.wormbase.org/>, WormBase Parasite, <http://parasite.wormbase.org/>). In NGS-based studies involving de novo assembly, correct frames of transcript sequences can be detected by ‘BLAST’ comparisons (Altschul et al. 1990) using previously known protein data, and these sequences can be conceptually translated into protein sequences to provide a parasite specific protein database for LC-MS/MS applications (Robinson et al. 2009; Wilson et al. 2011; Haçarız et al. 2014, 2015). To date, genomes and

transcriptomes of a number of trematodes have been successfully sequenced with current NGS technologies (Berriman et al. 2009; The *Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium 2009; Logan-Klumpler et al. 2012; Zerlotini et al. 2013; Young et al. 2012, 2014; Cwiklinski et al. 2015; Haçarız et al. 2015).

## Important issues for NGS

As for peptide analysis, careful isolation, recovery and maintenance of the parasite are important. Extraction of DNA and RNA is most commonly completed using commercially available kits, typically membrane and spin column-based, which facilitate a reliable and practical isolation procedure for genetic material. For the wet laboratory stage of NGS, quality and quantity of genetic material are critical parameters (Haçarız and Sayers 2013). Purity of material is checked with the absorbance of the nucleic acid at 260, 230 and 280 nm, using a spectrophotometer. For purity, the ratio of A280/A260 is suggested to be approximately 1.8 for DNA and approximately 2.0 for total RNA, and the ratio of A260/A230 is proposed to be 2.0–2.2 for both DNA and total RNA (T009-Technical bulletin, NanoDrop, ThermoScientific, USA). Integrity of DNA or total RNA could be checked with gel-based methods, but current approaches using microfluidics capillary-based electrophoresis would give more accurate assessment of the intactness, particularly for total RNA (Haçarız and Sayers 2013). Fluorescence-based quantification and quantitative PCR (qPCR) applications are recommended for starting material and DNA library quantification, respectively (Haçarız and Sayers 2013; Haçarız et al. 2015).

The dry lab stage of NGS currently necessitates the use of Linux operating systems (RedHat in particular) in a computer with large RAM memory (~128 gigabytes for RNA-seq/1 terabyte is usually enough for general purposes) for the assembly of sequence reads. Read quality analysis and other necessary processes (e.g. removal of undesired sequences) need to be completed using bioinformatic softwares (e.g. FASTQC; <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> and FASTX-Toolkit; [http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)) before sequence assembly is carried out (Young et al. 2010a,b; Haçarız et al. 2015). The accuracy at both base and sequence level which can be determined by Phred quality score (Q score) (a value of  $\geq 33$  is usually acceptable), an adequate sequencing coverage ( $\geq 30\times$  is preferred in general), and sequence duplication with a low ratio are important components of NGS (Ewing and Green 1998; Sims et al. 2014; Haçarız et al. 2015).

Knowledge about the trematode of interest is essential for effective evaluation of data. For example, a good understanding of the general biology of the trematode will maximise results from in silico work. In some cases, coordinators of studies may not be familiar with in silico data and related

applications, thereby potentially missing important findings and/or making incorrect interpretation of the data. An efficient correspondence between the coordinator and the expertise carrying out the omic scale experiment is necessary to prevent this outcome.

Whole transcriptome profiling with RNA-Seq is an important approach to prepare protein databases for LC-MS/MS. In turn, the RNA-Seq procedure is highly dependent on the quality of extracted total RNA. Poor quality of starting RNA material would decrease the read number, leading to less coverage of transcriptome and hence a limited protein database for LC-MS/MS. Assessment of the integrity of extracted total RNA is routinely examined by microfluidics capillary-based electrophoresis, e.g. Bioanalyzer (Agilent, USA) and, as part of this analysis, a heat-denaturation step is recommended to prevent loop formation of nucleic acid strands. However, the 28S ribosomal RNA of trematode parasites (e.g. *F. hepatica*, *S. japonicum*) is ruptured at the middle point of the RNA molecule by this heat-denaturation step [similar to other organisms such as insects (Ishikawa and Newburgh 1972; Winnebeck et al. 2010)] and prevents accurate assessment of quality using the RIN indicator (RNA integration number) (Haçariz and Sayers 2013; Gobert et al. 2009). Therefore, total RNA of trematode parasites should not be exposed to heat-denaturation before RNA quality assessment analysis (Haçariz and Sayers 2013). This approach has been successfully applied within a transcriptome study (Haçariz et al. 2015).

For more detailed and specific guidelines to these instruments (both LC-MS and NGS), readers are referred to relevant books (McMaster 2005; Alzate 2010 for LC-MS and Valencia et al. 2013; Wong 2013 for NGS) and technical protocols (from commercial providers such as Waters, [www.waters.com](http://www.waters.com), for LC-MS and Illumina, [www.illumina.com](http://www.illumina.com), for NGS) in the area.

### Current status of omic approaches in trematode research

Trematodes belong to the subclass digenea of the class trematode which branches off from the phylum platyhelminthes (flatworms). Taxonomical status of the flatworms, as shown in Fig. 2, could be a useful guide to researchers for comparative purposes in omic-based trematode research. The diversity amongst the parasites stems from the taxonomic stages of order, superfamily or family. In general, the life cycle of trematode parasites includes an intermediate host (gastropods (snails)), an additional intermediate host in some cases (e.g. ants for *Dicrocoelium sp.*, crustaceans for *Paragonimus sp.*, and fish for *Clonorchis sp.* and *Opisthorchis sp.*), and the main (definitive) host (mammals) (Toledo et al. 2011). The family

difference in the same suborder (e.g. *Clonorchis sp.* (liver fluke) versus *Metagonimus sp.* (intestinal fluke); *Fasciola sp.* (liver fluke) versus *Hypoderaeum sp.* (intestinal fluke)) and the genus difference in the same family (e.g. *Fasciola sp.* (liver fluke) vs. *Fasciolopsis sp.* (intestinal fluke)) could be associated with disparity of the infected organ in the definitive host.

The current status in protein/nucleotide sequences in trematode parasites and specialised data resources are demonstrated in Tables 1 and 2, respectively. In this section, omic-based studies (particularly profiling studies after approximately 2010) for common parasitic trematodes are outlined based on organ tropism under four categories (blood, liver, intestinal and other flukes).

#### Blood flukes

One of the most studied trematode genera at the omic scale is *Schistosoma* spp., including *S. mansoni*, *S. japonicum* and *S. haematobium* species which cause pathologies in various organs (e.g. liver, bladder, lungs) in mammals by entering the host circulatory system. There have been numerous proteomic studies particularly for *S. mansoni* since 2000, to which readers are referred to for further information (DeMarco and Verjovski-Almeida 2009; Toledo et al. 2011; Gobert et al. 2014; Mickum et al. 2014; van der Ree and Mutapi 2015). Previous studies have identified antigenic proteins, some of which are now proposed as candidates for both vaccines and diagnostics (Toledo et al. 2011; van der Ree and Mutapi 2015; Ludolf et al. 2014). In large-scale omic studies, the proteome profiles of teguments of *S. mansoni* and *S. japonicum* (Braschi et al. 2006a, b; Mulvenna et al. 2010a) and ES proteins of *S. japonicum* (Liu et al. 2009) have been investigated, an integrated immunoproteomic and bioinformatic approach has been applied to identify *S. japonicum* tegument proteins (Chen et al. 2014) and a differential proteomic analysis of different stages (schistosomula and adult) of *S. japonicum* has been described (Hong et al. 2013). In addition, tegumental and ES proteins of another *Schistosoma* species with veterinary importance, *S. bovis*, have been identified by LC-MS-based studies (Pérez-Sánchez et al. 2006, 2008; Ramajo-Hernández et al. 2007; Higón et al. 2011; de la Torre Escudero et al. 2011, 2013) and a recombinant version of *S. bovis* 22.6 antigen has been suggested to be used in diagnostics (de la Torre-Escudero et al. 2012).

Significant advances have been made recently using omic technologies to enhance the biological understanding of *Schistosoma* spp. (Toledo et al. 2011; van der Ree and Mutapi 2015). The NGS technology made it possible to sequence genomes of three *Schistosoma* species (*S. mansoni*, *S. japonicum* and *S. haematobium*) (Young et al. 2012; Berriman et al. 2009; The *Schistosoma japonicum* Genome

Taxonomy	Root								
No rank	cellular organisms								
Superkingdom	Eukaryota								
No rank	Opisthokonta								
Kingdom	Metazoa								
No rank	Eumetazoa								
No rank	Bilateria								
Phylum	Platyhelminthes								
Class	Trematoda								
Subclass	Digenea								
Order	Strigeidida	Opisthorchiida	Plagiorchiida	Plagiorchiida	Plagiorchiida	Plagiorchiida	Strigeidida	Digenea incertae sedis	
Suborder	-	Opisthorchiata	Echinostomata	Trogloremata	Pronocephalata	Xiphidiata	-	-	
Superfamily	Schistosomatoidea	-	Echinostomatoidea	-	Paramphistomoidea	Plagiorchioidea	Diplostomoidea	-	
Family	Schistosomatidae	/Opisthorchiidae //Heterophyidae	/Fasciolidae //Echinostomatidae	Troglorematidae	/Paramphistomidae //Gastrothylacidae ///Paramphistomatidae	Dicrocoeliidae	Diplostomatidae	Gymnophallidae	
Genus	Schistosoma	/Clonorchis /Opisthorchis //Metagonimus	/Fasciola /Fascioloides /Fasciolopsis //Echinostoma //Hypoderaeum	Paragonimus	/Paramphistomum //Fischoederius ///Watsonius	Dicrocoelium	Alaria Fibricola	Gymnophalloides Parvatrema Gymnophallus Bartolius	

**Fig. 2** The taxonomic classification of known trematode parasites (<http://www.ncbi.nlm.nih.gov/taxonomy/>). Based on current knowledge, a total of 20 different genera (stem from different families ( $n = 12$ ) of orders ( $n =$

4) from the subclass digenea, the class trematode) appeared to be responsible for various trematode infections. The relationship between family and genus is indicated with the slash symbol (/)

Sequencing and Functional Analysis Consortium 2009). In addition, a well-established annotation database, GeneDB (Logan-Klumpler et al. 2012) and an updated genome resource, SchistoDB (Zerlotini et al. 2013), as well as data from the Chinese National Human Genome Center at Shanghai, the *Schistosoma japonicum* Genome Project (The *Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium 2009), are now publicly available. A genome data for *S. bovis* is available at the single read archive of NCBI (SRA) but the annotated genome has not yet been published. Although a number of proteins have been identified at the peptide level by previous efforts, the integration of all current data with further proteomic studies (LC-MS/MS-based) is still thought to be useful to confirm previously non-reported proteins of all *Schistosoma* species at peptide level, thereby further improving the effectiveness of current strategies (van der Ree and Mutapi 2015).

Despite the genome of *Schistosoma sp.* being well studied, there is as yet no report about the comprehensive metabolome profiling of any species of this genus, or any other trematode genera, according to our search. However, there have been a number of studies investigating metabolite changes in hosts with schistosomiasis, which have been reviewed previously (Wang et al. 2010; O'Sullivan et al.

2013). In brief, alterations in metabolites related to energy metabolism (such as down regulation of tricarboxylic acid cycle) and gut microbiota are observed, which have been proposed to be important in the diagnosis of infection (Wang et al. 2010; Li et al. 2011; O'Sullivan et al. 2013).

### Liver flukes

Opisthorchiidae family includes two important genera, *Clonorchis* and *Opisthorchis*. *C. sinensis*, *O. viverrini* and *O. felineus* have two intermediate hosts (snail and fish) and infect the liver of humans. These parasites are known to be significant risk factors of cholangiocarcinoma in infected patients. Many studies have focused on this aspect of the parasite and several proteins have been suggested for serological diagnosis of the parasite.

