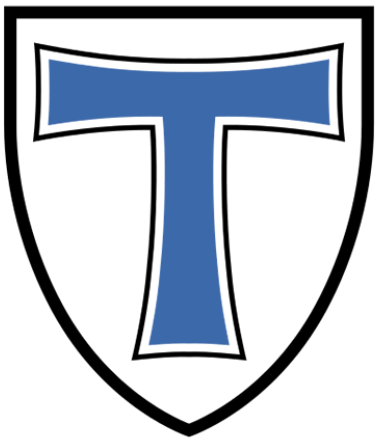
Uncovering the cellular diversity of *Fasciola hepatica* utilizing single-cell transcriptomics



Inaugural-Dissertation
zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)
im Fachbereich Biologie und Chemie der
Justus-Liebig-Universität Gießen

vorgelegt von

Oliver Puckelwaldt
(M. Sc. Biologie)

aus Idar-Oberstein

Oktober, 2024

The present work was carried out at the Institute of Parasitology (Faculty 10), Justus Liebig University Giessen, between January 2020 and October 2024 under the supervision of Prof. Dr. Nikola-Michael Prpic-Schäper and PD Dr. Simone Häberlein.

First supervisor: **Prof. Dr. Nikola-Michael Prpic-Schäper**

Institute of Zoology and Developmental Biology, Justus Liebig University Giessen, Heinrich-Buff-Ring 38, 35392 Giessen, Germany

Second supervisor: **PD Dr. Simone Häberlein**

Institute of Parasitology, BFS, Justus Liebig University Giessen,
Schubertstraße 81, 35392 Giessen, Germany

„Der Kampf gegen Gipfel vermag ein Menschenherz auszufüllen”

-

Albert Camus

List of Contents

List of abbreviations

Abbreviation	Full name
7b2	Neuroendocrine protein 7b2
ABAM	Antibiotic-antimycotic
ABZ	Albendazole
ACME	Acetic methanol
AP	Alkaline phosphatase
APC	Adenomatous polyposis coli protein
BAAT	ß-alanyl-tryptamine
BCIP	5-bromo-4-chloro-3-indolylphosphate
BLAST	Basic local alignment search tool
BMP-1	Bone morphogenic protein
Bmpg	Bone marrow proteoglycan
BSA	Bovine serum albumin
cDNA	Complementary DNA
CISH	Colorimetric *in situ* hybridization
DBD	DNA binding domain
DEPC	Diethyl pyrocarbonate
dH_2O	Destilled water
DIG	Digoxigenin
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

Abbreviation	Full name
Dsh	Dishevelled
dsRNA	Double-stranded RNA
DTT	Dithiothreitol
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EdU	5-ethynyl 2´-deoxyuridine
EST	Expressed sequence tag
FABP	Fatty acid binding proteins
FACS	Fluorescence activated cell sorting
FBT	Food-borne trematodiases
Fiji	Fiji is just imageJ
FISH	Fluorescent *in situ* hybridization
FITC	Fluorescein
GEM	Gel beads-in-emulsion
GO	Gene ontology
GPCR	G-protein coupled receptor
H2B	Histone 2b
H3P	Phosphorylated histone 3
HBSS	Hanks balanced salt solution
HCR	Hybridization chain reaction
HEPES	2-4-(2-hydroxyethyl)-1-piperazinyl ethane sulfonic acid

Abbreviation	Full name
Hoechst 33342	2'-[4-ethoxyphenyl]-5-[4-methyl-1-piperazinyl]-2,5'-bi-1H-benzimidazole trihydrochloride trihydrate
IgG	Immunoglobulin G
IHC	Immunohistochemistry
ISH	*In situ* hybridization
iToL	Interactive tree of life
LB	Lysogeny broth
MAB-T	Maleic acid buffer
MAFFT	Multiple alignment using fast fourier transform
MAR	Muscarinergic acetycholine receptor
MCM	Minichromosome maintenance
mRNA	Messenger RNA
MUSCLE	Multiple sequence comparison by log-expectation
NBT	4-Nitro blue tetrazolium chloride
NEJ	Newly excysted juveniles
NTD	Neglected tropical disease
NTP	Nucleoside triphosphate
O.C.T	Tissue-Tek® Compound
PAK	p21-activated kinase
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCR	Polymerase chain reaction

Abbreviation	Full name
PDB	p21-binding domain
PK	Protein kinase
PKC	Protein kinase C
POD	Peroxidase
PZQ	Praziquantel
qPCR	Quantitative PCR
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
rRNA	Ribosomal RNA
RT-qPCR	Reverse transcription quantitative PCR
scRNA-seq	Single-cell RNA sequencing
snRNA-seq	Single-nuclei RNA sequencing
SMART	Simple Modular Architecture Research Tool
SSC	Saline sodium citrate
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
TAE	Tris-Acetat-EDTA
TCBZ	Triclabendazole
TPM	Transcripts per million
TSA	Tyramide Signal Amplification
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier

Abbreviation	Full name
UTR	Untranslated region
VB1	Vitelline protein B1
VF1	Vitellogenic factor 1
WHO	World Health Organisation
Wnt	Wingless-related integration site

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Neglected tropical diseases by helminths

Infections caused by parasitic worms (helminths) cause a significant disease burden in humans and animals alike. Some of these parasites have impacted human health for a long time, as parasitic material has been found in Egyptian mummies and at various other archeological sites [1]. In present days, diseases caused by helminths affect millions of people worldwide, and numerous of these diseases show exceptionally high prevalence in developing regions such as sub-Saharan Africa, Asia, and South America [2,3]. Several of these diseases are categorized as “Neglected Tropical Diseases” (NTDs), a group of diseases that are recognized by the World Health Organization (WHO) as having significant medical, social, and economic implications. There has been an insufficient amount of attention paid to the development of strategies to overcome these diseases [3].

Helminths are divided into two phyla: roundworms (Nematoda) and flatworms (Platyhelminthes). In addition to the free-living Turbellaria, the Platyhelminthes contain the largest clade of obligate parasites, the Neodermata [4]. The Neodermata are characterized by a surface-spanning syncytial tissue called tegument or neodermis. Three classes are recognized within the clade: the Monogenea, Cestoda, and Trematoda. The Monogenea are ectoparasites of fish found on the skin, gills, and fins of fish [5,6]. A feature common in all Monogenea is a specialized attachment organ called the haptor, which can include hooks, anchors, or clamps and is used to attach to the host [5]. Cestodes or tapeworms have an anterior scolex, used as an attachment organ to anchor the worms within the host's intestinal wall [7]. Contrary to the Monogenea, tapeworms are known to cause major diseases in humans, such as *Taenia solium* or pork tapeworm, the causative agent of cysticercosis [8], or *Echinococcus multilocularis* causing alveolar echinococcosis [9]. Trematodes are the final and largest group within the Platyhelminthes. Trematodes, especially the Digenea, are known for having complex lifecycles with at least two hosts, mainly consisting of a mollusk intermediate host and a mammalian final host. Two diseases caused by trematodes are recognized as NTDs, namely schistosomiasis and a more diverse set of diseases summarised as “Foodborne Trematodiases” (FBT) [3]. These diseases share their way of infection, by consumption of fish, crustaceans, or vegetables. The parasites causing these diseases are the carcinogenic liver flukes *Clonorchis sinensis* and *Ophisthorchis viverrini,* the oriental lung fluke *Paragonimus westermani,* and two *Fasciola* species, the common liver fluke *Fasciola hepatica* and giant liver fluke *Fasciola gigantica.* Together, these diseases are

estimated to cause 7,000 deaths per year, and they are associated with 2,000,000 disability-adjusted life years [10].

1.2 The common liver fluke: *Fasciola hepatica*

The species *F. hepatica* and *F. gigantica* cause the disease known as fascioliasis. The parasites have differing geographical distributions; while *F. hepatica* is found worldwide, *F. gigantica* is found in more tropical areas [11]. In some regions where both species are found, hybridization/introgression events between the two species are suspected, which has important implications for the spread of resistance or increase of virulence [12]. The parasites mainly infect livestock, and the infection has a huge economic impact by reducing growth and milk yield, where the estimated loss is about 3 billion US dollars each year [13].

The lifecycle of *F. hepatica* involves two hosts, as is typical for trematodes of the Digenea ([14], Fig 1.1). The adult worms reside in the bile duct of the final host and release their eggs via bile into the duodenum. Here, the eggs are passed into the environment together with the feces. Released eggs develop over two to three weeks, depending on temperature, moisture, or oxygen [15]. The first larval stage, called miracidium, actively move with their surroundings, searching for the intermediate mollusk host. The lifespan of miracidia is relatively short (between 8 - 24 h) [16,17] as this larval stage uses internal glycogen and lipid storage as its sole energy source [18]. Intermediate hosts are snails of the family of Lymnaeidae, where more than 20 species have been identified to be susceptible to *Fasciola* infection [19]. Within the snail, the parasite undergoes an asexual reproduction process by which a mother sporocyst first gives rise to several daughter rediae. This temperature-dependent development takes 6 - 8 weeks [20]. Additionally, it has been shown that parasites can survive within hibernating snails during the winter, resulting in seasonal infection patterns [21]. During the clonal expansion of the parasite within the snail host, a single miracidium can give rise to 500 - 3200 cercariae [22]. These cercariae exit the snail, and free cercaria can attach to vegetation using their sucker and form a cyst. This encysted metacercaria is covered by a complex layer of proteins and mucopolysaccharides and can remain infective for more than a year, depending on moisture and temperature [23]. The final mammalian host is infected by ingesting metacercaria, which, after passage through the digestive tract, gives rise to the newly excysted juveniles (NEJ). The NEJs then penetrate the intestinal wall, and after around three days, depending on the host, they can be found in the liver tissue. There, juvenile worms feed on the tissue and grow rapidly, thereby causing substantial damage. The adult worms finally reside in the bile duct, releasing up to 20,000

eggs per day [24]. Within the bile duct, these worms can persist for several years or even decades [25].

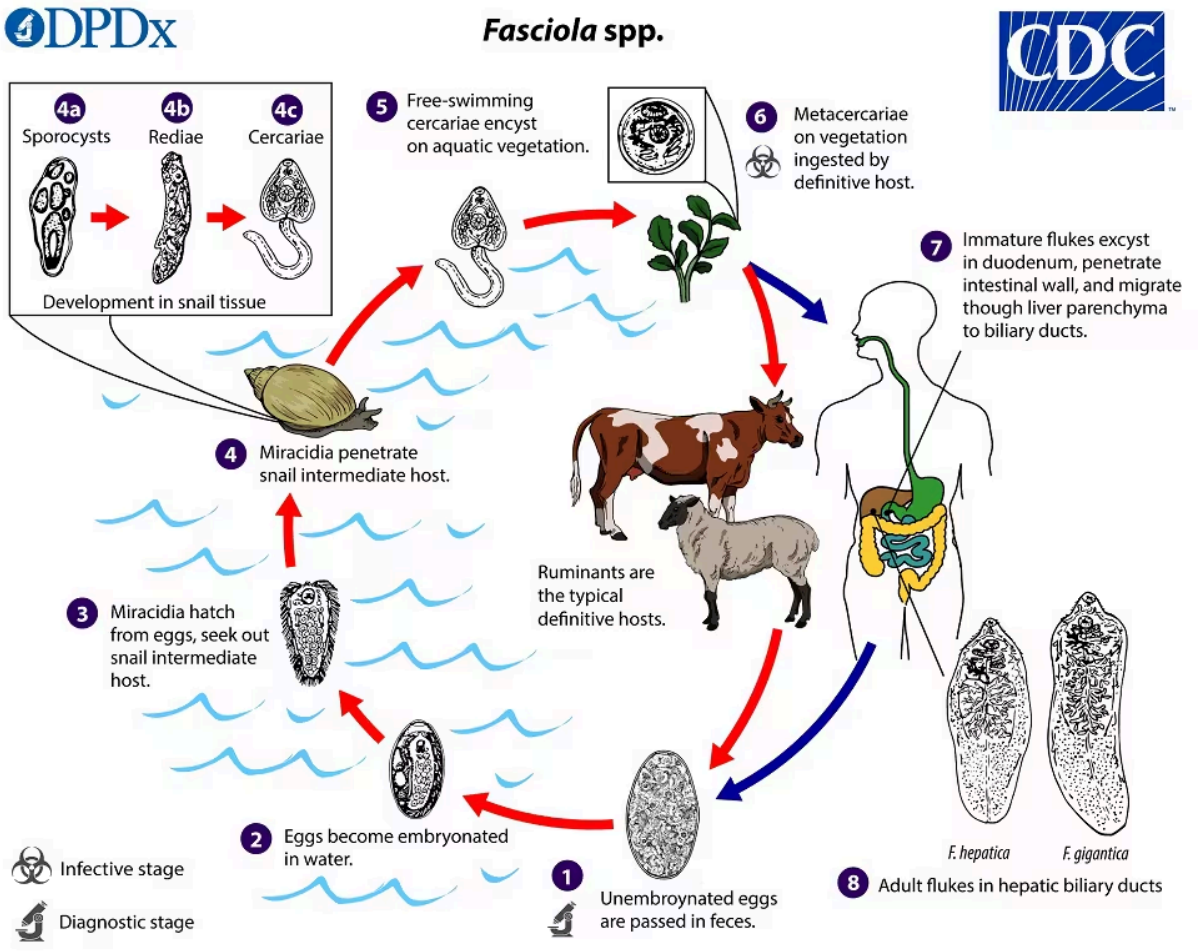


Figure 1.1 Lifecycle of *F. hepatica:* Source Centers for Disease Control and Prevention [26]

1.3 Epidemiology and pathology of fascioliasis

The zoonotic disease, fasciolosis, is divided into two phases, which are tightly linked to the lifecycle of the worms within the final host. The acute phase of the infection is caused by migrating larval stages, where the worms cause necrotic and hemorrhagic lesions during their feeding behavior. This leads to a strong inflammatory response linked with eosinophilia and granuloma formation. The chronic phase is caused by adult worms residing within the bile duct of the host. However, it must be noted that different stages of worms can be present at any point. Pathology and clinical symptoms can vary drastically depending on factors like the infective dose or parasite strain. Additionally, host species, age, and health status can change the course of infection [27].

1.3.1 Human disease

Human fascioliasis is a global health problem with an estimated number of infections around two million [28,29]. However, these numbers present rough estimates, and a lack of thorough surveys might lead to underestimating the disease burden [28,29]. Areas where humans and livestock are in close contact are most severely affected by the disease, with high reported prevalences for selected regions in South-America [30]. In humans, the acute phase of infection can present itself with high fever, abdominal pain, and hepatomegaly. Usually, chronic infection causes hyperplasia of the epithelium, cholangitis, or cholecystitis [20].

1.3.2 Animal disease

Owing to their economic significance, sheep and cattle are the most frequently mentioned animals concerning the disease. Moreover, the parasites have a broad mammalian host range, and wildlife infections are frequently reported [31]. Usually, sheep are more severely affected by the disease, with heavy infections leading to sudden death. Other species, like cattle, are less affected, and the chronic stage of the disease is more prevalent [32]. Additionally, differences in the susceptibility and duration of the disease have been observed between sheep and cattle, as the latter was shown to develop some natural resistance to the infection while sheep do not, which has been suggested to be related to differences in the immune response [33–36]. The disease can persist for several years in sheep [37], while infection in cattle has not been shown to persist for longer than two years [38].

1.4 Treatment of fascioliasis and resistance

1.4.1 Treatment

Although several drugs are available to treat fascioliasis, the benzimidazole triclabendazole (TCBZ) is the most effective option and the only approved drug for treatment in humans [39]. This is due to the differing efficiency of available treatment options against different parasite stages, as TCBZ is the only drug able to kill juvenile stages as early as two weeks post-infection or even one day post-infection at higher doses [40,41]. TCBZ treatment disrupts the tegument and causes apoptosis in the testes, ovary, and vitelline tissue [42,43]. In line with this was the observation that egg production is drastically reduced following TCBZ treatment after 24 h. The mechanism by which TCBZ acts against *F. hepatica* is not fully understood. It is implicated that it binds to ß-tubulin due to the targeting of microtubules [43]. Another drug of the class of benzimidazoles is albendazole (ABZ), which is effective against mature worms and has also been shown to be able to bind to the parasites' ß-tubulin

[44], but there have been implications that binding sites might differ between ABZ and TCBZ [45,46]. Similar to TCBZ, morphological changes were the most prominent in the testes, ovary, and vitelline cells [47].

Other treatment options include closantel, which disrupts the parasites' energy metabolism by uncoupling the oxidative phosphorylation [48], a function that was also suspected for another compound called nitroxynil [48]. Additionally, clorsulon, another drug, is used in combination therapies with TCBZ or ivermectin with some success [49,50]. Finally, an extract prepared from the plant myrrh is marketed in Egypt as Mirazid, and initial studies have shown promising results for its use against the infection.

1.4.2 Resistance

The abundance of use and reliance on TCBZ has led to the spread of resistance in various regions. Since the first report on resistance came from a farm in Australia [51], resistance to TCBZ has been reported, especially in sheep and cattle [52], but human cases have also been described [53,54]. Although not widespread, recently, the first report on TCBZ-resistant worms was made from a farm in Germany, in which a substantial number of sheep died as a result of treatment failure [55]. Resistance was also shown to develop against other compounds such as ABZ [56] or closantel [57] and even more alarming descriptions of fluke populations that showed resistance to several drugs [58].

The mechanisms underlying resistance to TCBZ are not clearly understood. At the time, three main hypotheses are discussed [59]. Typical for resistance to benzimidazoles like TCBZ or ABZ are mutations in the potential target ß-tubulin as described for nematodes [60]. However, the exact binding mechanism for TCBZ has not yet been elucidated. Therefore, it is difficult to conclude on mutations as cause of resistance. Other discussed options are changes in the drug uptake mechanism based on the observation that resistant parasites showed lower concentrations of TCBZ [61]. The role of drug efflux pumps or ABC transporters in the resistance has been indicated [62–64], and change in drug metabolism might lead to a lower drug concentration [65,66]. Finally, a genomic comparison of resistant and susceptible strains identified a single locus containing 30 genes as the basis for resistance [67], which might lead to a better understanding of the resistance in the future.

1.4.3 Development of novel treatment strategies

Developing a new treatment strategy, especially identifying new drug targets and effective compounds, requires much time and effort. The lack of an effective vaccination [68] and the threat of TCBZ resistance urge the need for novel treatment options. Strategies include, for example, the screening of plant-derived compounds [69,70], the inhibition of proteolytic

enzymes [71], or the testing of derivatives of benzimidazoles [72] or a derivative of TCBZ directly [73].

A milestone was recently achieved for schistosomal research as the target for praziquantel (PZQ) was discovered, the only effective drug to treat schistosomiasis [74]. PZQ was found to target a transient receptor potential ion channel termed $TRPM_{PZQ}$. While effective against many trematodes, PZQ is inactive against *Fasciola.* Identifying the orthologue in *F. hepatica* uncovered a single amino acid change as the basis for this resistance [75]. By structural activity assay, the Marchant lab disovered new potent agonists of the *Fasciola* $TRPM_{PZQ}$ channel for which our lab confirmed potent flukicide activity surpassing TCBZ [76].

Additionally, protein kinases (PKs) have been discussed as promising drug targets, showing efficacy in parasitic worms such as cestodes, filaria, or other trematodes [77]. PKs are enzymes that act by phosphorylating other proteins, which can change interaction patterns, functions, or localization and are involved in most, if not all, major biological processes [78]. Furthermore, some PKs have been found to play critical roles in multiple diseases, including cancer, which has led to the development of numerous small-molecule inhibitors for human PKs. A number of 80 inhibitors were already approved for use in humans [79]. We recently reported promising *in vitro* activity of two kinase inhibitors against *F. hepatica* [80]. The study assessed the impact of inhibitors on worm motility, egg production as well as uptake and drug metabolism. We could illustrate the promise of kinases as a research object for drug target discovery against this parasite.

1.5 Omics-technologies in liver fluke research

The urgent need for drug target discovery in liver fluke research has led to many open questions concerning various biological topics. These include host-parasite interaction, developmental biology, and genomics. Pressing questions include understanding how the parasite interacts with the host environment, particularly how it influences the immune system [81]. Another question is how the parasite maintains longevity, as the disease can persist for several years [82,83]. Significant research interest has been focused on stem cells, which have been described as the main driver of the disease in related worms [83–85]. These big biological questions cannot be answered by focusing on a single gene or protein; instead, they need to be investigated from a broader perspective. Omics technologies can provide an opportunity for research in this area.

Analysis regarding the systemic study of biologically relevant molecules such as DNA, RNA, or proteins is collectively called “omics”-technologies. These technologies enable researchers to study the complex structure of biological systems rather than focusing on

single aspects. Liver fluke research has benefited greatly from omics methods, and various datasets have been generated and enhanced in recent years (Fig 1.2).

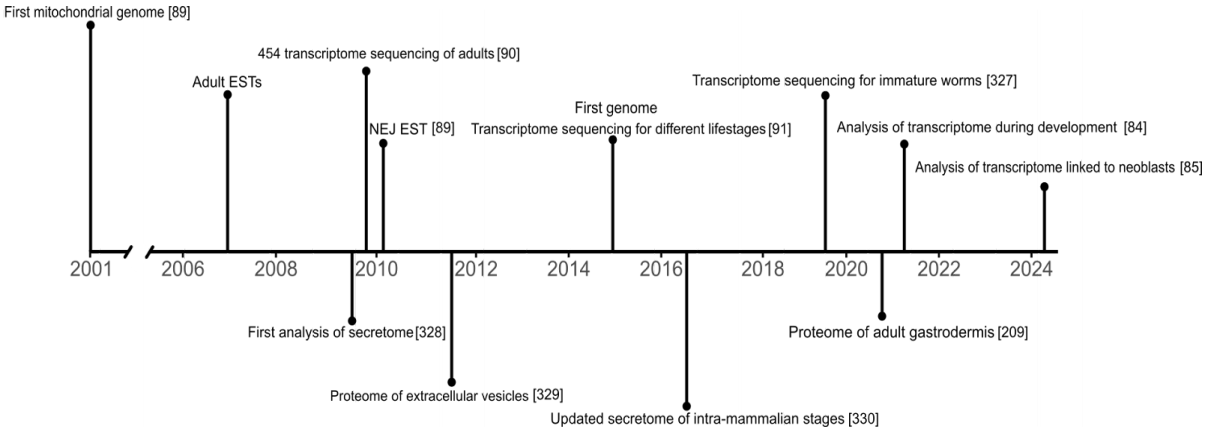


Figure 1.2 Selected milestones of "omics"-related research over time. Schematic modified after [86]. References are indicated.

