

Accepted Manuscript

Invited Review

Advances in *Fasciola hepatica* research using 'omics' technologies

Krystyna Cwiklinski, John P. Dalton

PII: S0020-7519(18)30023-7

DOI: <https://doi.org/10.1016/j.ijpara.2017.12.001>

Reference: PARA 4033

To appear in: *International Journal for Parasitology*

Received Date: 26 October 2017

Revised Date: 5 December 2017

Accepted Date: 19 December 2017

Please cite this article as: Cwiklinski, K., Dalton, J.P., Advances in *Fasciola hepatica* research using 'omics' technologies, *International Journal for Parasitology* (2018), doi: <https://doi.org/10.1016/j.ijpara.2017.12.001>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Invited Review

Advances in *Fasciola hepatica* research using ‘omics’ technologies

Krystyna Cwiklinski^{a,*}, John P. Dalton^{a,b}

^a *School of Biological Sciences, Medical Biology Centre, Queen's University Belfast, Belfast, Northern Ireland, UK*

^b *Institute for Global Food Security (IGFS), Queen's University Belfast, Belfast, Northern Ireland, UK*

*Corresponding Author.

E-mail address: k.cwiklinski@qub.ac.uk

Abstract

The liver fluke *Fasciola hepatica* is an economically important pathogen of livestock worldwide, as well as being an important neglected zoonosis. Parasite control is reliant on the use of drugs, particularly triclabendazole (TCBZ), which is effective against multiple parasite stages. However, the spread of parasites resistant to TCBZ has intensified the pursuit for novel control strategies. Emerging –omic technologies are helping advance our understanding of liver fluke biology, specifically the molecules that act at the host-parasite interface and are central to infection, virulence and long-term survival within the definitive host. This review discusses the technological sequencing advances that have facilitated the unbiased analysis of liver fluke biology, resulting in an extensive range of ‘omics’ datasets. In addition, we highlight the ‘omics’ studies of host responses to *F. hepatica* infection that, when combined with the parasite datasets, provide the opportunity for integrated analyses of host-parasite interactions. These extensive datasets will form the foundation for future in-depth analysis of *F. hepatica* biology and development, and the search for new drug or vaccine interventions.

Keywords: *Fasciola hepatica*, Helminth, Trematode, Genome, Transcriptome, Proteome

1. Introduction

DNA sequencing technologies have evolved rapidly over the past few decades, stemming from the traditional Sanger methodology used to map the first human genome (Lander et al., 2001; Venter et al., 2001), to the recent high-throughput sequencing technologies such as Roche 454 and Illumina (Reuter et al., 2015) that we use today. More recently, single cell sequencing has emerged, pioneered by Pacific Biosciences and Oxford Nanopore Technologies, through the PacBio and MinION platforms, respectively (Reuter et al., 2015). As the technology for DNA sequencing has progressed, so too have the routine protocols for the extraction of nucleic acids and library preparation (Price et al., 2009); this has allowed sequencing projects to be carried out on even the most challenging species to propagate in the laboratory and those for which it was previously difficult to obtain adequate quantities of nucleic acids. Consequently, the number of sequencing projects undertaken has exploded, including recent ambitious proposals to sequence 10,000 vertebrate genomes (Genome 10K project; Koepfli et al., 2015), 5000 arthropods (i5K project; Poelchau et al., 2015) and all 10,500 species of birds (B10K project; Jarvis, 2016), to name but a few.

In the area of parasitology, a similar large-scale collaboration was initiated with the aim of sequencing 50 helminth genomes from human and veterinary parasites of global importance (50 Helminth Genomes Project, 50HGP; <http://www.sanger.ac.uk/science/collaboration/50hgp>). The advances in sequencing technologies enabled the number of genomes sequenced under this directive to be exceeded. Now in its ninth release, the database housing these genomes, WormBase ParaSite (<https://parasite.wormbase.org>), comprises 134 genomes, representing 114 species (Howe et al., 2017). In addition to acting as a central repository and publically accessible database for the wider research community, WormBase ParaSite integrates all available genomic and transcriptomic data to provide functional annotation and expression information for each species, and thus facilitates comparative genomics analysis.

How we profile the repertoire of transcripts expressed by an organism, at a particular time-point or in response to external cues, has also evolved with advances in sequencing technology. Studies first focused on analysing partial sequences,

known as expressed sequence tags (ESTs), derived from libraries of cDNA clones (Parkinson and Blaxter, 2009). In conjunction, serial analysis of gene expression (SAGE) methodology facilitated differential or temporal gene expression studies, as well as the detection and analysis of low abundance transcripts (Sun et al., 2004). However, it was the development of gene expression microarrays that initially instigated high throughput transcriptomic analyses that are still used today (Schena et al., 1995; Malone and Oliver, 2011). Since microarrays only detect known gene transcripts immobilised on microchips, they are less useful for gene discovery. By contrast, the emergence of RNA sequencing (RNAseq) allowed the analysis of all gene transcripts present within a given sample and now, advanced through the development of next generation sequencing (NGS) technologies, has largely replaced microarrays for gene transcription analysis.

This emerging array of transcriptomic profiling tools has been applied extensively to helminth parasites. Approximately 508,000 ESTs have been generated from Platyhelminth parasites and are housed in the NCBI database dbEST (dbEST release 130101; <https://www.ncbi.nlm.nih.gov/dbEST/>). SAGE methodology has also been employed for the analysis of gene expression across different lifecycle stages (Knox and Skuce, 2005; Williams et al., 2007; Taft et al., 2009). More recently, large scale RNAseq analyses have been completed for a range of Platyhelminth parasites, several of which have been disseminated through the site Helminth.net (Martin et al., 2015). These freely accessible datasets have complemented ongoing genome projects.

In parallel with techniques to analyse nucleic acids, advances in modern proteomic technologies have allowed the high throughput identification and characterization of complex protein preparations (Yarmush and Jayaraman, 2002; Brewis and Brennan, 2010). Progress has also been made in developing extraction protocols for soluble and membrane-bound proteins, as well increasing the sensitivity of proteomic technologies, including gel-free protocols that can be carried out on very small amounts of proteins (micrograms) (Scherp et al., 2011; Nature Methods Editorial, 2013). By integrating proteomic data with genomic/transcriptomic data, functional annotation is more precise and can provide qualitative and quantitative information regarding the expression of genes and their products, as

well as data such as the existence of splice variants or the nature of post-translational modifications.

Parasite-host interaction is a complex phenomenon involving molecules produced by both partners. The ability of helminth parasites to invade, migrate and survive within their hosts is expedited by the range of proteins they excrete/secrete (ES proteins). The roles these released proteins play during infection have been investigated in many studies using proteomic tools and have provided a rich source of immunomodulators, diagnostic reagents and vaccine candidates that can be cherry-picked at will to bring forward into commercialisable biotherapeutics. The available genomic/transcriptomic data, including those present in WormBase ParaSite, complement these proteomic studies, providing publically available databases that can be used during the identification/annotation process to further our understanding of helminth parasites and their interaction with their hosts.

In this review, we focus on the datasets available for the liver fluke parasite, *Fasciola hepatica*, and in particular how they are currently analysed and interrogated to enhance our knowledge of liver fluke biology, with a particular emphasis towards elucidating how these parasites invade and survive within their hosts. The lifecycle of this digenean trematode includes a snail intermediate host, within which the parasite undergoes a clonal expansion, and a mammalian definitive host, where the parasite develops into sexually mature adults, releasing 20,000–24,000 eggs per fluke per day (Boray, 1969). Infection of the mammalian host occurs following the ingestion of the infective encysted stage, the metacercariae. Within the intestine, parasites excyst, as newly excysted juveniles (NEJs) that migrate across the intestinal wall, then across the peritoneal cavity to the liver and bile ducts. *Fasciola hepatica* is known to infect a broad range of mammalian hosts including rodents, ruminants, ungulates, kangaroos and primates (Robinson and Dalton, 2009), implying the parasite has evolved a universal process(es) of infection. As a hermaphroditic parasite, *F. hepatica* has the ability to self- and cross-fertilise. In addition, studies have shown that hybridisation with the sister species, *Fasciola gigantica*, can occur, resulting in intermediate or hybrid forms as determined by analysis of mitochondrial genes and intergenic genome sequences (Le et al., 2008; Itagaki et al., 2011; Ichikawa-Seki et al., 2017).

The extensive collection of 'omics' datasets now available for *F. hepatica* includes the draft genome, stage-specific transcriptomes, and proteomic datasets for the somatic proteome, secretome, extracellular vesicles and glycoproteome of the outer tegumental surface. These datasets can now be used to investigate the complex features of the *Fasciola* lifecycle, particularly their effects on life history traits that directly impact on gene flow within liver fluke populations, influencing the spread of drug resistance and virulence/pathogenicity traits.

2. Genomics

2.1. The *F. hepatica* mitochondrial genome

The characterisation and differentiation of *Fasciola* spp. using morphological features is often unreliable and can only be used for the differentiation of adult parasites found within the bile ducts. Molecular identification based on nuclear ribosomal and mitochondrial genes is a more robust method of species classification. These molecular tools also provide markers for population genetic studies and epidemiological analysis of *Fasciola* spp. The complete *F. hepatica* mitochondrial genome was the first to be sequenced from a trematode species (Le et al., 2001) and has since been used for several population genetics studies of *F. hepatica* (Walker et al., 2007, 2011, 2012; Bargues et al., 2017). Similarly, the complete mitochondrial genome of *F. gigantica* has been reported (Liu et al., 2014), which now provides species-specific references that can be used in species characterization studies. For example, Liu and colleagues (2014) sequenced the complete mitochondrial genome of an intermediate form of *F. hepatica* and *F. gigantica* found in the Heilongjiang province, China (Peng et al., 2009). Based on intergenic spacer regions (ITS-1 and ITS-2) this isolate is indeed inferred to be a hybrid of *F. hepatica* and *F. gigantica*, although comparative analyses between *Fasciola* spp. mitochondrial genomes revealed that the intermediate form was more closely related to *F. gigantica* than to *F. hepatica*. This study shows that hybridisation is not uniform across the genome and that sequence variations at different sites can occur, in this case within the nuclear ribosomal genes and the maternally inherited mitochondrial genes. Thus, the study also highlighted the complexity incurred during hybridization of *Fasciola* spp. and challenges that their subsequent characterization presents.

2.2. Nuclear genome

To date 33 platyhelminth genomes are publically available within WormBase ParaSite, comprising species from the Classes Trematoda, Cestoda, Monogenea and Rhabditophora. Analyses of the genome assembly sizes show that although individual species vary in respect to their genome size, trends can be observed. In general, the cestode tapeworms have considerably smaller genomes compared with other members of the Phylum Platyhelminthes. The major exception to this statement is *Spirometra erinaceieuropaei*, which has one of the largest platyhelminth genomes (1.3 Gb; Bennett et al., 2014). Concerning the Class Trematoda, the blood flukes of the species *Schistosoma* have smaller genomes compared with other members of the Class.