Publications on omic-scale studies for *C. sinensis* and *O. viverrini* have been steadily increasing since 2010. Studies investigating ES proteins of *C. sinensis* have been published and antigenic proteins have been detected (such as fructose-1,6-bisphosphatase, methionine aminopeptidase 2 and acid phosphatase) (Zheng et al. 2011, 2013). However, in these studies, the raw proteomic data is searched against *Caenorhabditis elegans* or *Schistosoma* databases, which may

**Table 1** Current status of the availability of trematode protein/nucleotide sequences at two global data resources (UniProt and NCBI)

Categorisation by organ tropism	Trematode parasite	UniProt	NCBI (pt)	NCBI (SRA)	NCBI (nt)	Date
Blood flukes	<i>Schistosoma mansoni</i>	103 (r), 12922 (ur)	14454	206 (136 DNA, 70 RNA)	16779	15. May. 2015
	<i>Schistosoma japonicum</i>	49 (r), 16264 (ur)	25596	14 (14 RNA)	54898	15. May. 2015
	<i>Schistosoma haematobium</i>	8 (r), 11336 (ur)	11393	10 (DNA)	4760	15. May. 2015
Liver flukes	<i>Clonorchis sinensis</i>	3 (r), 13917 (ur)	14310	17 (7 DNA, 10 RNA)	3469	08. May. 2015
	<i>Opisthorchis viverrini</i>	16408 (ur)	32954	23 (17 DNA, 6 RNA)	124132	07. May. 2015
	<i>Opisthorchis felineus</i>	1 (r), 75 (ur)	585	9 (microRNA transcriptome analysis)	13428	07. May. 2015
	<i>Fasciola hepatica</i>	23 (r), 522 (ur)	1225	46 (23 DNA, 23 RNA)	59563	14. May. 2015
	<i>Fasciola gigantica</i>	1 (r), 266 (ur)	652	4 (3 DNA, 1 RNA)	907	14. May. 2015
	<i>Fascioloides magna</i>	19 (ur)	61	1 (RNA)	95	15. May. 2015
	<i>Dicrocoelium dendriticum</i>	1 (r), 35 (ur)	118	3 (DNA)	478709	26. May. 2015
	Lung flukes	<i>Paragonimus kellicotti</i>	2 (ur)	2	10 (9 DNA, 1 RNA)	7
<i>Paragonimus westermani</i>		116 (ur)	151	3 (3 DNA)	293	21. May. 2015
<i>Paragonimus miyazakii</i>		6 (ur)	30	3 (3 DNA)	45	21. May. 2015
Rumen flukes	<i>Paramphistomum cervi</i>	12 (ur)	24	3 (RNA)	15	21. May. 2015
	<i>Fischoederius elongatus</i>	1	4	None	20	26. May. 2015
Intestinal flukes	<i>Echinostoma caproni</i>	45 (ur)	10	2 (1 DNA, 1 RNA)	86102	21. May. 2015
	<i>Hypoderaeum</i> sp.	16 (ur)	35	None	55	07. May. 2015
	<i>Fasciolopsis buski</i>	1 (ur)	1	3 (2 DNA, 1 RNA)	14	15. May. 2015
	<i>Watsonius watsoni</i>	None	None	None	2	27. May. 2015

Each number in NCBI (pt) or NCBI (nt) represents single query whilst each number in NCBI (SRA) corresponds millions of reads obtained by NGS. Data from NCBI (pt) and NCBI (nt) may include redundant sequences and some of nucleotide sequences may contain untranslated regions. Numbers of protein (in UniProt, NCBI (pt)) and nucleotide sequences (nucleus, mitochondrial) (in NCBI (nt)), and number of NGS entries (NCBI (SRA)) for various trematode species are shown

*pt* protein, *nt* nucleotide, *SRA* single read archive, *r* reviewed sequences, *ur* un-reviewed sequences, *NCBI* national centre for biotechnology information, *UniProt* universal protein resource

limit the protein identification data obtained. An organised expressed sequence tag (EST) database (ClonorESTdb;  $n = 13,305$ , consists of 6497 contigs and 6808 singletons) of *C. sinensis* is publicly available (<http://pathod.cdc.go.kr/clonorestdb/>) (Kim et al. 2014). Furthermore, the draft genome and transcriptome of the parasite have been published (Young et al. 2010a; Wang et al. 2011; Huang et al. 2013a).

The NGS derived data could also be useful for further LC-MS/MS-based studies to identify surface or ES proteins that have not been detected by previous efforts.

The secreted and surface proteins of the adult stage of *O. viverrini* have been previously investigated using an EST database. Immunomodulators such as annexins and orthologs of the schistosomiasis vaccine antigens Sm29 and



**Table 2** Current and publicly available important data resources for trematode research at omic level

Data resource	Internet link	Content
NCBI	<a href="http://www.ncbi.nlm.nih.gov">http://www.ncbi.nlm.nih.gov</a>	Global nucleotide and protein data
UniProt	<a href="http://www.uniprot.org">http://www.uniprot.org</a>	Global protein data
Query database	<a href="http://gasser-research.vet.unimelb.edu.au">http://gasser-research.vet.unimelb.edu.au</a>	cDNA/predicted protein library of various liver flukes for blast comparisons
Vaccine Investigation and Online Information Network	<a href="http://www.violinet.org">http://www.violinet.org</a>	Database of various micro- and macro- infectious organisms
GeneDB	<a href="http://www.genedb.org">http://www.genedb.org</a>	Database of various helminths including <i>Schistosoma</i> spp.
Helminth secretome database (HSD)	<a href="http://estexplorer.bioinfo.org/hsd">http://estexplorer.bioinfo.org/hsd</a>	Predicted secretory proteins of helminths
SchistoDB	<a href="http://SchistoDB.net">http://SchistoDB.net</a>	Nucleotide/protein data of <i>Schistosoma</i> spp.
Chinese National Human Genome Center (CHGC)	<a href="http://www.chgc.sh.cn/japonicum/Resources.html">http://www.chgc.sh.cn/japonicum/Resources.html</a>	Nucleotide/protein data of <i>S. japonicum</i>
DNA Data Bank of Japan (DDBJ)	<a href="http://www.ddbj.nig.ac.jp">http://www.ddbj.nig.ac.jp</a>	Nucleotide data of free-living trematodes ( <i>DugesIIDae</i> family)
SmedGD: the Schmidtea mediterranea genome database	<a href="http://smedgd.neuro.utah.edu">http://smedgd.neuro.utah.edu</a>	Nucleotide data of <i>S. mediterranea</i>
ClonorESTdb	<a href="http://pathod.cdc.go.kr/clonorestdb">http://pathod.cdc.go.kr/clonorestdb</a>	Nucleotide data of <i>C. sinensis</i>
Helminth.net	<a href="http://www.helminth.net">http://www.helminth.net</a> (including Trematode.net)	Nucleotide data of various trematode parasites
50 Helminth Genomes Initiative (Sanger Institute)	<a href="https://www.sanger.ac.uk/resources/downloads/helminths/">https://www.sanger.ac.uk/resources/downloads/helminths/</a>	Genome sequences of various helminths including <i>Schistosoma</i> spp., <i>Dicrocoelium</i> sp.
WormBase	<a href="http://www.wormbase.org">http://www.wormbase.org</a>	Nucleotide/protein data of nematodes, including the model organism <i>C. elegans</i>
WormBase ParaSite	<a href="http://parasite.wormbase.org/index.html">http://parasite.wormbase.org/index.html</a>	Genome sequences of some trematode species

The data resources are useful for omic-scale studies of trematode organisms, including annotation for de novo genome/transcriptome assemblies and more specific purposes including the identification of virulence-related genes/proteins, antigenic proteins

tetraspanin-2 have been suggested for the tegument-host interface in this parasite (Mulvenna et al. 2010b). More recently, a draft genome of the parasite has been demonstrated by Young et al. (2014). This study also provides protein sequences ( $n=16,379$ ; predicted from nucleotide sequences) of the parasite. Therefore, it would be helpful to conduct further LC-MS/MS integrated studies to identify other proteins (at peptide level) which could be important for vaccine development. Another important *Opisthorchis* species is *O. felinus*; however, both proteomic and genomic data for this species is not publicly available.

The other family, fasciolidae, includes *Fasciola hepatica* and *F. gigantica*, which cause liver infections in various mammals (e.g. ruminants and humans). So far, a number of omic-based studies have been done for *F. hepatica*. Initial attempts focused on the identification of ES proteins of the parasites (Jefferies et al. 2001; Gourbal et al. 2008). Robinson et al. (2009) studied the secretome of the parasite with LC-MS/MS proteomic technique and Hernández-González et al. (2010) studied the newly excysted juveniles of *F. hepatica* using the same technique. However, both searched the raw proteomic data against a preliminary EST database (published by the Sanger Institute); therefore, the findings were limited. This was also the case for other LC-MS/MS-based studies for more specific purposes, such as comparison of the effects of different conditions (in vivo/in vitro) on *F. hepatica*

(Morphew et al. 2007) and analysing the response of adult *F. hepatica* to triclabendazole (Chemale et al. 2010).

Large omic-scale studies have been started with initial nucleotide sequencing of the parasite with a NGS platform (454, Roche) (Young et al. 2010b). In 2011, the tegument of *F. hepatica* has been studied with an LC-MS/MS approach where the database from the 454 sequencer (Young et al. 2010b) and an in-house EST database were used (Wilson et al. 2011). In another study, surface and internal fractions of the parasite have been studied with LC-MS/MS, and compared against a number of databases (Haçarız et al. 2012). A more detailed analysis of these fractions have been studied using LC-MS/MS integrated with different acquisition methods and *in silico* database search against various protein databases and a new assembly of publicly available EST data (Haçarız et al. 2014), yielding a total of 776 proteins. Currently, the total number of detected proteins of the parasite at peptide level is approximately 1000 parasitic proteins. Recently, the draft genome of the parasite (Cwiklinski et al. 2015) and a detailed transcriptome study (Haçarız et al. 2015) have been published. Using combined approaches including the available nucleotide database and fractionation, coupled with the use of different acquisition methods, would be useful in revealing the whole proteome of this parasite.

According to our knowledge, *F. hepatica* is the only trematode to which a profiling metabolome analysis has been

applied. A total of 142 metabolites of the parasite have been identified (Saric et al. 2012); however, this identification number appears to represent a small percentage of the parasite's entire metabolome, considering the current *Escherichia coli* and *Saccharomyces cerevisiae* metabolome databases (each containing more than 2000 metabolite entries) and the human metabolome database (containing more than 40000 metabolite entries) (Jewison et al. 2012; Guo et al. 2013; Wishart et al. 2013).

In addition to the biological analyses of the parasite itself, variations in serum protein profiles during a *F. hepatica* infection in sheep have been demonstrated, reflecting the ability of a trematode to influence the host serum profile (Rioux et al. 2008).

The other important *Fasciola* species is *F. gigantica*, however, omic-based studies for this species are rare in comparison to *F. hepatica*. Young et al. (2011) studied the transcriptome of this parasite using NGS (Genome analyzer II, Illumina). Proteomic studies for this parasite have been completed for specific purposes, such as investigations for cathepsin I protease family and glutathione transferase superfamily (Morphew et al. 2011, 2012), but not whole profiling studies. The *Fasciolidae* family also contains other species, such as *Fascioloides magna* (generally detected in wild ruminants, cervids and bovids). The transcriptome and ES proteome of adult *F. magna* have been studied (Cantacessi et al. 2012b).

The other trematode species that infects the liver of the definitive host is *Dicrocoelium dendriticum* (belong to the *Dicrocoeliidae* family). An EST database of this parasite has been previously generated and surface/ES proteins of the parasite have been analysed by the LC-MS/MS approach (Martínez-Ibeas et al. 2013a, b; Bernal et al. 2014). Whole proteome profiling of this parasite remains to be studied, as a protein database reflecting whole transcriptome of the parasite has not been used in previous studies.

### Intestinal flukes

Other trematode species, which are intestinal inhabitants of various mammalian and non-mammalian hosts, are members of the family *Diplostomidae* (including *Alaria* sp., *Fibricola* sp.), of the family *Gymnophallidae* (including *Gymnophalloides* sp., *Parvatrema* sp., *Gymnophallus* sp., and *Bartolius* sp.), of the family *Heterophyidae* (such as *Metagonimus* sp.) and of the family *Echinostomatidae* (including *Echinostoma* sp., *Hypoderaeum* sp.) (Toledo et al. 2011).