1.5.1 Proteomics

The field of proteomics involves the identification and classification of proteins. In liver fluke research, proteomics mainly focused on analyzing the secretome of different life stages, which is made up of proteins that the parasite excretes or secretes into the environment including its host. Proteomics has been successfully applied to several life stages and worm tissue. Based on those results, several interesting aspects of the fluke's biology could be uncovered. This included the finding that metacercaria are, contrary to the belief at the time, metabolically inactive and that 70% of the secretome of NEJs could be attributed to 10 proteins [87]. It also enhanced our understanding of immunomodulation during the early stages of infection, and analysis of the secretome has had a major influence on the development of a vaccination against *F. hepatica* [88]

1.5.2 Transcriptomics and genomics

The first study, which can be grouped under the term "omics", was the sequencing of the *F. hepatica* mitochondrial genome [89]. Since then, mitochondrial genes have been used as markers for phylogenetic classifications of specimens in various regions and the differentiation between *F. hepatica* and *F. gigantica.*

The field of transcriptomics is concerned with analyzing gene expression by measuring transcripts, most often mRNA. Insights into gene expression over different life stages, environmental stimuli, or tissue types allow us to uncover valuable biological insights. The Wellcome Trust Sanger Center conducted initial transcriptomic studies, describing 6,819 expressed sequence tags (EST) derived from adult worms. While not annotated, this data aided in the initial description of several genes and the annotation of the first secretome. A

follow-up on NEJs detected 1,684 ESTs, detecting the first differences between the life stages, including the differential expression of cathepsins.

The first large-scale analysis of transcriptome data was performed using Roche 454 sequencing on adults [90], where the sequences were annotated by searching homologous sequences in other organisms and functionally annotated by tools such as gene ontology (GO). The groundbreaking development of Illumina sequencing had a transformative impact on understanding the biology of liver flukes. Together with the first genome, Cwiklinski and colleagues [91] reported transcriptome data for several life stages (metacercaria, NEJs, immature worms, and adult worms), and another genome assembly was added later [92]. The genome of the liver fluke was found to be the largest trematode genome sequenced to date, with a size of 1.3 Gb. The large genome size was found to be a trait acquired by the *Fasciola* species and was caused by longer repetitive elements and intergenic sequences like transposable elements [93]. Another major insight from genome sequencing was the expansion of several major key gene families. These expanded families were, for example, papain-like cysteine peptidases and peptidase inhibitors. Differential expression analysis between the different life stages showed highest differences in the described gene families, as well as the tubulins that were described earlier to be an extended gene family in the liver fluke [94,95]. In recent years, transcriptome analysis has been used to elucidate the molecular mechanisms during the development of *in vitro* grown worms, thereby identifying genes linked to proliferation [84], which was further expanded by RNA-seq of irradiated worms [83].

1.6 Single-cell transcriptomics

Although classical RNA-seq has transformed molecular biology, a major limitation of so-called bulk approaches is that these techniques can only analyze multiple cells or whole tissue samples. This limitation leads to the loss of heterogeneity, averaging the transcriptomics differences between single cells or masking the detection of transcripts only present in rare cell types.

The first unbiased RNA-seq approach in single-cell resolution was published in 2009 using blastomeres and oocytes [96]. Since then, technologies have rapidly developed that increase the throughput and sensitivity of these assays [97]. Methods can be divided into two popular approaches: plate-based, where cells are separated into individual wells, and protocols using microfluidic devices that encapsulate cells into droplets. Regardless of the protocol, the cells are lysed after capture, and mRNA is usually captured using poly(T)-containing primers. These primers include a barcode later used for cell deconvolution and a unique molecular identifier (UMI) accounting for amplification bias [98].

Both major approaches have their advantages. While plate-based approaches have a favorable sensitivity and can produce full-length transcript information, droplet-based approaches are superior in throughput, capturing thousands of cells in a single experiment [99]. The most widely used protocol for droplet-based systems is the commercially available 10x Genomics Chromium, which improved the two initial methods, DropSeq and inDrop [100–103].

Single-cell sequencing has had a transformative impact on biology, measuring the cellular heterogeneity of diseases, particularly cancer, or allowing the undertaking of cell atlas projects [97]. These projects aim to profile the cellular heterogeneity of tissue types or even whole organisms, enabling an unprecedented understanding of their cellular composition. The possibility has especially impacted research on non-model organisms due to their independence of molecular tools like antibodies. Transcriptomes were generated for various organisms, including zebrafish [104], frogs [105], and cnidarians like *Nematostella [106].*

For Plathelminthes, single-cell transcriptomics has been successfully utilized in the planaria *Schmidtea mediterranea,* identifying most cell types and describing previously unknown cells [107,108]. For parasitic trematodes, single-cell RNA sequencing (scRNA-seq) was only used in works characterizing *Schistosoma mansoni.* The first report identified stem-cell subtypes within the larval stages of the parasite using the Fluidigm C1 system [109,110]. In more recent years, droplet-based systems have been adopted for various life stages of the parasite to gain a more detailed picture of the cellular diversity within the blood fluke. A total of 43,642 cells from adult worms of both sexes were profiled, and 68 clusters were identified [111]. Following this milestone work, other studies were published profiling the single-cell transcriptomes of schistosomula [112], sporocysts [113], juvenile worms [114], and miracidia [115]. These studies served as a tool to elucidate the biology of these worms in diverse aspects such as drug discovery [116], the function of non-coding RNAs [117] or tegument differentiation [118]. The impact these technologies had on schistosome research illustrates the promise they hold for related parasitic flatworms, including *F. hepatica.*

1.7 Aims

Technologies like bulk RNA-seq have advanced the understanding of the unique biology of *F. hepatica* [86,88]. Despite their continued impact on the field, they lack cellular resolution, which limits the insights gained from the generated data. The introduction and continued development of scRNA-seq has revolutionized molecular biology and offers a new avenue to investigate a multicellular organism. This thesis aimed to create a scRNA-seq based atlas that profiles the cellular landscape of the liver fluke. A successful scRNA-seq project depends on several factors, such as the efficient dissociation of cells from fluke tissue, cell viability, and careful validation of the gained data. Therefore, the following milestones were formulated:

1) Establishing a protocol to generate a viable cell suspension from *F. hepatica* adults.
2) Verification of the suspension's suitability for scRNA-seq by determining RNA quality and validation of presence for selected cell types by imaging approaches and polymerase chain reaction (PCR).
3) Executing a scRNA-seq experiment using the 10x Genomics Chromium system and the subsequent data analysis.
4) Validation of the analysis by localizing selected marker genes by RNA *in situ* hybridization.
5) Application of the dataset by identifying potential drug targets in combination with drug target validation by inhibitor testing *in vitro.*

2. Material and Methods

2.1 Materials

2.1.1 List of chemicals

The chemicals used in this work are listed in Table 2.1

Table 2.1 Chemicals used in this work

Chemical	Manufacturer	Reference
Acetic acid	Roth	3738.5
Agarose	Roth	2267.3
Bovine serum albumin (BSA)	Roth	8076.1
BSA (ultrapure)	Sigma	126609
Blocking reagent	Roche	11096176001
5-bromo-4-chloro-3-indolylphosphate (BCIP)	Roche	11681451001
Calcium chloride ($CaCl_2$)	Roth	CN93.1
Diethylpyrocarbonate (DEPC)	Roth	K028.3
Deionized formamide	Roth	P040.2
Dextran sulfate	Roth	5956.4
Dimethyl sulfoxide (DMSO)	Roth	A994.1
Dithiothreitol (DTT)	Roth	6908.1
5-ethynyl 2´-deoxyuridine (EdU)	Invitrogen	A10044
Ethylenediaminetetraacetic acid (EDTA)	Roth	8043.2
Ethanol	Roth	9065.4
Formaldehyde	Roth	4979.1

Chemical	Manufacturer	Reference
ROTIMountFluorCare	Roth	HP19.1
Formamide	Roth	6749.1
Glycerol	Roth	7530.1
Goat serum	Gibco	16210064
Horse serum, heat-inactivated	Gibco	16050130
Hydrochloric acid (HCl)	Roth	T134.1
2-4-(2-hydroxyethyl)-1-piperazinyl ethane sulfonic acid (HEPES)	Roth	6763.3
Lithium chloride (LiCl)	Roth	P007.1
Magnesium chloride ($MgCl_2$)	Roth	KK36.3
Maleic acid	Sigma	N5625
Methanol	Roth	4627.1
Nitro blue tetrazolium chloride (NBT)	Roche	11383213001
Orange G	Roth	0318.2
O.C.T (Tissue-Tek® Compound)	Sakura	SA62550-01
Sodium hypochlorite	Roth	9062.3
Sodium chloride (NaCl)	Roth	3957.1
Triton X-100	Roth	3051.3
Trisodium citrate 2-hydrate	Roth	3580.3
Tween20	Roth	9127.2
Trypan Blue	Gibco	15250061
Yeast RNA	Roche	10109223001

2.1.2 Buffers and solutions

Buffers and solutions used in this work can be found in Table 2.2

Table 2.2 Buffers and solutions used in this work

Buffer/solution	Composition	Application
Alkaline Phosphatase (AP) buffer	100 mM Tris-HCl 100 mM NaCl 50 mM $MgCl_2$ in DEPC-H_2O pH 9.5	*In situ* hybridization
Bleaching solution	1 ml 6-14 % NaClO 9 ml dH_2O	Excystment
Blocking buffer ISH	0.5% Blocking reagent 1% BSA in MAB-T	*In situ* hybridization
Blocking buffer IHC	5% Goat serum in 1xPBS + 0.3% Tween20	Immunhisto-chemistry
Denhardt's (100x)	2x SSC 2% Ficoll 2% Polyvinylpyrolidon 2% BSA in dH_2O	*In situ* hybridization
Developing buffer	450 µg/ml NBT 175 µg/ml BCIP in AP buffer	*In situ* hybridization
DIG-NTP mix	10 mM ATP 10 mM CTP 10 mM GTP 7 mM UTP 3.5 mM DIG-UTP in DEPC-H_2O	*In situ* hybridization
FITC-NTP mix	10 mM ATP 10 mM CTP 10 mM GTP 7 mM UTP 3.5 mM FITC-UTP in DEPC-H_2O	*In situ* hybridization
dNTP mix	10 mM each dNTP	PCR
Hatching solution	0.6% $NaHCO_3$ 0.4% Sodium tauroglycocholate 0.025 M HCl 0.4% L-cysteine	Excystment

Buffer/solution	Composition	Application
Hank's Balanced Salt Solution (HBSS)	136 mM NaCl 5.365 mM KCl 1.261 mM $CaCl_2$ 0.4 mM $MgSO_4$ 0.491 mM $MgCl_2$ 0.337 mM Na_2HPO_4 0.44 mM KH_2PO_4 1 g/l glucose 4.166 mM $NaHCO_3$	Cell dissociation
Hybridization buffer	5x SSC 50% deionized formamide 10% dextran sulfate 1 mg/ml Yeast RNA 1x Denhardt's Solution in DEPC-H_2O	*In situ* hybridization
Hybridization wash buffer	5x SSC 50% formamide 0.1% Tween-20 in DEPC-H_2O	*In situ* hybridization
Maleic acid buffer (MAB-T)	0.1 M Maleic acid 0.15 M NaCl 0.1% Tween20 in DEPC-H_2O pH 7.5	*In situ* hybridization
10x Phosphate buffered saline (PBS)	1.37 M NaCl 27 mM KCl 100 mM Na_2HPO_4 17.5 mM KH_2PO_4 in dH_2O ph 7.4	General
PBSTx	1x PBS 0,5% Triton-X100	*In situ* hybridization
PBSTw	1x PBS 0,5% Tween-20	Immunhisto-chemistry
Peroxidase inactivating solution	0.003% H_2O_2 in 4x SSC	*In situ* hybridization
20x SSC (Saline sodium citrate)	3 M NaCl 0.3 M $Na_3C_6H_5O_7$ in DEPC-H_2O pH 7.0	*In situ* hybridization

Buffer/solution	Composition	Application
50x TAE (Tris-Acetat-EDTA) buffer	2 M Tris 50 mM EDTA 5.71% Acetic acid in dH_2O pH 8.0	Gel electrophoresis
10x Transcription buffer	0.4 M Tris 0.1 M $MgCl_2$ 20 mM Spermidine 0.1 mM DTT in DEPC-H_2O	dsRNA synthesis

2.1.3 Media and supplements

The media and supplements used in this work are listed in Table 2.3

Table 2.3 Media and media supplements used in this work

Media/supplement	Manufacturer	Reference	Working concentration
ABAM	C.C. pro	2-18-M	1%
Agar Agar	Roth	5210.1	1%
Ampicillin	Sigma	A9518	100 µg/ml
Chicken serum	Gibco	16110082	5 - 50%
Kanamycin	Sigma	K4000	50 µg/ml
LB Media	Roth	X964.2	20 g/l
M199	Roth	31100027	1x
Newborne Calf serum	Sigma	N4637	10%
1x RPMI Medium 1640	Gibco	11835	1x

Special media were used to maintain the worms for *in vitro* culture experiments using *S. mansoni* or *S. mediterranea.* The composition is listed in Table 2.4 below.

Table 2.4 Media used for worm culture of schistosomes and planarians

Buffer/solution	Composition	Application
M199 3+	1% ABAM 1% HEPES (ph 7.4) 10% NCS Add M199	*S. mansoni* culture
Montjuic Water	1.6 mM NaCl 1 mM $CaCl_2$ 1 mM $MgSO_4$ 0.1 mM $MgCl_2$ 0.1 mM KCl 1.2 mM $NaHCO_3$ Add ddH_2O	*S. mediterranea* culture

2.1.4 Dyes and conjugates

Dyes used in this study with their applicable working concentration can be found in Table 2.5

Table 2.5 Dyes and conjugates used in this work

Dye	Manufacturer	Reference	Concentration/Dilution
Calcein-AM	Invitrogen	65-0853-39	5 - 25 nM
Concanavalin A	Invitrogen	C11252	2 µg/ml
DRAQ5	Abcam	ab108410	20 µM
GelRed nucleic acid gel stain	Biotium	41003	1: 40,000
Hoechst 33342	Invitrogen	H1399	1:10,000
Lipidspot	Biotium	70065	1:1000
Zombie Red	BioLegends	423109	1:1000

2.1.5 Enzymes

The enzymes used in this work are listed in Table 2.6

Table 2.6 Enzymes used in this work

Enzyme	Manufacturer	Reference	Application
AccuPrime Taq DNA-Polymerase	Invitrogen	12346086	PCR
AdHI	NEB	R0584S	Cloning
Bromelain	Sigma	B4882	Cell dissociation
Collagenase I	Sigma	SCR103	Cell dissociation
Collagenase II	Sigma	C2-22-1G	Cell dissociation
Dispase II	Roche	04942078001	Cell dissociation
DNase I (RNase-free)	NEB	M0303L	dsRNA/probe synthesis
DNAse I	Sigma	DN25-100MG	Cell dissociation
Elastase	Sigma	E0258	Cell dissociation
FirePol DNA Polymerase	Solis Biodyne	01-01-00500	PCR
Hyaluronidase	Sigma	H3506-5G	Cell dissociation
Papain	Roche	10108014001	Cell dissociation
Proteinase K	Roche	03115801001	*in situ* hybridisation
Protease from *Bacillus licheniformis*	Sigma-Aldrich	P5380	Cell dissociation
Protector RNase Inhibitor	Roche	RNAINH-RO	Cell dissociation
Q5 High Fidelity DNA Polymerase	NEB	M0491S	PCR
SP6 RNA Polymerase	Roche	RPOLSP6-RO	*In situ* hybridization

Enzyme	Manufacturer	Reference	Application
T3 RNA Polymerase	Roche	RPOLT3-RO	*In situ* hybridization
T7 RNA Polymerase	Selfmade		dsRNA synthesis
Trypsin	Sigma-Aldrich	T4549	Cell dissociation
T4 DNA Ligase	NEB	M0202L	Cloning
RNase inhibitor	NEB	M0314S	dsRNA/probe synthesis

2.1.6 Kits

The kits used in this work are listed in Table 2.7.

Table 2.7 Kits used in this work

Product	Manufacturer	Reference	Application
Agilent RNA 6000 Nano Kit	Agilent Technologies	5067-1511	RNA Isolation
Agilent RNA 6000 Pico Kit	Agilent Technologies	5067-1513	RNA Isolation
Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1	10x Genomics	PN-1000128	Single-cell RNA sequencing
Chromium Next GEM Chip G Single Cell Kit	10x Genomics	PN-1000127	Single-cell RNA sequencing
Click-iT Plus EdU Alexa 488 Imaging Kit	Invitrogen	C10637	Imaging of proliferating cells
Monarch DNA Gel Extraction Kit	NEB	T1020	Isolation of DNA
Monarch PCR & DNA Cleanup Kit	NEB	T1030	Isolation of DNA
Monarch Plasmid Miniprep Kit	NEB	T1010	Isolation of DNA

Product	Manufacturer	Reference	Application
Monarch Total RNA Miniprep Kit	NEB	T2010	Isolation of RNA
QuantiTec Reverse Transcription Kit	Qiagen	205313	cDNA synthesis
Single Index Kit T Set A	10x Genomics	1000213	Single-cell RNA sequencing
SuperScript IV Reverse Transcriptase Kit	Invitrogen	18090010	cDNA synthesis
TSA Plus Cyanine 3	Akoya Bioscience	NEL704A001KT	*In situ* hybridization
TSA Plus Cyanine 5	Akoya Bioscience	NEL705A001KT	*In situ* hybridization
TSA Plus Fluorescein	Akoya Bioscience	NEL741001KT	*In situ* hybridization

2.1.7 Antibodies

Antibodies used in this work are listed in Table 2.8.

Table 2.8 Antibodies used in this work

Antibody	Supplier	Reference	Source	Working dilution
Anti-Digoxigenin-POD, Fab fragments	Roche	11207733910	Sheep	1:200
Anti-Digoxigenin-AP, Fab fragments	Roche	11093274910	Sheep	1:1000
Anti-Fluorescein-POD, Fab-fragments	Roche	11426346910	Sheep	1:200
Goat anti-Mouse IgG Alexa Fluor 488	Invitrogen	A32723	Goat	1:1000
Goat anti-Mouse IgG Alexa Fluor 647	Invitrogen	A21235	Goat	1:1000

Antibody	Supplier	Reference	Source	Working dilution
Anti-Histone H3 (phospho S10 + T11) antibody	Abcam	ab32107	Rabbit	1:1000
6G10-2C7 (muscle antibody)	Developmental Studies Hybridoma Bank	6G10-2C7	Mouse	1:500
3C11 (anti synapsin)	Developmental Studies Hybridoma Bank	3C11	Mouse	1:1000

2.1.8 Molecular weight standards

DNA ladders used in DNA gel electrophoresis can be found in Table 2.9

Table 2.9 DNA ladders used in this work

Ladder	Supplier	Reference
HyperLadder 50 bp	meridian bioscience	BIO-33054
HyperLadder 1 kb	meridian bioscience	BIO-33053

2.1.9 Plasmids

Table 2.10 contains the plasmids generated in this work. pJC53.2 [119] was used as a backbone for all generated plasmids. All plasmids were utilized for either riboprobe or dsRNA synthesis. Primers used for cloning can be found in table 2.11

Table 2.10 Plasmids used/generated in this work

Name	Reference	Insert
pJC53.2	[119]	not applicable
Cathepsin_WISH	This work	PCR fragment of D915_011077
H2B_WISH	This work	PCR fragment of D915_007751
Bmpg_ISH	This work	PCR fragment of D915_005862
Tyrosinase_ISH	This work	PCR fragment of D915_002718

Name	Reference	Insert
NomuscleMyosin_ISH	This work	PCR fragment of D915_003945
VB1_ISH	[120]	PCR fragment of D915_010963
Tektin_1_ISH1	This work	PCR fragment of D915_008401
Meiob_ISH	This work	PCR fragment of D915_001989
Pak4_ISH1	This work	PCR fragment of D915_004414
Collagen_ISH	This work	PCR fragment of D915_008507
Phospholipase B-like_ISH	This work	PCR fragment of D915_003832
Elf5_ISH1	This work	PCR fragment of D915_002050

2.1.10 Primers

The primers used in this study are listed in the following chapter and are separated by application. Numbers represent given numbers for all ordered primers in the project.