Surprisingly, *F. hepatica* has the largest trematode genome sequenced to date (1.3 Gb; Cwiklinski et al., 2015a). For a parasite such as *Fasciola* that ensures its own species survival through the daily generation of large numbers of eggs, the evolution of a large genome appears counter-intuitive as it potentially imposes a cost on egg production. The reason for the large genome size has yet to be determined, but our studies indicate that it has not arisen through genome duplication or an increase in the percentage of the genome that is comprised of repeat regions. Although an equivalent number of genes have been identified across the trematode genomes, comparative analyses reveal that increases in genome size are reflective of increases in average exon and intron lengths, although this alone does not fully explain the increased size of the *F. hepatica* genome. Further analysis of the non-coding regions is required to determine their function and, in particular, their importance in gene regulation (ENCODE Project Consortium, 2012).

The recent genome sequencing of *F. hepatica* isolates from the Americas by McNulty and colleagues (2017) confirmed that the large genome size is comparable between fluke isolates (McNulty et al., 2017). Interestingly, the analysis of these American isolates revealed the presence of a *Neorickettsia* endobacterium within the parasite, which was further demonstrated by immunolocalisation studies that found the bacterium within the eggs, reproductive system and the oral suckers of adult flukes. Consistent with other studies of trematode-*Neorickettsia* interactions,

Neorickettsia could also be detected in the *Fasciola* eggs by PCR methods. To date no other liver fluke isolates from other geographical locations have reported the presence of any *Neorickettsia* endobacteria, indicating that the acquisition of this endobacteria may have occurred since the introduction of *F. hepatica* to the Americas. The study by McNulty and colleagues (2017) highlights the potential interaction between *Fasciola* and endosymbionts/endobacteria, and warrants further investigation.

Single nucleotide polymorphism (SNP) analysis of UK *F. hepatica* isolates, including isolates resistant to the frontline anthelmintic, triclabendazole (Hodgkinson et al., 2013), has revealed high levels of sequence polymorphism in the *F. hepatica* genome (Cwiklinski et al., 2015a). In particular, a marked over-representation of genes with high levels of non-synonymous polymorphism was associated with axonogenesis and chemotaxis, reflecting the changing environments the parasite encounters during its migration in the host. This data has recently been complemented by microsatellite analysis that revealed high levels of genetic diversity and gene flow within field isolates in the UK (Beesley et al., 2017). High levels of genetic diversity and gene flow may be important to counter the decline of allelic diversity as a result of self-fertilisation (Noel et al., 2017).

The current *F. hepatica* genome assembly (PRJEB6687; Cwiklinski et al., 2015a) is comprised of a large number of scaffolds and contigs (20,158 scaffolds and 195,709 contigs, with a scaffold N50 of 204 kb), mainly due to the size of the genome and the high percentage of repeat regions, which has hindered the assembly. In the future, utilising sequencing platforms that generate longer reads, as well as technologies such as optical mapping, should resolve this problem. The sequencing reads can then be mapped to the 10 *F. hepatica* chromosomes (Sanderson, 1953), allowing analysis of genomic structure and genomic comparison of platyhelminth genome organisation.

3. Transcriptomics

The development of novel control strategies, vaccines and diagnostic tools aimed at specific *F. hepatica* lifecycle stages, requires an understanding of the

genes that are transcribed at each time-point in development as well as their specific transcriptional abundance. Initial studies of gene identification and analysis were based on a limited number of unannotated expressed sequence tags (ESTs; 6819 sequences) generated from adult *F. hepatica* parasites by the Wellcome Trust Sanger Institute (UK; <ftp://ftp.sanger.ac.uk/pub/pathogens/Fasciola/hepatica/ESTs/>). This EST database was also an essential resource for analysing peptide sequences for *F. hepatica* proteomic studies (Chemale et al., 2006, 2010; Robinson et al., 2009; Hacariz et al., 2014; Morphew et al., 2014).

The formative analysis of these EST sequences identified several key molecules of interest for further characterisation including glutathione transferases (GSTs; Chemale et al., 2006), calcium binding proteins (Banford et al., 2013), mucin-like proteins (Cancela et al., 2015) and the helminth defence molecule (Robinson et al., 2011; Martinez-Sernandez et al., 2014). Enhancing our understanding of the *F. hepatica* lifecycle, Robinson and colleagues (2009) utilised an integrated transcriptomic and proteomic approach based on these adult-specific *Fasciola* ESTs, to profile the expression of proteins secreted by *Fasciola* parasites as they migrate through the host. However, this analysis was based on the premise that similarities could be drawn between the proteins expressed by the adult parasites residing in the bile ducts and those expressed by the migrating NEJ parasites. Utilising an adult-specific database, especially one with a limited number of sequences, likely resulted in NEJ-specific proteins being overlooked.

In 2010, Cancela and colleagues (2010) reported the generation of 1684 ESTs from the NEJ stage at excystment. The limited number of ESTs is reflective of the amount of total RNA that could be extracted from 1200 NEJs and subsequently used for cDNA synthesis (200 ng). Nevertheless, analysis of these sequences identified several sequences that had not been previously reported within the adult ESTs, implying that they were NEJ-specific. Specifically, several cathepsin cysteine proteases and antioxidant enzymes were characterised and showed that *F. hepatica* has adapted stage-specific proteases and enzymes to utilise throughout its lifecycle. The identification of novel stage-specific genes within this study highlighted the need for more extensive lifecycle stage-specific transcriptomes to further *Fasciola* research.

Led by the developments in sequencing technologies, Young and colleagues (2010) reported the first extensive adult *F. hepatica* transcriptome sequenced using 454 sequencing technology. In comparison to the 6819 unannotated adult-specific EST sequences available, this study generated a total of 590,927 high quality reads that were clustered into approximately 48,000 sequences, of which 15,423 supercontigs of 745 bp (\pm 517 bp) were enriched for open reading frames (ORFs). These sequences were subjected to extensive homology searches and protein prediction, using tools such as InterProScan, gene ontology (GO) and KOBAS (KEGG Orthology-Based Annotation System) to annotate the predicted proteins. Based on the publically available datasets at the time, approximately 44% of the sequences were classified, identifying proteins representative of the adult stage parasite. In keeping with the fact that *F. hepatica* expresses a range of cathepsin cysteine proteases, several cysteine peptidase family members were identified within the adult transcriptome. The predicted protein sequences were also screened for signal peptide and transmembrane domains to profile those proteins secreted by classical pathways within the ES proteins by the adult parasites; this analysis identified all the 160 ES proteins reported by Robinson et al. (2009). Importantly, comparing the Robinson et al. (2009) proteomic dataset with this more extensive adult *F. hepatica* database resulted in the annotation of previously unclassified proteins, including a group of fatty acid binding proteins and redox antioxidant enzymes. A further 18,347 contigs have been generated using 454 sequencing of adult fluke cDNA by Wilson et al. (2011) during their interrogation of the adult tegument. This more extensive dataset for adult *F. hepatica* has been interrogated by various research groups and has led to the identification of a range of proteins including SCP/TAPS proteins (Cantacessi et al., 2012), glutathione transferases (Morphew et al., 2012) and cathepsin cysteine proteases (Morphew et al., 2011).

The available transcriptomic data and subsequent analysis of *Fasciola* spp. has since been further improved with the development of short read Illumina sequencing that has increased sequence depth and coverage (Reuter et al., 2015). Investigation of the similarities between *F. hepatica* and *F. gigantica*, particularly those molecules important at the host-parasite interface, has been carried out following the first characterisation of the *F. gigantica* adult transcriptome (Young et al., 2011). Similarly, in-depth Illumina sequencing has been applied to the study of

virulence- and immunomodulation-related genes of adult *F. hepatica*, identifying 62 previously uncharacterised virulence-related genes. In silico characterisation subsequently implied that these genes have immunomodulatory properties since they were comparable to various immune-related molecules including cytokines and immune receptors (Hacariz et al., 2015).

In particular, the development of Illumina sequencing technology has advanced our knowledge of other *F. hepatica* lifecycle stages that have been previously difficult to analyse. We have reported the sequencing of several early lifecycle stages, namely the infective metacercarial stage, the NEJ parasites 1 h, 3 h and 24 h post-excystment, as well as juvenile parasites at 21 days p.i. and adult parasites which has provided a transcriptional profile of *F. hepatica* during infection (Cwiklinski et al., 2015a). Differential gene transcription analysis showed that the parasite regulates the transcription of many of its genes with progressively more genes being highly transcribed as the parasite rapidly grows and develops in preparation for migration through the host liver (>8000; Cwiklinski et al., 2015a).

The integration of transcriptomic data with the *F. hepatica* genome has also revealed that gene family expansion is a key feature of *F. hepatica* adaptation and survival; we have shown that *F. hepatica* transcribes different members of these gene families during different stages of the lifecycle (Cwiklinski et al., 2015a). Key examples of such expanded gene families are the cathepsin cysteine proteases and the microtubule-related alpha and beta tubulin genes. Biochemical analysis of the family of cathepsin proteases has shown that the different clades have evolved distinctive peptidolytic activity specific to the requirements of different lifecycle stages (Robinson et al., 2008). Similarly, the transcription of the diverse range of beta tubulin isoforms that are temporally regulated could explain the stage-specific efficacy of benzimidazole anthelmintics (Sanabria et al., 2013).

Coupled with comprehensive proteomic analyses, a current focus of our work is to investigate the infective and invasive lifecycle stages, namely the metacercariae and NEJs, to elucidate how the parasite prepares for infection and undergoes alterations to ensure its own survival (Cwiklinski et al., 2018). In-depth interrogation of the transcriptomic data available for these lifecycle stages has shown that the infective stages, the metacercariae, are metabolically active and that early juvenile

stages regulate the transcription of metabolic pathways, particularly those related to aerobic energy metabolism (Cwiklinski et al., 2018).

McNulty et al. (2017) reported a transcriptomic analysis of *F. hepatica* eggs, metacercariae and adult stages, as part of their genome characterisation of American *F. hepatica* isolates, and identified several gene sets that were over-expressed by specific lifecycle stages. In particular, consistent with our analysis (Cwiklinski et al., 2015a), the cathepsin proteases were found to be highly regulated; different clade isotypes were over-expressed by the metacercariae compared with the adult parasites. The most significantly over-expressed gene in eggs was found to be the rate-limiting enzyme of the pentose phosphate pathway, glucose-6-phosphate dehydrogenase.

Recently, there has been an interest in extracellular vesicles (EVs) within parasite secretomes and the role they play at the host-parasite interface (Marcilla et al., 2014; Coakley et al., 2015). EVs enable cell-to-cell communication by transferring proteins, lipids and microRNA (miRNA) (El Andaloussi et al., 2013; Record et al., 2014; Huang-Doran et al., 2017). At least two sub-populations of EVs with different protein content have been shown by centrifugation methods to be secreted by *F. hepatica*, including large EVs released from the parasite gut (collected after centrifugation at 15,000 g, referred to hereinafter as 15K EVs) and smaller exosome-like vesicle released from the tegumental surface (collected after centrifugation at 120,000 g, referred to hereinafter as 120K EVs) (Marcilla et al., 2012; Cwiklinski et al., 2015b). Transcriptomic analysis of the genes involved in the EV biogenesis pathway suggests that the synthesis of these two sub-populations of EVs occur via separate pathways, namely the Endosomal sorting complex required for transport (ESCRT) and lipid-related/ESCRT-independent pathways, respectively (Cwiklinski et al., 2015b; de la Torre-Escudero et al., 2016).