The transcriptome of *E. caproni* has been described by Garg et al. (2013). Predictably, some of the proteins may not have been identified by the proteomic studies available before 2013 (Guillou et al. 2007; Sotillo et al. 2010) due to lack of

NGS data. Other studies, which have focused on metabolite changes in hosts infected with *E. caproni*, have been reviewed by O'Sullivan et al. (2013). Various metabolite alterations in plasma and urine have been detected at 1 and 8 days at post-infection in mice, respectively (Saric et al. 2008). Another species of the *Echinostomatidae* family is *Hypoderaeum conoideum*, which is found in intestines of birds, has led to significant economic losses to the poultry industry, particularly in some Asian countries (Yang et al. 2015a). Although the mitochondrial genome of this parasite has been studied (Yang et al. 2015a), no omic studies have yet been completed.

Two other genera, *Fasciolopsis* sp. and *Watsonius* sp., can cause intestinal infection. Although *Fasciolopsis* sp. is taxonomically classified under *Fasciolidae* family (mostly including liver flukes), a species of this genus, *Fasciolopsis buski*, is an intestinal fluke of humans and only one omic-based study has been published (Biswal et al. 2013). *Watsonius watsoni* (belong to the *Paramphistomatidae* family) is a trematode parasite, attached to intestine wall of primates and humans, causing diarrhoea (Mas-Coma et al. 2005), but this species has not been well studied.

### Other flukes

*Paramphistomum* species such as *P. cervi* infects the rumen of ruminants (such as cattle and sheep), causing diarrhoea, anaemia, and lethargy (Toledo et al. 2011). Therefore, this parasite is important for animal health with associated economical costs. According to our literary review, there has not been a detailed omic scale study for this genus. However, transcriptome data of *P. cervi* (isolated from various main hosts) from NGS studies have been deposited at SRA. *Fischoederius elongatus* is observed in the rumen of ruminants, causing significant economic losses to the livestock industry (Yang et al. 2015b), however there is no omic scale study for this parasite to date.

A comprehensive omic-based analysis has been done for *Paragonimus kellicotti*, a neglected trematode, causing inflammation in lungs of humans (McNulty et al. 2014). In this study, LC-MS/MS has been integrated with the parasite's transcriptome to identify immuno-dominant antigens. Similar studies could be carried out for other *Paragonimus* species (*P. westermani*, *P. miyazakii*), of which, a large nucleotide sequences can be utilised from SRA.

### From 'many to few'—identifying the proteins of interest

Identifying vaccine/drug targets for trematode parasites by determining those proteins that play specific roles in establishment and maintenance of infection, so-called virulence-related proteins, is a difficult but essential step. Most of the parasite

proteins identified through omic studies will relate to physiological pathways of the parasite and are not directly associated with virulence. To eliminate these ‘house-keeping’ proteins, *in silico* comparative approaches are useful. Recently, a large number of nucleotide sequences of free-living trematodes (*DugesIIDae* family including *Dugesia* sp. and *Schmidtea* sp.) have been made publicly available at DNA Data Bank of Japan and a genome database of *S. mediterranea* has been deposited at SmedGD (Robb et al. 2008, 2015). Nucleotide and protein data of another well studied organism, *C. elegans*, is freely accessible from WormBase (Harris et al. 2014). Both trematode and nematode organisms have similar proteins; therefore, *C. elegans* database may serve as an important resource to observe the degree of similarity between these species (Haçarız et al. 2015).

Comparative analysis of the parasite proteins to databases containing sequence data of these non-pathogenic organisms may be a first step approach to simplify the search in identifying potential proteins of virulence, prior to the use of pathogen related databases such as Vaccine Investigation and Online Information Network (He et al. 2014) and Helminth Secretome Database (Garg and Ranganathan 2011, 2012). The other data resource, Helminth.net including Trematode.net, providing collective information for trematodes, could be useful for comparative studies (Martin et al. 2015). In addition, evolutionary changes of trematode proteins under positive selection pressure can be investigated. For this purpose, orthologous sequences (parasite vs. non-parasite) can be subjected to an evolutionary analysis, nonsynonymous/synonymous (Ka/Ks) substitution rate analysis, where Ka/Ks greater than a value of 1.0 is considered as parasitic related diversity (Hurst 2002).

Protein sequences of interest could be subjected to a series of bioinformatical analyses such as motif search (e.g. family, domain, conserved site) using InterProScan as an example (Jones et al. 2014). Biological pathways could be predicted by KEGG (Kyoto Encyclopedia of Genes and Genomes) (Ogata et al. 1999) and gene ontology (GO) information (Ashburner et al. 2000; Gene Ontology Consortium 2015) could also be investigated. However, some protein motifs of trematode parasites (e.g. domain of unknown function DUF1968 of *F. hepatica*, Haçarız et al. 2015) do not have GO information in the current GO database, so protein motifs can manually be inspected online (<http://www.ebi.ac.uk/interpro/>). Although this would be a time-consuming approach, it could provide significant results that could be missed by automated database searches.

The subcellular localisation of virulence-related proteins, which interact directly with the host, could be predicted using bioinformatical software such as WoLF PSORT (Horton et al. 2007). In addition, protein interactions, commonly referred to as the protein ‘interactome’ (Kandpal et al. 2009), reflecting

interactions both within a parasite and between parasite and its host, could be investigated using bioinformatic approaches and molecular techniques. For example, the application of co-immunoprecipitation, fluorescence resonance energy transfer (FRET) and molecular imaging can be used to define those parasite protein of interest (Kandpal et al. 2009; Lahiri et al. 2014; Jia et al. 2014).

Analysis of differential gene expression levels in comparing different trematode organisms (e.g. infectious versus non-infectious, infectious versus infectious), different conditions (in vivo/in vitro) and/or different stages (newly excysted juvenile versus adult) of same trematode parasite, could be important to identify virulence-related genes. These could be investigated by NGS (e.g. using Cufflinks in RNA-Seq) (Trapnell et al. 2012) and then confirmed by real-time PCR (for chosen genes) (Haçarız et al. 2009a). However, the exact confirmation of differentially expressed proteins (in the different aspects) with quantitative LC-MS/MS is recommended (Haçarız et al. 2014).

To enhance the detection of antigenic virulence-related proteins of trematodes, two-dimensional gel-based techniques, including the use of positive host serum taken at appropriate time point(s) of infection (Haçarız et al. 2011) can be integrated with LC-MS/MS (McNulty et al. 2014). Possible B or T cell epitopes of predicted virulence-related proteins could be investigated using epitope databases (Immune Epitope Database and Analysis Resource; <http://www.iedb.org/>) (Vita et al. 2015), which would be important for vaccine studies, in particular.

Proteins identified by LC-MS/MS can be searched against drug target databases such as Potential Drug Target Database (<http://www.dddc.ac.cn/pdtd/>) (Li et al. 2006; Gao et al. 2008), Therapeutic Target Database (<http://bidd.nus.edu.sg/group/ttd/ttd.asp>) (Zhu et al. 2012) and DrugBank 4.0 (Law et al. 2014) for drug development studies. Information from these data resources could help find the related drug category and drug structure, which may aid researchers in choosing drugs as new therapeutics or assist with novel drug design.

Apart from the investigation of novel molecules that are potentially important in combatting fluke infections, omic approaches are also useful to enhance molecular understanding of anthelmintic drugs (e.g. triclabendazole), their mechanism of action and the rise in trematode resistance to these drugs (Chemale et al. 2010). As current omic technologies allow ‘global’ data screening of genes/proteins of organisms, it is possible to detect significant differences of the genes/proteins (of importance) that may be missed by less advanced technologies. With current omic approaches, differences in genes/proteins in comparative experimental models (such as the comparison of parasites which are susceptible or resistant to anthelmintics) could be detected, even to single nucleotide/amino acid change level, highlighting potential issues for drug metabolism/resistance studies (Chemale et al. 2010; Haçarız

et al. 2012, 2014, 2015). This information could be used to re-synthesise the existing drug with the necessary chemical alterations to accommodate these molecular changes, at least for a period of time - as drug resistance could be developed for any drug design. In addition, omic approaches could also be utilised to detect important variations (e.g. drug resistance) in genes/proteins amongst parasite populations from different geographical locations (Hodgkinson et al. 2013).

The identification of novel trematode-based immunomodulatory molecules is another potentially important outcome of omic studies. It is well known that parasitic helminths modulate the host immune system (Flynn et al. 2007; Robinson et al. 2013; Khan et al. 2015). As an example, the use of porcine whipworm (a nematode helminth) eggs (*Trichuris suis* ova) has been suggested in the treatment of ulcerative colitis (Summers et al. 2005). Recent studies have indicated that proteins secreted by trematode parasites (such as cathepsin L1, peroxiredoxin and helminth defence molecule of *F. hepatica*) are potentially important in dealing with immune-related diseases (Robinson et al. 2009, 2013). A recent NGS-based study has suggested that some of the predicted proteins of *F. hepatica* imitate host cytokine and cytokine receptors suggesting a possible therapeutic approach to dealing with immune-related diseases (Haçarız et al. 2015). It is thought that trematodes release specific molecules (such as cytokine-like) which only interact with unique targets of the host, hence inducing a series of unique pathways (Robinson et al. 2013; Wammes et al. 2014; Haçarız et al. 2015; Helmby 2015). Overall, the integration of LC-MS/MS strategies with growing nucleotide data will help to identify trematodes' 'molecules of interest', at peptide level, which could be utilised to treat immune-related diseases (e.g. auto-immune diseases).

### Production of important parasite proteins for further experimental studies

In general terms, LC-MS/MS and RNA-Seq do not provide full sequences of identified proteins or encoding genes, respectively, as molecule identification is based on the detected sequences even if incomplete but of sufficient length. If the gene sequence of the protein of interest is partially known, full length of that sequence (derived from RNA-seq studies) can be obtained using RACE methods or deduced from genome data of the parasite if it is available (Japa et al. 2015). Gene sequences can be amplified by PCR and, if genetic material cannot be obtained, the gene sequence can be synthesised artificially (Hughes et al. 2011). In general, the 'his-tag' sequence encoding at least six histidine residues is incorporated with the gene sequence of interest and the his-tagged protein (either at N or C terminal) is purified using affinity

chromatography (Pedersen et al. 1999; Cass et al. 2005). If necessary, the his-tag peptide could be removed after the protein is purified (Pedersen et al. 1999; Andberg et al. 2007).

In order to investigate a protein as a therapeutic or prophylactic target, a recombinant of the identified parasite proteins (partial or full) will need to be produced, commonly using other organisms' cells. In general, expression vectors (such as virus or plasmid) are used to deliver the gene encoding the desired protein into protein expressing cells (Pedersen et al. 1999; Cass et al. 2005). In doing so, it is important that chemical properties (e.g. hydrophobicity) of the protein of interest are examined, and determine whether these properties are appropriate for the planned protein production protocol. A number of bioinformatic tools in on-line resources (e.g. ExPasy; <http://www.expasy.org/>) are helpful to predict such important chemical properties (Artimo et al. 2012).

As trematode parasites are eukaryotic organisms, accurate post-translation modifications of the parasite proteins will, most likely, be required. Production of trematode proteins in prokaryotic cells (such as *E. coli*) may lead to differences in three dimensional structure (e.g. protein folding) and post-translation modifications (e.g. phosphorylation) relative to the original. Although prokaryotic cell systems could initially be tested (Khow and Suntrarachun 2012), adopting yeast (such as *S.cerevisiae*), insect cell systems (baculovirus + insect cell) or mammalian cells systems may be required to have better results in terms of similarity to the original protein (Endo and Sawasaki 2006). However, protein expression in heterologous systems (even eukaryotic) may not always guarantee the production of an identical parasite protein in some components, carbohydrate motifs, for example.