Table 2.11 Primers used for cloning. Upper sequence represents forward, lower sequence reverse primer

Gene Name	Gene ID	Number	Sequence 5’ - 3’
Histone H2B	D915_007751	21/22	GTGGACAAGAGGAAGAAGCG
			CTTCGTCCCCTCCGAAACT
Cathepsin L	D915_011077	5/2	GCCAAAGTGATTAACTTTCAAACG
			CAACCGAATGTAACCAGACTC
Nonmuscle myosin II	D915_003945	335/336	TCCAACGTGTTTGGCATGTTC
			TGGTTGCCACGTTTGATCAGT
Tyrosinase	D915_002718	141/142	GAGCAACATCACGGTGACC
			GATCAAGTACTGATAGTCTACATC
Vitelline protein BI	D915_010963	195/196	TCGATATGTATGGAAATGTGAAGG
			GTAACTGTATTTGTCATATGGAGTC
Tektin	D915_008401	209/210	TGCGATTAGCCAGACACAGC

Gene Name	Gene ID	Number	Sequence 5’ - 3’
			CGTTCCAGTTCGGTTTTCCA
Pak4	D915_004414	283/284	CCGAAAGCCACACAGAAATGA
			CCCGTCTTTGTTGTTTGAATATGT
Meiob	D915_001989	237/238	TTGGACATTCAGTCGTGTTGC
			GGGGAAGTCTGGCAAATGTC
Collagen	D915_008507	361/362	CTTGACGGGTCGTTTGGGAA
			CCTGGACTACCGTCTTTTCC
Bmpg	D915_005862	41/42	CAAGTTCAAGCGCTGGTCGT
			GGTACGGCAATCCTTTTCTCA
Phospholipase B-like	D915_003832	303/304	CAGTAAAAACGCATCGAACAGAG
			CAGTTAAAGGACGACTTGAGTG
Elf5	D915_002050	437/438	CTAGCTCAGATGCTGATTTGAAG
			TGTCCTGCTTCATCTATAACAAATC
pJC53.2_T7_extended	not applicable		CCTAATACGACTCACTATAGGGAG
pJC53_seq_s	not applicable		TTCTGCGGACTGGCTTTC TAC

Table 2.12 Primers used for RT-qPCR

Gen Name	Gene ID	Number	Sequence 5’ - 3’
PAK4	D915_004414	425/426	GAGATGGTTGACGGTGAGC
			GTTGGGCTGCCGTTGCC
Tbcd	D915_002510	[121]	CAGCAGCCGCATTTCAGG
			AGCCAAATGGTCAATCATCGC
Eprs	D915_002574	[121]	TACACCACAACCATCGAG
			GTGGTCAATCCCCACGAGT

Table 2.13 Primers used for cell type validation

Name	Gene ID	Number	Sequence 5’ - 3’
Cathepsin L	D915_011077	5/2	GCCAAAGTGATTAACTTTCAAACG
			CAACCGAATGTAACCAGACTC
Histone H2B	D915_007751	21/22	GTGGACAAGAGGAAGAAGCG
			CTTCGTCCCCTCCGAAACT
Fascilin	D915_002002	47/48	GCTCCACTTGGCTTTGTATAC
			CAATCTTCTCGGCTGGAATAC
Tetraspanin	D915_004328	78/79	GATTGTCACGGCAATTTGCTTCT
			GTATTCACATTTGAGTTCTCCTC
7b2	D915_008396	61/62	CCGATATGGGGTGAACATAGA
			CGCAACAAGTCGTTTTCTTAAGG
MAR	D915_000580	79/80	GTACGGGTGATAGCGAGATG
			CCTAAAATGAATGTGATAGTGCG
Troponin	D915_005504	73/74	CACTCTCTTCACAATGACTTTATG
			CTGCATGATCTTCTCAGAACTAC
Bmpg	D915_005862	41/42	CAAGTTCAAGCGCTGGTCGT
			GGTACGGCAATCCTTTTCTCA

2.1.11 Instruments

Besides the standard laboratory equipment, special instruments were used that are listed in Table 2.14

Table 2.14 Instruments used in this work

Device	Model	Supplier
Bioanalyzer	Agilent 2100 Bioanalyzer	Agilent
Fluoresence microscope	DMIL LED Fluo	Leica
Fluoresence microscope	Olympus IX81	Olympus
Inverse Confocal Laser Scanning Microscope	Leica TCS SP5 VIS	Leica

Device	Model	Supplier
Bright field microscope	TL3000Ergo	Leica
Flow cytometer	S3e	Bio-Rad
Flow cytometer	FACSymphony S6	BD Bioscience
Flow cytometer	FACSMelody	BD Bioscience
Quantitative RT-qPCR	Rotor-Gene Q	Qiagen
Photometer	BioSpectrometer basic	Eppendorf
scRNA-seq	10x Chromium Controller	10x Genomics

2.1.12 Database and software

The databases and software used in this work are listed in Table 2.15 and Table 2.16.

Table 2.15 Online tools and websites used in this work

Website URL
Tree visualization
https://itol.embl.de/ iToL [122]
Plasmid design
https://www.benchling.com/
Primer design
http://biotools.nubic.northwestern.edu/OligoCalc.html Oligo Calc [123]
https://eu.idtdna.com/pages OligoAnalyzer, IDT
Protein sequences and analysis
https://www.uniprot.org UniProt [124]
http://smart.embl-heidelberg.de/ SMART domain analysis [125]
https://www.ebi.ac.uk/interpro/search/sequence/ InterProScan [126]
https://guolab.wchscu.cn/AnimalTFDB4//#/ AnimalTFDB 4.0 [127]
https://planosphere.stowers.org/ Planosphere [128]
https://parasite.wormbase.org/index.html WormBase ParaSite [129]

***Schistosoma* transcriptome data**

www.collinslab.org/schistocyte/ Single-cell RNAseq data of adult *S. mansoni [111]*

www.schisto.xyz *S. mansoni* gene expression data [130,131]

Multiple sequence alignment
https://www.ebi.ac.uk/jdispatcher/ (Clustal Ω, [132])

Table 2.16 Software used in this work

Software	Version	Source	Reference
RNA velocity			
samtools	1.10	http://www.htslib.org/download/	[133]
velocyto	0.17alpa	https://velocyto.org/velocyto.py/index.htmld	[134]
scVelo	0.3.0	https://scvelo.readthedocs.io/en/stable/installation.html	[135]
python	3.9.13	https://www.python.org/	[136]
Jupyter Notebook	6.4.8	https://docs.jupyter.org/en/latest/index.html	[137]
FACS data analysis			
FCSalyzer	0.9.22-alpha	https://sourceforge.net/projects/fcsalyzer/	
FCS Express	7.0	https://denovosoftware.com/full-access/download-landing/	De Novo Software
Sequence analysis			
cellranger	7.0.0	https://www.10xgenomics.com/support/software/cell-ranger/latest	10xGenomics
Interproscan	5.60-92.0	https://ftp.ebi.ac.uk/pub/software/unix/iprscan/5/5.60-92.0	[126]

Software	Version	Source	Reference
BLAST+	2.13.0	https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/2.13.0/	[138]
HMMER	3.3.2	https://www.ebi.ac.uk/Tools/hmmer/	[139]
Image editing and processing			
Inkscape	0.92	https://inkscape.org/release/inkscape-1.3.2/	[140]
Fiji	2.14.0	https://imagej.net/software/fiji/downloads	[141]
qRT-PCR			
Q-Rex	1.0.2	https://www.qiagen.com/de-de/products/discovery-and-translational-research/epigenetics/dna-methylation/methylation-specific-pcr/rotor-gene-q	Qiagen
Q-Rex Gene expression Plugin	2.0.0.4	https://www.qiagen.com/de-de/products/discovery-and-translational-research/epigenetics/dna-methylation/methylation-specific-pcr/rotor-gene-q	Qiagen
Phylogenetic tree construction			
MUSCLE	5.1.0	https://github.com/rcedgar/muscle/releases/tag/5.1.0	[142]
MAFFT	7.505	https://mafft.cbrc.jp/alignment/software/linux.html	[143]
MrBayes	3.2.7a	https://nbisweden.github.io/MrBayes/download.html	[144]
Data Analysis with R			
R	4.3.2	https://cloud.r-project.org/	[145]
RStudio	RStudio 2023.09.1+494	https://posit.co/downloads/	[146]
dplyr		https://cran.r-project.org/web/packages/dplyr/index.html	[147]

Software	Version	Source	Reference
stringr	1.5.1	https://cran.r-project.org/web/packages/stringr/index.html	[148]
tidyr	1.3.1	https://cran.r-project.org/web/packages/tidyr/index.html	[149]
Seurat	5.0.3	https://cran.r-project.org/web/packages/Seurat/index.html	[150]
ggplot2	3.5.0	https://cran.r-project.org/web/packages/ggplot2/index.html	[151]
topGO	2.54.0	https://bioconductor.org/packages/release/bioc/html/topGO.html	[152]
pheatmap	1.0.12	https://cran.r-project.org/web/packages/pheatmap/index.html	[153]
ggVennDiagram	1.5.2	https://cran.r-project.org/web/packages/ggVennDiagram/index.html	[154]

2.2 Methods

2.2.1 Ethics statement

Animal experiments using rats (*Rattus norvegicus)* as model hosts were performed by Directive 2010/63/EU on the protection of animals used for scientific purposes and the German Animal Welfare Act. The experiments were approved by the Regional Council (Regierungspraesidium) Giessen (V54-19c20 15 h 02 GI 18/10 Nr. A16/2018).

2.2.2 Maintenance and culture of worms

2.2.2.1 Newly excysted juveniles (NEJs)

To obtain NEJs, metacercaria of the Italian stain were purchased from Ridgeway Research (UK). Excystment was performed as outlined before [121,155]. Following the mechanical removal of the outer wall, the cyst was treated with a 10% bleaching solution for 3-5 min. Afterward, the metacercaria were washed with H_2O 3 times and incubated in a prewarmed hatching solution at 37 °C and 5% CO_2. The NEJs were collected throughout 1-2 h and placed in RPMI1640 containing 1% ABAM and 10% chicken serumand kept at 37 °C in a CO_2 incubator. The remaining metacercariae were left in the hatching solution overnight, and residual NEJs were collected the following day.

2.2.2.2 Immature and adult fluke

The immature and adult stages were obtained by infection of rats with 20 - 25 metacercariae. To avoid strenuous infections of rats via gavage tube, rats were accustomed to an alternative infection procedure to achieve voluntary feeding, which involved feeding them sugar solution few days before infection by means of a pipette. The rats were infected by feeding them sugar solution containing the infectious metacercariae using a pipette. To obtain immature worms, rats were sacrificed by cervical dislocation after euthanization with CO_2 after 4 weeks of infection. Afterward, livers were collected in prewarmed 0.9 % NaCl containing 1% ABAMl. The livers were chopped into equal pieces, and in 30 min intervals, worms were collected, and the liver was cut again into smaller pieces and 0.9% NaCl solution was replaced. The whole process was performed over a timeframe of around 4 h during which the liver preparation was kept at 37°C in a CO_2 incubator for most of the time. The worms were transferred to RPMI1640, which contained 1% ABAM and 10% chicken serum. For adult worms, rats were kept for at least 12 weeks of infection. After sacrificing the rats, worms could be collected from the rats' bile ducts. For any downstream analysis, the worms were kept in RPMI1640 containing 1% ABAM for at least 1 h to regurgitate their gut

content, thereby avoiding contamination. Both immature and adult worms were maintained at 37°C and 5% CO_2 in RPMI1640 containing 1% ABAM and 10% chicken serum until further use.

2.2.2.3 *In vitro* culture experiments with *F. hepatica*

For experiments testing the efficacy of inhibitors, worms were placed in wells within 12-well or 48-well plates for adult and immature worms, respectively. Adult worms were placed in individual wells containing 3 ml of RPMI containing 10% CS with 1% ABAM and inhibitor in the used concentrations. For immature worms, 650 µl of the same culture media containing 10% CS with 1% ABAM was used, with 3 worms placed in the same well. The solvent of the inhibitor DMSO was used as a control using a volume equivalent to the amount used in the highest concentration of the tested inhibitor. The media and inhibitor were refreshed daily after scoring the worms according to their motility every 24 h. Worm motility was scored using the a brightfield microscope at 4x magnification. The system used 3 for normal motility, 2 for reduced motility, 1 for minimal and sporadic movements, and 0 if flukes were dead. Worms were counted as being dead when no movement even upon gentle squeezing with a plastic forceps was observed. At the end of the experiment, worms were either snap-frozen in 150 µl Monarch RNA Protect Buffer or frozen in O.C.T until further use.

2.2.2.4 Knockdown experiments with dsRNA

In RNA interference (RNAi) experiments, immature worms were placed individually in a 48-well plate containing 3 ml RPMI1640 containing 1% ABAM with 10% heat inactivated CS. The dsRNA was used in a 50 µg/µl concentration per worm. To control for the general effects of dsRNA, a non-helminth dsRNA specific for the bacterial neomycin resistance gene was used in the same concentration [156]. The dsRNA was added with fresh medium on two consecutive days at the start of the experiment, following 48-h intervals for 21 days. At the end of the experiment, worms were either fixed for imaging applications or frozen in Monarch RNA Protect Buffer at -80°C for later RNA isolation.

2.2.2.5 *In vitro* culture experiments with *S. mansoni*

Schistosome couples were obtained by perfusion of Syrian hamsters (*Mesocricetus auratus*), which and were gratefully provided by the Grevelding Lab. The worms were maintained at 37 °C and 5% CO_2 in M199 3+. For inhibitor testing, 10 couples were placed in a 6-well plate containing M199 3+ medium and cultivated in inhibitor-containing media for 72 h. The inhibitor was refreshed every 24 h after scoring the worms according to their motility following [157]. Worm motility was scored using the a brightfield microscope at 4x magnification.

2.2.2.6 *In vitro* culture experiments with *S. mediterranea*

The planarian worms were gratefully provided by the Stowers Institute for Medical Research in Kansas City. The worms were maintained in Montjuic Water at 20 °C, and 10 worms were sorted in 6-well plates with 5 ml water per well. The inhibitor was tested for 72 h and refreshed every 24 h.

2.2.3 Work with *F. hepatica* cells

2.2.3.1 Dissociation of *F. hepatica* adults into single cells

To generate a single-cell suspension of adult worms, they were first allowed to clear their gut content by keeping them in RPMI supplemented with 10% chicken serum and 1% ABAM for at least half an hour after removal from the rat bile duct. The worms were between 12 and 14 weeks old. The worms were washed briefly with HBSS two times and placed into a small petri dish. Finally, the worms were chopped into small pieces using a razor blade and flushed into a 15 ml falcon tube using the digestion solution (5x trypsin in HBSS w/o calcium and magnesium supplemented with 400 U/ml DNAse I and 1 mM EDTA). The use of enzyme was carefully evaluated and tested conditions are shown in Table 3.1. The dissociation was conducted for 30 - 40 min at 37 °C under slight agitation in an incubator. For 5 min intervals, the suspension was carefully pipetted up and down using a 1 ml pipette and wide bore filter tips. The progression of cell dissociation was monitored by checking aliquots using a microscope. The reaction was stopped by adding an equal amount of cold 1x HBSS with 2% BSA and 1 mM EDTA. Afterwards, the cells were filtered using a 100 µm and 40 µm filter. Residual cells were washed from the filter with 5 ml of wash buffer. The cells were then spun down at 300 x g at 4 °C for 10 min following two washing steps using 5 - 10 ml of wash buffer depending on the size of the pellet and centrifugation at 300 x g at 4 °C for 5 min. Finally, the vitality of cells was checked using a 1:10 dilution of cells with trypan blue, and the viable cell count was determinded using a counting chamber. Depending on the later application, the cells were spun down again and frozen in Monarch RNA Protect buffer for RNA isolation, dehydrated and fixed in ACME for long-term storage at 20 °C, or directly stained for flow cytometry. For some applications within this work, worms were divided into anterior and posterior fragments before dissociation. The parts were treated equally following the described protocol.

2.2.3.2 Flow cytometry of cells

Cells were sorted via fluorescence-activated cell sorting (FACS) to enrich viable cells or separate certain cell fractions. Cells were stained using calcein-AM at a final concentration

of 1 µg/ml for 30 min at room temperature. The dye was excited using a 488 nm laser and detected using standard settings. Sorting was done using a 130 µm nozzle and a slow flow rate to avoid damaging sensitive cells due to shear stress.

Fixed cells were sorted using the BD FACSymphony S6 system. Fixed cells were swiftly thawed and centrifuged at 1000 x g at 4 °C for 5 min. The cells were resuspended in 1 ml of washing buffer (1x PBS, 2% BSA, and RNAse inhibitor). Finally, the cells were stained with 1 µl/ml of a 5 mM stock of DRAQ5 to stain nuclei and 2 µl of a 1 mg/ml stock of Concanvalin-A conjugated with Alexa-488 to stain the cytoplasm of cells. Staining was done on ice for 30 min, and the cells were centrifuged again at 1000 x g for 5 min at 4 °C. The pellet was washed again using wash buffer and loaded on the flow cytometer. Cells were sorted in 4-way purity mode into tubes with wash buffer. Depending on the later application, the cells were spun down again and frozen in Monarch RNA Protect buffer for RNA isolation, or used for a scRNA-seq experiment.

2.2.3.3 Fixation of cells

In order to optimize a protocol for the fixation and later transport, different protocols for fixation were evaluated. Cells were fixed using a fixation protocol utilizing acetic methanol (ACME) outlined for *S. mediterranea* [158] and a classical fixation using formaldehyde. The cells were prepared from adult worms, as described previously. Next, 300 - 600 µl of the solution was mixed with 6.5 ml 1x PBS with 1% BSA, 1 ml glycerol, and 1 ml of 100% acetic acid. This solution was placed on a shaker at room temperature. After 20 min, 1.5 ml of methanol was added and placed under slight agitation for 20 min. The solution was centrifuged for 5 minutes at 4 °C at 1000 x g to remove the ACME. The cells were washed afterward with 7 ml 1x PBS with 1% BSA, and after another washing step, the cells were resuspended in 1 ml of 1x PBS with 1% BSA in a 1.5 ml reaction tube. After adding 100 µl DMSO, the cells were stored at -80 °C until further use. For fixation with formaldehyde, cells were resuspended in 8 ml of buffer (PBS + 1 % BSA), and 2 ml of formaldehyde were added for a final concentration of 0.1%, 0.5%, 1%, and 5%. The mixture was gently mixed and placed at 4 °C at room temperature for 10 min. Next, the cells were pelleted by centrifugation at 1000 x g for 5 min at 4 °C using 1x PBS with 1% BSA and finally stored in the same buffer at -80 °C.

2.2.4 Determination of total nuclei number

To determine the total nuclei of adult worms, 10 µm thick cryosections were used. Those were prepared as described earlier using frontal or transversal section planes. Staining of nuclei was done with 1 µg/ml Hoechst diluted in 1x PBS for 10 min at room temperature. For

transversal sections, different regions were sampled. Sections were taken from the most anterior part between the oral and ventral sucker, and between the ventral sucker and uterus. The rest of the sections originated from an area posterior to the uterus. The image analysis software Fiji [141] was used to count nuclei. In detail, the background was subtracted first using a sliding paraboloid method. Following this, a threshold was determined to select single nuclei, and the image was binarized. Water-shedding was applied to split those objects to avoid merging close nuclei into one object. Counting was done using the “Analyze Particles” option in Fiji. Samples in each described region contained replicates from four separate sections. To calculate the total number of nuclei, the mean number of each region was calculated and multiplied by a thickness factor derived from the thickness of the respective tissue area divided by the section thickness of 10 µm. The nuclei number for the frontal section was calculated similarly, taking the worm's thickness in this plane into account.

2.2.5 Staining methods

For all the methods presented, microscopy was performed using the Leica DM IL LED for basic microscopy and the TCS SP5 confocal laser scanning microscope for confocal microscopy.

2.2.5.1 Preparation of cryosections for stainings

For cryosectioning, worms were sectioned using a cryostat Epredia CryoStar NX50 (Thermo Fisher Scientific) at a 10 µm thickness. The tissue sections were mounted on Super Frost™ Slides (Thermo Fisher Scientific) and stored at -80 °C until further use. Before sectioning the worms, they were embedded in cryomolds in the appropriate format. To increase the sectioning quality, worms were first flattened within the cryomold by placing them in a small amount of O.C.T. Afterward, a second cryomold was added to the top to apply pressure, and the specimen was frozen on a metal plate precooled on dry ice. After the tissue was frozen, the second mold was removed, and additional O.C.T was added.

2.2.5.2 EdU staining of adult *F. hepatica*

A commercially available Click-iT plus 5-Ethynyl-2′-deoxyuridine (EdU) assay kit was used to stain proliferating cells. The nucleotide homolog EdU was added to the culture media 16 h before fixation in a final concentration of 500 mM. Worms were frozen in blocks using O.C.T as described prior and stored at -80 °C until further use. For staining, the cryosections were thawed at room temperature for 30 min, following a fixation using 4% formaldehyde in 1x PBS for 15 min. The slides were then transferred to a glass cuvette, permeabilized for 10

min in PBSTx and washed twice with 1x PBS. The staining reaction was carried out in the dark within a humid chamber. The reaction buffer was prepared according to the manufacturer and prepared fresh before being used for staining. For each section, a total volume of 200 µl reaction buffer was used, and the reaction was performed for 30 min. Afterward, sections were washed three times with 1x PBS for 10 min, and for the final washing step, Hoechst 33342 was added in a concentration of 1 µg/ml to counterstain nuclei. Finally, the sections were mounted with Rotimount FluorCare, and the coverslip was sealed using nail polish to prevent evaporation. Slides were stored at 4 °C until further use.

2.2.5.3 Immunostaining

Specific primary antibodies used in other organisms were applied to detect certain proteins (Tabel 2.8). The cryosections were thawed at room temperature for 30 min and fixed with 4% formaldehyde in 1x PBS for 15 min on the slide after which the slides were transferred to a cuvette. After washing twice with 1x PBS, the sections were blocked with blocking buffer (5% goat serum in PBS with 0.5% Tween20) for 1 h within a humid chamber at RT. The primary antibody was incubated overnight at 4°C using the appropriate concentration for each antibody (Tabel 2.8). The following day, sections were washed 6 times for 10 min each in 1x PBS with 0.5% Tween20. Afterward, the secondary antibody was added in a dilution of 1:1000 in a blocking buffer for 2 h at room temperature. Again, the sections were washed 6 times for 10 min using 1x PBS with 0.5% Tween20, and Hoechst 33342 in a concentration of 1 µg/ml was added for the final washing step. Finally, the sections were mounted with Rotimount FluorCare, and the coverslip was sealed using nail polish to avoid evaporation. Slides were stored at 4°C until further use.

2.2.5.4 RNA *in situ* hybridization

The protocol for *in situ* hybridization was modified from a protocol described by Castella and Koziol 2020 [159]. The detailed protocol is described below.

2.2.5.4.1 Synthesis of riboprobes

The principle of RNA *in situ* hybridization is the sequence-dependent pairing of RNA probes to a specific target sequence [160]. This can be achieved by labeling the probes with haptens like digoxigenin or fluorescein, which can be detected using immunolocalization. To generate the labeled RNA probes, sequences matching selected genes were cloned into the plasmid pJC53.2. Before synthesizing the probe, the sequence was amplified using T7_extended primers using the NEB Q5 polymerase PCR reaction outlined in chapter 2.2.9. The *in vitro* transcription reaction was prepared according to Table 2.16. The reaction was kept at 37 °C for at least 4 h or overnight. Residual DNA was removed using a DNase I (2 U)

for 20 min at 37°C. Finally, the RNA was precipitated by adding an equal amount of 7.5 M LiCl to the reaction. The reaction was placed at -80 °C for at least 30 min to enhance the precipitation. The samples were centrifuged at 16 000 x g for 45 min at 4 °C, washed with 70% cold ethanol, and finally, the RNA pellet was resuspended in 20 µl DEPC-water. The concentration of RNA (riboprobes) was determined using a photometer.
Additionally, the riboprobes were checked by gel electrophoresis to validate their length and status of degradation. Between 100 - 200 ng of the riboprobe were diluted in RNA loading dye and heated to 80 °C for 3 min to remove any secondary structure. After heating the probes, they were immediately loaded onto a 1.5% agarose gel.