Further analyses of *Fasciola* miRNAs has been carried out following the generation of three small RNAseq libraries from adult parasites (Xu et al., 2012; Fromm et al., 2015), the NEJ stage (Fontenla et al., 2015) and EVs isolated from adult parasite secretion (Fromm et al., 2015). These studies have identified 52 non-coding miRNAs corresponding to 32 metazoan-conserved miRNA families (Fromm et al., 2017). In addition, five *F. hepatica*-specific sequences were identified.

Comparative analysis with *F. gigantica* indicates that these five sequences are specific to *F. hepatica* and are not shared across the *Fasciola* genus. Correspondingly, eight miRNAs have been identified as *F. gigantica*-specific (Xu et al., 2012). Whether or not these *Fasciola*-specific miRNAs are important for infection of the mammalian host has yet to be determined. Throughout the lifecycle the abundance of the miRNAs expressed by *F. hepatica* varies, indicating stage-specific roles (Fromm et al., 2017), with those miRNAs present within the EVs most likely important for host-parasite interactions. In particular, the predicted targets of five immuno-regulatory miRNAs found to be enriched within the EV warrant further investigation (Fromm et al., 2015, 2017).

4. Proteomics

Molecules that are excreted/secreted from liver flukes, also known as the ES proteins, are considered necessary for their migration through the tissues of the host and evasion of immune responses. While the early migrating stages of *F. hepatica* are mainly tissue feeders, adult parasites residing in the bile ducts are obligate blood feeders. The adult flukes are readily recovered from the bile ducts of infected livestock and ES proteins are released in abundance when the adult parasites are maintained in culture medium (even microgram quantities can be obtained from 10 adult parasites in vitro). Thus, the ES proteins of adult parasites have been extensively studied using proteomic tools. Early studies of *F. hepatica* proteins used radio-metabolic labelling to differentiate between the proteins of the various lifecycle stages (Irving and Howell, 1982; Dalton et al., 1985). Isoelectric focusing and densitometry were also carried out to characterise the ES proteins secreted by flukes maintained in different mammalian systems, namely llamas, rats, mice and cattle, which showed a different banding pattern (Lee et al., 1992a, 1992b). Jefferies and colleagues (2000, 2001) improved this analysis using two-dimensional (2D) gel electrophoresis and subsequent characterisation and annotation of protein spots to identify a range of cathepsin L proteases, superoxide dismutase, peroxiredoxin, glutathione S-transferases and fatty acid binding proteins. This study formed the basis for further in-depth analyses of these protein families using modern proteomic techniques and phylogenetic tools to elucidate how these protein families have

diverged and adapted (Chemale et al., 2006; Robinson et al., 2008; Marcilla et al., 2008; Morphew et al., 2011, 2012, 2013, 2016; Cwiklinski et al., 2015b; Di Maggio et al., 2016). Furthermore, proteomic analysis of the proteins within the extracellular vesicles released within the ES proteins has revealed that the 15K and 120K subpopulations of EVs released by *F. hepatica* vary in their protein cargo composition (Cwiklinski et al., 2015b).

Analysis of the adult liver fluke secretome has been used to assess the mode of action of the anthelmintic drug triclabendazole (TCBZ), and suggest that TCBZ broadly affects liver fluke metabolism (Chemale et al., 2010). In particular, protein signatures of liver fluke parasites susceptible and putatively resistant to TCBZ can be discerned based on the parasite's response to the TCBZ metabolite triclabendazole sulphoxide (TCBZ-SO) (Morphew et al., 2014). Parasite susceptibility to TCBZ characterised by lethal activity was indicated by the presence of actin, gelsolin, DJ-1 and triose phosphate isomerase, whereas putative resistance characterised by sub-lethal activity was indicated by the presence of calreticulin, cathepsin L proteases and enolase. These highly specific protein profiles provide potential markers that can be used for future TCBZ efficacy studies.

In contrast to the large amount of protein secreted by the adult parasites, analysis of the early developmental and migratory stages of *F. hepatica* is more challenging given their small size and the difficulty in locating them in host tissues. Accordingly, fewer proteomic studies have been reported for these stages. The molecular investigation of egg embryonation, however, characterised 28 proteins within the somatic proteome from 200,000 eggs (Moxon et al., 2010), and revealed that protein complexity increases as eggs mature, consistent with the development of the miracidial stage. This study also demonstrated that eggs have a substantially different protein profile from the other developmental stages of *F. hepatica* (Moxon et al., 2010). Similarly, a study of the secretome of an intra-molluscan stage, in vitro transformed mother sporocysts, required 388,000 parasites to generate sufficient protein for analysis (Gourbal et al., 2008). Seventeen of the most abundant proteins were analysed, in particular two antioxidant enzymes, Cu/Zn superoxide dismutase and thioredoxin (Gourbal et al., 2008). These enzymes were previously reported within the adult secretome (Jefferies et al., 2001; Robinson et al., 2009), suggesting a uniform process of detoxification of reactive oxygen species.

The development of proteomic tools and the accessibility of *F. hepatica* parasites have facilitated the expansion of the available proteome datasets for the NEJ migratory stages. Profiling the proteins secreted by the early infective stages, namely the NEJ 24 h post-excystment and the juvenile fluke 21 days p.i., with those of adult parasites has allowed stage-specific proteins to be determined (Robinson et al., 2009). A greater level of protein complexity was observed within the juvenile secretome (45 proteins) compared with the NEJ 24 h (29 proteins) and adult (22 proteins) secretomes, with a wider range of cathepsin isotypes and antioxidant enzymes being secreted. This is consistent with the migratory and feeding traits of this stage, and the upregulation of gene transcription observed during this stage (Andrews, 1999; Cwiklinski et al., 2015a). The NEJ 24 h secretome profile also confirmed the initial characterisation by N-terminal sequencing of the NEJ secreted proteins carried out in the Meeusen laboratory (Tkalcevic et al., 1995), which described an abundance of cathepsin L and asparaginyl endopeptidase cysteine proteases.

Facilitated by the sequencing of the *F. hepatica* genome, Di Maggio and colleagues (2016) recently reported a comprehensive analysis of the secreted proteins of adult liver fluke and NEJ 48 h post-excystment, and compared these with the somatic proteome of the NEJ 48 h post-excystment. Using gel-free proteomic techniques, this study identified 202 proteins within the adult secretome, 90 proteins within the NEJ 48 h secretome and 575 proteins in the somatic proteome of the NEJ 48 h. Consistent with other secretome analyses, a range of proteases and protease inhibitors were detected by both developmental stages, representing >70% and <10% of the total protein secreted, respectively. Furthermore, previously unreported proteins were identified within the NEJ somatic proteome, including structural proteins and proteins related to metabolism, expanding our knowledge of this lifecycle stage.

Untangling the complexities of host-parasite interactions is key to furthering our understanding of how this helminth evades the host immune system. The adult liver fluke parasites reside within the bile ducts, immersed in bile composed of bile acids, phospholipids, cholesterol, bilirubin and inorganic salts (Farina et al., 2009). Safe from the host's immune response (Andrews, 1999; Correia et al., 2001), proteomic analysis has shown that the adult secretome in the bile of sheep infected

with *F. hepatica* is dominated by cathepsin L proteases (Morphew et al., 2007), similar to that shown in vitro (Robinson et al., 2009; Morphew et al., 2011).

Key to survival within the mammalian host is the parasite tegument that can be rapidly turned over to prevent attachment of immune effector cells. Proteomic characterisation of the adult tegument was found to be enriched in structural proteins, transporters, proteins involved in secretory pathways and antioxidant enzymes (Wilson et al., 2011; Hacariz et al., 2012). A similar range of proteins was identified within the somatic proteome of the outer tegumental surface of NEJs (Hernandez-Gonzalez et al., 2010). Furthermore, recent analysis has been carried out on the tegumental immunoprecipitate formed following the incubation of live adult *F. hepatica* flukes in purified IgG from *F. gigantica*-infected Thin Tailed sheep (Cameron et al., 2017). This study identified several molecules consistent with previous analyses of the tegument (Wilson et al., 2011; Hacariz et al., 2012), as well as a range of proteins associated with *F. hepatica* exosomes (15K EVs; Cwiklinski et al., 2015b). In addition, it highlighted the cross-reactivity between antibodies elicited against *F. gigantica* during infection of Thin Tailed sheep and *F. hepatica* tegumental proteins, and raised the interesting question of whether different proteins and EV components are secreted/released under different 'environmental' conditions.

5. Glycomics

To date *F. hepatica* glycomic analyses have focussed on the outer surface of the parasite, the glycocalyx, that is rich in glycoproteins and glycolipids (Threadgold, 1976). Analyses have shown that the tegumental surface is highly glycosylated, with an abundance of mannose-rich N-linked glycoproteins present on the surface, spines and suckers (Garcia-Campos et al., 2016; Ravidá et al., 2016a, 2016b). The exact role these N-glycoproteins play at the host-parasite interface is currently unknown, although parasite glyco-conjugates have been implicated in evading the host immune response (van Die and Cummings, 2010). Moreover, blocking of the N-glycans on the surface of the NEJs using lectins has been shown to inhibit their migration across the intestinal wall (Garcia-Campos et al., 2017). Studies of the *F. hepatica* glycans have shown that they have immune modulatory properties, modulating toll-like receptor-induced maturation of dendritic cells through

carbohydrate-specific receptors (CLR) (Rodriguez et al., 2015, 2017; Ravidá et al., 2016a). In contrast to the analysis of the *F. hepatica* N-glycans, the composition of the O-glycans has yet to be determined, although potential sites have been identified within the glycoproteins of the tegument (Ravidá et al., 2016b).

The glycocalyx is also rich in glycolipids that are highly antigenic (Wuhrer et al., 2003) and share terminal Gal α 1-4Gal and Gal β 1-6Gal motifs with cestodes that result in serological cross-reactivity (Wuhrer et al., 2004). As well as being cross-reactive with other members of the Phylum Platyhelminthes, these glycolipids mimic mammalian-type glycolipids (Wuhrer et al., 2001, 2004), facilitating parasite survival.

6. Proteomic and transcriptomic analyses of host responses to *F. hepatica*

Several recent studies have utilised 'omics' approaches to investigate the responses of the host during infection with *F. hepatica*, with the aim to elucidate host responses that mirror the stage of infection and the developmental changes that occur within the migrating parasite. These large-scale investigations of the host responses give an unbiased global view of the effects of fasciolosis on host immune tissues, and have revealed novel aspects of pathogenesis associated with infection. In addition, these approaches are being used to evaluate potential vaccine candidates, identifying the genes involved in conferring protection (Wesołowska et al., 2013; Rojas-Caraballo et al., 2017).