Primary structure (peptide sequence) and purity of produced protein can be investigated with LC-MS/MS or MALDI-TOF systems (Chen and Pramanik 2009; Webster and Oxley 2012). Endotoxin levels can be analysed with Limulus Amebocyte Lysate Assay (LAL assay) and possible endotoxin residues could be removed using a phase separation approach (Haçarız et al. 2011). Defining primary structure (peptide sequence) of homogenous-produced protein may be satisfactory in some cases. However, further structural analysis may be of benefit for more promising protein targets. Secondary structure (alpha and beta sheets) of the produced protein can be detected with circular dichroism (Kelly and Price 2000). Tertiary or quaternary structures of the produced protein can be demonstrated by NMR or X-ray crystallography (Littlechild et al. 1987; Spencer et al. 2015).

After proteins are expressed in vitro, comparative vaccine trials will assess the protective nature of the protein. Monitoring the effect of adjuvants is also important for trematodes, as some adjuvants themselves (such as Quil A) may influence parasitic burden/egg number, as described for

*F. hepatica* (Haçarız et al. 2009b), masking the specific effect of the antigen.

For the development of anthelmintics, in silico approaches could be utilised to identify inhibitors of trematode molecules, as it has been shown for cathepsin L3 of *F. hepatica* (Hernández Alvarez et al. 2015). High-throughput in vitro models can be designed if the number of testing molecules is high before using in vivo models (Tsai et al. 2013).

Apart from developing strategies to deal with infection, recombinant trematode molecules could be suitable for diagnostic tests (Gonzales Santana et al. 2013; Teimoori et al. 2015) and other in vitro/in vivo disease models in terms of therapeutic applications of immune-related diseases (Dalton et al. 2013; Robinson et al. 2013; Hasby et al. 2015; Heylen et al. 2015).

Finally, for commercialisation of novel biological molecules, GLP (Good Laboratory Practice) and GMP (Good Manufacturing Practice) procedures (Leblanc et al. 2014) need to be adopted.

### Translation of omic research into reality—commercialisation of novel trematode-derived molecules

Whilst the omic approach to researching parasitic mechanisms must now be favoured as the best current approach to unlocking control mechanisms, any newly designed control measure for trematode parasitism in human or veterinary medicine, prophylactic or therapeutic, will be benchmarked against the practical benefits of anthelmintics. These commonly used products have been hugely successful in the veterinary field for specific reasons - the cheap price relative to the cost of production losses, the wide spectrum of activity including a number of seasonal co-infecting trematodes (and nematodes), and the ease of integration into farm management practices and easily stored. Vaccine-based concepts carry additional criteria of requiring the induction of memory, effective for at least one season. Stricter food residue safety concerns, animal welfare and environmental regulations will place additional requirements on bringing any new product to market in the field of veterinary medicine, in particular.

The omic approach has been successfully applied to vaccine development and commercialisation. The sequence-based ‘reverse vaccinology’ approach has been utilised to develop an innovative commercialised vaccine against *Neisseria meningitidis* serogroup B (Heinson et al. 2015), with a similar omic approach being applied to *Streptococcus agalactiae*, *S. pyogenes*, *S. pneumoniae* and pathogenic *E. coli* (Seib et al. 2012).

The status of the omic-based approach for many pathogens, including the causative agents of tick-borne diseases, malaria and hookworm disease, has been previously reviewed and concluded that the approach contributes to the better understanding of the parasites’ infection mechanisms and identification of the parasites’ important molecules for control strategies, but further efforts (wet lab-based research in particular) are needed to utilise the omic-based findings for the development of the practical tools (such as vaccine) (Loukas et al. 2011; Marcelino et al. 2012; Nóbrega de Sousa et al. 2013).

For trematode parasites, the application of omic technologies into developing commercialised vaccines and drug targets is at an early stage; however, examples are emerging. A 28 kDa glutathione S-transferase of *S. haematobium* has been detected (Higón et al. 2011; van der Ree and Mutapi 2015), and a recombinant version of this molecule (rSh28GST), along with an aluminium hydroxide (alum) adjuvant, has been tested in a phase I trial for its safety, tolerability and immunogenicity for the purpose of reducing egg production and subsequent pathology (Riveau et al. 2012). The beneficial effect against schistosomiasis is currently thought to be improved by considering other important antigens detected by proteomic studies (Ludolf et al. 2014; van der Ree and Mutapi 2015). Protein kinase C and extracellular signal-regulated kinase, proteins that were identified through omic analysis, have been recently suggested as potential targets for chemotherapeutic treatments against human schistosomiasis (Ressurreição et al. 2014; Walker et al. 2014). The identification of an insulin receptor of *S. japonicum* (The *Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium 2009) has led to an experimental study where the use of the ligand domain of *S. japonicum* insulin receptor 2 (SjLD2) (produced in *E. coli*) provided a high level decrease (75 %) in the number of mature intestinal eggs in a mice challenge model (You et al. 2012; You and McManus 2015).

Following efforts on omic-scale work for another trematode parasite, *F. hepatica* (Young et al. 2010b; Wilson et al. 2011), further research work has indicated that there are different FhCD59-like sequences which play distinct roles in the development of the parasite (Shi et al. 2014) and FhCD59-like proteins are now considered vaccine candidates (Toet et al. 2014). Additionally, the identification of potential *S. mansoni* vaccine candidates with computational vaccinology has been reviewed (Pinheiro et al. 2011).

The omic research approach may be useful to identify essential virulence genes that can be targeted by RNA interference (RNAi), thereby silencing the phenotypic effect of these genes and negatively effecting the survival of the parasite (McGonigle et al. 2008; McVeigh et al. 2014; Hagen et al. 2015). Pereira et al. (2008) have shown that injection with small interfering RNAs (siRNAs) (designed and produced against HGPRTase gene of the parasite)

through tail vein of *S. mansoni* infected mice leads to a reduction in worm burden (approximately 27 %, compared to control) and proposed that the treatment could be improved by molecule delivery and siRNA dose. Recently, Cao et al. (2014) showed that injection of siRNAs against lethal giant larvae (Lg1) gene of *S. japonicum* through tail vein results in a reduction of up to 85 % in hatching rates of eggs for the parasite-infected mice. Apart from their gene silencing effect, siRNAs may also be used to stimulate the host immune system to enhance resistance to infectious organisms (Marques and Williams 2005; Gantier 2013). Overall, the RNAi approach is a promising tool to deal with infection for trematode parasites. However, more studies are necessary to develop commercial products which are effective and safe in trematode infections.

In addition to vaccine, drug or RNAi development, omic approaches show promise in the development/improvement of diagnostic kits via the identification of species-specific antigens to address cross-reactivity (van der Ree and Mutapi 2015). Fast and correct diagnosis of infection is important for rapid targeted treatment, reduced pathology and decreased prevalence of the parasite through reduced infectious potential.

As the expansion of omic techniques and technologies provide larger volumes of information, tunnelling this information from omic studies into a commercialised product requires several steps from a number of disciplines, including research, development and commercialisation experts. Whilst individual researchers, or ad hoc research groups, may not have the resources or time to enter into the commercialisation of products, it is the responsibility of first stage researcher to ensure that research into strategies for parasitic control are carried out mindful of the practical applications required of such a strategy. Such focused goals are needed to address the mounting challenges in animal and human health (Hood et al. 2012) and the application of technical advances, such as omic technologies, should be done appropriately (McShane et al. 2013). Integration of industrial, commercial and discovery research partners, through intellectual property rights must be regarded as the most likely approach to drug and vaccine discovery, which should be reflected in project plans from the outset (Guerrero et al. 2012).

### Challenges and perspectives for omic-scale studies in trematode research

One of the major difficulties in omic-based trematode research is the identification capacity of LC-MS/MS. Recent LC-MS/MS systems without sample pre-fractionation cannot reach a protein identification number greater than 10,000 (Thakur et al. 2011; Nagaraj et al. 2012; Branca et al. 2014). Advances in LC-MS/MS technology, for the identification

of genome-size proteomes of higher eukaryotes including trematodes, without a pre-fractionation step will be an important future development in simplifying the wet lab procedure.

To date, in general terms, omic-based strategies have been applied to a single trematode organism isolated from a chosen location. However, genetic/proteomic heterogeneity has been shown to influence drug resistance (such as the case for *F. hepatica*) (Chemale et al. 2010). In the future, omic-based strategies should focus on deciphering genetic differences among the variants within a trematode species, as the heterogeneity could define the effectiveness of treatment/prevention strategies. These strategies will be significant in the development of multi-valent vaccines for the variants within a species.

The other difficulty is the annotation for some of newly found trematode sequences from omic-based studies. The comparative BLAST analysis of the sequences using public databases does not always provide significant hits. Therefore, these sequences need be defined further. Additionally, a significant number of trematode sequences are called 'unknown protein' in the databases, which remain to be studied further for functional annotation. Developments in bioinformatic approaches may help to solve this problem.

The existing difficulties in the development of preventive vaccines against the trematode parasites could be summarised as follows: lack of full protective effect, undesired inflammation of the antigen/adjuvant combinations, and individual diversity in both parasite and host in terms of magnitude of immune responses leading to difficulty in assessing the effectiveness of the vaccine (Haçarız et al. 2009a, b; McNeilly and Nisbet 2014). Omic-based studies will eventually be helpful to enhance understanding of the biological characteristics of both parasite and host, at individual level and host-pathogen interaction, which will guide the development of effective preventive vaccines (Molina-Hernández et al. 2015). This will be addressed in over a longer time period. Nevertheless, the omic-based results would contribute to the development of other control approaches such as the use of therapeutics within a shorter timescale.

Each promising experimental vaccine is generally proposed for use as a preventative strategy for a single disease in the field of parasitology. As the feasibility of this 'one vaccine - one disease' strategy comes under increasing financial scrutiny, the use of omic-based studies, which can identify important proteins for the development of combination vaccines, could incentivise the commercialisation of such broad spectrum vaccines.

### Conclusions

Five years ago, a large number of genes/proteins from omic-based studies were only reported for *Schistosoma* spp., particularly *S. mansoni*, in the field trematodology.

The omic-based strategy has led to the discovery of molecules (e.g. 28 kDa glutathione S-transferase, protein kinase C, extracellular signal-regulated kinase, SjLDs fusion proteins) which are now proposed to be promising for the immunotherapy of schistosomiasis and/or transmission of the disease (Knudsen et al. 2005; Riveau et al. 2012; You et al. 2012; You and McManus 2015; Ressurreição et al. 2014; Walker et al. 2014; van der Ree and Mutapi 2015). We now have a large volume of new data for many other trematode species. However, results from the omic-based studies for these genera are relevantly new, compared to *S. mansoni* and much newer in comparison with most of infectious microorganisms. With well-designed experiments, LC-MS/MS will identify essential parasite proteins at peptide level and open avenues to better understand parasite biology, including virulence mechanisms, parasite protein interactions (proteins that are correlated in expression or additive in their effect), interactions between the parasite and its host, and identifying better protective strategies such as vaccine and therapeutic approaches such as drug or RNAi.

Integrated approaches, with bioinformatic tools in particular, will be helpful to strategically identify particular proteins of interest from a whole trematode proteome which can then be used as protective commercialised product, as in the case of the meningococcal vaccine (Heinson et al. 2015).

To date, metabolome profiling of bacteria has led to enhanced understanding antibiotic-resistance and discovery of novel drug targets (Birkenstock et al. 2012; du Preez and Loots 2012; Milshteyn et al. 2014). Compared to genomic and proteomic profiling of trematodes, omic-scale metabolome profiling of these organisms is still at its infancy. Understanding the anthelmintic resistance to commonly used drugs and finding alternative drug targets would be the first target in the trematode research in applying metabolome analyses. We foresee that integration of metabolome information with genome and proteome data will be possible in the future and enlighten our understanding of trematode biology.