Table 2.17 Reaction mix for *in vitro* transcription of riboprobes

Component	**Volume**
Template DNA	5 µl
T3/SP6 RNA-Polymerase	1 µl
10x Transcription buffer	2 µl
DIG-rNTP Mix	2 µl
RNase inhibitor	1 µl
DEPC-H_2O	9 µl

2.2.5.4.2 Hybridization

The protocol for *in situ* hybridization was slightly adjusted based on the final detection method, either colorimetric (CISH) or fluorescent (FISH). Changes were noted as necessary.
As described, the sections were thawed for 30 min at room temperature before fixation with 4% formaldehyde in 1x PBS for 15 min. Following the fixation, the sections were transferred into glass cuvettes and permeabilized twice in PBSTx for 10 min each. Following permeabilization, residual peroxidase activity was quenched for fluorescent detection of transcripts using 0.03 % H_2O_2 in 4x SSC for 30 min. After washing twice for 5 min in 4x SSC, the slides were transferred into a humid chamber and prehybridized in a prewarmed hybridization buffer at 55 °C. After 10 min, the hybridization buffer was refreshed and kept in the section for an additional 90 min. Finally, the probes were diluted in prewarmed hybridization buffer to a final volume of 1 ng/µl and heated at 80 °C for 3 min before being given on the tissue section. The sections were covered with parafilm for overnight hybridization at 55 °C to avoid evaporation. The next day, the parafilm was removed by washing the slides in prewarmed 2x SSC, a procedure by which the parafilm floats off the slide. Afterward, the sections were washed in a series of buffers prewarmed to 55 °C in the

following order: twice for 20 min each with hybridization wash buffer, twice for 10 min with 2x SSC, twice for 10 min with 0.2x SSC, and finally twice for 10 min with 0.1x SSC. After washing, the sections were blocked with an *in situ* blocking buffer for 1 h.

The antibodies were added in appropriate dilutions: 1:1000 for Anti-DIG-AP (CISH), 1:200 for Anti-DIG-POD or Anti-FITC-POD (FISH), and placed at room temperature for 2 h within a humid chamber. Subsequently, the sections were washed 3 times with MAB-T buffer for 10 min each. Before development of sections following the CISH protocol, slides were incubated with 5 mM levamisol in AP buffer for 15 min before development in order to reduce endogenous alkine phosphatase activity. Development for CISH was started by covering the sections with developmental solution (3.3 µl/ml 5-bromo-4-chloro-3-indolyl phosphate, 3.3 µl/ml nitro blue tetrazolium chloride in AP buffer) and carried out under microscopic observation until a satisfactory signal was reached. The TSA Plus Kit was used to develop FISH samples after washing with MAB-T. The substrate was diluted 1:100 in Amplification Diluent, and this development solution was put on the slide for 10-15 min in the dark. The development was stopped by washing the slides in 1x PBS 3 times for 10 min each. Hoechst 33342 was added in a 1 µg/ml concentration for atleast 10 min at RT to counterstain the nuclei for the final wash. Additional stainings like EdU or immunostaining were added following FISH development starting from the post-permeabilization steps. The sections were mounted either with FluorCare for FISH samples or with 80% glycerol for CISH samples.

2.2.5.5 Staining of lipid droplets in single-cells

Lipids within the cells were stained using the LipidSpot dye (Biotium). As methanol fixation washed away lipids, only living cells were stained using this procedure. Cells were stained using a 1:1000 dilution of the stock solution in 1x PBS with 2% BSA for 30 min at room temperature. Hoechst 33342 in a 1 µg/ml concentration was added to the staining solution to counterstain nuclei.

2.2.5.6 Detection of proliferation in single-cells

A modified version of the EdU assay was used, which had already been described to label proliferating cells after dissociation [85]. Worms were kept in a culture with 500 mM EdU (EdU stock dissolved in dH_2O) as described before and were later dissociated so that stained cells could be spun onto a gelatine-coated coverslip. To this end, cells were suspended in 1 ml of cold 1x PBS, and the gelatine-coated coverslip was placed into a well of 24 well plates with the cell suspension. The cells were spun onto the coverslip by centrifugation at 200 x g for 5 min in a swing-bucked rotor without brakes. The cells were fixed afterwards with 4% formaldehyde in 1x PBS at RT for 4h. Cells were centrifuged again

for 2 min at 200 x g. After removing the fixation solution, 100% methanol was added, and the cells were stored at -20 °C until further use. Staining was carried out following the manufacturer's instructions.

2.2.5.7 Staining of dead cells

Zombie Red was used to stain dead cells. This amine fluorescent dye cannot pass through the cell membrane of viable cells and can therefore only stains dead cells. Staining was achieved by incubating cells in a staining solution containing a 1:1000 dye dilution in 1x HBSS in 2% BSA. Staining was carried out over 30 min at room temperature and washed using 1x HBSS in 2% BSA. Afterward, the cells were fixed using formaldehyde

2.2.6 Work with bacteria

All work concerning bacteria was performed using the *Escherichia coli* strain NEB 10ß

2.2.5.1 Cultivation of *E. coli* and preparation of glycerol stocks

Bacteria were grown at 37°C on agar plates or in liquid LB media with either kanamycin or ampicillin. For culturing bacteria on plates, 50 - 100 µl of liquid culture was plated evenly on the agar plate and grown in an incubator at 37°C overnight. Liquid cultures were inoculated with a single colony and incubated at 37°C within a thermoshaker at 170 rpm overnight.
Bacteria were stored in 20% glycerol. For this, 400 µl of an overnight culture was mixed with 100 µl glycerol, vortexed briefly, and frozen in liquid nitrogen. Finally, the frozen bacteria were stored at -80°C.

2.2.6.2 Generation of competent cells

An aliquot of *E. coli* was thawed on ice and inoculated in LB media without antibiotics overnight at 37°C and 170 rpm to generate competent cells. Of this overnight culture, 2 ml was inoculated in fresh 200 ml LB media and grown to an OD_{600} of 0.5. The culture was then cooled on ice for 5 min and finally spun down in a centrifuge bottle by centrifugation at 4.000 x g at 4°C for 5 min. The pellet was resuspended in cold transformation buffer I and kept on ice for 90 min. The bacteria were centrifuged again and resuspended in 8 ml transformation buffer II. Aliquots of 100 µl cells were prepared in 1.5 ml reaction tubes and frozen in liquid nitrogen. For long-term storage, the bacteria were kept at -80°C.

2.2.7 Cloning

Gene fragments were cloned into the cloning vector pJC3.2 for dsRNA and riboprobe synthesis [119]. This vector harbors two T7 promotor sequences as well as T3 and SP6 promotors on the opposing strands, all flanking the multiple cloning site

2.2.7.1 Probe selection for riboprobes and RNAi

For both methods, RNA *in situ* hybridization and RNAi sequences were selected based on certain criteria. The fragments were selected from coding sequence data of the *F. hepatica* genome [92] and downloaded from WormBase ParaSite. Probes were designed to have a size between 300 - 500 bp in length and were finally blasted against the genome to ensure specificity. As outlined before, the dsRNA sequence was also checked for potential off-target activity with the tool siRNAi-Finder (siFi) [161,162]. For the amplification of gene fragments, primers were designed as outlined before.

2.2.7.2 Digestion of pJC53.2

All fragments in this work were cloned into digested pJC53.2 vector using the same basic cloning scheme. To this end, the plasmid was digested using the restriction enzyme *AdhI* following the manufacturer's recommendation. The reaction was carried out for 1 h at 37°C. The digestion of the plasmid was validated using gel electrophoresis, and the digestion product was purified as outlined previously.

2.2.7.3 Amplification and modification of gene fragments

The gene fragments were amplified using a basic Q5 PCR, as described previously. The specificity of the PCR was validated by gel electrophoresis, and the fragment was purified from a gel fragment, as described. As the digested plasmid contains thymine overhangs, the Accuprime DNA polymerase was used in an A-tailing reaction to add adenines to the blunt end of the fragment. The product of this reaction was purified from the PCR reaction mix as described before, and the DNA concentration was determined photometrically.

2.2.7.4 Ligation

A T4 DNA ligase from NEB was used to ligate the purified DNA fragments into the cloning vector. The DNA fragment was mixed in a molar ratio of 7:1 with 50 ng of the digested cloning vector, as calculated using the NEB Ligation Calculator (https://nebiocalculator.neb.com/#!/ligation). Finally, the reaction was prepared and incubated overnight at 4°C.

2.2.7.5 Transformation of bacteria

To introduce endogenous DNA in *E. coli*, the chemically competent cells were thawed on ice, and DNA in the form of a 5 µl ligation mixture was mixed with the cells by flicking. The mixture was kept on ice for 30 minutes and placed at 42°C for a heat shock of 90 seconds. After the cells were cooled down for 3 min on ice, 600 µl of LB media was added, and the cells were incubated at 37°C on a thermoshaker with slight agitation (300 rpm). Following

incubation, the bacteria were centrifuged for 5 min at 5000 x g, and the supernatant was discarded. The pellet was resuspended in the residual media, and 50 µl was plated on an LB agar plate with either kanamycin or ampicillin and the plate incubated overnight at 37°C.
Colony PCR determined the success of the transformation, and one to two colonies of a cloning reaction were inoculated for an overnight culture in LB media containing either kanamycin or ampicillin. This culture was grown at 37°C at 170 rpm.

2.2.7.6 Sequencing of plasmids

Plasmids were isolated from overnight culture, as described before. For sequencing, plasmids were diluted to a concentration between 80 - 120 ng/µl and sequenced using a premixed sequencing (table 2.11) primer by the Microsynth AG. Sequencing was validated by blasting the results against the *F. hepatica* genome. For a detailed view, the sequence was validated by aligning the sequence of the gene of interest against the sequencing results. Bacteria transformed with the validated plasmid were stored as glycerol stocks at -80°C.

2.2.8 Work with nucleic acids

2.2.8.1 Isolation of RNA from whole worms

The commercially available Monarch Total RNA Miniprep Kit was used to isolate RNA. Single worms were individually frozen in Monarch RNA Protect Buffer before use. The manufacturer's recommendations for up to 10 mg of homogenized tissue were followed for the procedure. To this end, the worms were homogenized using a pistil. The optional on-column DNA digestion was performed. The final RNA was eluted with 30 - 50 µł nuclease-free water.

2.2.8.2 Isolation of RNA from single cells

RNA from cells was isolated using the Monarch Total RNA Miniprep Kit. Atleast 200,000 cells were frozen in Monarch RNA Protect Buffer and stored at -80°C until use. The cells were lysed by 5 freeze-thaw cycles using liquid nitrogen. Afterward, the manufacturer's recommendations for cells were followed. The final RNA was eluted with 30 - 50 µl nuclease-free water.

2.2.8.3 Analysis of RNA integrity

According to the manufacturer's instructions, the integrity and concentration of the isolated RNA were checked by electropherogram analysis using the Agilent RNA 6000 Nano or Pico kit (Agilent 2100 Bioanalyzer; Agilent Technologies, USA). A good RNA quality is usually expressed by a RIN value computed by electrophoretic features of the 28S and 18S

ribosomal RNAs. For *F. hepatica,* the 28S rRNA is known [163] to be composed of two fragments that separate easily, thereby leading to a failure to calculate a valid RIN value. Despite this limitation, the control of RNA integrity by the presence of an 18S ribosomal RNA was deemed a valid readout for high-quality RNA.

2.2.8.4 Isolation of plasmid DNA from *E. coli*

The Monarch Plasmid Miniprep Kit isolated plasmid DNA from bacterial cells. To this end, 2 ml of an overnight culture was pelleted by centrifugation at 5000 x g for 5 min. The plasmid isolation was carried out following the manufacturer protocol, and plasmids were eluted using 30 µl nuclease-free water.

2.2.8.5 Isolation of DNA from agarose gel or after PCR

Cleanup of DNA from either agarose gels after gel electrophoresis or from PCR reactions was performed using the Monarch Gel Extraction Kit or the Monarch PCR and DNA Extraction Kit, respectively. The DNA was eluted using 30 µl of nuclease-free water.

2.2.8.6 cDNA synthesis from whole worm RNA

For the synthesis of cDNA from total mRNA samples, the Quantitect Reverse transcription Kit from Qiagen was used. For synthesis, 100-1000 ng of RNA was used as a template. The reaction was carried out following the manufacturer's protocol.

2.2.8.7 cDNA synthesis from single-cell RNA

Due to the lower RNA yield from single cells and a need for better sensitivity, cDNA was synthesized by using the SuperScript™ II Reverse Transcriptase. The reaction mix was prepared following the manufacturer's instructions using Oligo$(dT)_{16}$ primer to capture mRNA transcripts.

2.2.8.8 Synthesis of dsRNA

For synthesis of dsRNA a *in vitro* transcription was performed using fragments containing a T7 promoter sequence. The fragments were synthesized using PCR with a pJC53 plasmid using a Q5 PCR. A dilution to 1 ng/µl of the plasmid was used as the template. For the *in vitro* transcription, the reaction was prepared using homemade T7 RNA polymerase, as outlined below.

Table 2.18 Reaction volume used for *in vitro* transcription

Component	Volume
PCR product	5 µl
10x transcription buffer	10 µl
rNTP mix	20 µl
T7 RNA polymerase	5 µl
Inorganic pyrophosphatase	1 µl
DEPC-H_2O	59 µl
Sum	100 µl

The reaction was incubated at 37°C overnight following removal of the DNA template using 2 U/µl DNAse I for 20 min at 37°C. The dsRNA was precipitated using an equal amount of 7 M LiCl at -80°C for at least 30 min. Afterwards, the dsRNA samples were centrifuged for 45 min at 4°C at maximum speed. The resulting pellet was washed using cold 70% ethanol and centrifuged again for 2 min at 4°C using full speed. The pellet was then air-dried and finally resuspended in DEPC-H_2O. The dsRNA was checked on an agarose gel and stored at -20°C.

2.2.9 Polymerase chain reaction (PCR)

PCR has been used in this work for various applications, including:

a) Amplification of cell marker genes for validation in RT-PCR
b) Amplification of DNA templates for the synthesis of dsRNA
c) Amplification of DNA templates for the synthesis of RNA probes for WISH
d) Amplification of gene fragments from transformed *E. coli*

For applications that did not demand a proofreading activity, the FirePol DNA Polymerase was used (a and e), while for applications more sensitive to mutations introduced by PCR, the Q5 High Fidelity DNA polymerase was used. Reaction volumes and cycle conditions can be found Table 2.19 - Table 2.22

Table 2.19 PCR reaction mix for PCR with FirePol DNA Polymerase

Component (stock conc.)	Volume
10x reaction buffer BD	2.5 µl
$MgCl_2$ (25 mM)	2.5 µl
dNTP mix (10 mM)	0.5 µl
Primer forward (10 µM)	0.75 µl
Primer reverse (10 µM)	0.75 µl
Firepol DNA polymerase	0.25 µl
Template DNA	2 µl*
PCR-H_2O	15.75 µl*
Sum	25 µl

*Volumes were adjusted based on template concentration

Table 2.20 PCR cycling steps PCR with FirePol DNA Polymerase

Temperature	Time	
95°C	3 min	
95°C	30 sec	35x
60°C	30 sec	
72°C	30 sec	
72°C	5 min	

Table 2.21 PCR reaction mix for PCR with Q5 DNA Polymerase

Component (stock conc.)	Volume
5x Q5 buffer	5 µl
dNTP mix (10 mM)	0.5 µl
Primer forward (10 µM)	1.25 µl
Primer reverse (10 µM)	1.25 µl
Q5 High-Fidelity DNA polymerase	0.25 µl
Template DNA	1 µl
PCR-H_2O	15.75 µl
Sum	25 µl

*As adjustment for generating PCR products with T7 promotor
3 µl T7_extended primer was used for each forward and reverse and
4 µl plasmid (1 ng/µl) used as template

Table 2.22 PCR cycling steps PCR with Q5 DNA Polymerase

Temperature	Time	
98°C	3 min	
98°C	30 sec	35x
60°C*	30 sec	
72°C	30 sec	
72°C	5 min	

*As adjustment for generating PCR products with T7 promotors 57.4°C was used

2.2.9.1 Colony PCR

A colony PCR was performed to validate the successful transformation of *E. coli* and check if the transformed plasmid contained the valid insert. To this end, single colonies were picked and resuspended in 50 µl LB media in 1.5 ml reaction tubes. A standard FirePol PCR mix was prepared with T7_extended primers, and 3 µl of the suspensions were added to the PCR.

2.2.9.2 RT-qPCR

A quantitative PCR was performed with cDNA generated using a reverse transcription (RT-qPCR) to quantify the amount of RNA in a given sample. Before PCR, cDNA was synthesized from RNA. For RT-qPCR, the SYBR Green-based KAPA SYBR Fast mix from Sigma was used as described in Table 2.23, with cycling conditions listed in Table 2.24.

Table 2.23 PCR reaction mix for RT-qPCR

Component (stock conc.)	Volume
2x KAPA SYBR Fast Mix	10 µl
Primer forward (10 µM)	0.8 µl
Primer reverse (10 µM)	0.8 µl
Template DNA	5 µl
PCR-H_2O	4.2 µl
Sum	20 µl

Table 2.24 RT-qPCR cycling conditions

Temperature	Time	
95°C	3 min	
95°C	10 sec	45x
60°C	15 sec	
72°C	20 sec	
60 - 95°C	20 sec	

2.2.9.2.1 Primer design and validation for RT-qPCR

As RT-qPCR applications based on the dye SYBR Green are sensitive to the formation of all types of double-stranded DNA, special care was taken to design and validate primers used for those applications. The coding sequence for a gene of interest was downloaded from WormBase ParaSite [129]. When applicable, the primers were desinged spanning over exon-exon junctions in order to distinguish amplification products derived from genomic DNA instead of cDNA. Using the online tool OligoAnalyzer, the primers were checked for the potential formations of hairpin structures and the stability of potential homo/heterodimers.

Primers were designed to have potential hairpins and dimers with a low stability, i.e. delta G values greater than -9 kcal/mol.
Additionally, the primers were blasted against the *F. hepatica* genome [92] to ensure specificity. The melting temperature was adjusted to 60°C within a range of 0.5°C with OligoCalc (salt-adjusted value) [123], and the amplicon length was set to 140-200 bp. The primers were ordered from Integrated DNA Technologies (IDT) and used in a final concentration of 400 nM.

To validate the suitability of primers *in vitro,* these primers were used in a FirePol PCR and checked for one specific product, as well as the absence of dimers via gel electrophoresis. The PCR product was purified from the gel, as outlined before. In order to determine the primer efficiency, the amplicon concentration was first determined by photometry. Afterwards, the molecular weight of the amplicon was determined, as shown in equation 2.1. From the determined molecular weight, the copy number per µl was calculated, and the product was used in a serial dilution to receive 10^5, 10^4, 10^3, 10^2, and 10^1 amplicons which were finally utilized as a template for qRT-PCR in a Qiagen Rotor-Gene device using the reaction mix and setup as outlined in table 2.23 and 2.24. The primer efficiency was determined by Qiagen Q-Rex Gene expression Plugin v2.0.0.4 and deemed suitable when reaching 0.9 - 1.1.

$$MW_{Ia} = \frac{Length_{Ia}\ x\ G/C\ content\ [\%]\ x\ MW_{G/C}}{100} + \frac{Length_{Ia}\ x\ A/T\ content\ [\%]\ x\ MW_{A/T}}{100} + 2\,MW_{H_2O} - 2\,MW_{HPO_3}$$

Equation 2.1 Determination of molecular weight

2.2.9.2.2 Determination of gene expression

The same PCR setup as for primer efficiencies was used to quantify transcript abundance for any gene of interest. The term “gene expression“ is used throughout this work and refers to transcript abundance. The reaction was carried out in 20 µl reaction volume, and at the end of each run, the presence of a specific product was checked by melting point analysis. Water control was included for each primer mix to control for potential contamination within the reaction mix. The relative expression of a gene of interest was calculated by the ΔΔCt Method [164]. The relative expression was calculated by normalization to the geometric mean of Ct values obtained for two selected reference genes, *fheprs* and *fhtbcd* [121]. To determine the fold change of expression between two samples, the formula $2^{-\Delta\Delta Ct}$ was used, which is described in detail below. In brief, the difference between the gene of interest and the housekeeping gene is calculated as ΔCt for the sample and controls. In the next step,

the ΔΔCt is calculated as the difference of ΔCt_{sample} and $\Delta Ct_{control}$. Finally, the fold change is calculated as 2 to the power of negative ΔΔCt.

$$\Delta Ct = Ct(gene\ of\ interest) - Ct(housekeeping\ gene)$$

$$\Delta\Delta Ct = \Delta Ct_{sample} - \Delta Ct_{control}$$

$$fold\ change = 2^{-\Delta\Delta Ct}$$

Equation 2.2 Determination of gene expression fold change using the ΔΔCt method

2.2.10 Analysis of scRNA-seq data

2.2.10.1 Library preparation and sequencing

For scRNAseq, the commercially available Chromium Next GEM Single Cell V3.1 product from 10x Genomics was utilized. All steps from cell dissociation until library preparation were done in the Christos Samakovlis lab with help by Janine Koepke. Cells were dissociated as described earlier, sorted to include only viable cells in the analysis, and centrifuged at 400 g and 4°C for 5 min. The cells were then resuspended in 0.2% BSA in PBS and loaded targeting a final cell count of 62,000 cells. Barcoding and library preparation were done following the manufacturer's protocol. The resulting libraries were sequenced on a NextSeq 2000 or NovaSeq 6000 system to a depth of 50,000 reads per cell.

2.2.10.2 Analysis of scRNA-seq data

As to analyzing the scRNAseq data, the sequenced reads were mapped onto the *Fasciola* genome obtained from WormBase Parasite release 17 (PRJNA179522) using Cell Ranger (version 7.0.0). The reference genome was modified to include the mitochondrial genome of *F. hepatica* [89] (available under GeneBank accession number: AF216697). The cellranger outputs were further analyzed using the software package Seurat version 4.1.1 (https://satijalab.org/seurat/). Each sample was filtered to only contain cells with more than 1000 UMI and less than 50.000 unique molecular identifiers (UMI), more than 750 Features, and less than 5 percent mitochondrial reads. The data was then normalized using log normalization option in the NormalizeData function. A total of 1500 variable features was determined using the FindVariableFeatures function with the “vst” method. Next, samples from separate groups (anterior, posterior, or whole worm) were integrated using

FindIntegrationAnchors and IntegrateData functions. After integrating the samples groupwise, a final dataset was merged using the same methodology. Finally, the ScaleData function, RunPCA, and RunUMAP(reduction = "pca," dims = 1:25, n.neighbors = 30, seed.use = 42) were used to further process the data. The number of principal components was determined by running an Elbowplot and using the JackStraw function implemented in Seurat. The cells were clustered based on their transcriptomics identity by running FindNeighbors(reduction = "pca," dims = 1:25) and FindClusters (resolution = 0.4). A suitable resolution was determined using the clustree package [165], identifying the breakdown of stable clustering. Enriched marker genes were predicted with FindAllMarkers(cells_all, test.use = "roc", only.pos = TRUE, logfc.threshold = 0.0, min.pct = 0.0). Following this, the clusters were annotated using known markers genes from literature of related species. Using our approach, the following clusters did not have biologically significant markers and were merged: testes-related (clusters 5, 4) and late male germ cells (clusters 2,7,6,0). The known gut cluster could not be resolved as a cluster, so it was manually defined using the expression of conserved markers for this tissue.