Transcriptomic responses within macroscopic lesions of *F. hepatica*-infected liver at 8 weeks p.i. in sheep revealed that gene expression is highly regulated (Alvarez Rojas et al., 2015), consistent with comparable microarray studies of mice (Rojas-Caraballo et al., 2015). Several processes characteristic of acute fasciolosis, that are upregulated in response to the damage caused by the parasite, were identified. Genes corresponding to fibrosis and tissue repair were found to be upregulated, consistent with the subsequent tissue regeneration required following the invasive migration of the parasite. In keeping with observations that helminth infections typically skew host immune responses towards a Th2 type, genes associated with Th2 differentiation and B cell activation were found to be upregulated, while Th1 type responses were down-regulated. Interestingly, this

study also reported that an increased abundance of circulating reticulocytes is associated with the blood feeding activity of *F. hepatica*, which can cause anaemia. Increased transcription of haemoglobin-related genes and four genes putatively associated with Fanconi anaemia were also observed.

In other investigations of the host responses to fasciolosis, two recent studies analysed the transcriptomic responses of ovine peripheral blood mononuclear cells (PBMCs) at stages throughout infection (Alvarez Rojas et al., 2016; Fu et al., 2016). Despite different protocols being used for sample preparation, RNA extraction and subsequent analyses of the RNAseq data, both studies reveal that gene transcription is highly regulated during *F. hepatica* infection of sheep, particularly during acute infection (1-2 weeks p.i.). Both studies also observed the upregulation of genes associated with transforming growth factor (TGF) beta signalling, including the genes TGF beta, collagen type 1 and the downstream SMAD signalling genes. These genes play a major role in fibrosis, and as such were also observed in the transcriptome analysis of infected liver described above (Alvarez Rojas et al., 2015). Upregulation of genes associated with the complement and coagulation cascades, chemokine signalling pathway and cytokine-cytokine receptor interaction pathway were also reported by Alvarez Rojas et al. (2016).

Consistent with the polarization of immune responses towards a Th2 type, the gene encoding inducible nitric oxide synthase (iNOS) was shown to be down-regulated in response to infection with *F. hepatica* in both acute and chronic stages in the study by Fu et al. (2016). Transcription of Th17 related genes were also down-regulated, suggesting that *F. hepatica* is able to inhibit the differentiation and stability of Th17 cells. In contrast with the study by Fu and colleagues (2016), however, genes encoding interleukins, particularly those related to Th2 type responses such as IL-4, were not reported by Alvarez Rojas et al. (2016) to be differentially expressed at statistical levels. This difference may reflect the different protocols used to process the PBMCs for RNA extraction, namely processing fresh cells versus storage in RNAlater, which may have had an effect on the stability of gene expression profiles (Debey-Pascher et al., 2011; Eikmans et al., 2013). Equally, the types of strategies used for sequence analyses may contribute to the differences observed.

As infection by *F. hepatica* progresses, the amount and composition of immune cells present both within the peritoneal cavity and circulating in the peripheral blood shifts to an abundance of eosinophils, which is associated with a polarisation to Th2 type immune responses. Differential eosinophil cell counts were only reported in the Alvarez-Rojas et al. study (2016), and showed the infected animals had substantially more eosinophils than the control non-infected animals. At 4 weeks p.i., the eosinophil count in the infected group ranged from 12-39% compared with the control group counts of 1-5%. Therefore, the changes in transcriptomic response are reflective of both a change in transcription during infection as well as a change in the number and type of cells within the PBMC fraction, and must be interpreted accordingly.

Compared with inbred mouse strains, large animal mammalian hosts of helminth parasites, in particular sheep and cattle, are genetically more variable at both the individual animal level and between breeds. This can have a significant effect on how an animal or animal breed responds to infection with *F. hepatica* (Ardia et al., 2011). In the case of the two studies analysing transcriptomic responses in PBMCs, each used different sheep breeds sorted into groups of four animals. To address the possible between-animal variation, Alvarez Rojas et al. (2016) employed two strategies to analyse their data; (i) assessment of each animal as an independent experiment and (ii) treating each group (control non-infected and infected) as biological replicates. Differentially expressed genes identified by both strategies were then used for further analysis, resulting in the identification of 183 and 76 genes differentially expressed at 2 weeks p.i. and 8 weeks p.i., respectively. In comparison, the study by Fu et al. (2016), which compared animals as biological replicates, identified 6490 differentially expressed genes at 1 week p.i., indicating that many genes of interest were overlooked by the stringent process employed by Alvarez-Rojas et al. (2016). However, the sheep breed-specific responses may also be a factor in the differences observed between these trials. Therefore, it is important for studies of host-parasite interactions in ruminants that sufficient numbers of animals are used and that the type of analysis utilised is appropriately considered and validated.

The host responses to the migrating parasites within the peritoneal compartment during early infection and in the bile ducts consistent with chronic

infection have also been investigated using proteomic tools. Early fasciolosis is characterised by the migration of *F. hepatica* through the intestinal wall to the liver via the peritoneal cavity. Proteomic analysis of the peritoneal fluid from sheep infected by *F. hepatica* at 18 days p.i. identified an abundance of proteins associated with the complement system and proteins associated with the liver extracellular matrix (ECM) (Ruiz-Campillo et al., 2017). The presence of proteins associated with the liver ECM, including collagen VI, fibronectin and fibrocystin, is likely the result of the damage caused by the parasite as it invades and migrates through the liver. Intriguingly, this study also detected two ECM-related molecules, periostin and vascular cell adhesion protein 1 (VCAM-1), that mediate leukocyte infiltration and are associated with marked eosinophilia, which warrant further investigation as biomarkers of infection.

Systemic responses have been analysed using proteomic tools to identify biomarkers of infection within host serum. Rioux et al. (2008) showed that there were significant changes within the sera beginning within 3 weeks of infection, consistent with the transcriptomic analysis of PBMCs that highlighted a greater level of differential gene expression during acute infection (Fu et al., 2016; Alvarez Rojas et al., 2016). These striking changes also coincide with marked expression of >8000 genes that accompany the rapid growth and development of *F. hepatica* during the first 3 weeks of infection (Cwiklinski et al., 2015a). Two markers of particular interest, namely transferrin and apolipoprotein A-IV (Apo A-IV), were upregulated during this early period (Rioux et al., 2008). Transferrin is associated with anaemia caused by the blood feeding parasites, whereas Apo A-IV is associated with regulation of appetite within the intestine of mammals (although studies in rats and mice insinuate a possible role within the liver; VerHague et al., 2013). In comparison, levels of transferrin detected within the bile by Morphew et al. (2007) were reduced compared with levels in the serum, further highlighting that the data can vary significantly depending of the sample type (serum, bile, peritoneal fluid etc) and time of infection.

7. Concluding remarks

Over the last few decades, major advances that have been made through 'omics' technologies have provided the liver fluke research community with an extensive array of datasets that can be interrogated to further our understanding of liver fluke biology (Fig. 1). The number of genes encoded within the *F. hepatica* genome has been clarified. In particular, this information has been crucial in elucidating gene family organisations, which in the past have been complicated by the large number of gene sequences of similar classification present within the NCBI Genbank database. In addition, the genes transcribed by *F. hepatica* have been found to be highly regulated throughout the lifecycle within the mammalian host. This knowledge is vital to our continuing efforts to develop control strategies that, in particular, target the early parasite stages. Proteomic analysis of the ES proteins has highlighted key molecules that play an important role at the host-parasite interface. Biochemical characterization of these key molecules has also revealed stage-specific adaptations, including the activity of cathepsin L proteases that includes collagenolytic activity specific to the juvenile parasites and haemolytic activity restricted to the adult parasites (Robinson et al., 2008). It has also revealed some unexpected adaptations such as the kunitz-type serpin inhibitors that have inhibitory activity against cathepsin L cysteine proteases and not serine proteases (Smith et al., 2016).

However, the functions of a large proportion of *F. hepatica* genes and the proteins they encode still remain unknown. In general, these genes only share homology with uncharacterised genes from other Platyhelminthes, indicating that they are Phylum-specific. Further investigation is therefore required to decipher the functions of these genes and specifically their importance for host-parasite interactions. This can be achieved by utilising post-genomic tools such as RNA interference (RNAi) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), as well as protein structural analysis, to increase our knowledge of these uncharacterised genes, facilitating the annotation of Platyhelminthes datasets. Furthermore, the addition of this information into the various software packages available for 'omics' analyses such as STRING (Szklarczyk et al., 2015) and PANTHER (Mi et al., 2017) where there is a current lack of data relating to the Phylum Platyhelminthes, is essential to expand our knowledge of parasite protein-protein interaction networks.

How parasites regulate their genes, specifically in response to their environment, and particularly the host immune response, is becoming an area of intense interest. In particular, this analysis has encompassed understanding the epigenetic process of gene regulation, through DNA methylation, histone modification and non-coding RNA associated with gene silencing (Egger et al., 2004). The roles these epigenetic processes play in facilitating *Fasciola* invasion and survival are yet to be investigated. However, studies of closely related Platyhelminthes have indicated that these warrant further investigation. In particular, the study of DNA methylation across the Phylum Platyhelminthes has shown that cytosine methylation is a functionally conserved epigenetic feature (Geyer et al., 2013). Furthermore, recent analysis of the epigenome of *Schistosoma mansoni* cercariae revealed that histone modifications play an important role in regulating the transcription of genes, with the cercariae being transcriptionally inactive (Roquis et al., 2015). In-depth analysis of the *Fasciola* genome has already revealed an array of non-coding small RNAs that may play a part in the post-transcriptional regulation of *Fasciola* genes and/or be important for the regulation and manipulation of the mammalian host. Similar analysis of the epigenome of the different lifecycle stages will show if there are any lifecycle stage regulatory factors associated with liver fluke gene regulation.

For the future development of control strategies, a greater understanding of host-helminth interactions is paramount. This review has discussed the large-scale datasets available for the study of liver fluke infection, from both the parasite and the mammalian host. Going forward the analyses of these data should be integrated to elucidate the delicate interplay that occurs during infection and to determine if the pathogenicity/virulence of liver fluke isolates within field populations plays a role in this interaction.

Acknowledgements

KC and JPD are funded by a European Research Council Advanced Grant (HELIVAC, 322725) awarded to JPD and are members of the Horizon 2020-funded Consortium PARAGONE.