Overall, it is reasonable to assume that omic approaches are key to providing insights into the biological characteristics of the parasite, and will contribute to the development of innovative and effective strategies against trematodes, with the possibility of helminth-based therapeutic strategies for autoimmune-related diseases. The development of rapid, effective and accurate procedures to allow researchers to pinpoint proteins of interest from the trematode proteome pool will be key to maximising the benefits of this omic era. Based on previous success of omic-based studies in different fields, we expect that results from the relevant studies will turn into useful products to deal with trematode infections in the foreseeable future.

## References

- Acosta D, Cancela M, Piacenza L, Roche L, Carmona C, Tort JF (2008) *Fasciola hepatica* leucine aminopeptidase, a promising candidate for vaccination against ruminant fasciolosis. *Mol Biochem Parasitol* 158:52–64
- Alonso A, Marsal S, Julià A (2015) Analytical methods in untargeted metabolomics: state of the art in 2015. *Front Bioeng Biotechnol* 3:23
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Alzate O (2010) *Neuroproteomics*. CRC Press/Taylor & Francis, Boca Raton (FL)
- Andberg M, Jääntti J, Heilimo S, Pihkala P, Paananen A, Koskinen AM, Söderlund H, Linder MB (2007) Cleavage of recombinant proteins at poly-His sequences by Co(II) and Cu(II). *Protein Sci* 16:1751–1761
- Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H (2012) ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res* 40:W597–W603
- 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 (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25:25–29
- Becker S, Kortz L, Helmschrodt C, Thiery J, Ceglarek U (2012) LC-MS-based metabolomics in the clinical laboratory. *J Chromatogr B Anal Technol Biomed Life Sci* 883–884:68–75
- Bernal D, Trellis M, Montaner S, Cantalapiedra F, Galiano A, Hackenberg M, Marcilla A (2014) Surface analysis of *Dicrocoelium dendriticum*. The molecular characterization of exosomes reveals the presence of miRNAs. *J Proteomics* 105:232–241
- Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC, Mashiyama ST, Al-Lazikani B, Andrade LF, Ashton PD, Aslett MA, Bartholomeu DC, Blandin G, Caffrey CR, Coghlan A, Coulson R, Day TA, Delcher A, DeMarco R, Djikeng A, Eyre T, Gamble JA, Ghedin E, Gu Y, Hertz-Fowler C, Hirai H, Hirai Y, Houston R, Ivens A, Johnston DA, Lacerda D, Macedo CD, McVeigh P, Ning Z, Oliveira G, Overington JP, Parkhill J, Perteau M, Pierce RJ, Protasio AV, Quail MA, Rajandream MA, Rogers J, Sajid M, Salzberg SL, Stanke M, Tivey AR, White O, Williams DL, Wortman J, Wu W, Zamanian M, Zerlotini A, Fraser-Liggett CM, Barrell BG, El-Sayed NM (2009) The genome of the blood fluke *Schistosoma mansoni*. *Nature* 460:352–358
- Birkenstock T, Liebeke M, Winstel V, Krismer B, Gekeler C, Niemiec MJ, Bisswanger H, Lalk M, Peschel A (2012) Exometabolome analysis identifies pyruvate dehydrogenase as a target for the antibiotic triphenylbismuthdichloride in multiresistant bacterial pathogens. *J Biol Chem* 287:2887–2895
- Biswal DK, Ghatani S, Shylla JA, Sahu R, Mullapudi N, Bhattacharya A, Tandon V (2013) An integrated pipeline for next generation sequencing and annotation of the complete mitochondrial genome of the giant intestinal fluke, *Fasciolopsis buski* (Lankester, 1857) Looss, 1899. *Peer J* 1:e207
- Branca RM, Orre LM, Johansson HJ, Granholm V, Huss M, Pérez-Bercoff Á, Forshed J, Käll L, Lehtiö J (2014) HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. *Nat Methods* 11:59–62
- Braschi S, Borges WC, Wilson RA (2006a) Proteomic analysis of the schistosome tegument and its surface membranes. *Mem Inst Oswaldo Cruz* 101:205–212

- Braschi S, Curwen RS, Ashton PD, Verjovski-Almeida S, Wilson A (2006b) The tegument surface membranes of the human blood parasite *Schistosoma mansoni*: a proteomic analysis after differential extraction. *Proteomics* 6:1471–1482
- Brockwell YM, Elliott TP, Anderson GR, Stanton R, Spithill TW, Sangster NC (2013) Confirmation of *Fasciola hepatica* resistant to triclabendazole in naturally infected Australian beef and dairy cattle. *Int J Parasitol Drugs Drug Resist* 4:48–54
- Cantacessi C, Campbell BE, Jex AR, Young ND, Hall RS, Ranganathan S, Gasser RB (2012a) Bioinformatics meets parasitology. *Parasite Immunol* 34:265–275
- Cantacessi C, Mulvenna J, Young ND, Kasny M, Horak P, Aziz A, Hofmann A, Loukas A, Gasser RB (2012b) A deep exploration of the transcriptome and “excretory/secretory” proteome of adult *Fascioloides magna*. *Mol Cell Proteomics* 11:1340–1353
- Cao Y, Shi Y, Qiao H, Yang Y, Liu J, Shi Y, Lin J, Zhu G, Jin Y (2014) Distribution of lethal giant larvae (Lgl) protein in the tegument and negative impact of siRNA-based gene silencing on worm surface structure and egg hatching in *Schistosoma japonicum*. *Parasitol Res* 113:1–9
- Cass B, Pham PL, Kamen A, Durocher Y (2005) Purification of recombinant proteins from mammalian cell culture using a generic double-affinity chromatography scheme. *Protein Expr Purif* 40:77–85
- Chemale G, Perally S, LaCourse EJ, Prescott MC, Jones LM, Ward D, Meaney M, Hoey E, Brennan GP, Fairweather I, Trudgett A, Brophy PM (2010) Comparative proteomic analysis of triclabendazole response in the liver fluke *Fasciola hepatica*. *J Proteome Res* 9:4940–4951
- Chen G, Pramanik BN (2009) Application of LC/MS to proteomics studies: current status and future prospects. *Drug Discov Today* 14:465–471
- Chen JH, Zhang T, Ju C, Xu B, Lu Y, Mo XJ, Chen SB, Fan YT, Hu W, Zhou XN (2014) An integrated immunoproteomics and bioinformatics approach for the analysis of *Schistosoma japonicum* tegument proteins. *J Proteomics* 98:289–299
- Cwiklinski K, Dalton JP, Dufresne PJ, La Course J, Williams DJ, 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
- Dalton JP, Robinson MW, Mulcahy G, O’Neill SM, Donnelly S (2013) Immunomodulatory molecules of *Fasciola hepatica*: candidates for both vaccine and immunotherapeutic development. *Vet Parasitol* 195:272–285
- de Hoffmann E (1996) Tandem mass spectrometry: a primer. *J Mass Spectrom* 31:129–137
- De la Torre Escudero E, Manzano-Román R, Valero L, Oleaga A, Pérez-Sánchez R, Hernández-González A, Siles-Lucas M (2011) Comparative proteomic analysis of *Fasciola hepatica* juveniles and *Schistosoma bovis* schistosomula. *J Proteomics* 74:1534–1544
- de la Torre-Escudero E, Manzano-Román R, Pérez-Sánchez R, Barrera I, Siles-Lucas M, Oleaga A (2012) Molecular cloning, characterization and diagnostic performance of the *Schistosoma bovis* 22.6antigen. *Vet Parasitol* 190:530–540
- de la Torre-Escudero E, Pérez-Sánchez R, Manzano-Román R, Oleaga A (2013) In vivo intravascular biotinylation of *Schistosoma bovis* adult worms and proteomic analysis of tegumental surface proteins. *J Proteomics* 94:513–526
- DeMarco R, Verjovski-Almeida S (2009) Schistosomes—proteomics studies for potential novel vaccines and drug targets. *Drug Discov Today* 14:472–478
- du Preez I, Loots DT (2012) Altered fatty acid metabolism due to rifampicin-resistance conferring mutations in the rpoB Gene of *Mycobacterium tuberculosis*: mapping the potential of pharmacometabolomics for global health and personalized medicine. *OMICS* 16:596–603
- Elliott TP, Kelley JM, Rawlin G, Spithill TW (2015) High prevalence of fasciolosis and evaluation of drug efficacy against *Fasciola hepatica* in dairy cattle in the Maffra and Bairnsdale districts of Gippsland, Victoria, Australia. *Vet Parasitol* 209:117–124
- Endo Y, Sawasaki T (2006) Cell-free expression systems for eukaryotic protein production. *Curr Opin Biotechnol* 17:373–380
- Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8:186–194
- Fischer F, Wolters D, Rögner M, Poetsch A (2006) Toward the complete membrane proteome: high coverage of integral membrane proteins through transmembrane peptide detection. *Mol Cell Proteomics* 5:444–453
- Flynn RJ, Mannion C, Golden O, Hacariz O, Mulcahy G (2007) Experimental *Fasciola hepatica* infection alters responses to tests used for diagnosis of bovine tuberculosis. *Infect Immun* 75:1373–1381
- Gantier MP (2013) Strategies for designing and validating immunostimulatory siRNAs. *Methods Mol Biol* 942:179–191
- Gao Z, Li H, Zhang H, Liu X, Kang L, Luo X, Zhu W, Chen K, Wang X, Jiang H (2008) PDTD: a web-accessible protein database for drug target identification. *BMC Bioinformatics* 9:104
- Garg G, Ranganathan S (2011) *In silico* secretome analysis approach for next generation sequencing transcriptomic data. *BMC Genomics* 12(Suppl 3):S14
- Garg G, Ranganathan S (2012) Helminth secretome database (HSD): a collection of helminth excretory/secretory proteins predicted from expressed sequence tags (ESTs). *BMC Genomics* 13(Suppl 7):S8
- Garg G, Bernal D, Trelis M, Forment J, Ortiz J, Valero ML, Pedrola L, Martinez-Blanch J, Esteban JG, Ranganathan S, Toledo R, Marcilla A (2013) The transcriptome of *Echinostoma caproni* adults: further characterization of the secretome and identification of new potential drug targets. *J Proteomics* 89:202–214
- Gene Ontology Consortium (2015) Gene ontology consortium: going forward. *Nucleic Acids Res* 43:D1049–D1056
- Gingras AC, Aebersold R, Raught B (2005) Advances in protein complex analysis using mass spectrometry. *J Physiol* 563:11–21
- Gobert GN, McManus DP, Nawaratna S, Moertel L, Mulvenna J, Jones MK (2009) Tissue specific profiling of females of *Schistosoma japonicum* by integrated laser microdissection microscopy and microarray analysis. *PLoS Negl Trop Dis* 3:e469
- Gobert GN, You H, McManus DP (2014) Gaining biological perspectives from schistosome genomes. *Mol Biochem Parasitol* 196:21–28
- Golden O, Flynn RJ, Read C, Sekiya M, Donnelly SM, Stack C, Dalton JP, Mulcahy G (2010) Protection of cattle against a natural infection of *Fasciola hepatica* by vaccination with recombinant cathepsin L1 (rFhCL1). *Vaccine* 28:5551–5557
- Gonzales Santana B, Dalton JP, Vasquez Camargo F, Parkinson M, Ndao M (2013) The diagnosis of human fascioliasis by enzyme-linked immunosorbent assay (ELISA) using recombinant cathepsin L protease. *PLoS Negl Trop Dis* 7:e2414
- Gourbal BE, Guillou F, Mitta G, Sibille P, Théron A, Pointier JP, Coustau C (2008) Excretory-secretory products of larval *Fasciola hepatica* investigated using a two-dimensional proteomic approach. *Mol Biochem Parasitol* 161:63–66
- Grebe SK, Singh RJ (2011) LC-MS/MS in the clinical laboratory—where to from here? *Clin Biochem Rev* 32:5–31
- Guerrero FD, Miller RJ, Pérez de León AA (2012) Cattle tick vaccines: many candidate antigens, but will a commercially viable product emerge? *Int J Parasitol* 42:421–427
- Guillou F, Roger E, Moné Y, Rognon A, Grunau C, Théron A, Mitta G, Coustau C, Gourbal BE (2007) Excretory-secretory proteome of larval *Schistosoma mansoni* and *Echinostoma caproni*, two parasites of *Biomphalaria glabrata*. *Mol Biochem Parasitol* 155:45–56
- Guo AC, Jewison T, Wilson M, Liu Y, Knox C, Djoumbou Y, Lo P, Mandal R, Krishnamurthy R, Wishart DS (2013) ECMDB: the *E. coli* metabolome database. *Nucleic Acids Res* 41:D625–D630
- Gupta N, Tanner S, Jaitly N, Adkins JN, Lipton M, Edwards R, Romine M, Osterman A, Bafna V, Smith RD, Pevzner PA (2007) Whole