2.2.10.3 RNA velocity

The RNA velocity is an approach for scRNA-seq data that allows the prediction of directed information about the differentiation state between different cell types. The basis for this approach is the assumption that splicing kinetics of certain genes can be used as predictors of differentiation state [135]. To obtain this kind of analysis, the sequencing data was mapped to the reference genome with cellranger version 7.0, which includes intronic reads by default. The resulting bam files were sorted based on their genomic location using the samtools sort command. Afterward, a loom file for each sample was constructed, containing count matrices with spliced and unspliced counts using velocyto [134]. Barcodes of cells applicable for the analysis were extracted using Seurat [166] and used to filter the loom files. RNA velocities were obtained by running scVelo [135] using the dynamical model. Genes that did appear to have expression patterns that were shown to lead to faulty velocity estimates were removed following the workflow suggested in Barile et al [167].

2.2.10.4 Cell-cycle scoring

For scoring the cell-cycle state within the scRNA-seq data, the CellCycleScoring function implemented in Seurat was used. This marker gene-based method takes the average expression of the given marker genes for phase S and G2M and compares them to the control expression. Marker genes were identified by identifying orthologs to the default gene set in Seurat, and can be found in supplementary table 1 [168].

2.2.10.5 Label transfer

The single-cell data was mapped into the spatial transcriptome [169] using an anchoring approach in Seurat [170]. This method identifies “anchors” between the two datasets to pair cells or spots together, this information is used to transfer the cell labels into the spatial reference. To this end the scRNA-seq data was normalised using SCTransform, and anchors were identified using the singel-cell dataset as a reference and spatial data as a query. Prediction scores were calculated based on the identified anchors using the TransferData function using the first nine principle components to further weigh the importance of anchors.

2.2.10.6 GO Term analysis

In order to predict potential functions associated with different clusters, Gene Ontology (GO) enrichment analysis was performed with the topGO package [152]. A list of all genes and their associated terms was obtained from WormBase ParaSite Version 17. This list was modified by adding terms obtained by running InterProScan v 5.60-92-0 [126]. Analysis was performed using the weight01 method for two categories (BP, MF). The node size for each category was restricted to >= 4, and significance was determined using Fisher’s exact test against all expressed genes.

2.2.10.7 STRINGdb analysis

For predicting protein-protein interaction networks enriched within a selected set of genes, STRING online tool (v12) [171] was utilized. The *F. hepatica* proteome (PRJNA179522) was uploaded to the database and used as an input dataset.

2.2.11 Sequence analysis

2.2.11.1 BLAST

A local blast was used to determine cloning success and perform extensive sequence similarity searches (BLAST+ version 2.13.0). Sequences were blasted against a database built from transcript sequences obtained from Wormbase Parasite release 18 [129]. Results were obtained using the default output format 6.

2.2.11.2 Transcription factor prediction

Potential transcription factors encoded within the *F. hepatica* genome were predicted largely following a recently outlined workflow for planarians [172]. As such, Pfam Hidden Markov models were downloaded from InterPro https://www.ebi.ac.uk/interpro/download/Pfam/,

accessed 02.2024) and searched for Pfam profiles associated with DNA binding domain (DBD) using hmmscan command from HMMER v3 (https://www.ebi.ac.uk/Tools/hmmer). The protein sequences of *F. hepatica* obtained from WombaseParasite [129] were searched for DBD domains using hmmsearch. All sequences with a hit in this prediction were selected and blasted against known human transcription factors (obtained from AnimalTFDB 4.0 [127] or planarian transcription factors [172]. Only sequences with a match of (evalue < 1e-5) were retained. Finally, transcription factor families for each sequence were predicted using AnimalTFDB 4.0.

2.2.11.3 Phylogenetic analysis

For phylogenetic analysis of *F. hepatica* p21-activated kinase (PAK) family members, those were identified based on the presence of a p21-binding domain (PDB) and kinase domains, which was validated by SMART v 9.0 of selected amino acid sequences [125]. Other sequences from selected organisms of PAK family members were downloaded from varying databases.

All the sequences were first aligned using MAFFT (version 7.505) with the L-INS-i setting [143]. The resulting alignment was refined using MUSCLE (version 5.1) [142]. Bayesian tree construction was done with MrBayes version 3.2.7a [144]. Posterior probabilities were calculated as recommended, with 1,000,000 trees sampling every 100th tree. The first 25% was discarded as burn-in, and a majority rule tree was constructed. The finalized tree was drawn using the Interactive Tree of Life [122].

2.2.11.4 Remapping of bulk RNAseq data

The available RNAseq data of the liver fluke lifestages [91] available under accession number PRJEB6904 was mapped to the used genome using salmon (v 1.3.0 [173]). To this end the transcript sequences were downloaded from WormBase ParaSite [129] and an index was built using the default options.

2.2.12 Statistics

Data received from RT-qPCR analysis were statistically analyzed by comparative testing of the normalized transcription levels as previously described in Moescheid and Puckelwaldt *et al.* [161]. Statistical analyses were carried out using the programming language R in RStudio [145,146]. For statistical analysis, the two-tailed t test was applied for parametric distributed data. P-values < 0.05 were considered as statistically significant.

3.Results

This work is structured in three main parts. The first section concerns establishing and validating a protocol to dissociate *F. hepatica* into vital single cells. The second part will focus on analyzing, validating, and improving a scRNA-seq atlas. Finally, the third section will illustrate examples of leveraging this dataset regarding drug target discovery.

3.1 Establishment of a high-quality single-cell dissociation for adult *F. hepatica*

The advent of new high throughput methods like scRNA-seq enables unprecedented insights into previously neglected organisms. A significant hurdle in utilizing these methods is the complex organization of a whole organism, which makes generating a single cell suspension difficult. The quality of the cell dissociation and the cell viability are essential for every aspect of the following analysis. Before this project, no suitable protocol existed for dissociating adult *F. hepatica* into single cells. This work aimed to create a fast and efficient dissociation method without impacting cells more than necessary.

3.1.1 Determination of optimal conditions for cell dissociation

Tissue dissociation describes the release of cells from their surrounding tissue. A sufficient quality dissociation produces a suspension of single cells with a low amount of cell clumps, and a high percentage of viable cells. The recommended cell viability for scRNA-seq is over 70% [174]. The success of optimal dissociation relies on many conditions, including parameters like dissociation buffer, temperature, and dissociation time [175]. Most dissociations utilize proteases to liberate cells. The choice of enzyme can significantly impact how efficient a dissociation is and what cells can be efficiently isolated. This process was not without its challenges. Multiple dissociation protocols were evaluated based on their ability to generate single cells (outlined in Table 3.1). These dissociation protocols utilized one of three proteases after a mechanical tissue disruption. The three enzymes trypsin, collagenase I, and collagenase II were selected as they have been frequently used in cell dissociation protocols [175]. Collagenases are known to be effective in the digestion of extracellular matrix (ECM) due to their breakdown of collagen. At the same time, trypsins break down proteins in the cell-cell junctions between cells [176]. DNAse I was added to most dissociation reactions to digest DNA, which can cause excessive clumping. Under the chosen buffer conditions, dissociation with collagenase II was unsuitable for *F. hepatica*, as the dissociation resulted either in cell death or severe cell aggregation.

Adding other ECM degrading enzymes like hyaluronidase or dispase II did not improve the outcome. While collagenase II was not successful in producing good single cell suspension reactions, using collagenase I was more promising, resulting in a higher yield. However, attempts to increase the reaction's efficacy by longer incubation times or adding other proteases failed. Of note was the complete absence of dissociation using ACME dissociation on the worms, which was highly successful for planarians [158]. Dissociation using trypsin showed a significant improvement to the other dissociation protocols, having good cell viability and yield, though the initial results also showed cells still clumping within the suspension. Based on these promising results using trypsin, the dissociation was further optimized by adding EDTA, which removes calcium ions essential for maintaining cell adhesions. Besides this, a significant improvement was achieved by eliminating magnesium and calcium from dissociation to further prevent cell clumping. The final protocol used 5% trypsin in HBSS without calcium and magnesium ions. To avoid clumping, DNAse I and EDTA were added to the solution. The best results were achieved by dissociating for around 30 min. The viability of the cells directly after dissociation was ~ 80% and around 200,000 to 1 million cells could be dissociated from one adult worm.

Table 3.1 List of conditions tested for cell dissociation of *F. hepatica* adults. The best dissociation condition is marked by a (*)

Enzyme	Buffer	Duration	Temperature	Result
Collagenase II				
Collagenase II (250 U/ml) DNAse I (400 U/ml)	1x HBSS 5 mM $MgCl_2$	15 min	37°C	Good yield Medium cell death Moderate Clumping
Collagenase II (250 U/ml) DNAse I (400 U/ml)	1x HBSS 5 mM $MgCl_2$	30 min	37°C	Good yield Medium cell death Moderate Clumping
Collagenase II (400 U/ml) Dispase II (2 U/ml)	1x HBSS	30 min	37°C	Low yield High cell death
Collagenase II (200 U/ml) Dispase II (1 U/ml) DNAse I (100 µg/ml)	1x HBSS 5 mM $MgCl_2$	20 min	37°C	Low yield Low cell death

Enzyme	Buffer	Duration	Temperature	Result
Collagenase II (200 U/ml) Dispase II (1 U/ml) DNAse I (400 U/ml)	1x HBSS 5 mM $MgCl_2$	25 min	37°C	Low yield Low cell death
Collagenase II (200 U/ml) Dispase II (1 U/ml) DNAse I (400 U/ml)	1x HBSS 5 mM $MgCl_2$	30 min	37°C	Medium yield High cell death
Collagenase II (200 U/ml) Dispase II (10 U/ml) Hyaluronidase (400 U/ml) DNAse I (400 U/ml)	1x HBSS	30 min	37°C	Low yield Strong clumping Moderate cell death
Collagenase II (200 U/ml) Dispase II (1 U/ml) Hyaluronidase (400 U/ml) DNAse I (400 U/ml)	1x RPMI	40 min	37°C	Low yield Strong clumping High cell death
Collagenase II (200 U/ml) Dispase II (1 U/ml) Hyaluronidase (400 U/ml)	1x HBSS 5 mM $MgCl_2$ 2 mM $CaCl_2$	30 min	37°C	Medium yield Medium clumping Moderate cell death
Collagenase II (200 U/ml) Dispase II (1 U/ml) Elastase (5 U/ml) DNAse I (100 µg/ml)	1x HBSS	30 min	37°C	Low yield Strong clumping Moderate cell death
Collagenase I				
Collagenase I (200 U(ml) DNAse I (400 U/ml)	1x HBSS	35 min	37°C	Medium yield Low clumping Low cell death
Collagenase I (200 U/ml) Dispase II (10 U/ml) DNAse I (400 U/ml)	1x HBSS	40 min	37°C	Medium yield High clumping High cell death

Enzyme	Buffer	Duration	Temperature	Result
Collagenase I (200 U/ml) Dispase II (10 U/ml) Hyaluronidase (400 U/ml) DNAse I (400 U/ml)	1x HBSS	30min	37°C	Low yield High clumping High cell death
Trypsin				
Trypsin (3.5x)	1x PBS	2 h	37°C	High Yield Moderate Clumping Cell death Long Duration
Trypsin (3.5x) DNAse I (400 U/ml)	1x HBSS 5 mM $MgCl_2$ 2 mM $CaCl_2$	30 min	37°C	Medium yield Moderate clumping Low cell death
Trypsin (3.5x) DNAse I (400 U/ml)	1x HBSS 1 mM EDTA	15 min	37°C	Medium yield Moderate clumping Low cell death
Trypsin (3.5 x) DNAse I (400 U/ml)	1x HBSS 1 mM EDTA w/o Ca^{2+} w/o $MgCl_2$	30 min	37°C	High yield Low clumping Low cell death
Trypsin (5 x) * DNAse I (400 U/ml)	1x HBSS 1 mM EDTA w/o Ca^{2+} w/o $MgCl_2$	30 min	37°C	High yield Low clumping Low cell death
Trypsin (3.5x) Papain (13 U/ml)	1x HBSS 1 mM EDTA	30 min	37°C	Medium Yield Good viability
Trypsin (3.5x) Bromelain (50 U/ml)	1x HBSS 1 mM EDTA	30 min	37°C	High yield Moderate clumping Medium cell death
Trypsin (0.5x) Bromelain (20 U/ml)	1x HBSS 1 % BSA	30 min	37°C	Low yield High clumping Medium cell death
Trypsin (0.5x) Bromelain (20 U/ml)	1x HBSS 5 % BSA	30 min	37°C	Low yield High clumping Medium cell death

Enzyme	Buffer	Duration	Temperature	Result
Trypsin (10x)	1x HBSS	30 min	37°C	Low yield Moderate clumping Low cell death
Misc				
Papain (20 U/ml)	1x HBSS	30 min	37°C	Low yield Low clumping Low cell death
Bromelain (20 U/ml)	1x HBSS	30 min	37°C	Medium yield Moderate clumping Low cell death
Bromelain (50 U/ml)	1x HBSS	30 min	37°C	Low yield High clumping High cell death
ACME (Acetic methanol)	10 % Acetic acid 15 % Methanol 10 % Glycerol in H_2O	1 h	RT	no dissociation

3.1.2 Evaluation of single cell quality

To further evaluate the suitability of the cell suspension for a scRNA-seq experiment, the viability of the cells was further checked by viability staining with two separate dyes. First, viability was assessed using calcein-AM, which functions by being processed by intracellular esterases. The product of this reaction is a green fluorescent fluorophore, which can be detected by microscopy or FACS. Staining the cells validated their high viability (Fig 3.1A), showing almost no dead cells. The cells within the suspension were heterogeneous in size and form, ranging from 5 µm to 35 µm in diameter. This observation was noteworthy as the planned scRNA-seq assay is limited to cells smaller than 50 µm due to size limitations of the Chromium controller [177]. The high viability was further confirmed by staining with the fixable Zombie dye, which cannot penetrate the membrane of living cells. In agreement with the high number of cells staining positive for calcein, only a few cells stained positive for Zombie (Fig 3.1B). Cells with a positive Zombie signal mostly had a high granular content. The next logical step to validate the suitability of sample preparation for an RNA sequencing experiment was to check for RNA integrity, as a high amount of RNA degradation might lead to poor data quality. Therefore, the total RNA was isolated from cells and tested using the Agilent BioAnalyzer system. It was possible to detect the intact band characteristic for the

18S rRNA (Fig 3.1C). Due to a gap region within the 28S rRNA of *F. hepatica [163]*, which leads to a cleavage of this ribosomal RNA, this band can not be detected. As the typical calculation of a RIN value [178] relies on the presence of said 28S rRNA, such a measurement could not be given.

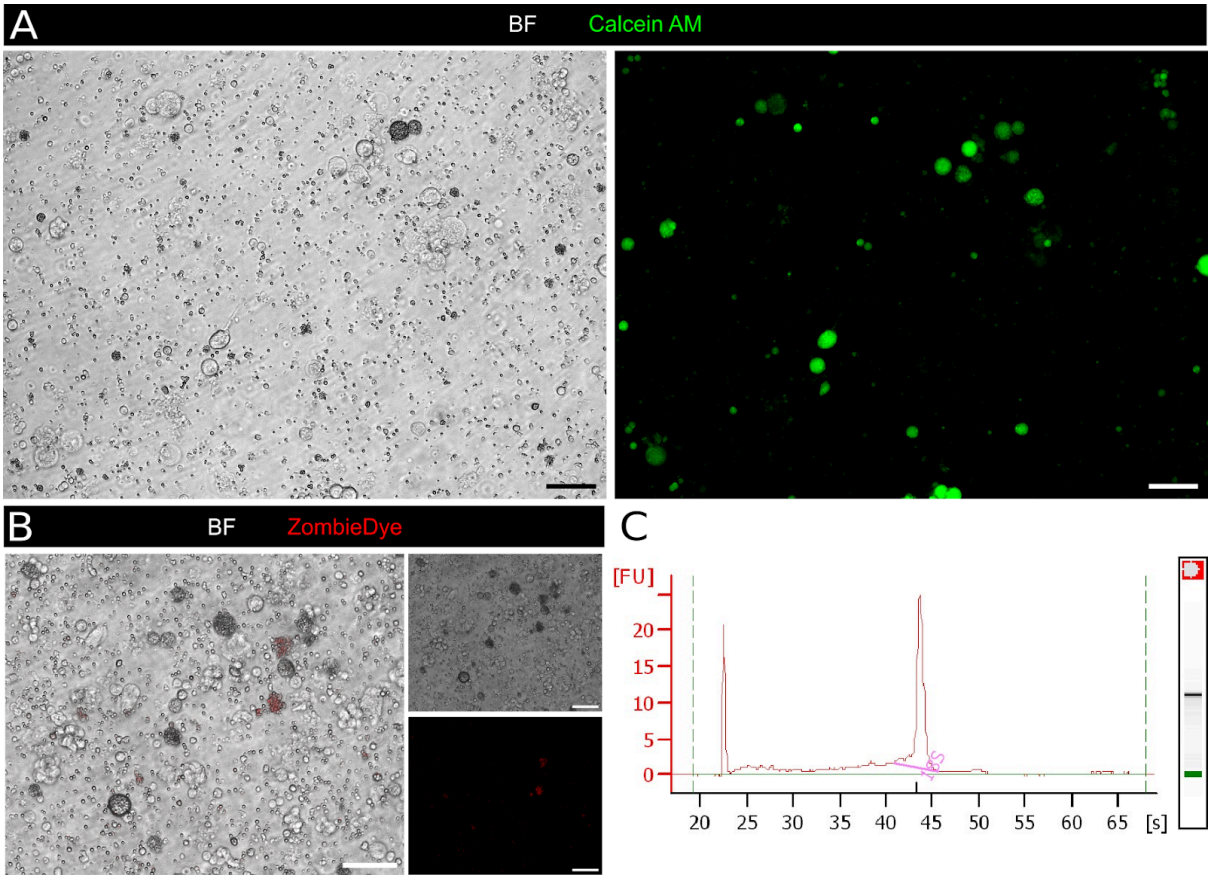


Figure 3.1 Dissociation with trypsin produces a highly viable cell suspension. Representative microscopic images of dissociated liver fluke cells **A.** Cells stained with the viability dye calcein-AM in brightfield (left) and stained in green (right). **B.** Cells stained with the fixable viability dye ZombieRed. **C.** Bioanalyzer electropherogram of RNA from single cells showing high quality. Scale: 50 µm.

One major challenge in working with this cell suspension was the high amount of debris, which might complicate several steps within a scRNA-seq workflow, such as clogging the device used for barcoding. As such, precautions were taken, including FACS for removing debris and enriching viable cells. A protocol shown to be useful for planarian cells [179] was adapted for this work. The cells were stained using calcein-AM in combination with the far-red nucleic dye DRAQ5. Following the strategy outlined (Fig 3.2), the cells were first gated based on their properties in the SSC and FSC channels to remove cellular debris (Fig 3.2 A-C). Following cells were selected based on a positive DRAQ5 signal and finally gated on their calcein-AM positive signal (Fig 3.2 D-E). Overall, cell viability was determined to be above 75%, which met the manufacturer's recommended viability of at least 70% for cells in scRNA-seq experiments [174].

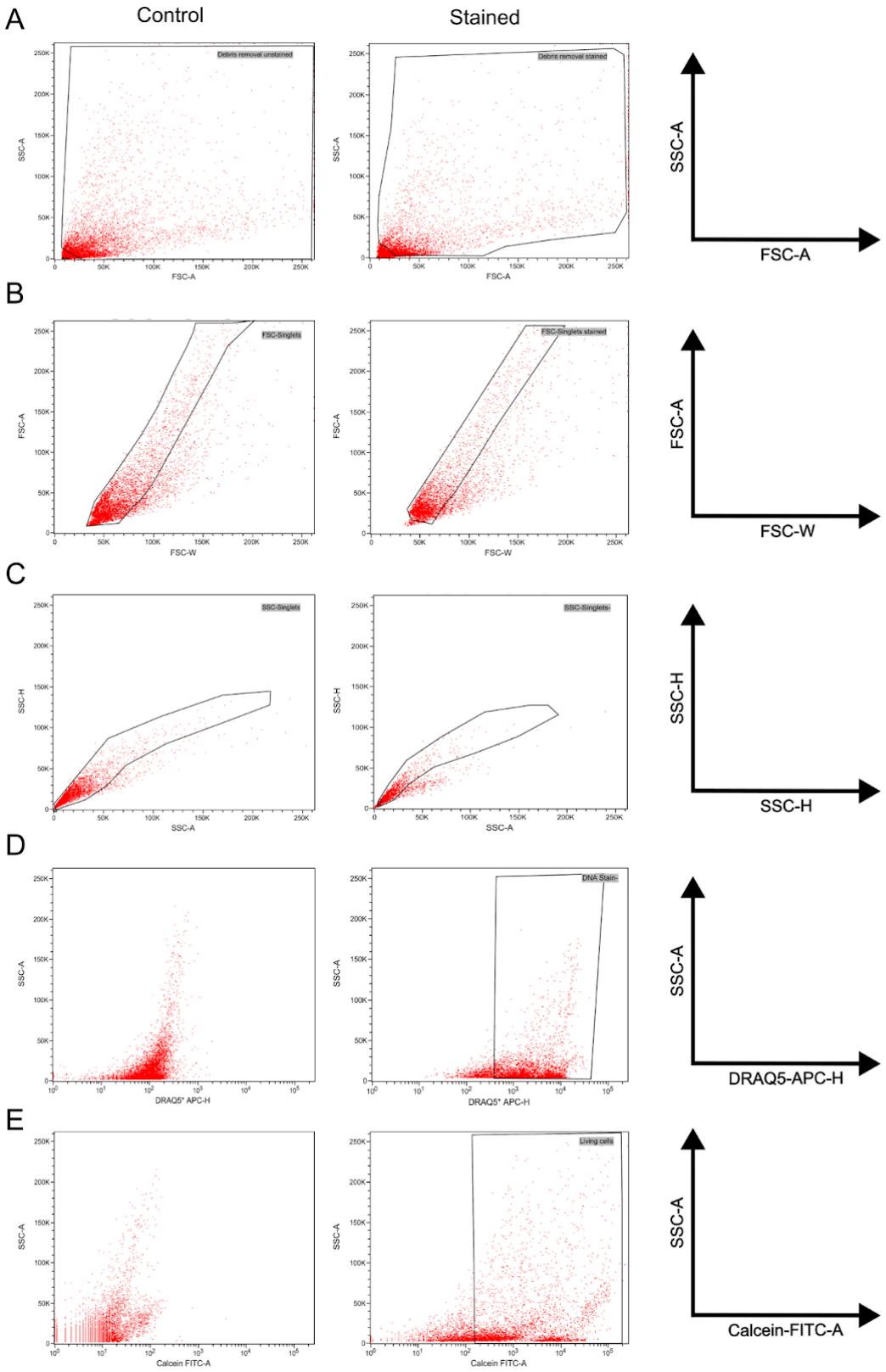


Figure 3.2 Gating strategy for cell sorting. Representative figures were generated by collecting 10,000 events. Unstained cells served as control. **A.** All events in the SSC vs. FSC gate were set to remove debris. **B.** Cells were analyzed in FSC-W vs. FSC-A to remove doublets in cells gated on A. **C.** An additional doublet discrimination step was done by visualizing the cells gated in B in SSC-A vs. SSC-H. **D**. Cells gated in **C** were analyzed in DRAQ5-APC to select cells based on their DNA content. **E.** To detect only living cells gated in D, the cells were visualized in Calcein-FITC-A vs. SSC-A.