678 **References**

- 679 Alvarez Rojas, C.A., Ansell, B.R., Hall, R.S., Gasser, R.B., Young, N.D., Jex, A.R.,
 680 Scheerlinck, J.P., 2015. Transcriptional analysis identifies key genes involved
 681 in metabolism, fibrosis/tissue repair and the immune response against *Fasciola*
 682 *hepatica* in sheep liver. *Parasit. Vectors* 8, 124-015-0715-7.
- 683 Alvarez Rojas, C.A., Scheerlinck, J.P., Ansell, B.R., Hall, R.S., Gasser, R.B., Jex,
 684 A.R., 2016. Time-course study of the transcriptome of peripheral blood
 685 mononuclear cells (PBMCs) from sheep infected with *Fasciola hepatica*. *PLoS*
 686 *One* 11, e0159194.
- 687 Andrews, S., 1999. The life cycle of *Fasciola hepatica*. In: Dalton, J.P. (Ed.),
 688 Fasciolosis. CABI Publishing, New York pp. 1-29.
- 689 Ardia, D.R., Parmentier, H.K., Vogel, L.A., 2011. The role of constraints and
 690 limitation in driving individual variation in immune response. *Funct. Ecol.* 25,
 691 61-73.
- 692 Banford, S., Drysdale, O., Hoey, E.M., Trudgett, A., Timson, D.J., 2013. FhCaBP3: A
 693 *Fasciola hepatica* calcium binding protein with EF-hand and dynein light chain
 694 domains. *Biochimie* 95, 751-758.
- 695 Bargues, M.D., Gayo, V., Sanchis, J., Artigas, P., Khoubbane, M., Birriel, S., Mas-
 696 Coma, S., 2017. DNA multigene characterization of *Fasciola hepatica* and
 697 *Lymnaea neotropica* and its fascioliasis transmission capacity in Uruguay, with
 698 historical correlation, human report review and infection risk analysis. *PLoS*
 699 *Negl Trop. Dis.* 11, e0005352.
- 700 Beesley, N.J., Williams, D.J., Paterson, S., Hodgkinson, J., 2017. *Fasciola hepatica*
 701 demonstrates high levels of genetic diversity, a lack of population structure and
 702 high gene flow: Possible implications for drug resistance. *Int. J. Parasitol.* 47,
 703 11-20.
- 704 Bennett, H.M., Mok, H.P., Gkrania-Klotsas, E., Tsai, I.J., Stanley, E.J., Antoun, N.M.,
 705 Coghlan, A., Harsha, B., Traini, A., Ribeiro, D.M., Steinbiss, S., Lucas, S.B.,
 706 Allinson, K.S., Price, S.J., Santarius, T.S., Carmichael, A.J., Chiodini, P.L.,
 707 Holroyd, N., Dean, A.F., Berriman, M., 2014. The genome of the sparganosis
 708 tapeworm *Spirometra erinaceieuropaei* isolated from the biopsy of a migrating
 709 brain lesion. *Genome Biol.* 15, 510.
- 710 Boray, J.C., 1969. Experimental fascioliasis in Australia. *Adv. Parasitol.* 7, 95-210.
- 711 Brewis, I.A., Brennan, P., 2010. Proteomics technologies for the global identification
 712 and quantification of proteins. *Adv. Protein Chem. Struct. Biol.* 80, 1-44.
- 713 Cameron, T.C., Cooke, I., Faou, P., Toet, H., Piedrafita, D., Young, N., Rathinasamy,
 714 V., Beddoe, R., Anderson, G., Dempster, R., Spithill, T.W., 2017. A novel ex
 715 vivo immunoproteomic approach characterising *Fasciola hepatica* tegumental

- 716 antigens identified using immune antibody from resistant sheep. Int. J.
717 Parasitol. 47, 555-567.
- 718 Cancela, M., Ruetalo, N., Dell'Oca, N., da Silva, E., Smircich, P., Rinaldi, G., Roche,
719 L., Carmona, C., Alvarez-Valin, F., Zaha, A., Tort, J.F., 2010. Survey of
720 transcripts expressed by the invasive juvenile stage of the liver fluke *Fasciola*
721 *hepatica*. BMC Genomics 11, 227-2164-11-227.
- 722 Cancela, M., Santos, G.B., Carmona, C., Ferreira, H.B., Tort, J.F., Zaha, A., 2015.
723 *Fasciola hepatica* mucin-encoding gene: Expression, variability and its potential
724 relevance in host-parasite relationship. Parasitology 142, 1673-1681.
- 725 Cantacessi, C., Hofmann, A., Young, N.D., Broder, U., Hall, R.S., Loukas, A.,
726 Gasser, R.B., 2012. Insights into SCP/TAPS proteins of liver flukes based on
727 large-scale bioinformatic analyses of sequence datasets. PLoS One 7, e31164.
- 728 Chemale, G., Morphew, R., Moxon, J.V., Morassuti, A.L., Lacourse, E.J., Barrett, J.,
729 Johnston, D.A., Brophy, P.M., 2006. Proteomic analysis of glutathione
730 transferases from the liver fluke parasite, *Fasciola hepatica*. Proteomics 6,
731 6263-6273.
- 732 Chemale, G., Perally, S., LaCourse, E.J., Prescott, M.C., Jones, L.M., Ward, D.,
733 Meaney, M., Hoey, E., Brennan, G.P., Fairweather, I., Trudgett, A., Brophy,
734 P.M., 2010. Comparative proteomic analysis of triclabendazole response in the
735 liver fluke *Fasciola hepatica*. J. Proteome Res. 9, 4940-4951.
- 736 Coakley, G., Maizels, R.M., Buck, A.H., 2015. Exosomes and other extracellular
737 vesicles: The new communicators in parasite infections. Trends Parasitol. 31,
738 477-489.
- 739 Correia, L., Podevin, P., Borderie, D., Verthier, N., Montet, J.C., Feldmann, G.,
740 Poupon, R., Weill, B., Calmus, Y., 2001. Effects of bile acids on the humoral
741 immune response: A mechanistic approach. Life Sci. 69, 2337-2348.
- 742 Cwiklinski, K., Jewhurst, H., McVeigh, P., Barbour, T., Maule, A.G., Tort, J., O'Neill,
743 S.M., Robinson, M.W., Donnelly, S., Dalton, J.P., 2018. Infection by the
744 helminth parasite *Fasciola hepatica* requires rapid regulation of metabolic,
745 virulence, and invasive factors to adjust to its mammalian host. Mol. Cell.
746 Proteomics. In press.
- 747 Cwiklinski, K., Dalton, J.P., Dufresne, P.J., La Course, J., Williams, D.J.,
748 Hodgkinson, J., Paterson, S., 2015a. The *Fasciola hepatica* genome: Gene
749 duplication and polymorphism reveals adaptation to the host environment and
750 the capacity for rapid evolution. Genome Biol. 16, 71-015-0632-2.
- 751 Cwiklinski, K., de la Torre-Escudero, E., Trelis, M., Bernal, D., Dufresne, P.J.,
752 Brennan, G.P., O'Neill, S., Tort, J., Paterson, S., Marcilla, A., Dalton, J.P.,
753 Robinson, M.W., 2015b. The extracellular vesicles of the helminth pathogen,
754 *Fasciola hepatica*: Biogenesis pathways and cargo molecules involved in
755 parasite pathogenesis. Mol. Cell. Proteomics 14, 3258-3273.

- 756 Dalton, J.P., Tom, T.D., Strand, M., 1985. *Fasciola hepatica*: comparison of
757 immature and mature immunoreactive glycoproteins. *Parasite Immunol.* 7, 643-
758 57.
- 759 de la Torre-Escudero, E., Bennett, A.P., Clarke, A., Brennan, G.P., Robinson, M.W.,
760 2016. Extracellular vesicle biogenesis in helminths: More than one route to the
761 surface? *Trends Parasitol.* 32, 921-929.
- 762 Debey-Pascher, S., Hofmann, A., Kreusch, F., Schuler, G., Schuler-Thurner, B.,
763 Schultze, J.L., Staratschek-Jox, A., 2011. RNA-stabilized whole blood samples
764 but not peripheral blood mononuclear cells can be stored for prolonged time
765 periods prior to transcriptome analysis. *J. Mol. Diagn.* 13, 452-460.
- 766 Di Maggio, L.S., Tirloni, L., Pinto, A.F., Diedrich, J.K., Yates Iii, J.R., Benavides, U.,
767 Carmona, C., da Silva Vaz, I., Jr, Berasain, P., 2016. Across intra-mammalian
768 stages of the liver fluke *Fasciola hepatica*: A proteomic study. *Sci. Rep.* 6,
769 32796.
- 770 Egger, G., Liang, G., Aparicio, A., Jones, P.A., 2004. Epigenetics in human disease
771 and prospects for epigenetic therapy. *Nature* 429, 457-463.
- 772 Eikmans, M., Rekers, N.V., Anholts, J.D., Heidt, S., Claas, F.H., 2013. Blood cell
773 mRNAs and microRNAs: Optimized protocols for extraction and preservation.
774 *Blood* 121, e81-9.
- 775 El Andaloussi, S., Mager, I., Breakefield, X.O., Wood, M.J., 2013. Extracellular
776 vesicles: Biology and emerging therapeutic opportunities. *Nat. Rev. Drug*
777 *Discov.* 12, 347-357.
- 778 ENCODE Project Consortium, 2012. An integrated encyclopedia of DNA elements in
779 the human genome. *Nature* 489, 57-74.
- 780 Farina, A., Dumonceau, J.M., Lescuyer, P., 2009. Proteomic analysis of human bile
781 and potential applications for cancer diagnosis. *Expert Rev. Proteomics* 6, 285-
782 301.
- 783 Fontenla, S., Dell'Oca, N., Smircich, P., Tort, J.F., Siles-Lucas, M., 2015. The
784 miRnome of *Fasciola hepatica* juveniles endorses the existence of a reduced
785 set of highly divergent microRNAs in parasitic flatworms. *Int. J. Parasitol.* 45,
786 901-913.
- 787 Fromm, B., Ovchinnikov, V., Hoyer, E., Bernal, D., Hackenberg, M., Marcilla, A.,
788 2017. On the presence and immunoregulatory functions of extracellular
789 microRNAs in the trematode *Fasciola hepatica*. *Parasite Immunol.* 39,
790 10.1111/pim.12399.
- 791 Fromm, B., Trelis, M., Hackenberg, M., Cantalapiedra, F., Bernal, D., Marcilla, A.,
792 2015. The revised microRNA complement of *Fasciola hepatica* reveals a
793 plethora of overlooked microRNAs and evidence for enrichment of immuno-
794 regulatory microRNAs in extracellular vesicles. *Int. J. Parasitol.* 45, 697-702.