- proteome analysis of post-translational modifications: applications of mass-spectrometry for proteogenomic annotation. *Genome Res* 17:1362–1377
- Haçarız O, Baykal AT (2014) Investigation of the abundance of proteins secreted by *Fasciola hepatica*, which is exposed to environmental change in experimental studies, with an advanced proteomic approach. *Turk Parazitol Derg* 38:106–110
- Haçarız O, Sayers G (2013) *Fasciola hepatica*—where is 28S ribosomal RNA? *Exp Parasitol* 135:426–429
- Haçarız O, Sayers G, Flynn RJ, Lejeune A, Mulcahy G (2009a) IL-10 and TGF-beta 1 are associated with variations in fluke burdens following experimental fasciolosis in sheep. *Parasite Immunol* 31:613–622
- Haçarız O, Sayers G, McCullough M, Garrett M, O'Donovan J, Mulcahy G (2009b) The effect of Quil A adjuvant on the course of experimental *Fasciola hepatica* infection in sheep. *Vaccine* 27:45–50
- Haçarız O, Sayers G, Mulcahy G (2011) A preliminary study to understand the effect of *Fasciola hepatica* tegument on naïve macrophages and humoral responses in an ovine model. *Vet Immunol Immunopathol* 139:245–249
- Haçarız O, Sayers G, Baykal AT (2012) A proteomic approach to investigate the distribution and abundance of surface and internal *Fasciola hepatica* proteins during the chronic stage of natural liver fluke infection in cattle. *J Proteome Res* 11:3592–3604
- Haçarız O, Baykal AT, Akgün M, Kavak P, Sağıroğlu MŞ, Sayers GP (2014) Generating a detailed protein profile of *Fasciola hepatica* during the chronic stage of infection in cattle. *Proteomics* 14:1519–1530
- Haçarız O, Akgün M, Kavak P, Yüksel B, Sağıroğlu MŞ (2015) Comparative transcriptome profiling approach to glean virulence and immunomodulation-related genes of *Fasciola hepatica*. *BMC Genomics* 16:366
- Hagen J, Scheerlinck JP, Gasser RB (2015) Knocking down schistosomes—promise for lentiviral transduction in parasites. *Trends Parasitol* 31:324–332
- Harris TW, Baran J, Bieri T, Cabunoc A, Chan J, Chen WJ, Davis P, Done J, Grove C, Howe K, Kishore R, Lee R, Li Y, Muller HM, Nakamura C, Ozersky P, Paulini M, Raciti D, Schindelman G, Tuli MA, Van Auken K, Wang D, Wang X, Williams G, Wong JD, Yook K, Schedl T, Hodgkin J, Berriman M, Kersey P, Spieth J, Stein L, Sternberg PW (2014) WormBase 2014: new views of curated biology. *Nucleic Acids Res* 42:D789–D793
- Hasby EA, Hasby Saad MA, Shohieb Z, El Noby K (2015) FoxP3+ T regulatory cells and immunomodulation after *Schistosoma mansoni* egg antigen immunization in experimental model of inflammatory bowel disease. *Cell Immunol* 295:67–76
- He Y, Racz R, Sayers S, Lin Y, Todd T, Hur J, Li X, Patel M, Zhao B, Chung M, Ostrow J, Sylora A, Dungarani P, Ulysse G, Kochhar K, Vidri B, Strait K, Jourdan GW, Xiang Z (2014) Updates on the web-based VIOLIN vaccine database and analysis system. *Nucleic Acids Res* 42:D1124–D1132
- Heinson AI, Woelk CH, Newell ML (2015) The promise of reverse vaccinology. *Int Health* 7:85–89
- Helmy H (2015) Human helminth therapy to treat inflammatory disorders—where do we stand? *BMC Immunol* 16:12
- Hernández Alvarez L, Naranjo Feliciano D, Hernández González JE, Soares RO, Barreto Gomes DE, Pascutti PG (2015) Insights into the interactions of *Fasciola hepatica* Cathepsin L3 with a substrate and potential novel inhibitors through *in silico* approaches. *PLoS Negl Trop Dis* 9:e0003759
- Hernández-González A, Valero ML, del Pino MS, Oleaga A, Siles-Lucas M (2010) Proteomic analysis of *in vitro* newly excysted juveniles from *Fasciola hepatica*. *Mol Biochem Parasitol* 172:121–128
- Heylen M, Ruysers NE, Nullens S, Schramm G, Pelckmans PA, Moreels TG, De Man JG, De Winter BY (2015) Treatment with egg antigens of *Schistosoma mansoni* ameliorates experimental colitis in mice through a colonic T-cell-dependent mechanism. *Inflamm Bowel Dis* 21:48–59
- Higón M, Cowan G, Nausch N, Cavanagh D, Oleaga A, Toledo R, Stothard JR, Antúnez O, Marcilla A, Burchmore R, Mutapi F (2011) Screening trematodes for novel intervention targets: a proteomic and immunological comparison of *Schistosoma haematobium*, *Schistosoma bovis* and *Echinostoma caproni*. *Parasitology* 138:1607–1619
- Hodgkinson J, Cwiklinski K, Beesley NJ, Paterson S, Williams DJ (2013) Identification of putative markers of triclabendazole resistance by a genome-wide analysis of genetically recombinant *Fasciola hepatica*. *Parasitology* 140:1523–1533
- Hong Y, Sun A, Zhang M, Gao F, Han Y, Fu Z, Shi Y, Lin J (2013) Proteomics analysis of differentially expressed proteins in schistosomula and adult worms of *Schistosoma japonicum*. *Acta Trop* 126:1–10
- Hood LE, Omenn GS, Moritz RL, Aebersold R, Yamamoto KR, Amos M, Hunter-Cevera J, Locascio L, Workshop Participants (2012) New and improved proteomics technologies for understanding complex biological systems: addressing a grand challenge in the life sciences. *Proteomics* 12:2773–2783
- Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35:W585–W587
- Hotez PJ, Bethony JM, Diemert DJ, Pearson M, Loukas A (2010) Developing vaccines to combat hookworm infection and intestinal schistosomiasis. *Nat Rev Microbiol* 8:814–826
- Huang TC, Pinto SM, Pandey A (2013a) Proteomics for understanding miRNA biology. *Proteomics* 13:558–567
- 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, Zhou J, Hu Y, Li R, Zhang F, Lei H, Li W, Hu X, Liang C, Xu J, Li X, Yu X (2013b) The carcinogenic liver fluke, *Clonorchis sinensis*: new assembly, reannotation and analysis of the genome and characterization of tissue transcriptomes. *PLoS One* 8:e54732
- Hughes RA, Miklos AE, Ellington AD (2011) Gene synthesis: methods and applications. *Methods Enzymol* 498:277–309
- Hurst LD (2002) The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet* 18:486
- Ishikawa H, Newburgh RW (1972) Studies of the thermal conversion of 28 S RNA of *Galleria mellonella* (L.) to an 18 S product. *J Mol Biol* 64:135–144
- Japa O, Hodgkinson JE, Emes RD, Flynn RJ (2015) TGF- $\beta$  superfamily members from the helminth *Fasciola hepatica* show intrinsic effects on viability and development. *Vet Res* 46:29
- Jefferies JR, Campbell AM, van Rossum AJ, Barrett J, Brophy PM (2001) Proteomic analysis of *Fasciola hepatica* excretory-secretory products. *Proteomics* 1:1128–1132
- Jewison T, Knox C, Neveu V, Djoumbou Y, Guo AC, Lee J, Liu P, Mandal R, Krishnamurthy R, Sinelnikov I, Wilson M, Wishart DS (2012) YMDB: the yeast metabolome database. *Nucleic Acids Res* 40:D815–D820
- Jia X, Schulte L, Loukas A, Pickering D, Pearson M, Mobli M, Jones A, Rosengren KJ, Daly NL, Gobert GN, Jones MK, Craik DJ, Mulvenna J (2014) Solution structure, membrane interactions, and protein binding partners of the tetraspanin Sm-TSP-2, a vaccine antigen from the human blood fluke *Schistosoma mansoni*. *J Biol Chem* 289:7151–7163
- Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240
- Kandpal R, Saviola B, Felton J (2009) The era of 'omics unlimited. *Biotechniques* 46:351–355
- Kelly SM, Price NC (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr Protein Pept Sci* 1:349–384