While being an important tool, cell sorting can have a severe impact on viability and cell composition. Careful validation was performed to confirm the suitability of sorting for this project. Initially, sufficient RNA quality was confirmed before and after sorting (Fig 3.3 A). Also, a strategy was devised to test for the presence of a selection of major cell types within the cell suspension. Using scRNA-seq data of the related blood fluke *S. mansoni [111]*, cell marker genes were selected (Fig 3.3 B), and orthologues of those genes were identified in the *F. hepatica* genome. The selected genes were *neuroendocrine protein 7b2* (*7b2*) and a putative *muscarinic acetylcholine receptor* (*mar*) for nerve cells, *bone marrow proteoglycan* (*bmpg*) for female germ cells, *troponin* for muscle cells, as well as *histone 2b* (*h2b*) as a marker for proliferating cells. Another addition to the potential marker panel was a *cathepsin* known to be expressed in intestinal cells of the liver fluke [180]. The presence of marker gene transcripts was tested using an RT-PCR with cDNA from cells before and after sorting (Fig 3.3 C). Results indicated that none of the transcripts was entirely lost by sorting. However, the reduced signal strength of PCR products for *h2b* and *cathepsin* indicated that sorting might impact cells containing these transcripts.

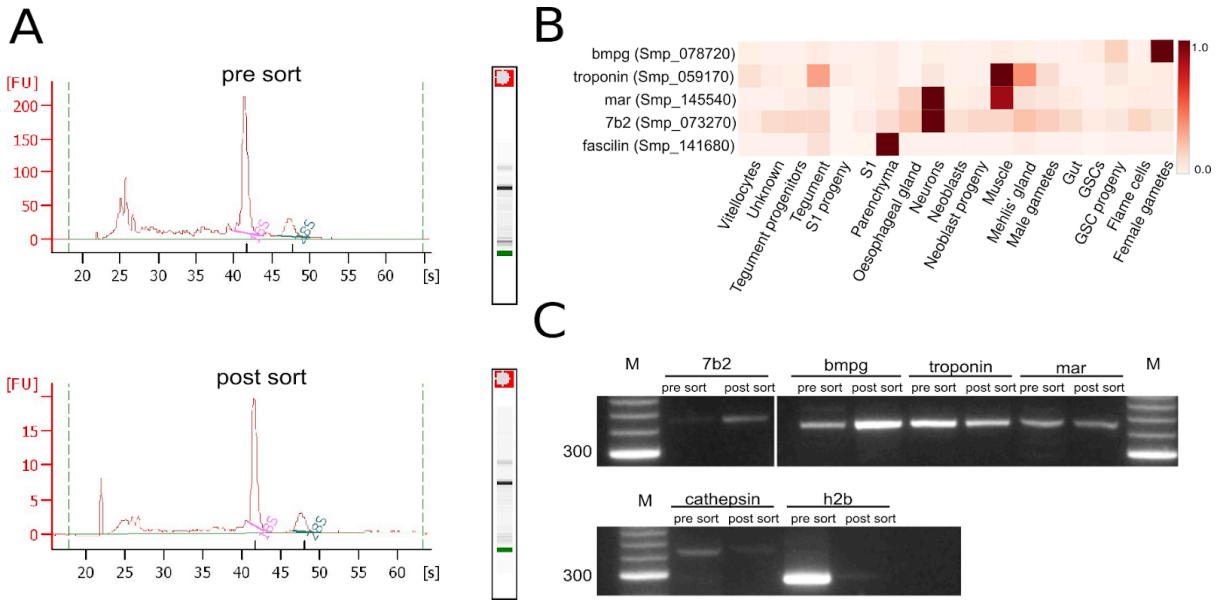


Figure 3.3 Validation of cell suspension quality and composition before and after sorting. A. Bioanalyzer electropherogram of RNA from single cells before (pre) and after (post) sorting. **B.** Averaged expression of selected cell marker genes from the adult *Schistosoma* scRNA-seq data [111]; values were scaled and centered per gene individually. **C.** By RT-PCR using cDNA prepared from RNA of cell suspension, transcripts of the following genes were detected before and after sorting: *cathepsin* (intestinal marker), putative orthologs of *h2b* (stem cell marker), *fascilin* (parenchymal marker), *neuroendocrine protein 7b2* (neuronal marker), putative *muscarinic acetylcholine receptor/mar* (neuronal marker), *troponin* (muscle marker), *bone marrow proteoglycan/bmpg* (germline marker). Legend: M: marker; pre: before sorting; post: after sorting

Though the presence of transcripts for *h2b* still indicated the presence of proliferating cells, it was decided to confirm the presence of those cells. Therefore, worms were incubated before dissociation with EdU, sorted, and finally, the presence of EdU was visualized (Fig 3.4). This approach was used to further validate the presence of a population of proliferating cells within the sorted cells. The EdU positive cells showed the characteristic morphology of stem cells, with a large prominent nucleus and little cytoplasm.

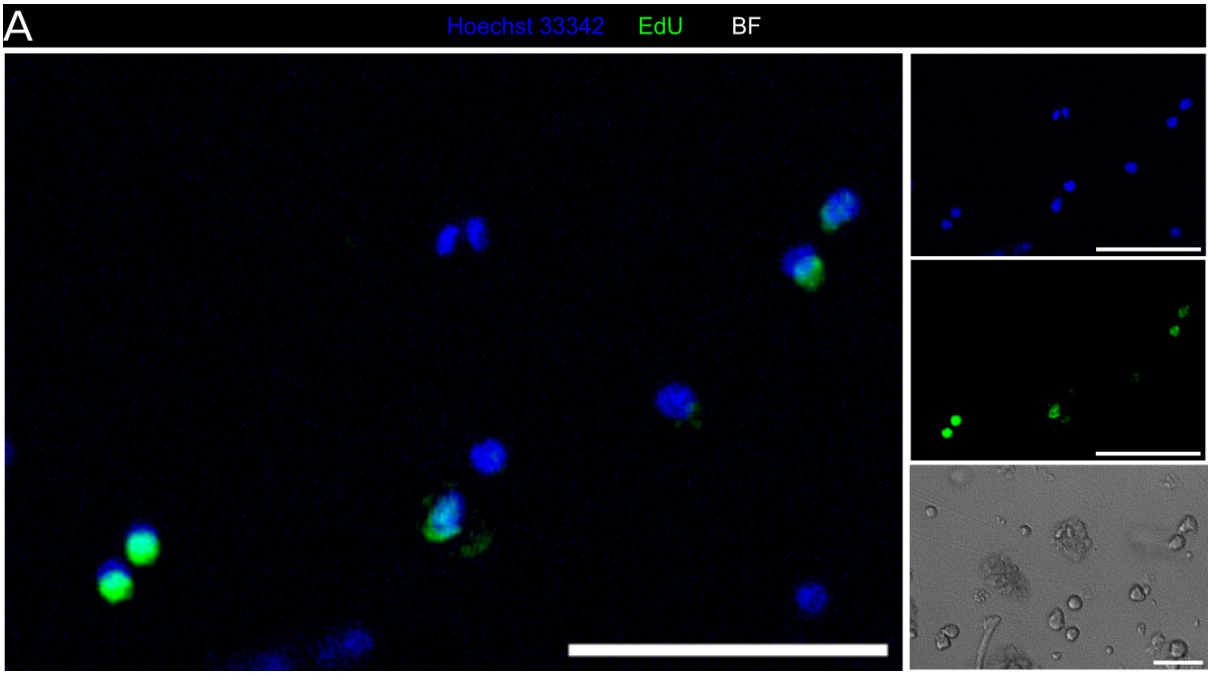


Figure 3.4 Presence of proliferating cells after cell dissociation. Confocal image showing EdU positive cells. EdU is visualized in green, and nuclei were stained using Hoechst 33342 in blue. Scale: 50 µm

Another cell type that required extra validation was the vitelline cells. These cells serve as the support cells of the developing embryo within the egg, contribute during eggshell formation, and provide nutrients for the developing embryo. These cells have been shown to differentiate through four developmental stages [181,182], whereby the final S4 cell stage is known to contain large accumulations of lipids and other nutrients [182]. Due to their sensitive nature, these cells are fragile and have been shown to break apart easily, especially during FACS [183]. An approach similar to a previous study was used to validate that these sensitive S4 vitelline cells were present within the cell suspension, especially after FACS [183]. To this end, these cells were stained with LipidSpot, which stains lipid droplets within living and fixed cells. It could be confirmed that a population of large cells containing lipid droplets was still present within the suspension (Fig 3.5), which resembled the description of those cells in schistosomes [183].

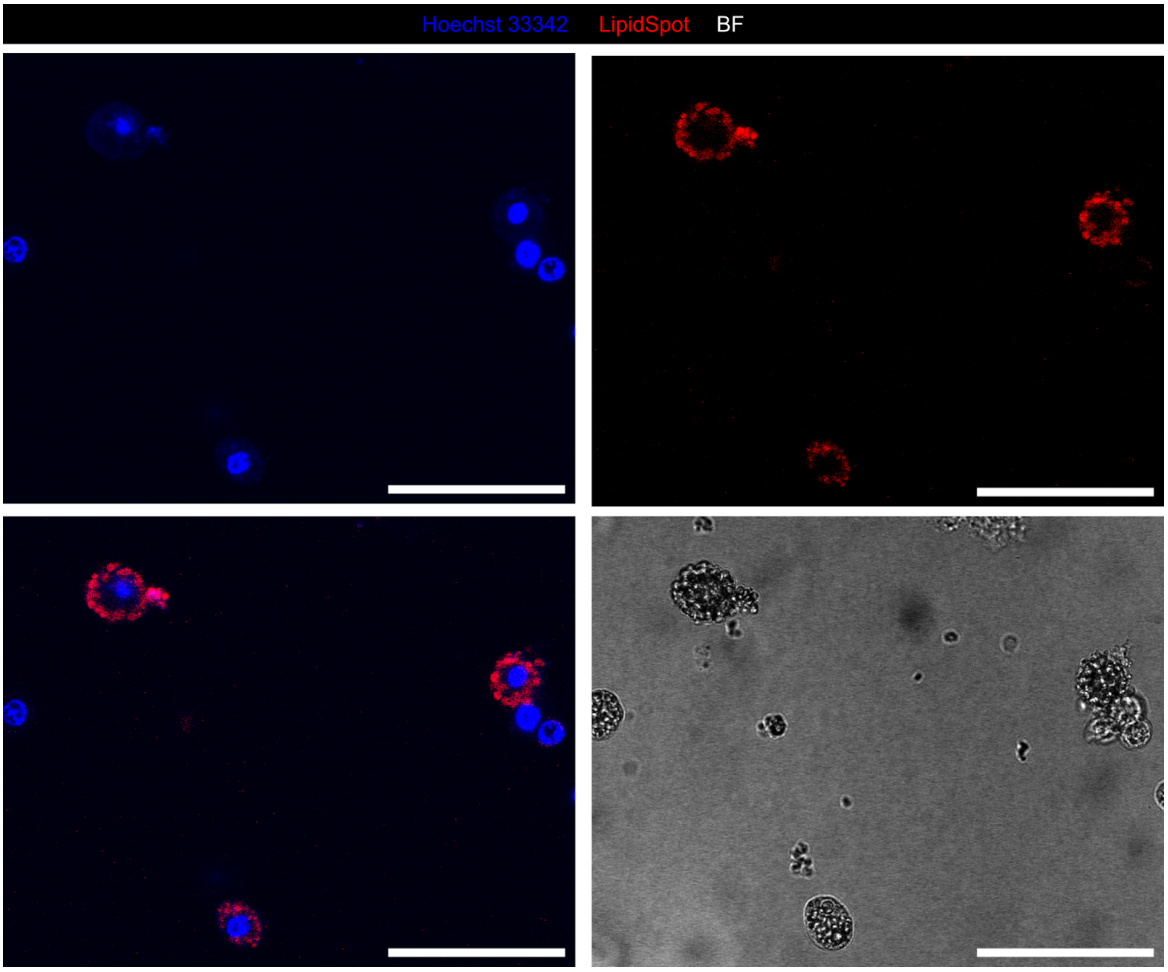


Figure 3.5 Cells showing characteristic lipid droplet distribution in cell suspension. Confocal image showing cells stained with the dye LipidSpot (red) with nuclei stained with Hoechst 33342 in blue. Scale: 50 µm

Many lipid droplets were also present within the cell suspension, mainly contributing to a large amount of debris, indicating the suspected rupture of vitelline cells (Fig 3.6).

Even though vitelline cells seemed to rupture significantly during the dissociation process and there might be a negative impact of FACS on the cell suspension, the high grade of heterogeneity, high cell viability, and confirmed integrity of RNA after dissociation and FACS made the cells suitable for a scRNA-seq experiment.

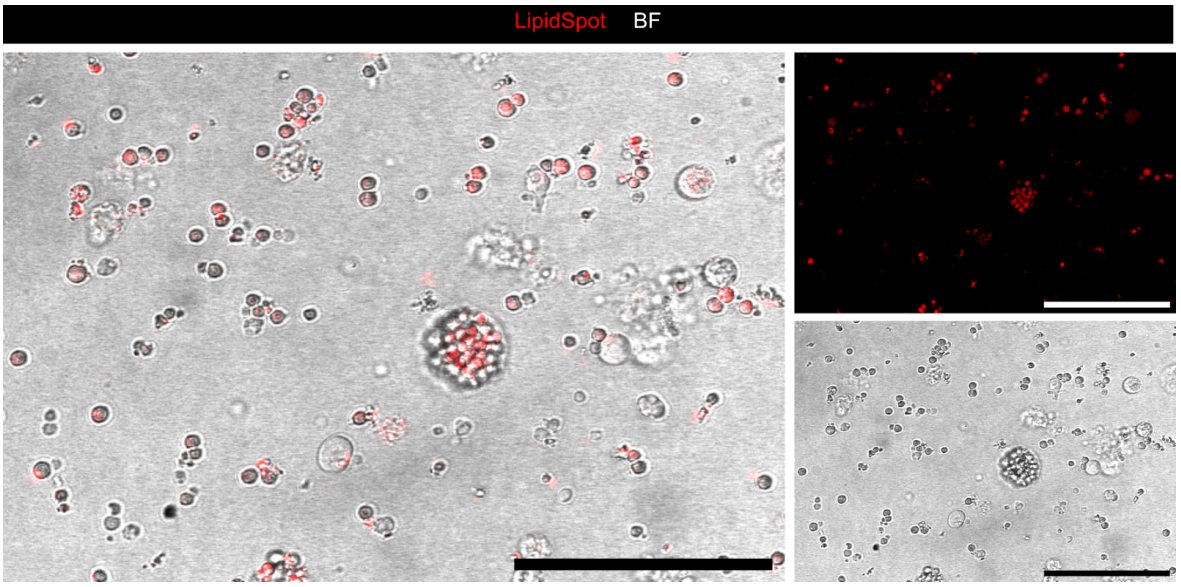


Figure 3.6 Small lipid droplets are the main components of debris after dissociation. A. Confocal image showing droplets stained with the dye LipidSpot (red). Scale: 50 µm

3.2 Single-cell transcriptomics for liver fluke cells

Having generated a high quality single cell suspension for use in scRNA-seq experiments, this chapter will describe the considerations for designing the initially performed scRNA-seq experiment and its thorough analysis. Finally, improvements made to the protocol will be presented later in this study.

3.2.1 Determination of total cell number for adult *F. hepatica*

Prior to a scRNA-seq experiment, some considerations had to be taken. Protocols like the 10x Genomics Chromium solution utilized here are droplet-based solutions adapted from the initially described Drop-seq system [100]. These systems core features is the co-capture of cells together with oligo-coated beads inside droplets. The great advantage of those systems is their throughput, with a maximum number of 10,000 cells that can be loaded on the 10x system. Even though droplet-based scRNA-seq assays have a high throughput, capturing all cells of a complex organism is nearly impossible. Due to this restriction, the cells examined in such an experiment are only a random sample from the total heterogeneous cell population. Knowledge of the cell number and rough composition within the liver fluke is, therefore, essential to make assumptions on how many cells need to be sampled to capture rarer cell populations. As there was no prior knowledge in terms of total cell count for the liver fluke, a prior estimation was gained from counting and extrapolating nuclei numbers from different regions of the worm. Initially, tests proved that whole worm samples were unsuitable as the worm size complicated the imaging procedure (data not shown). As a

workaround, worms were cross sectioned in two regions anterior to the uterus and the region posterior to the uterus, as well as sectioned in frontal orientation (Fig 3.7 A-B). Frontal sections were gratefully provided by Svenja Gramberg. The extrapolated nuclei number per worm was around 17 million cells (Fig 3.7 C). This number was independent of the section plane. However, this number can only approximate the actual number of cells, as the tegument is a syncytial tissue, and fused cells are described during spermatogenesis [184]. It still provided important information for the later experiments. An important observation was the difference between the total nuclei count between anterior and posterior body parts, which originated from the densely populated male reproductive tissues and the vitellarium found in the posterior part, apart from the noticeable size difference of areas observed. These tissue types are highly proliferating parts of the worm's body, essential for the worm's reproductive output. In terms of profiling cells by scRNA-seq, this posed a challenge because the overabundance of cells in these tissues could hinder the detection of rarer cell types that might not be captured during the barcoding procedure.

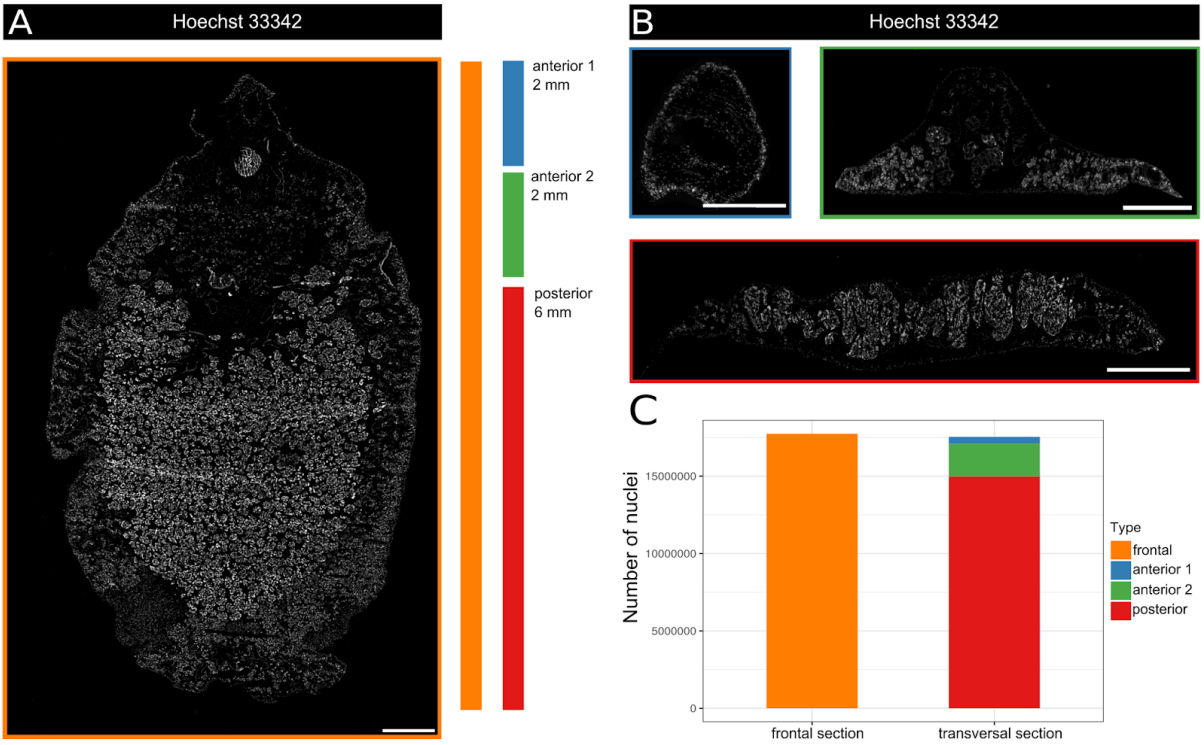


Figure 3.7 Determination of nuclei number results in areas with differing cellular density. A. Staining of nuclei in a representative frontal section from an adult worm using Hoechst 33342. Colored bars indicate regions for transversal sectioning. Scale bar: 1000 µm. **B.** Transversal sections with nuclei stained with Hoechst 33342. Section from anterior 1 (top left), anterior 2 (top right) and posterior (bottom). Scale bar: 1000 µm. **C.** Extrapolated total cell number from frontal or transversal sections illustrated as a stacked bar plot.