- 795 Fu, Y., Chryssafidis, A.L., Browne, J.A., O'Sullivan, J., McGettigan, P.A., Mulcahy,
796 G., 2016. Transcriptomic study on ovine immune responses to *Fasciola*
797 *hepatica* infection. PLoS Negl Trop. Dis. 10, e0005015.
- 798 Garcia-Campos, A., Baird, A.W., Mulcahy, G., 2017. Migration of *Fasciola hepatica*
799 newly excysted juveniles is inhibited by high-mannose and oligomannose-type
800 N-glycan-binding lectins. Parasitology 144, 1708-1717.
- 801 Garcia-Campos, A., Ravidá, A., Nguyen, D.L., Cwiklinski, K., Dalton, J.P., Hokke,
802 C.H., O'Neill, S., Mulcahy, G., 2016. Tegument glycoproteins and cathepsins of
803 newly excysted juvenile *Fasciola hepatica* carry mannosidic and
804 paucimannosidic N-glycans. PLoS Negl. Trop. Dis. 10, e0004688.
- 805 Geyer, K.K., Chalmers, I.W., Mackintosh, N., Hirst, J.E., Geoghegan, R., Badets, M.,
806 Brophy, P.M., Brehm, K., Hoffmann, K.F., 2013. Cytosine methylation is a
807 conserved epigenetic feature found throughout the phylum Platyhelminthes.
808 BMC Genomics 14, 462-2164-14-462.
- 809 Gourbal, B.E., Guillou, F., Mitta, G., Sibille, P., Theron, A., Pointier, J.P., Coustau,
810 C., 2008. Excretory-secretory products of larval *Fasciola hepatica* investigated
811 using a two-dimensional proteomic approach. Mol. Biochem. Parasitol. 161, 63-
812 66.
- 813 Hacariz, O., Akgun, M., Kavak, P., Yuksel, B., Sagiroglu, M.S., 2015. Comparative
814 transcriptome profiling approach to glean virulence and immunomodulation-
815 related genes of *Fasciola hepatica*. BMC Genomics 16, 366-015-1539-8.
- 816 Hacariz, O., Baykal, A.T., Akgun, M., Kavak, P., Sagiroglu, M.S., Sayers, G.P., 2014.
817 Generating a detailed protein profile of *Fasciola hepatica* during the chronic
818 stage of infection in cattle. Proteomics 14, 1519-1530.
- 819 Hacariz, O., Sayers, G., Baykal, A.T., 2012. A proteomic approach to investigate the
820 distribution and abundance of surface and internal *Fasciola hepatica* proteins
821 during the chronic stage of natural liver fluke infection in cattle. J. Proteome
822 Res. 11, 3592-3604.
- 823 Hernandez-Gonzalez, A., Valero, M.L., del Pino, M.S., Oleaga, A., Siles-Lucas, M.,
824 2010. Proteomic analysis of in vitro newly excysted juveniles from *Fasciola*
825 *hepatica*. Mol. Biochem. Parasitol. 172, 121-128.
- 826 Hodgkinson, J., Cwiklinski, K., Beesley, N.J., Paterson, S., Williams, D.J., 2013.
827 Identification of putative markers of triclabendazole resistance by a genome-
828 wide analysis of genetically recombinant *Fasciola hepatica*. Parasitol. 140,
829 1523-1533
- 830 Howe, K.L., Bolt, B.J., Shafie, M., Kersey, P., Berriman, M., 2017. WormBase
831 ParaSite - a comprehensive resource for helminth genomics. Mol. Biochem.
832 Parasitol. 215, 2-10.

- 833 Huang-Doran, I., Zhang, C.Y., Vidal-Puig, A., 2017. Extracellular vesicles: Novel
834 mediators of cell communication in metabolic disease. Trends Endocrinol.
835 Metab. 28, 3-18.
- 836 Ichikawa-Seki, M., Peng, M., Hayashi, K., Shoriki, T., Mohanta, U.K., Shibahara, T.,
837 Itagaki, T., 2017. Nuclear and mitochondrial DNA analysis reveals that
838 hybridization between *Fasciola hepatica* and *Fasciola gigantica* occurred in
839 China. Parasitology 144, 206-213.
- 840 Irving D. O., Howell, M.J., 1982. Characterization of excretory-secretory antigens of
841 *Fasciola hepatica*. Parasitology 85, 179-188.
- 842 Itagaki, T., Ichinomiya, M., Fukuda, K., Fusyuku, S., Carmona, C., 2011.
843 Hybridization experiments indicate incomplete reproductive isolating
844 mechanism between *Fasciola hepatica* and *Fasciola gigantica*. Parasitology
845 138, 1278-1284.
- 846 Jarvis, E.D., 2016. Perspectives from the avian phylogenomics project: Questions
847 that can be answered with sequencing all genomes of a vertebrate class. Annu.
848 Rev. Anim. Biosci. 4, 45-59.
- 849 Jefferies, J.R., Brophy, P.M., Barrett, J., 2000. Investigation of *Fasciola hepatica*
850 sample preparation for two-dimensional electrophoresis. Electrophoresis 21,
851 3724-3729.
- 852 Jefferies, J.R., Campbell, A.M., van Rossum, A.J., Barrett, J., Brophy, P.M., 2001.
853 Proteomic analysis of *Fasciola hepatica* excretory-secretory products.
854 Proteomics 1, 1128-1132.
- 855 Knox, D.P., Skuce, P.J., 2005. SAGE and the quantitative analysis of gene
856 expression in parasites. Trends Parasitol. 21, 322-326.
- 857 Koepfli, K.P., Paten, B., Genome 10K Community of Scientists, O'Brien, S.J., 2015.
858 The genome 10K project: A way forward. Annu. Rev. Anim. Biosci. 3, 57-111.
- 859 Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon,
860 K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K.,
861 Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P.,
862 McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J.,
863 Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-
864 Thomann, Y., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J.,
865 Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter,
866 N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin,
867 R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A.,
868 Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S.,
869 Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S.,
870 Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A.,
871 Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chisoe,
872 S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B.,
873 Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T.,

Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J., Szustakowki, J., International Human Genome Sequencing Consortium, 2001. Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.

Le, T.H., Blair, D., McManus, D.P., 2001. Complete DNA sequence and gene organization of the mitochondrial genome of the liverfluke, *Fasciola hepatica* L. (Platyhelminthes; Trematoda). *Parasitology* 123, 609-621.

Le, T.H., De, N.V., Agatsuma, T., Thi Nguyen, T.G., Nguyen, Q.D., McManus, D.P., Blair, D., 2008. Human fascioliasis and the presence of hybrid/introgressed forms of *Fasciola hepatica* and *Fasciola gigantica* in Vietnam. *Int. J. Parasitol.* 38, 725-730.

Lee, C.G., Zimmerman, G.L., Bishop, J.K., 1992a. Host influence on the banding profiles of whole-body protein and excretory-secretory product of *Fasciola hepatica* (Trematoda) by isoelectric focusing. *Vet. Parasitol.* 41, 57-68.

Lee, C.G., Zimmerman, G.L., Mulrooney, D.M., 1992b. Isoelectric focusing of soluble proteins from *Fasciola hepatica* L, 1758 and *Fascioloides magna* B, 1875. *Am. J. Vet. Res.* 53, 246-250.

- 918 Liu, G.H., Gasser, R.B., Young, N.D., Song, H.Q., Ai, L., Zhu, X.Q., 2014. Complete
919 mitochondrial genomes of the 'intermediate form' of *Fasciola* and *Fasciola*
920 *gigantica*, and their comparison with *F. hepatica*. *Parasit. Vectors* 7, 150-3305-
921 7-150.
- 922 Malone, J.H., Oliver, B., 2011. Microarrays, deep sequencing and the true measure
923 of the transcriptome. *BMC Biol.* 9, 34-7007-9-34.
- 924 Marcilla, A., De la Rubia, J.E., Sotillo, J., Bernal, D., Carmona, C., Villavicencio, Z.,
925 Acosta, D., Tort, J., Bornay, F.J., Esteban, J.G., Toledo, R., 2008. Leucine
926 aminopeptidase is an immunodominant antigen of *Fasciola hepatica* excretory
927 and secretory products in human infections. *Clin. Vaccine Immunol.* 15, 95-100.
- 928 Marcilla, A., Martin-Jaular, L., Trelis, M., de Menezes-Neto, A., Osuna, A., Bernal,
929 D., Fernandez-Becerra, C., Almeida, I.C., Del Portillo, H.A., 2014. Extracellular
930 vesicles in parasitic diseases. *J. Extracell. Vesicles* 3, 25040.
- 931 Marcilla, A., Trelis, M., Cortes, A., Sotillo, J., Cantalapiedra, F., Minguez, M.T.,
932 Valero, M.L., Sanchez del Pino, M.M., Munoz-Antoli, C., Toledo, R., Bernal, D.,
933 2012. Extracellular vesicles from parasitic helminths contain specific
934 excretory/secretory proteins and are internalized in intestinal host cells. *PLoS*
935 *One* 7, e45974.
- 936 Martin, J., Rosa, B.A., Ozersky, P., Hallsworth-Pepin, K., Zhang, X., Bhonagiri-
937 Palsikar, V., Tyagi, R., Wang, Q., Choi, Y.J., Gao, X., McNulty, S.N., Brindley,
938 P.J., Mitreva, M., 2015. Helminth.net: Expansions to nematode.net and an
939 introduction to trematode.net. *Nucleic Acids Res.* 43, D698-706.
- 940 Martinez-Sernandez, V., Mezo, M., Gonzalez-Warleta, M., Perteguer, M.J., Muino,
941 L., Guitian, E., Garate, T., Ubeira, F.M., 2014. The MF6p/FhHDM-1 major
942 antigen secreted by the trematode parasite *Fasciola hepatica* is a heme-
943 binding protein. *J. Biol. Chem.* 289, 1441-1456.
- 944 Nature Methods Editorial, 2013. Method of the year 2012. *Nat. Methods* 10, 1.
945 doi:10.1038/nmeth.2329
- 946 McNulty, S.N., Tort, J.F., Rinaldi, G., Fischer, K., Rosa, B.A., Smircich, P., Fontenla,
947 S., Choi, Y.J., Tyagi, R., Hallsworth-Pepin, K., Mann, V.H., Kammili, L.,
948 Latham, P.S., Dell'Oca, N., Dominguez, F., Carmona, C., Fischer, P.U.,
949 Brindley, P.J., Mitreva, M., 2017. Genomes of *Fasciola hepatica* from the
950 Americas reveal colonization with *Neorickettsia* endobacteria related to the
951 agents of Potomac horse and human Sennetsu fevers. *PLoS Genet.* 13,
952 e1006537.
- 953 Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., Thomas, P.D.,
954 2017. PANTHER version 11: expanded annotation data from Gene Ontology
955 and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids*
956 *Res.* 45, D183-D189.