- Khan YA, Umar S, Abidi SM (2015) Somatic antigens of tropical liver flukes ameliorate collagen-induced arthritis in wistar rats. *PLoS One* 10:e0126429
- Khow O, Suntrarachun S (2012) Strategies for production of active eukaryotic proteins in bacterial expression system. *Asian Pac J Trop Biomed* 2:159–162
- Kim D, Perteu G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36
- Kim DW, Yoo WG, Lee S, Lee MR, Kim YJ, Cho SH, Lee WJ, Ju JW (2014) ClonorESTdb: a comprehensive database for *Clonorchis sinensis* EST sequences. *BMC Res Notes* 7:388
- Knudsen GM, Medzihradsky KF, Lim KC, Hansell E, McKerrow JH (2005) Proteomic analysis of *Schistosoma mansoni* cercarial secretions. *Mol Cell Proteomics* 4:1862–1875
- Lahiri C, Pawar S, Sabarinathan R, Ashraf MI, Chand Y, Chakravorty D (2014) Interactome analyses of *Salmonella* pathogenicity islands reveal SicA indispensable for virulence. *J Theor Biol* 363:188–197
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359
- Larance M, Lamond AI (2015) Multidimensional proteomics for cell biology. *Nat Rev Mol Cell Biol* 16:269–280
- Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, Maciejewski A, Amdt D, Wilson M, Neveu V, Tang A, Gabriel G, Ly C, Adamjee S, Dame ZT, Han B, Zhou Y, Wishart DS (2014) DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* 42:D1091–D1097
- Leblanc P, Moise L, Luza C, Chantalarawan K, Lezeau L, Yuan J, Field M, Richer D, Boyle C, Martin WD, Fishman JB, Berg EA, Baker D, Zeigler B, Mais DE, Taylor W, Coleman R, Warren HS, Gelfand JA, De Groot AS, Brauns T, Poznansky MC (2014) VaxCelerate II: rapid development of a self-assembling vaccine for Lassa fever. *Hum Vaccin Immunother* 10:3022–3038
- Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323
- Li H, Gao Z, Kang L, Zhang H, Yang K, Yu K, Luo X, Zhu W, Chen K, Shen J, Wang X, Jiang H (2006) TarFisDock: a web server for identifying drug targets with docking approach. *Nucleic Acids Res* 34:W219–W224
- Li JV, Saric J, Wang Y, Keiser J, Utzinger J, Holmes E (2011) Chemometric analysis of biofluids from mice experimentally infected with *Schistosoma mansoni*. *Parasit Vectors* 4:179
- Littlechild J, Malcolm A, Paterakis K, Ackermann I, Dijk J (1987) The tertiary structure of salt-extracted ribosomal proteins from *Escherichia coli* as studied by protonmagnetic resonance spectroscopy and limited proteolysis experiments. *Biochim Biophys Acta* 913:245–255
- Liu T, Qian WJ, Mottaz HM, Gritsenko MA, Norbeck AD, Moore RJ, Purvine SO, Camp DG 2nd, Smith RD (2006) Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol Cell Proteomics* 5:2167–2174
- Liu F, Cui SJ, Hu W, Feng Z, Wang ZQ, Han ZG (2009) Excretory/secretory proteome of the adult developmental stage of human blood fluke, *Schistosoma japonicum*. *Mol Cell Proteomics* 8:1236–1251
- Logan-Klumpler FJ, De Silva N, Boehme U, Rogers MB, Velarde G, McQuillan JA, Carver T, Aslett M, Olsen C, Subramanian S, Phan I, Farris C, Mitra S, Ramasamy G, Wang H, Tivey A, Jackson A, Houston R, Parkhill J, Holden M, Harb OS, Brunk BP, Myler PJ, Roos D, Carrington M, Smith DF, Hertz-Fowler C, Berriman M (2012) GeneDB—an annotation database for pathogens. *Nucleic Acids Res* 40:D98–D108
- Loukas A, Gaze S, Mulvenna JP, Gasser RB, Brindley PJ, Doolan DL, Bethony JM, Jones MK, Gobert GN, Driguez P, McManus DP, Hotez PJ (2011) Vaccinomics for the major blood feeding helminths of humans. *OMICS* 15:567–577
- Ludolf F, Patrocínio PR, Corrêa-Oliveira R, Gazzinelli A, Falcone FH, Teixeira-Ferreira A, Perales J, Oliveira GC, Silva-Pereira RA (2014) Serological screening of the *Schistosoma mansoni* adult worm proteome. *PLoS Negl Trop Dis* 8:e2745
- Marcelino I, de Almeida AM, Ventosa M, Pruneau L, Meyer DF, Martinez D, Lefrançois T, Vachiéry N, Coelho AV (2012) Tick-borne diseases in cattle: applications of proteomics to develop new generation vaccines. *J Proteomics* 75:4232–4250
- Mardis ER (2013) Next-generation sequencing platforms. *Annu Rev Anal Chem (Palo Alto, Calif)* 6:287–303
- Marques JT, Williams BR (2005) Activation of the mammalian immune system by siRNAs. *Nat Biotechnol* 23:1399–1405
- Martin J, Rosa BA, Ozersky P, Hallsworth-Pepin K (2015) Helminth.net: expansions to Nematode.net and an introduction to Trematode.net. *Nucleic Acids Res* 43:D698–D706
- Martínez-Ibeas AM, González-Lanza C, Manga-González MY (2013a) Proteomic analysis of the tegument and excretory-secretory products of *Dicrocoelium dendriticum* (Digenea) adult worms. *Exp Parasitol* 133:411–420
- Martínez-Ibeas AM, Perteguer MJ, González-Lanza C, Gárate T, Manga-González MY (2013b) Analysis of an expressed sequence tag library from *Dicrocoelium dendriticum*. *Exp Parasitol* 135:287–296
- Mas-Coma S, Bargues MD, Valero MA (2005) Fascioliasis and other plant-borne trematode zoonoses. *Int J Parasitol* 35:1255–1278
- McGonigle L, Mousley A, Marks NJ, Brennan GP, Dalton JP, Spithill TW, Day TA, Maule AG (2008) The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduces gut penetration. *Int J Parasitol* 38:149–155
- McMaster MC (2005) LC/MS: a practical user's guide. John Wiley & Sons Inc, Hoboken
- McNeilly TN, Nisbet AJ (2014) Immune modulation by helminth parasites of ruminants: implications for vaccine development and host immune competence. *Parasite* 21:51
- McNulty SN, Fischer PU, Townsend RR, Curtis KC, Weil GJ, Mitreva M (2014) Systems biology studies of adult paragonimus lung flukes facilitate the identification of immunodominant parasite antigens. *PLoS Negl Trop Dis* 8:e3242
- McShane LM, Cavenagh MM, Lively TG, Eberhard DA, Bigbee WL, Williams PM, Mesirov JP, Polley MY, Kim KY, Tricoli JV, Taylor JM, Shuman DJ, Simon RM, Doroshov JH, Conley BA (2013) Criteria for the use of omics-based predictors in clinical trials: explanation and elaboration. *BMC Med* 11:220
- McVeigh P, McCammick EM, McCusker P, Mophew RM, Mousley A, Abidi A, Saifullah KM, Muthusamy R, Gopalakrishnan R, Spithill TW, Dalton JP, Brophy PM, Marks NJ, Maule AG (2014) RNAi dynamics in Juvenile *Fasciola* spp. Liver flukes reveals the persistence of gene silencing *in vitro*. *PLoS Negl Trop Dis* 8:e3185
- Mickum ML, Prasanphanich NS, Heimburg-Molinaro J, Leon KE, Cummings RD (2014) Deciphering the glycogenome of schistosomes. *Front Genet* 5:262
- Milshteyn A, Schneider JS, Brady SF (2014) Mining the metabiome: identifying novel natural products from microbial communities. *Chem Biol* 21:1211–1223
- Molina-Hernández V, Mulcahy G, Pérez J, Martínez-Moreno Á, Donnelly S, O'Neill SM, Dalton JP, Cwiklinski K (2015) *Fasciola hepatica* vaccine: we may not be there yet but we're on the right road. *Vet Parasitol* 208:101–111
- Mophew RM, Wright HA, LaCourse EJ, Woods DJ, Brophy PM (2007) Comparative proteomics of excretory-secretory proteins released by the liver fluke *Fasciola hepatica* in sheep host bile and during *in vitro* culture ex host. *Mol Cell Proteomics* 6:963–972
- Mophew RM, Wright HA, Lacourse EJ, Porter J, Barrett J, Woods DJ, Brophy PM (2011) Towards delineating functions within the fasciola secreted cathepsin L protease family by integrating *in vivo*

- based sub-proteomics and phylogenetics. *PLoS Negl Trop Dis* 5: e937
- Morphew RM, Eccleston N, Wilkinson TJ, McGarry J, Perally S, Prescott M, Ward D, Williams D, Paterson S, Raman M, Ravikumar G, Khalid Saifullah M, Abbas Abidi SM, McVeigh P, Maule AG, Brophy PM, LaCourse EJ (2012) Proteomics and *in silico* approaches to extend understanding of the glutathione transferase superfamily of the tropical liver fluke *Fasciola gigantica*. *J Proteome Res* 11:5876–5889
- Morphew RM, Hamilton CM, Wright HA, Dowling DJ, O'Neill SM, Brophy PM (2013) Identification of the major proteins of an immune modulating fraction from adult *Fasciola hepatica* released by Nonidet P40. *Vet Parasitol* 191:379–385
- Mulvenna J, Moertel L, Jones MK, Nawaratna S, Lovas EM, Gobert GN, Colgrave M, Jones A, Loukas A, McManus DP (2010a) Exposed proteins of the *Schistosoma japonicum* tegument. *Int J Parasitol* 40: 543–554
- Mulvenna J, Srija B, Brindley PJ, Gorman J, Jones MK, Colgrave ML, Jones A, Nawaratna S, Laha T, Suttiprapa S, Smout MJ, Loukas A (2010b) The secreted and surface proteomes of the adult stage of the carcinogenic human liver fluke *Opisthorchis viverrini*. *Proteomics* 10:1063–1078
- Nagaraj N, Kulak NA, Cox J, Neuhauser N, Mayr K, Hoerning O, Vorm O, Mann M (2012) System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Mol Cell Proteomics* 11:M111.013722
- Ng PC, Kirkness EF (2010) Whole genome sequencing. *Methods Mol Biol* 628:215–226
- Nóbrega de Sousa T, de Menezes Neto A, Alves de Brito CF (2013) “Omics” in the study of the major parasitic diseases malaria and schistosomiasis. *Infect Genet Evol* 19:258–273
- O'Sullivan C, Fried B, Sherma J (2013) Metabolic profiling of *Echinostoma caproni* and *Schistosoma mansoni* in their definitive and intermediate hosts. *Acta Parasitol* 58:1–5
- Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 27:29–34
- Ortiz P, Scarcella S, Cerna C, Rosales C, Cabrera M, Guzmán M, Lamenza P, Solana H (2013) Resistance of *Fasciola hepatica* against Triclabendazole in cattle in Cajamarca (Peru): a clinical trial and an *in vivo* efficacy test in sheep. *Vet Parasitol* 195:118–121
- Pedersen J, Lauritzen C, Madsen MT, Weis Dahl S (1999) Removal of N-terminal polyhistidine tags from recombinant proteins using engineered aminopeptidases. *Protein Expr Purif* 15:389–400
- Pereira TC, Pascoal VD, Marchesini RB, Maia IG, Magalhães LA, Zanotti-Magalhães EM, Lopes-Cendes I (2008) *Schistosoma mansoni*: evaluation of an RNAi-based treatment targeting HGPRTase gene. *Exp Parasitol* 118:619–623
- Pérez-Sánchez R, Ramajo-Hernández A, Ramajo-Martín V, Oleaga A (2006) Proteomic analysis of the tegument and excretory-secretory products of adult *Schistosoma bovis* worms. *Proteomics* 6(Suppl 1): S226–S236
- Pérez-Sánchez R, Valero ML, Ramajo-Hernández A, Siles-Lucas M, Ramajo-Martín V, Oleaga A (2008) A proteomic approach to the identification of tegumental proteins of male and female *Schistosoma bovis* worms. *Mol Biochem Parasitol* 161:112–123
- Pinheiro CS, Martins VP, Assis NR, Figueiredo BC, Morais SB, Azevedo V, Oliveira SC (2011) Computational vaccinology: an important strategy to discover new potential *S. mansoni* vaccine candidates. *J Biomed Biotechnol* 2011:503068
- Preidis GA, Hotez PJ (2015) The newest “omics”—metagenomics and metabolomics—enter the battle against the neglected tropical diseases. *PLoS Negl Trop Dis* 9:e0003382
- Qian MB, Utzinger J, Keiser J, Zhou XN (2015) Clonorchiasis. *Lancet* 387:800–810
- Ramajo-Hernández A, Oleaga A, Ramajo-Martín V, Pérez-Sánchez R (2007) Carbohydrate profiling and protein identification of tegumental and excreted/secreted glycoproteins of adult *Schistosoma bovis* worms. *Vet Parasitol* 144:45–60
- Rauh M (2012) LC-MS/MS for protein and peptide quantification in clinical chemistry. *J Chromatogr B Anal Technol Biomed Life Sci* 883–884:59–67
- Ressurreição M, De Saram P, Kirk RS, Rollinson D, Emery AM, Page NM, Davies AJ, Walker AJ (2014) Protein kinase C and extracellular signal-regulated kinase regulate movement, attachment, pairing and egg release in *Schistosoma mansoni*. *PLoS Negl Trop Dis* 8: e2924
- Rioux MC, Carmona C, Acosta D, Ward B, Ndao M, Gibbs BF, Bennett HP, Spithill TW (2008) Discovery and validation of serum biomarkers expressed over the first twelve weeks of *Fasciola hepatica* infection in sheep. *Int J Parasitol* 38:123–136
- Riveau G, Deplanque D, Remoué F, Schacht AM, Vodougnon H, Capron M, Thiry M, Martial J, Libersa C, Capron A (2012) Safety and immunogenicity of rSh28GST antigen in humans: phase 1 randomized clinical study of a vaccine candidate against urinary schistosomiasis. *PLoS Negl Trop Dis* 6:e1704
- Robb SM, Ross E, Sánchez Alvarado A (2008) SmedGD: the *Schmidtea mediterranea* genome database. *Nucleic Acids Res* 36:D599–D606
- Robb SM, Gotting K, Ross E, Sánchez Alvarado A (2015) SmedGD 2.0: the *Schmidtea mediterranea* genome database. *Genesis* 53:535–546
- Robinson MW, Menon R, Donnelly SM, Dalton JP, Ranganathan S (2009) An integrated transcriptomics and proteomics analysis of the secretome of the helminth pathogen *Fasciola hepatica*: proteins associated with invasion and infection of the mammalian host. *Mol Cell Proteomics* 8:1891–1907
- Robinson MW, Dalton JP, O'Brien BA, Donnelly S (2013) *Fasciola hepatica*: the therapeutic potential of a worm secretome. *Int J Parasitol* 43:283–291
- Sandin M, Chawade A, Levander F (2015) Is label-free LC-MS/MS ready for biomarker discovery? *Proteomics Clin Appl* 9:289–294
- Saric J, Li JV, Wang Y, Keiser J, Bundy JG, Holmes E, Utzinger J (2008) Metabolic profiling of an *Echinostoma caproni* infection in the mouse for biomarker discovery. *PLoS Negl Trop Dis* 2:e254
- Saric J, Want EJ, Duthaler U, Lewis M, Keiser J, Shockcor JP, Ross GA, Nicholson JK, Holmes E, Tavares MF (2012) Systematic evaluation of extraction methods for multiplatform-based metabolotyping: application to the *Fasciola hepatica* metabolome. *Anal Chem* 84:6963–6972
- Schaefer O, Ohtsuki S, Kawakami H, Inoue T, Liehner S, Saito A, Sakamoto A, Ishiguro N, Matsumaru T, Terasaki T, Ebner T (2012) Absolute quantification and differential expression of drug transporters, cytochrome P450 enzymes, and UDP-glucuronosyltransferases in cultured primary human hepatocytes. *Drug Metab Dispos* 40:93–103
- Schulz MH, Zerbino DR, Vingron M, Birney E (2012) Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28:1086–1092
- Seib KL, Zhao X, Rappuoli R (2012) Developing vaccines in the era of genomics: a decade of reverse vaccinology. *Clin Microbiol Infect* 18(Suppl 5):109–116
- Shi Y, Toet H, Rathinasamy V, Young ND, Gasser RB, Beddoe T, Huang W, Spithill TW (2014) First insight into CD59-like molecules of adult *Fasciola hepatica*. *Exp Parasitol* 144:57–64
- Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP (2014) Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet* 15:121–132
- Smith RM, Webb A, Papp AC, Newman LC, Handelman SK, Suhy A, Mascarenhas R, Oberdick J, Sadee W (2013) Whole transcriptome RNA-Seq allelic expression in human brain. *BMC Genomics* 14:571