3.2.2 Workflow and QC

Based on the knowledge gained from investigating the worm's nuclei density, the workflow was optimized to capture the highest variation of cell types possible and to enrich rarer cells, derived from the more sparsely nucleated anterior regions. Therefore, the worm was cut and separated into anterior and posterior parts before it was dissociated and the cells were sorted for viability by calcein AM staining and FACS. A total of 10 samples, consisting of viable, unfixed cells, was analyzed using the 10x Genomics Chromium platform, separated into four samples from each the anterior and posterior part and two samples using whole worm cells as input. This approach enriched anterior cells that would have been undersampled due to the abundance of cells in the worm's posterior part. The resulting libraries were sent to sequencing, and reads were mapped to the liver flukes genome [92] using the cellranger pipeline [102].

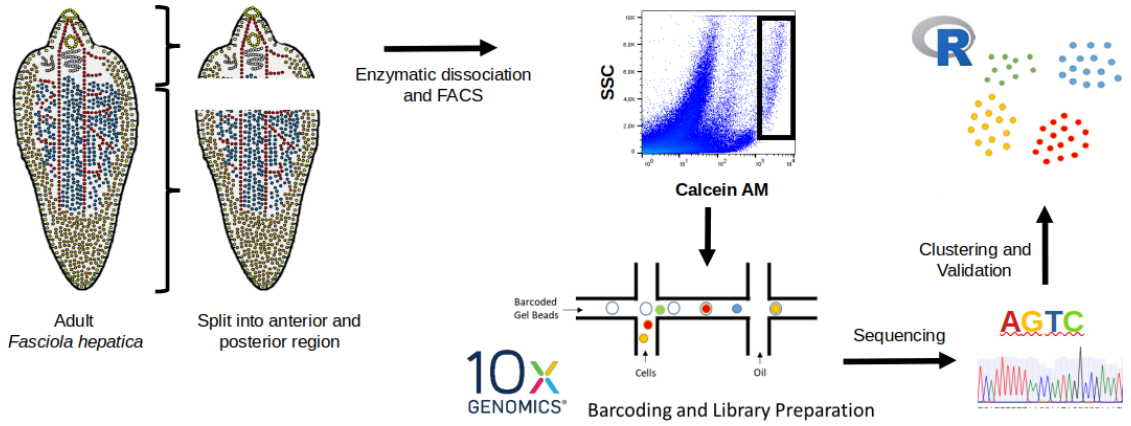


Figure 3.8 Schematic representation of the single cell RNA-seq workflow. This workflow shows the major steps for data generation. Adult worms were mechanically divided into anterior and posterior regions, following dissociation into single cells by enzymatic dissociation. Afterwards, cells were sorted based on their viability, which was determined by calcein-AM staining. Following barcoding using the 10x Genomics Chromium protocol, the libraries were sequenced. Finally, the resulting cells were clustered based on their transcriptomes.

Around 90% of the reads in the samples were confidently mapped to the genome. On average, cells were sequenced to 70,000 reads per cell, ranging from 25,000 to 280,000 reads per cell (Table 3.2). The median number of genes detected per cell was 2644 with an average of 14,087 UMIs per cell. That means that, on average, 30% of the 8716 total transcribed genes in adult worms were detected per cell (TPM > 2) [91].

Table 3.2 QC parameters of the used libraries

Sample	Reads per cell	Average genes per cell	Number of cells	Average UMI per cell
Body_1	40,882	33,245	3103	20,949
Body_2	40,020	3570	3134	23,955
Body_3	73,595	3062	1178	20,435
Body_4	111,491	2602	1017	15,643
Head_1	287,800	2504	670	10,720
Head_2	26,794	1605	1859	6136
Head_3	45,108	1708	1078	6191
Head_4	57,994	2404	1116	12,409
Whole_1	25,598	3010	3265	14,416
Whole_2	60,887	2602	3261	10,022

One initial challenge in scRNA-seq analysis and droplet-based approaches is distinguishing between actual cells and droplets containing ambient RNA. Though the mapping and preprocessing through the cellranger pipeline [102] includes a simplified cell calling step, this was found too broad to be used as a final set of cells for the analysis. Additionally, other metrics have seen common use, such as the percentage of mitochondrial reads associated with damaged cells or a cutoff for high gene and UMI counts to exclude doublets [98]. Plotting the number of UMIs against the number of genes using prefiltered cellranger output showed that while there was an expected linear relationship between UMIs and the detected number of genes within a cell, a proportion of cells did not adhere to this relationship (Fig 3.9 A). Upon closer inspection, most of these cells had high mitochondrial reads and were deemed insufficient for the analysis. Based on those observations, a conservative threshold was set so that cells with more than 5% of mitochondrial reads were excluded from the analysis (Fig 3.9 B). Finally, cells with at least 1000 UMIs but not exceeding 50,000 and with at least 750 genes detected were included.

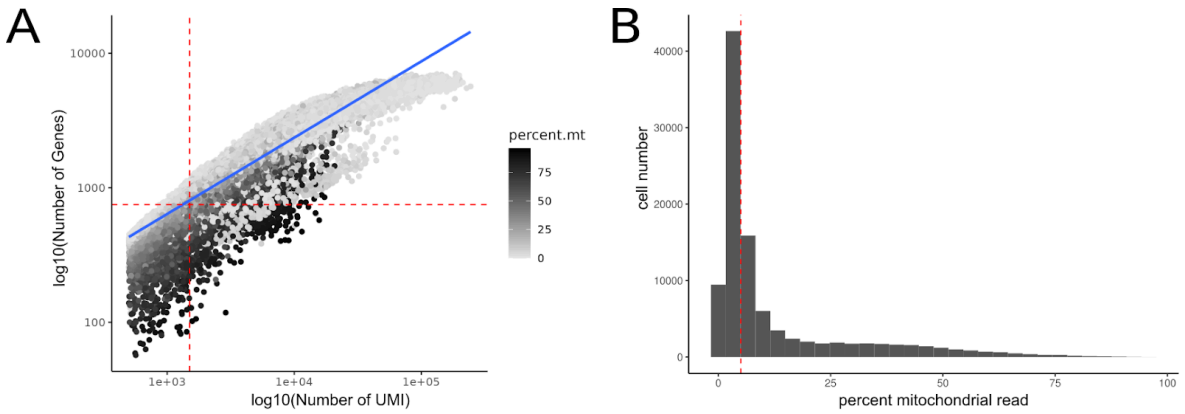


Figure 3.9 Applied thresholds to select cells with sufficient quality. Plots showing cellranger output data before filtering. **A.** The number of UMIs plotted against the number of genes per cell shows a linear relationship. Cells are shaded based on their percentage of mitochondrial genes. The dotted red lines indicate chosen thresholds. Only cells within the lower left quadrant were kept in the analysis. **B.** A histogram plotting the total number of cells based on their percentage of mitochondrial genes. The dotted red line shows the chosen threshold beyond which cells were not included in the analysis.

After filtering, 19,581 cells remained in the analysis. The different samples showed a comparable distribution with respect to the chosen QC metric (Fig 3.10 A-C). The subsequent analysis was performed using the R software package Seurat V4. Following the Seurat workflow, the data was preprocessed, which involved normalization scaling and, most importantly, dimensionality reduction. A linear dimensionality reduction was performed using principal component analysis to reduce the intensity of computation and lower the noise of the data. The principal components (PC) represent sets of genes later used to cluster the cells based on their transcriptional identity. Using both elbow plotting and JackStrawing [100], the first 25 PCs were found to carry the most information. Finally, the cells were clustered based on their transcriptomes.

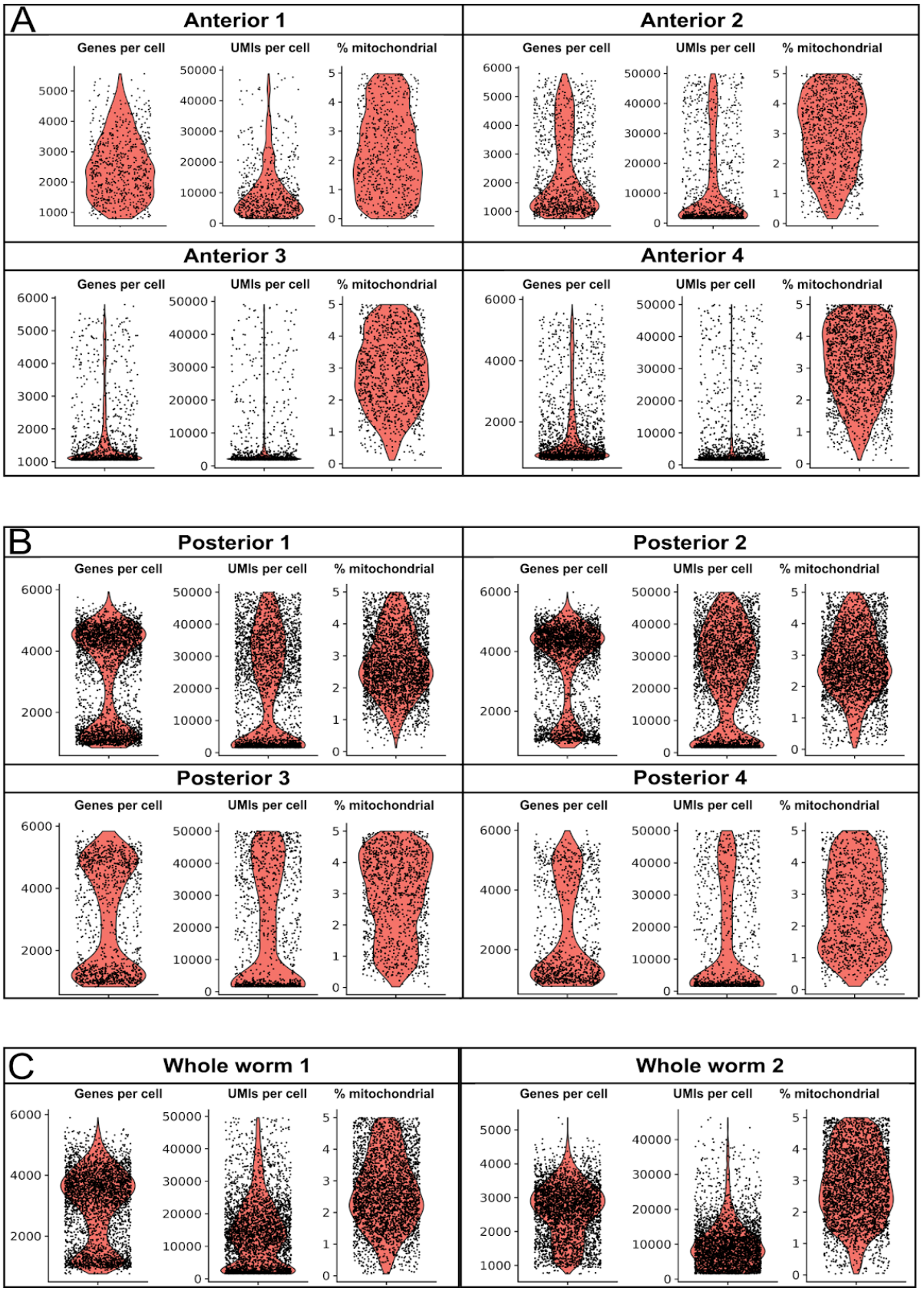


Figure 3.10 QC parameter distributions per sample. Plots showing data after filtering. Violin plots showing the distribution of the number of genes, number of UMIs, and percentage of mitochondrial genes for anterior (**A),** posterior **(B),** and whole worm samples **(C).**

3.2.3 Single-cell sequencing: clustering and validation

After preprocessing the data, the cells were clustered and visualized using the uniform manifold approximation and projection (UMAP). A total of 15 clusters were detected using the described workflow (Fig 3.11). Marker genes for these clusters were computed and annotated (Fig 3.12), which is further detailed below. A marker gene is defined as a gene that is expressed significantly higher in one cell type compared to the others. The predicted clusters could be mapped back to known tissues with cells belonging to muscle and muscle-associated (2 clusters), gut (1), testes (3), stem cells (1), ovary (2), vitellarium (2), eggs/uterus (1), Mehlis' gland (1), and two clusters for which no annotation was found.

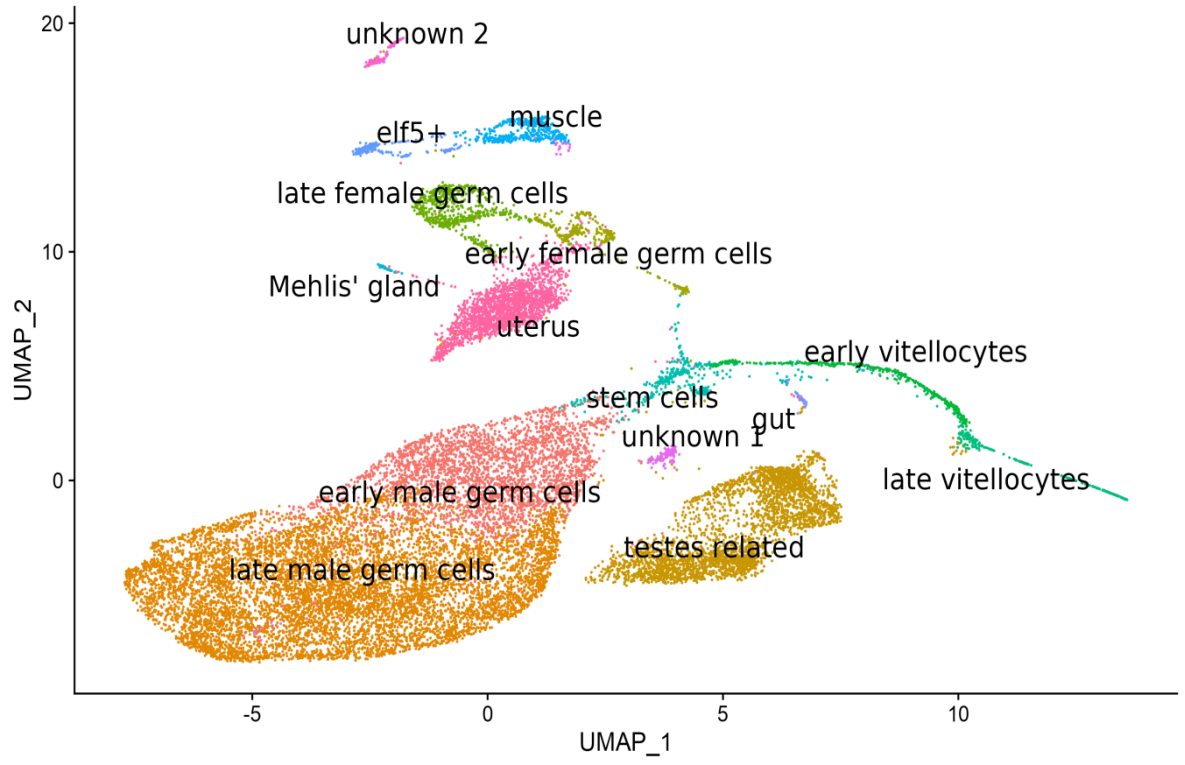


Figure 3.11 scRNA-seq analysis identifies 15 distinct transcriptome clusters. Uniform Manifold Approximation and Projection (UMAP) of 19,581 cells. Clusters are colored, and labels are added.

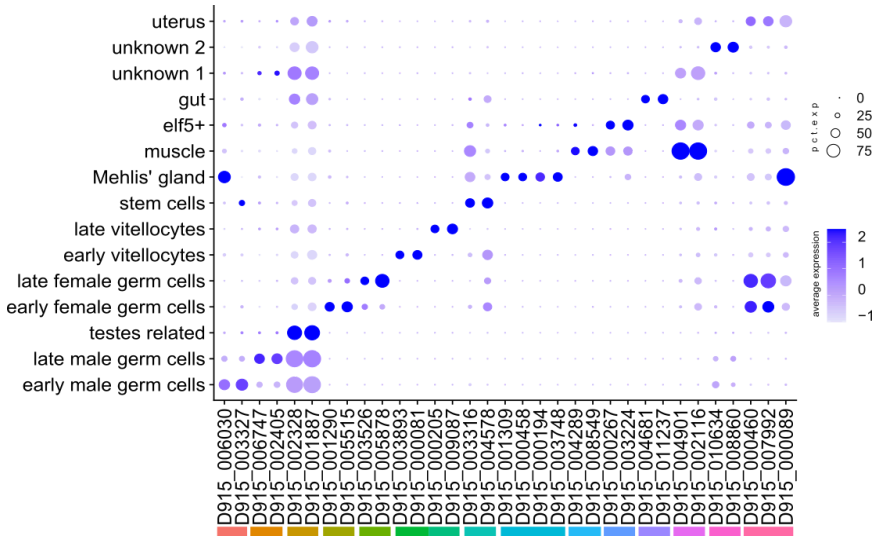


Figure 3.12 Identified clusters had distinct marker profiles. Expression profiles of selected genes over all clusters are illustrated as a dot plot. Shown is the expression of at least two marker genes for each cluster, and the cluster color is indicated below each marker pair. The expression level is indicated by color, from blue (high expression) to lavender (low expression). The size of the circles represents the percentage with which the cluster cells express the given gene.

The number of cells within a cluster and its marker genes differed significantly (Table 3.3). Here, an abundance of cells associated with the male reproductive tract was apparent, which was in line with earlier observations that this tissue contributed the majority of cells to the total cell count. Cells from the vitellarium and the ovary added to the high number of cells belonging to reproductive tissues. Eventually, 81% of the cells within the dataset could be directly attributed to the reproductive apparatus.

Based on this observation, the composition of cell types within the different sample types was plotted to validate the utility of the enrichment of anterior cells (Fig 3.13). A higher heterogeneity was detected in the anterior samples. In contrast, whole worm samples and the posterior-only samples had most cells stemming from the male germ cells. To assess whether the scRNA-seq data had a comparable cell composition as bulk transcriptomic data [91], the correlation of data from both techniques was calculated. The Pearson correlation coefficient was 0.142, which indicated no correlation (Fig 3.14).

Table 3.3 Total cell and marker number per cluster

Cluster	Cell number	Marker gene number
late male germ cells	80,75	560
testes related	3218	11
early male germ cells	3195	258
uterus	2059	37
late female germ cells	675	418
muscle	409	858
early vitellocytes	398	1429
stem cells	398	174
early female germ cells	323	538
late vitellocytes	231	89
elf5+	227	216
unknown 1	162	12
unknown 2	147	283
Mehlis' Gland	45	570
gut	19	211

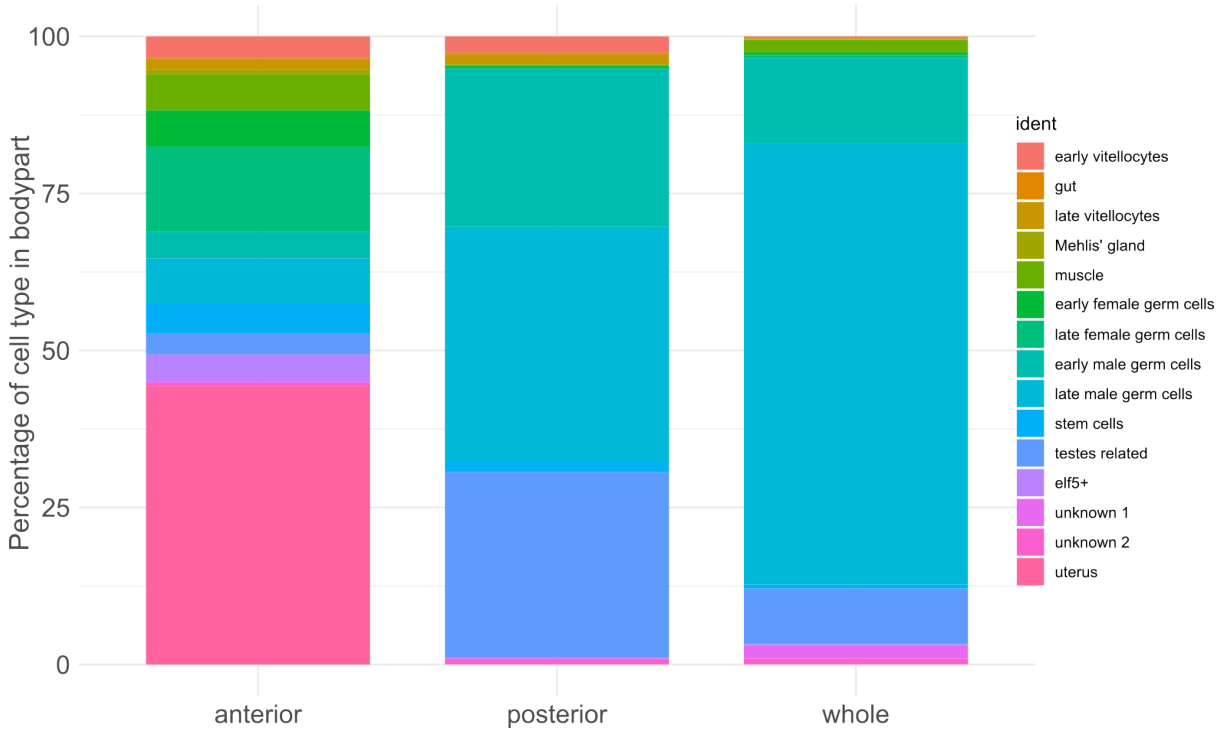


Figure 3.13 Composition of cell types is different between samples. Cell numbers for each sample type (anterior, posterior, or whole worm samples) were determined. Shown is the percentage that each cluster covers within the total cell number per sample type

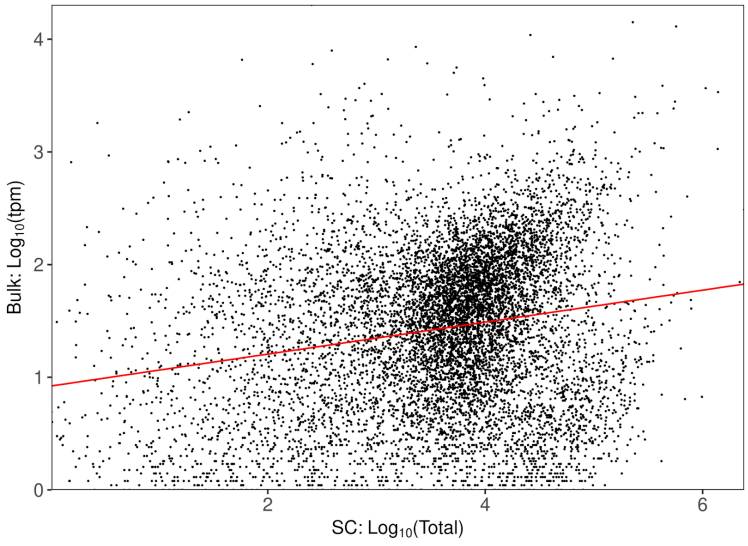


Figure 3.14 scRNA-seq data had no significant correlation with bulk RNA-seq datasets. Comparison of bulk transcript per million counts (tpm) and UMI counts from the single-cell RNA-seq counts. The Pearson correlation coefficient assessed the correlation between the bulk and single-cell RNA-seq datasets. The linear model regression line is indicated in red.