- 957 Morphew, R.M., Eccleston, N., Wilkinson, T.J., McGarry, J., Perally, S., Prescott, M.,
958 Ward, D., Williams, D., Paterson, S., Raman, M., Ravikumar, G., Khalid
959 Saifullah, M., Abbas Abidi, S.M., McVeigh, P., Maule, A.G., Brophy, P.M.,
960 LaCourse, E.J., 2012. Proteomics and in silico approaches to extend
961 understanding of the glutathione transferase superfamily of the tropical liver
962 fluke *Fasciola gigantica*. J. Proteome Res. 11, 5876-5889.
- 963 Morphew, R.M., Hamilton, C.M., Wright, H.A., Dowling, D.J., O'Neill, S.M., Brophy,
964 P.M., 2013. Identification of the major proteins of an immune modulating
965 fraction from adult *Fasciola hepatica* released by nonidet P40. Vet. Parasitol.
966 191, 379-385.
- 967 Morphew, R.M., Wilkinson, T.J., Mackintosh, N., Jahndel, V., Paterson, S., McVeigh,
968 P., Abbas Abidi, S.M., Saifullah, K., Raman, M., Ravikumar, G., LaCourse, J.,
969 Maule, A., Brophy, P.M., 2016. Exploring and expanding the fatty-acid-binding
970 protein superfamily in *Fasciola* species. J. Proteome Res. 15, 3308-3321.
- 971 Morphew, R.M., Wright, H.A., Lacourse, E.J., Porter, J., Barrett, J., Woods, D.J.,
972 Brophy, P.M., 2011. Towards delineating functions within the *Fasciola* secreted
973 cathepsin I protease family by integrating in vivo based sub-proteomics and
974 phylogenetics. PLoS Negl. Trop. Dis. 5, e937.
- 975 Morphew, R.M., Wright, H.A., LaCourse, E.J., Woods, D.J., Brophy, P.M., 2007.
976 Comparative proteomics of excretory-secretory proteins released by the liver
977 fluke *Fasciola hepatica* in sheep host bile and during in vitro culture ex host.
978 Mol. Cell. Proteomics 6, 963-972.
- 979 Morphew, R.M., MacKintosh, N., Hart, E.H., Prescott, M., LaCourse, E.J., Brophy,
980 P.M., 2014. In vitro biomarker discovery in the parasitic flatworm *Fasciola*
981 *hepatica* for monitoring chemotherapeutic treatment. EuPA Open Proteomics 3,
982 85-99.
- 983 Moxon, J.V., LaCourse, E.J., Wright, H.A., Perally, S., Prescott, M.C., Gillard, J.L.,
984 Barrett, J., Hamilton, J.V., Brophy, P.M., 2010. Proteomic analysis of embryonic
985 *Fasciola hepatica*: Characterization and antigenic potential of a
986 developmentally regulated heat shock protein. Vet. Parasitol. 169, 62-75.
- 987 Noel, E., Jarne, P., Glemin, S., MacKenzie, A., Segard, A., Sarda, V., David, P.,
988 2017. Experimental evidence for the negative effects of self-fertilization on the
989 adaptive potential of populations. Curr. Biol. 27, 237-242.
- 990 Parkinson, J., and Blaxter, M., 2009. Expressed sequence tags: An overview.
991 Methods Mol. Biol. 533, 1-12.
- 992 Peng, M., Ichinomiya, M., Ohtori, M., Ichikawa, M., Shibahara, T., Itagaki, T., 2009.
993 Molecular characterization of *Fasciola hepatica*, *Fasciola gigantica*, and
994 aspermic *Fasciola* sp. in China based on nuclear and mitochondrial DNA.
995 Parasitol. Res. 105, 809-815.

- 996 Poelchau, M., Childers, C., Moore, G., Tsavatapalli, V., Evans, J., Lee, C.Y., Lin, H.,
997 Lin, J.W., Hackett, K., 2015. The i5k Workspace@NAL--enabling genomic data
998 access, visualization and curation of arthropod genomes. *Nucleic Acids Res.*
999 43, D714-9.
- 1000 Price, C.W., Leslie, D.C., Landers, J.P., 2009. Nucleic acid extraction techniques
1001 and application to the microchip. *Lab. Chip* 9, 2484-2494.
- 1002 Ravida, A., Aldridge, A.M., Driessen, N.N., Heus, F.A., Hokke, C.H., O'Neill, S.M.,
1003 2016a. *Fasciola hepatica* surface coat glycoproteins contain mannosylated and
1004 phosphorylated N-glycans and exhibit immune modulatory properties
1005 independent of the mannose receptor. *PLoS Negl. Trop. Dis.* 10, e0004601.
- 1006 Ravida, A., Cwiklinski, K., Aldridge, A.M., Clarke, P., Thompson, R., Gerlach, J.Q.,
1007 Kilcoyne, M., Hokke, C.H., Dalton, J.P., O'Neill, S.M., 2016b. *Fasciola hepatica*
1008 surface tegument: Glycoproteins at the interface of parasite and host. *Mol. Cell.*
1009 *Proteomics* 15, 3139-3153.
- 1010 Record, M., Carayon, K., Poirot, M., Silvente-Poirot, S., 2014. Exosomes as new
1011 vesicular lipid transporters involved in cell-cell communication and various
1012 pathophysiological processes. *Biochim. Biophys. Acta* 1841, 108-120.
- 1013 Reuter, J.A., Spacek, D.V., Snyder, M.P., 2015. High-throughput sequencing
1014 technologies. *Mol. Cell* 58, 586-597.
- 1015 Rioux, M.C., Carmona, C., Acosta, D., Ward, B., Ndao, M., Gibbs, B.F., Bennett,
1016 H.P., Spithill, T.W., 2008. Discovery and validation of serum biomarkers
1017 expressed over the first twelve weeks of *Fasciola hepatica* infection in sheep.
1018 *Int. J. Parasitol.* 38, 123-136.
- 1019 Robinson, M.W., Dalton, J.P., 2009. Zoonotic helminth infections with particular
1020 emphasis on fasciolosis and other trematodiasis. *Philos. Trans. R. Soc. Lond.*
1021 *B. Biol. Sci.* 364, 2763-2776.
- 1022 Robinson, M.W., Donnelly, S., Hutchinson, A.T., To, J., Taylor, N.L., Norton, R.S.,
1023 Perugini, M.A., Dalton, J.P., 2011. A family of helminth molecules that modulate
1024 innate cell responses via molecular mimicry of host antimicrobial peptides.
1025 *PLoS Pathog.* 7, e1002042.
- 1026 Robinson, M.W., Menon, R., Donnelly, S.M., Dalton, J.P., Ranganathan, S., 2009.
1027 An integrated transcriptomics and proteomics analysis of the secretome of the
1028 helminth pathogen *Fasciola hepatica*: Proteins associated with invasion and
1029 infection of the mammalian host. *Mol. Cell. Proteomics* 8, 1891-1907.
- 1030 Robinson, M.W., Tort, J.F., Lowther, J., Donnelly, S.M., Wong, E., Xu, W., Stack,
1031 C.M., Padula, M., Herbert, B., Dalton, J.P., 2008. Proteomics and phylogenetic
1032 analysis of the cathepsin L protease family of the helminth pathogen *Fasciola*
1033 *hepatica*: Expansion of a repertoire of virulence-associated factors. *Mol. Cell.*
1034 *Proteomics* 7, 1111-1123.

- 1035 Rodriguez, E., Carasi, P., Frigerio, S., da Costa, V., van Vliet, S., Noya, V.,
1036 Brossard, N., van Kooyk, Y., Garcia-Vallejo, J.J., Freire, T., 2017. *Fasciola*
1037 *hepatica* immune regulates CD11c+ cells by interacting with the macrophage
1038 gal/GalNAc lectin. *Front. Immunol.* 8, 264.
- 1039 Rodriguez, E., Noya, V., Cervi, L., Chiribao, M.L., Brossard, N., Chiale, C., Carmona,
1040 C., Giacomini, C., Freire, T., 2015. Glycans from *Fasciola hepatica* modulate
1041 the host immune response and TLR-induced maturation of dendritic cells. *PLoS*
1042 *Negl Trop. Dis.* 9, e0004234.
- 1043 Rojas-Caraballo, J., López-Abán, J., Fernández-Soto, P., Vicente, B., Collía, F.,
1044 Muro, A., 2015. Gene expression profile in the liver of BALB/c mice infected
1045 with *Fasciola hepatica*. *PLoS One.* 10, e0134910.
- 1046 Rojas-Caraballo, J., López-Abán, J., Moreno-Pérez, D.A., Vicente, B., Fernández-
1047 Soto, P., Del Olmo, E., Patarroyo, M.A., Muro, A., 2017. Transcriptome profiling
1048 of gene expression during immunisation trial against *Fasciola hepatica*:
1049 identification of genes and pathways involved in conferring immunoprotection in
1050 a murine model. *BMC Infect. Dis.* 17, 94.
- 1051 Roquis, D., Lepesant, J.M., Picard, M.A., Freitag, M., Parrinello, H., Groth, M.,
1052 Emans, R., Cosseau, C., Grunau, C., 2015. The epigenome of *Schistosoma*
1053 *mansoni* provides insight about how cercariae poise transcription until infection.
1054 *PLoS Negl Trop. Dis.* 9, e0003853.
- 1055 Ruiz-Campillo, M.T., Molina Hernandez, V., Escamilla, A., Stevenson, M., Perez, J.,
1056 Martinez-Moreno, A., Donnelly, S., Dalton, J.P., Cwiklinski, K., 2017. Immune
1057 signatures of pathogenesis in the peritoneal compartment during early infection
1058 of sheep with *Fasciola hepatica*. *Sci. Rep.* 7, 2782-017-03094-0.
- 1059 Sanabria, R., Ceballos, L., Moreno, L., Romero, J., Lanusse, C., Alvarez, L., 2013.
1060 Identification of a field isolate of *Fasciola hepatica* resistant to albendazole and
1061 susceptible to triclabendazole. *Vet. Parasitol.* 193, 105-110.
- 1062 Sanderson, A.R., 1953. Maturation and probable gynogenesis in the liver fluke,
1063 *Fasciola hepatica* L. *Nature* 172, 110-112.
- 1064 Schena, M., Shalon, D., Davis, R.W., Brown, P.O., 1995. Quantitative monitoring of
1065 gene expression patterns with a complementary DNA microarray. *Science* 270,
1066 467-470.
- 1067 Scherp, P., Ku, G., Coleman, L., Kheterpal, I., 2011. Gel-based and gel-free
1068 proteomic technologies. *Methods Mol. Biol.* 702, 163-190.
- 1069 Smith, D., Tikhonova, I.G., Jewhurst, H.L., Drysdale, O.C., Dvorak, J., Robinson,
1070 M.W., Cwiklinski, K., Dalton, J.P., 2016. Unexpected activity of a novel kunitz-
1071 type inhibitor: Inhibition of cysteine proteases but not serine proteases. *J. Biol.*
1072 *Chem.* 291, 19220-19234.

- 1073 Sun, M., Zhou, G., Lee, S., Chen, J., Shi, R.Z., Wang, S.M., 2004. SAGE is far more
1074 sensitive than EST for detecting low-abundance transcripts. BMC Genomics 5,
1075 1.
- 1076 Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas,
1077 J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., Kuhn, M., Bork, P.,
1078 Jensen, L.J., von Mering, C., 2015. STRING v10: protein-protein interaction
1079 networks, integrated over the tree of life. Nucleic Acids Res 43, D447-52.
- 1080 Taft, A.S., Vermeire, J.J., Bernier, J., Birkeland, S.R., Cipriano, M.J., Papa, A.R.,
1081 McArthur, A.G., Yoshino, T.P., 2009. Transcriptome analysis of *Schistosoma*
1082 *mansoni* larval development using serial analysis of gene expression (SAGE).
1083 Parasitology 136, 469-485.
- 1084 Threadgold, L.T., 1976. *Fasciola hepatica*: Ultrastructure and histochemistry of the
1085 glycocalyx of the tegument. Exp. Parasitol. 39, 119-134.
- 1086 Tkalcovic, J., Ashman, K., Meeusen, E., 1995. *Fasciola hepatica*: Rapid identification
1087 of newly excysted juvenile proteins. Biochem. Biophys. Res. Commun. 213,
1088 169-174.
- 1089 van Die, I., Cummings, R.D., 2010. Glycan gimmickry by parasitic helminths: A
1090 strategy for modulating the host immune response? Glycobiology 20, 2-12.
- 1091 Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith,
1092 H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P.,
1093 Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng,
1094 X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor
1095 Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A.,
1096 Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M.,
1097 Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern,
1098 A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington,
1099 K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill,
1100 M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di
1101 Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan,
1102 W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R.,
1103 Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F.,
1104 Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam,
1105 B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J.,
1106 Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C.,
1107 Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong,
1108 F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter,
1109 C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H.,
1110 Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A.,
1111 Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S.,
1112 Dodson, K., Doup, L., Ferreira, S., Garg, N., Gluecksmann, A., Hart, B.,
1113 Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland,
1114 T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann,
1115 F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L.,
1116 Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H.,