- Sotillo J, Valero ML, Sánchez Del Pino MM, Fried B, Esteban JG, Marcilla A, Toledo R (2010) Excretory/secretory proteome of the adult stage of *Echinostoma caproni*. *Parasitol Res* 107:691–697
- Spencer RK, Kreutzer AG, Salveson PJ, Li H, Nowick JS (2015) X-ray crystallographic structures of oligomers of peptides derived from  $\beta$ 2-microglobulin. *J Am Chem Soc* 137:6304–6311
- Summers RW, Elliott DE, Urban JF Jr, Thompson RA, Weinstock JV (2005) *Trichuris suis* therapy for active ulcerative colitis: a randomized controlled trial. *Gastroenterology* 128:825–832
- Swaney DL, Wenger CD, Coon JJ (2010) Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *J Proteome Res* 9:1323–1329
- T009-Technical bulletin, NanoDrop, ThermoScientific, USA.
- Teimoori S, Arimatsu Y, Laha T, Kaewkes S, Sereerak P, Tangkawattana S, Brindley PJ, Sripa B (2015) Immunodiagnosis of opisthorchiasis using parasite cathepsin F. *Parasitol Res* 114:4571–4578
- Thakur SS, Geiger T, Chatterjee B, Bandilla P, Fröhlich F, Cox J, Mann M (2011) Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Mol Cell Proteomics* 10:M110.003699
- The Schistosoma japonicum Genome Sequencing and Functional Analysis Consortium (2009) The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature* 460:345–351
- Thiede B, Höhenwarter W, Krah A, Mattow J, Schmid M, Schmidt F, Jungblut PR (2005) Peptide mass fingerprinting. *Methods* 35:237–247
- Tian F, Hou M, Chen L, Gao Y, Zhang X, Ji M, Wu G (2013) Proteomic analysis of schistosomiasis japonica vaccine candidate antigens recognized by UV-attenuated cercariae-immunized porcine serum IgG2. *Parasitol Res* 112:2791–2803
- Toet H, Piedrafita DM, Spithill TW (2014) Liver fluke vaccines in ruminants: strategies, progress and future opportunities. *Int J Parasitol* 44:915–927
- Toledo R, Bernal MD, Marcilla A (2011) Proteomics of foodborne trematodes. *J Proteomics* 74:1485–1503
- Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25:1105–1111
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7:562–578
- Tsai IJ, Zarowiecki M, Holroyd N, Garcarrubio A, Sanchez-Flores A, Brooks KL, Tracey A, Bobes RJ, Fragoso G, Scitutto E, Aslett M, Beasley H, Bennett HM, Cai J, Camicia F, Clark R, Cucher M, De Silva N, Day TA, Deplazes P, Estrada K, Fernández C, Holland PW, Hou J, Hu S, Huckvale T, Hung SS, Kamenetzky L, Keane JA, Kiss F, Koziol U, Lambert O, Liu K, Luo X, Luo Y, Macchiaroli N, Nichol S, Paps J, Parkinson J, Pouchkina-Stantcheva N, Riddiford N, Rosenzvit M, Salinas G, Wasmuth JD, Zamanian M, Zheng Y, Taenia solium Genome Consortium, Cai X, Soberón X, Olson PD, Lacleite JP, Brehm K, Berriman M (2013) The genomes of four tapeworm species reveal adaptations to parasitism. *Nature* 496:57–63
- Valencia CA, Pervaiz MA, Husami A, Qian Y, Zhang K (2013) Next generation sequencing technologies in medical genetics. Springer, New York
- van der Ree AM, Mutapi F (2015) The helminth parasite proteome at the host-parasite interface—informing diagnosis and control. *Exp Parasitol* 157:48–58
- Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, Wheeler DK, Gabbard JL, Hix D, Sette A, Peters B (2015) The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* 43:D405–D412
- Walker AJ, Ressurreição M, Rothermel R (2014) Exploring the function of protein kinases in schistosomes: perspectives from the laboratory and from comparative genomics. *Front Genet* 5:229
- Walz Y, Wegmann M, Dech S, Raso G, Utzinger J (2015) Risk profiling of schistosomiasis using remote sensing: approaches, challenges and outlook. *Parasit Vectors* 8:163
- Wammes LJ, Mpairwe H, Elliott AM, Yazdanbakhsh M (2014) Helminth therapy or elimination: epidemiological, immunological, and clinical considerations. *Lancet Infect Dis* 14:1150–1162
- Wang P, Wilson SR (2013) Mass spectrometry-based protein identification by integrating de novo sequencing with database searching. *BMC Bioinformatics* 14(Suppl 2):S24
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63
- Wang Y, Li JV, Saric J, Keiser J, Wu J, Utzinger J, Holmes E (2010) Advances in metabolic profiling of experimental nematode and trematode infections. *Adv Parasitol* 73:373–404
- 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* 12:R107
- Webster J, Oxley D (2012) Protein identification by MALDI-TOF mass spectrometry. *Methods Mol Biol* 800:227–240
- Wilson RA, Wright JM, de Castro-Borges W, Parker-Manuel SJ, Dowle AA, Ashton PD, Young ND, Gasser RB, Spithill TW (2011) Exploring the *Fasciola hepatica* tegument proteome. *Int J Parasitol* 41:1347–1359
- Winnebeck EC, Millar CD, Warman GR (2010) Why does insect RNA look degraded? *J Insect Sci* 10:159
- Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, Djoumbou Y, Mandal R, Aziat F, Dong E, Bouatra S, Sinelnikov I, Arndt D, Xia J, Liu P, Yallou F, Bjorn Dahl T, Perez-Pineiro R, Eisner R, Allen F, Neveu V, Greiner R, Scalbert A (2013) HMDB 3.0—The human metabolome database in 2013. *Nucleic Acids Res* 41:D801–D807
- Wiśniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. *Nat Methods* 6:359–362
- Wong L-JC (2013) Next generation sequencing: translation to clinical diagnostics. Springer, New York
- Yang X, Gasser RB, Koehler AV, Wang L, Zhu K, Chen L, Feng H, Hu M, Fang R (2015a) Mitochondrial genome of *Hypoderaeum conoideum* - comparison with selected trematodes. *Parasit Vectors* 8:97
- Yang X, Zhao Y, Wang L, Feng H, Tan L, Lei W, Zhao P, Hu M, Fang R (2015b) Analysis of the complete *Fischoederius elongatus* (Paramphistomidae, Trematoda) mitochondrial genome. *Parasit Vectors* 8:279
- You H, McManus DP (2015) Vaccines and diagnostics for zoonotic schistosomiasis japonica. *Parasitology* 142:271–289
- You H, Gobert GN, Duke MG, Zhang W, Li Y, Jones MK, McManus DP (2012) The insulin receptor is a transmission blocking veterinary vaccine target for zoonotic *Schistosoma japonicum*. *Int J Parasitol* 42:801–807
- Young ND, Campbell BE, Hall RS, Jex AR, Cantacessi C, Laha T, Sohn WM, Sripa B, Loukas A, Brindley PJ, Gasser RB (2010a) Unlocking the transcriptomes of two carcinogenic parasites, *Clonorchis sinensis* and *Opisthorchis viverrini*. *PLoS Negl Trop Dis* 4:e719
- Young ND, Hall RS, Jex AR, Cantacessi C, Gasser RB (2010b) Elucidating the transcriptome of *Fasciola hepatica*—a key to fundamental and biotechnological discoveries for a neglected parasite. *Biotechnol Adv* 28:222–231
- Young ND, Jex AR, Cantacessi C, Hall RS, Campbell BE, Spithill TW, Tangkawattana S, Tangkawattana P, Laha T, Gasser RB (2011) A portrait of the transcriptome of the neglected trematode, *Fasciola gigantica*—biological and biotechnological implications. *PLoS Negl Trop Dis* 5:e1004
- Young ND, Jex AR, Li B, Liu S, Yang L, Xiong Z, Li Y, Cantacessi C, Hall RS, Xu X, Chen F, Wu X, Zerlotini A, Oliveira G, Hofmann A,

- Zhang G, Fang X, Kang Y, Campbell BE, Loukas A, Ranganathan S, Rollinson D, Rinaldi G, Brindley PJ, Yang H, Wang J, Wang J, Gasser RB (2012) Whole-genome sequence of *Schistosoma haematobium*. *Nat Genet* 44:221–225
- Young ND, Nagarajan N, Lin SJ, Korhonen PK, Jex AR, Hall RS, Safavi-Hemami H, Kaewkong W, Bertrand D, Gao S, Seet Q, Wongkham S, Teh BT, Wongkham C, Intapan PM, Maleewong W, Yang X, Hu M, Wang Z, Hofmann A, Sternberg PW, Tan P, Wang J, Gasser RB (2014) The *Opisthorchis viverrini* genome provides insights into life in the bile duct. *Nat Commun* 5:4378
- Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829
- Zerlotini A, Aguiar ER, Yu F, Xu H, Li Y, Young ND, Gasser RB, Protasio AV, Berriman M, Roos DS, Kissinger JC, Oliveira G (2013) SchistoDB: an updated genome resource for the three key schistosomes of humans. *Nucleic Acids Res* 41:D728–D731
- Zheng M, Hu K, Liu W, Hu X, Hu F, Huang L, Wang P, Hu Y, Huang Y, Li W, Liang C, Yin X, He Q, Yu X (2011) Proteomic analysis of excretory secretory products from *Clonorchis sinensis* adult worms: molecular characterization and serological reactivity of a excretory-secretory antigen-fructose-1,6-bisphosphatase. *Parasitol Res* 109:737–744
- Zheng M, Hu K, Liu W, Li H, Chen J, Yu X (2013) Proteomic analysis of different period excretory secretory products from *Clonorchis sinensis* adult worms: molecular characterization, immunolocalization, and serological reactivity of two excretory secretory antigens-methionine aminopeptidase 2 and acid phosphatase. *Parasitol Res* 112:1287–1297
- Zhu F, Shi Z, Qin C, Tao L, Liu X, Xu F, Zhang L, Song Y, Liu X, Zhang J, Han B, Zhang P, Chen Y (2012) Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. *Nucleic Acids Res* 40:D1128–D1136