- 1117 Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R.,
 1118 Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint,
 1119 N.N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M.,
 1120 Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigo,
 1121 R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva,
 1122 B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S.,
 1123 Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yoosheph, S., Allen,
 1124 D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P.,
 1125 Chiang, Y.H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M.,
 1126 Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek,
 1127 A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S.,
 1128 Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C.,
 1129 Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J.,
 1130 Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck,
 1131 J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T.,
 1132 Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D.,
 1133 Wu, M., Xia, A., Zandieh, A., Zhu, X., 2001. The sequence of the human
 1134 genome. *Science* 291, 1304-1351.
- 1135 VerHague, M.A., Cheng, D., Weinberg, R.B., Shelness, G.S., 2013. Apolipoprotein
 1136 A-IV expression in mouse liver enhances triglyceride secretion and reduces
 1137 hepatic lipid content by promoting very low density lipoprotein particle
 1138 expansion. *Arterioscler. Thromb. Vasc. Biol.* 33, 2501-2508.
- 1139 Walker, S.M., Johnston, C., Hoey, E.M., Fairweather, I., Borgsteede, F.,
 1140 Gaasenbeek, C., Prodohl, P.A., Trudgett, A., 2011. Population dynamics of the
 1141 liver fluke, *Fasciola hepatica*: The effect of time and spatial separation on the
 1142 genetic diversity of fluke populations in the netherlands. *Parasitology* 138, 215-
 1143 223.
- 1144 Walker, S.M., Prodohl, P.A., Fletcher, H.L., Hanna, R.E., Kantzoura, V., Hoey, E.M.,
 1145 Trudgett, A., 2007. Evidence for multiple mitochondrial lineages of *Fasciola*
 1146 *hepatica* (liver fluke) within infrapopulations from cattle and sheep. *Parasitol.*
 1147 *Res.* 101, 117-125.
- 1148 Walker, S.M., Prodohl, P.A., Hoey, E.M., Fairweather, I., Hanna, R.E., Brennan, G.,
 1149 Trudgett, A., 2012. Substantial genetic divergence between morphologically
 1150 indistinguishable populations of *Fasciola* suggests the possibility of cryptic
 1151 speciation. *Int. J. Parasitol.* 42, 1193-1199.
- 1152 Wesolowska, A., Jaros, S., Norbury, L.J., Jaros, D., Zygmier, W., Wędrychowicz, H.,
 1153 2013. Microarray analysis of rat immune responses to liver fluke infection
 1154 following vaccination with *Fasciola hepatica* phosphoglycerate kinase. *Exp.*
 1155 *Parasitol.* 134, 33-8.
- 1156 Williams, D.L., Sayed, A.A., Bernier, J., Birkeland, S.R., Cipriano, M.J., Papa, A.R.,
 1157 McArthur, A.G., Taft, A., Vermeire, J.J., Yoshino, T.P., 2007. Profiling
 1158 *Schistosoma mansoni* development using serial analysis of gene expression
 1159 (SAGE). *Exp. Parasitol.* 117, 246-258.

- 1160 Wilson, R.A., Wright, J.M., de Castro-Borges, W., Parker-Manuel, S.J., Dowle, A.A.,
1161 Ashton, P.D., Young, N.D., Gasser, R.B., Spithill, T.W., 2011. Exploring the
1162 *Fasciola hepatica* tegument proteome. *Int. J. Parasitol.* 41, 1347-1359.
- 1163 Wuhrer, M., Berkefeld, C., Dennis, R.D., Idris, M.A., Geyer, R., 2001. The liver flukes
1164 *Fasciola gigantica* and *Fasciola hepatica* express the leukocyte cluster of
1165 differentiation marker CD77 (globotriaosylceramide) in their tegument. *Biol.*
1166 *Chem.* 382, 195-207.
- 1167 Wuhrer, M., Grimm, C., Zahringer U., Dennis, R.D., Berkefeld, C.M., Idris, M.A.,
1168 Geyer, R., 2003. A novel GlcNAc α 1-HPO3-6Gal(1-1)ceramide antigen and
1169 alkylated inositol-phosphoglycerolipids expressed by the liver fluke *Fasciola*
1170 *hepatica*. *Glycobiol.* 13, 129-37.
- 1171 Wuhrer, M., Grimm, C., Dennis, R.D., Idris, M.A., Geyer, R., 2004. The parasitic
1172 trematode *Fasciola hepatica* exhibits mammalian-type glycolipids as well as
1173 Gal(beta1-6)Gal-terminating glycolipids that account for cestode serological
1174 cross-reactivity. *Glycobiol.* 14, 115-26.
- 1175 Xu, M.J., Ai, L., Fu, J.H., Nisbet, A.J., Liu, Q.Y., Chen, M.X., Zhou, D.H., Zhu, X.Q.,
1176 2012. Comparative characterization of microRNAs from the liver flukes *Fasciola*
1177 *gigantica* and *F. hepatica*. *PLoS One* 7, e53387.
- 1178 Yarmush, M.L., Jayaraman, A., 2002. Advances in proteomic technologies. *Annu.*
1179 *Rev. Biomed. Eng.* 4, 349-373.
- 1180 Young, N.D., Hall, R.S., Jex, A.R., Cantacessi, C., Gasser, R.B., 2010. Elucidating
1181 the transcriptome of *Fasciola hepatica* - a key to fundamental and
1182 biotechnological discoveries for a neglected parasite. *Biotechnol. Adv.* 28, 222-
1183 231.
- 1184 Young, N.D., Jex, A.R., Cantacessi, C., Hall, R.S., Campbell, B.E., Spithill, T.W.,
1185 Tangkawattana, S., Tangkawattana, P., Laha, T., Gasser, R.B., 2011. A portrait
1186 of the transcriptome of the neglected trematode, *Fasciola gigantica*--biological
1187 and biotechnological implications. *PLoS Negl Trop. Dis.* 5, e1004.
- 1188
- 1189

Figure Legend

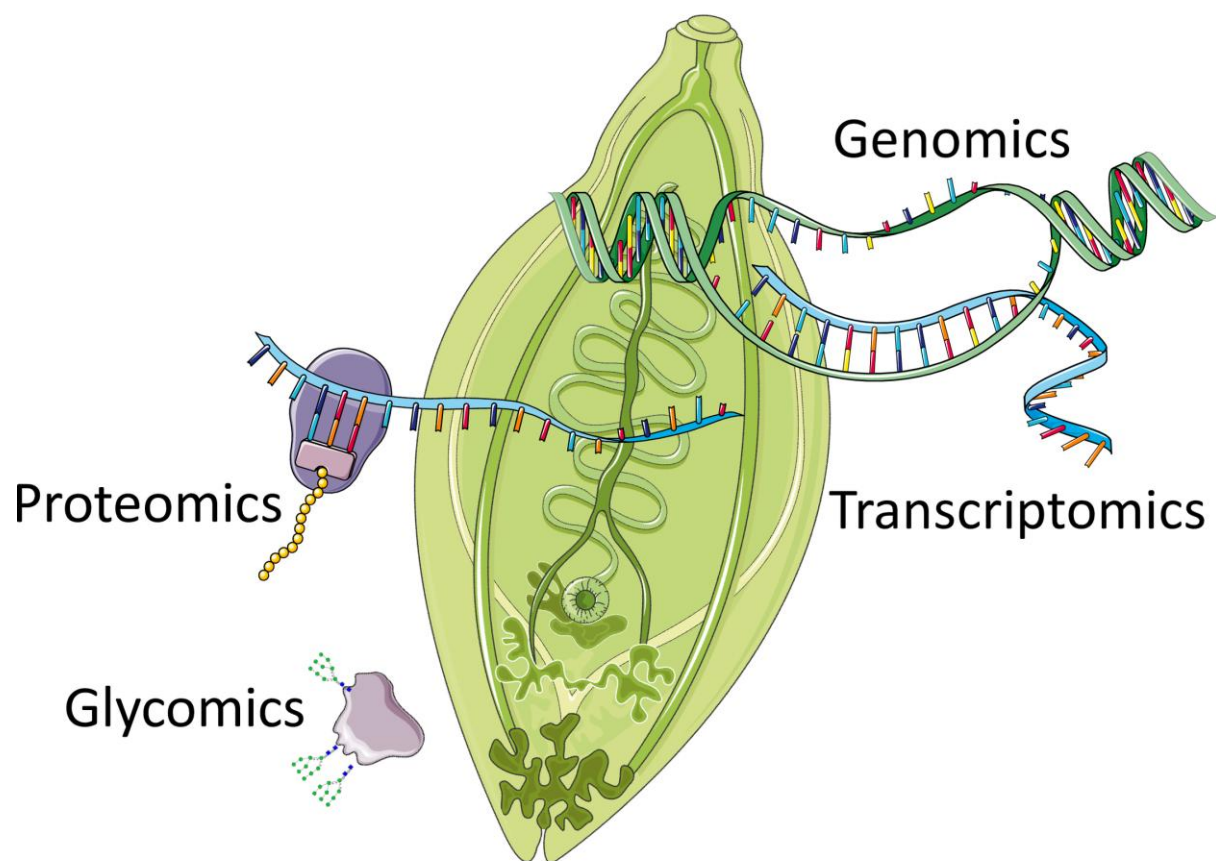
Fig. 1. Schematic of the major *Fasciola* 'omics' advances detailed over time. The principal references are denoted by numbers as follows: 1) Dalton et al., 1985; 2) Tkalcevic et al., 1995; 3) Jefferies et al., 2000; 4) Le et al., 2001; 5) <ftp://ftp.sanger.ac.uk/pub/pathogens/Fasciola/hepatica/ESTs/>; 6) Gourbal et al., 2008; 7) Robinson et al., 2009; 8) Young et al., 2010; 9) Cancela et al., 2010; 10) Moxon et al., 2010; 11) Xu et al., 2012; 12) Marcilla et al., 2012; 13) Cwiklinski et al., 2015a; 14) McNulty et al., 2017; 15) Garcia-Campos et al., 2016; 16) Ravida et al., 2016b; 17) Di Maggio et al., 2016. 2D, two-dimensional; ESTs, expressed sequence tags; EVs, extracellular vesicles; miRNA, microRNA; mt, mitochondrial; NEJ, newly excysted juveniles;

Highlights

- *Fasciola hepatica* is an economically important global pathogen of livestock and people.
- An extensive collection of 'omics' technology-derived datasets are now available.
- The *F. hepatica* genome is large (1.3 Gb) and exhibits high polymorphism.
- Stage-specific transcriptomic and proteomic analyses highlight key molecules of interest.
- These datasets can now be exploited to further our understanding of liver fluke biology.

1220 Graphical abstract
1221

Fasciola hepatica 'omics' analysis



Medical art provided by Les Laboratoires Servier
<https://smart.servier.com/>

1222

1223