3.3 Characteristics of cell clusters in the single-cell atlas

3.3.1 Stem cells show diverse marker profiles reflective of different cell states

Stem cells are a unique cell type capable of self-renewal and can give rise to progenitor cells. For flatworms, the stem cells are subdivided into the so called neoblast-like cells [185,186], which give rise to the somatic cells, and the germline stem cells that build the resource for the differentiation of the gametes. An enhanced understanding of these cells could lead to novel treatment strategies as these worms rely on these cells for critical biological processes like growth and reproduction. For the liver flukes, the massive proliferative activity of the germline stem cells lays the foundation for an immense reproductive output of thousands of eggs per day [24]. Additionally, the neoblasts have been identified as the main driver of growth in juvenile worms [83–85].

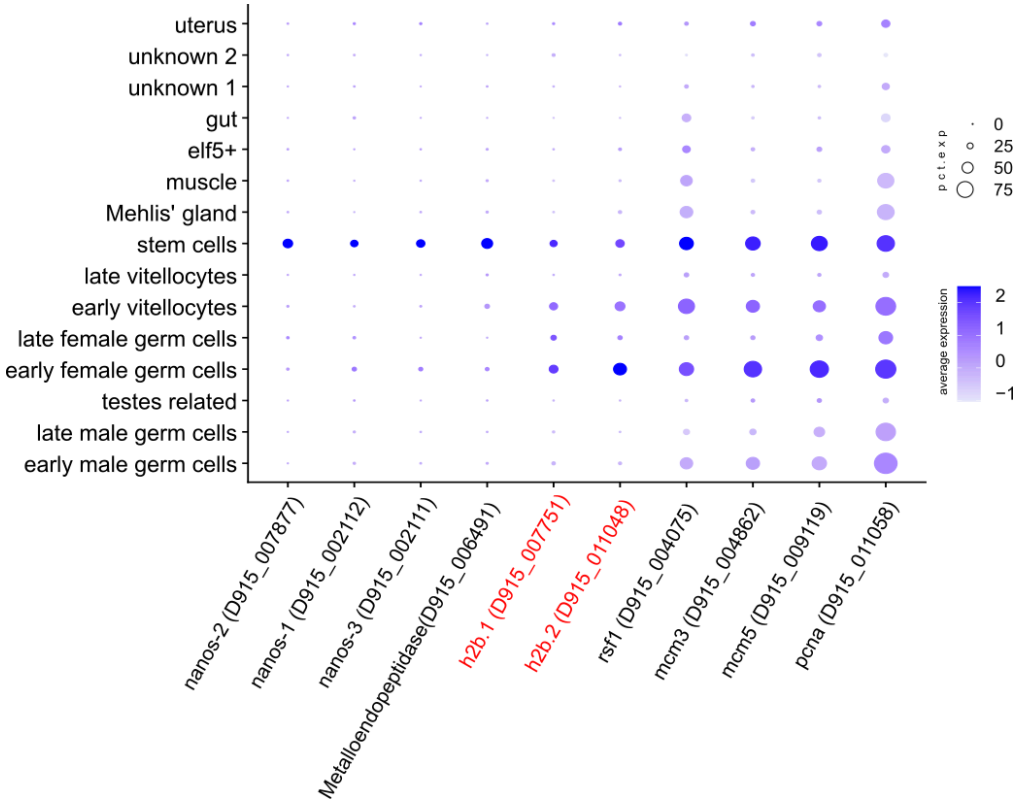


Figure 3.15 Predicted stem cells markers show two expression patterns. Profiles of selected marker gene expression for all clusters are illustrated as a dotplot. The expression level is indicated by color, from blue (high expression) to lavender (low expression). The percentage with which the cluster cells express the given gene is represented by the size of the circles. Genes for which *in situ* validation was performed are colored in red.

A cluster related to stem cells was identified based on three Nanos isoforms (D915_007877, D915_002111, D915_002112), which all showed specific expression profiles specific to the

stem cell cluster (Fig 3.15). As expression of different *nanos* genes is known to be a valuable tool to distinguish neoblast from germ line stem cells (GSC) in *S. mansoni [109]*, where Nanos-1 (Smp_055740) is a marker for GSCs, while Nanos-2 (Smp_051920) is expressed in all stem cells. In order to determine if Nanos might be a marker for different stem cells in *F. hepatica.* The three liver fluke Nanos proteins were compared to Nanos proteins from other parasitic flatworms via phylogenetic analysis (Fig 3.16). The parasitic flatworms formed a separate group, compared to Nanos from different organisms (supplementary Fig S1, 3.16). Within this group, subgroups formed around the schistosomal Nanos proteins. While D915_007877 was annotated as Nanos-1 *[129]*, based on their phylogenetic relationship, it was more similar to the schistosomal Nanos-2. Even though this relationship could be established, *nanos* expression patterns largely overlapped in the liver fluke scRNAseq data, and didn't show distinct profiles as in *S. mansoni* (Fig 3.15).

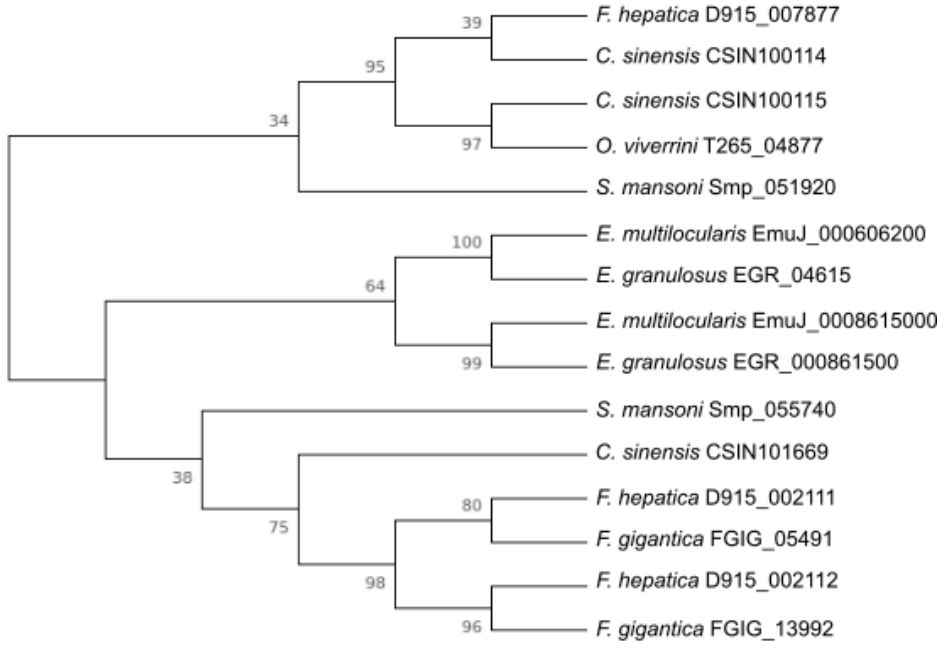


Figure 3.16 Phylogenetic analysis identifies nanos orthologues. The analysis is based on the amino acid sequences of selected Nanos proteins. A subtree of a broader analysis covering also non-parasitic organisms (supplementary Fig S1) is shown. For each node, the posterior probabilities are noted. The gene ID for each entry in WormBase ParaSite has been added.

A GO term analysis using marker genes for the stem cell cluster showed enrichment for genes associated with the molecular function of RNA binding (Fig 3.17), which is also a described function of Nanos [187]. This functional group of proteins is associated with essential processes like mRNA processing, binding, and splicing. By post-transcriptional regulation these processes are an additional layer of gene expression. Further looking into the genes associated with RNA binding, a functionally related network of genes involved with RNA splicing was identified, consisting of arginine-serine rich splicing factors (D915_007541,

D915_009045, D915_009797, D915_011156) and the *mRNA binding protein 25* (D915_009651), which are described to regulate alternative splicing and thereby processes like proliferation [188,189]. Besides the enrichment for genes involved with RNA binding, many enriched marker genes for the stem cell cluster were involved in translational activity, encoding the structural components of the ribosome and metabolic processes, highlighting the metabolic demand involved in proliferation.

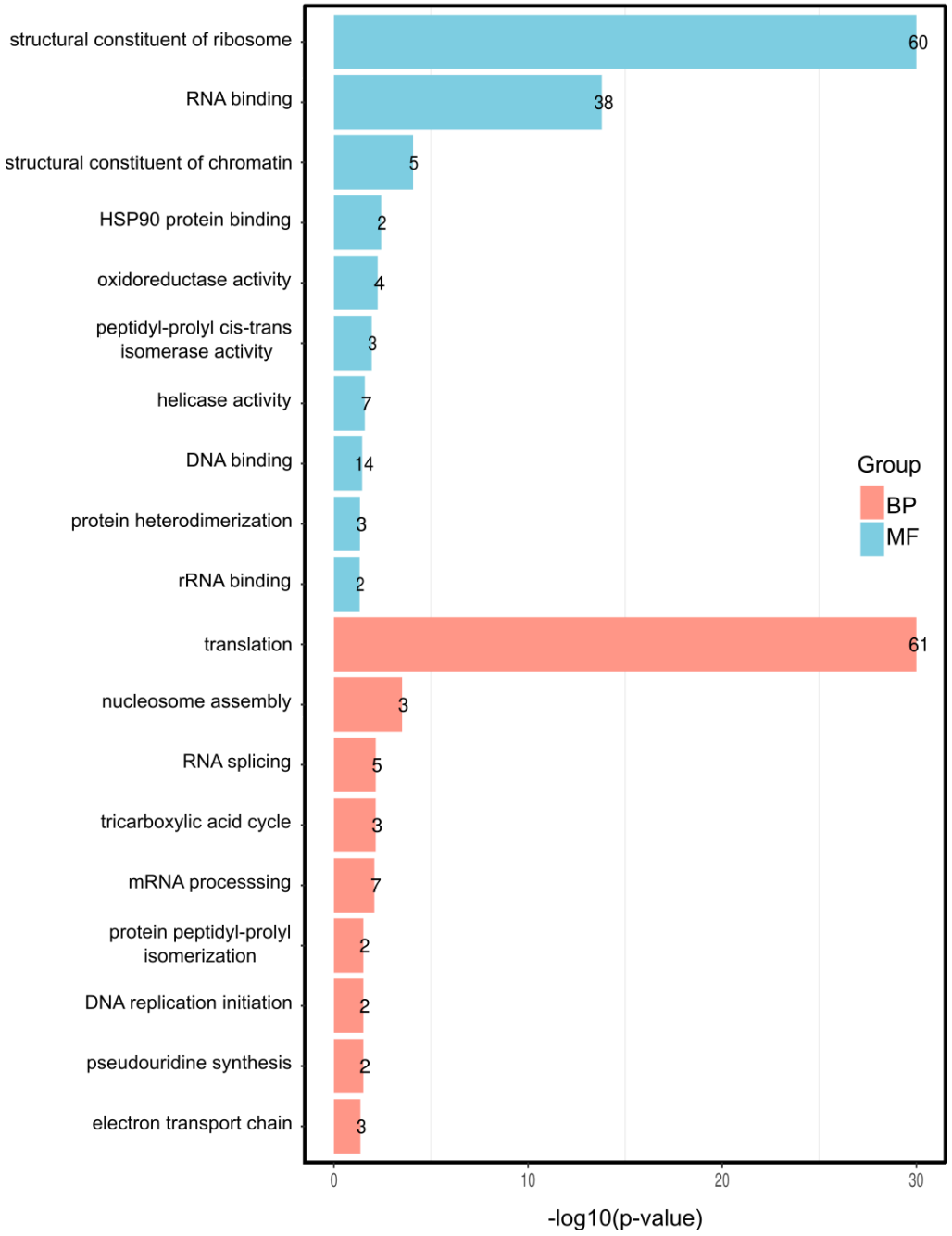


Figure 3.17 GO term analysis elucidates processes enriched in the stem cell cluster. GO enrichment for biological processes (BP) and molecular function (MF) is shown. Marker genes of the stem cell cluster were used as an input. For each term the false discovery rate is plotted on the x-axis. The number of genes involved with the term is indicated.

Markers confirming the proliferating activity of the cells in this cluster were genes like *histone 2b* or short *h2b* (D915_007751, D915_011048) as well as genes encoding the minichromosome maintenance (MCM) complex. Some of these marker genes for the stem cell cluster showed interesting patterns when investigating their expression over all additional clusters. While *nanos* expression was only specific for the stem cell cluster, *h2b* transcripts were detected in other potential clusters of progenitor cells like early vitellocytes or the early male germ cell cluster. Previous studies in helminths have already highlighted the value of histone expression as a method to label and measure active proliferation [185], and a novel work described the ablation of stem cells following *h2b* RNAi in larval stages of *F. hepatica* [83]. To validate these findings for adult worms, *h2b* transcripts were labeled via RNA FISH combined with an EdU (thymidine analog) cell proliferation assay. It has to be noted that *h2b* transcripts here refer to transcripts from two genes as their sequence is similar to a degree that is not distinguishable by the RNA probes used here. Despite this limitation, both genes show a comparable expression pattern across the different clusters (Fig 3.15) and are therefore proposed to follow similar transcription patterns. Overall, 70% of the cells positive for transcripts of *h2b* also stained positive for EdU. Double positive cells were located in regions known to be hotspots for proliferation, like the periphery of the testicular lobes or the ovary ([184], Fig 3.18). Additionally, cells near the tegument and, to a lesser extent, the intestine showed double positive staining, likely representing the neoblast-like cells. Unexpectedly, intense staining for *h2b* transcripts was detected in the mature part of the ovary, though these cells were negative for EdU.

Finally, it has to be highlighted that a metalloendopeptidase (D915_006491) was specifically detected in the stem cell cluster, similar to the *nanos* genes (Fig 3.15). Searching the protein sequence of this peptidase against the Hmmer3 database [139] annotated the enzyme as an astacin-like protease. Members of this family of proteases include the human *bone morphogenetic protein 1* (BMP-1), lovastatin, and others. These families are known to regulate extracellular matrix composition [190], as well as downstream of growth factors [191].

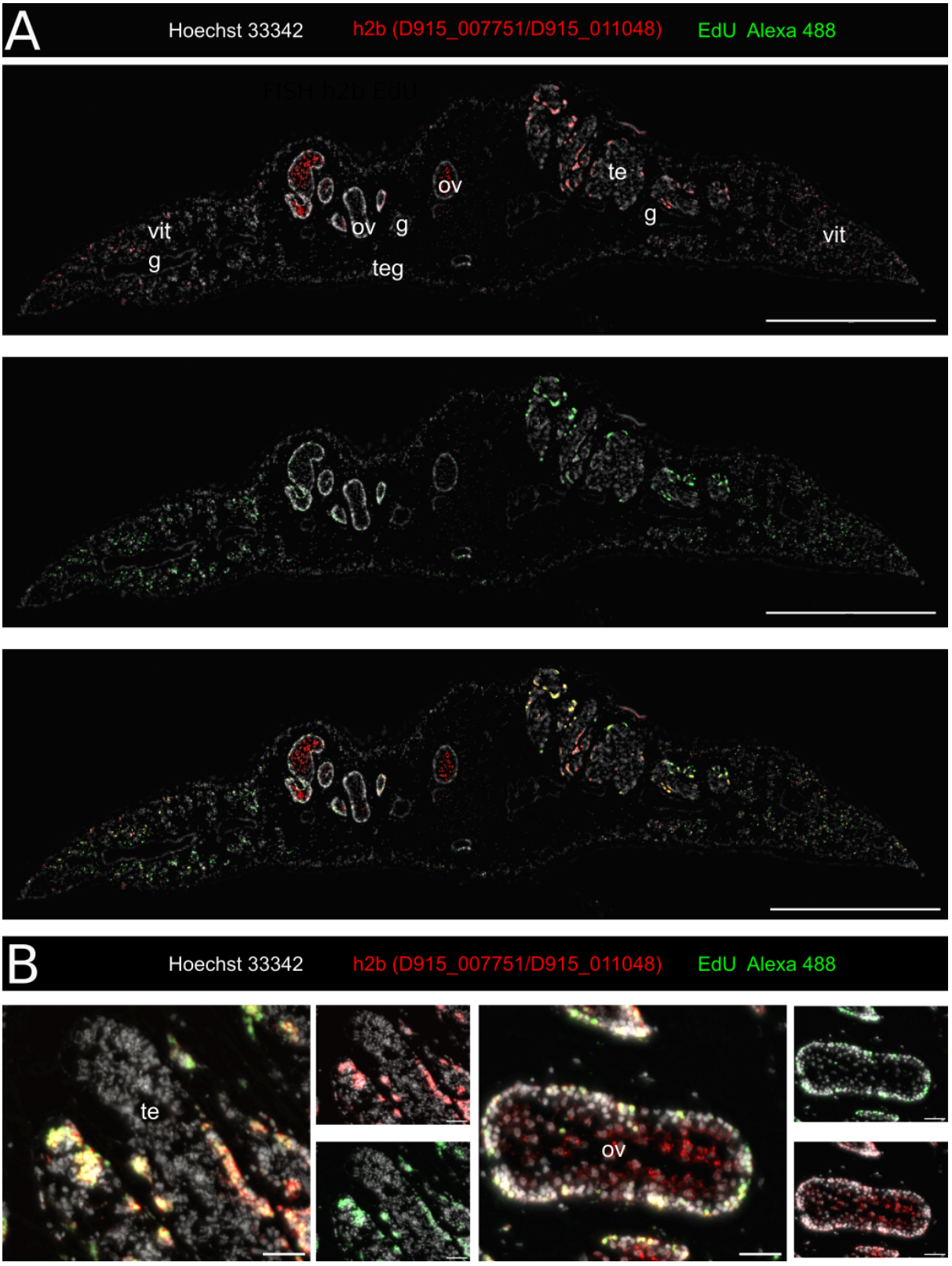


Figure 3.18 Transcripts of *h2b* have high overlap with EdU positive cells. A. Transversal section stained for *h2b* transcripts by FISH in, proliferating cells with EdU in green and nuclei stained with Hoechst in grey. Scale bar: 1000 µm. **B.** Close-up of the testis (left) and ovary (right) from B. Scale bar: 50 µm Legend: gastrodermis (g), tegument (teg), testes (te), ovary (ov), vitellarium (vit). Sense controls were found to be negative (Supplementary Fig. S8)

3.3.2 Male and female germ cell signatures during differentiation

As stated earlier, *F. hepatica* has an immense reproductive capacity fueled by both female and male reproductive tissues. In this dataset, clusters of cells that could be connected to both of these tissue types were annotated. Cells from both undifferentiated cells (early male/female germ cells) and more differentiated cells (late male/female germ cells) were captured and had distinct marker profiles (Fig 3.19).

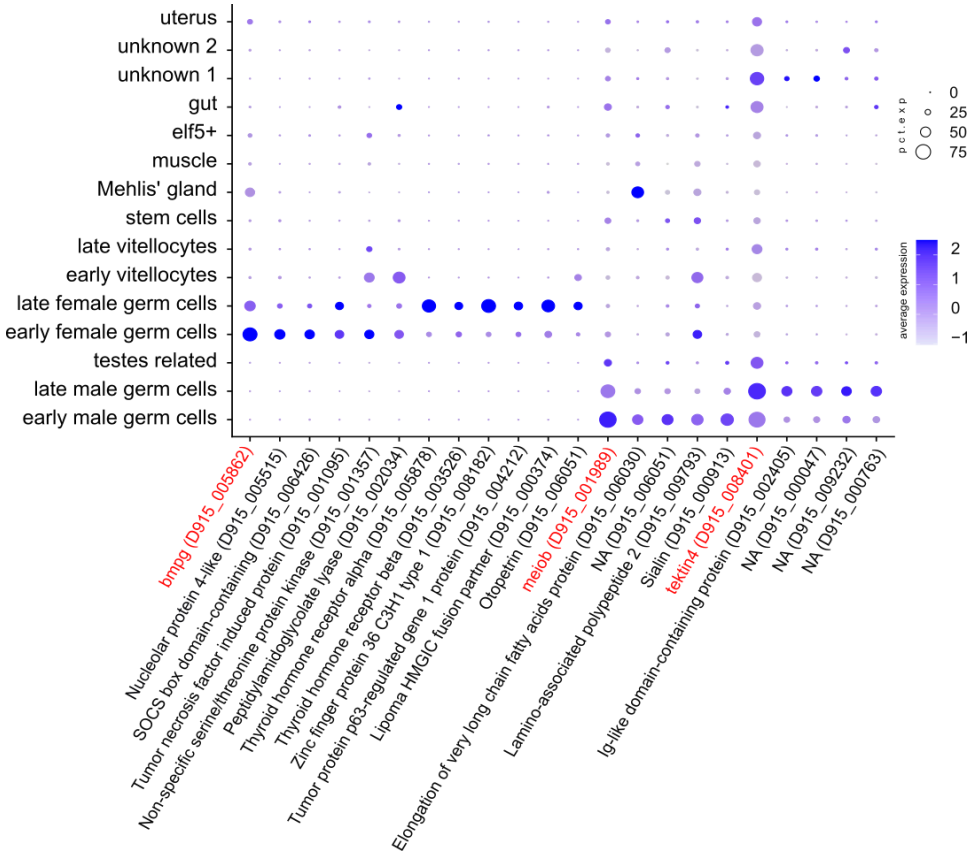


Figure 3.19 Clusters of male and female reproductive tracts had distinct markers. Profiles of gene expression over all clusters are illustrated as dotplot. The expression level is indicated by color, from blue (high expression) to lavender (low expression). The percentage with which the cluster cells express the given gene is represented by the size of the circles.

The cells of the early male germ cell cluster expressed genes like *boule* (D915_007531) and the transcription factor *one cut 1* (D915_002483) [114,192], which were previously identified in schistosomes where they serve function in early male germ cell development (Fig 3.20). The early cell clusters also showed a marked expression of the *synaptonemal complex protein 2* (D915_003691, Fig 3.20). Notably, the synaptonemal complex has been successfully stained by immunohistology in early prophase oocytes of *F. hepatica [193]*. The synaptonemal complex, while necessary for meiotic division, can be expressed as early as the premeiotic spermatogonia stage [194,195]. Additionally, the expression of

meiosis-specific with OB-fold (*meiob*) was found, which is essential for meiotic recombination in humans [196,197].

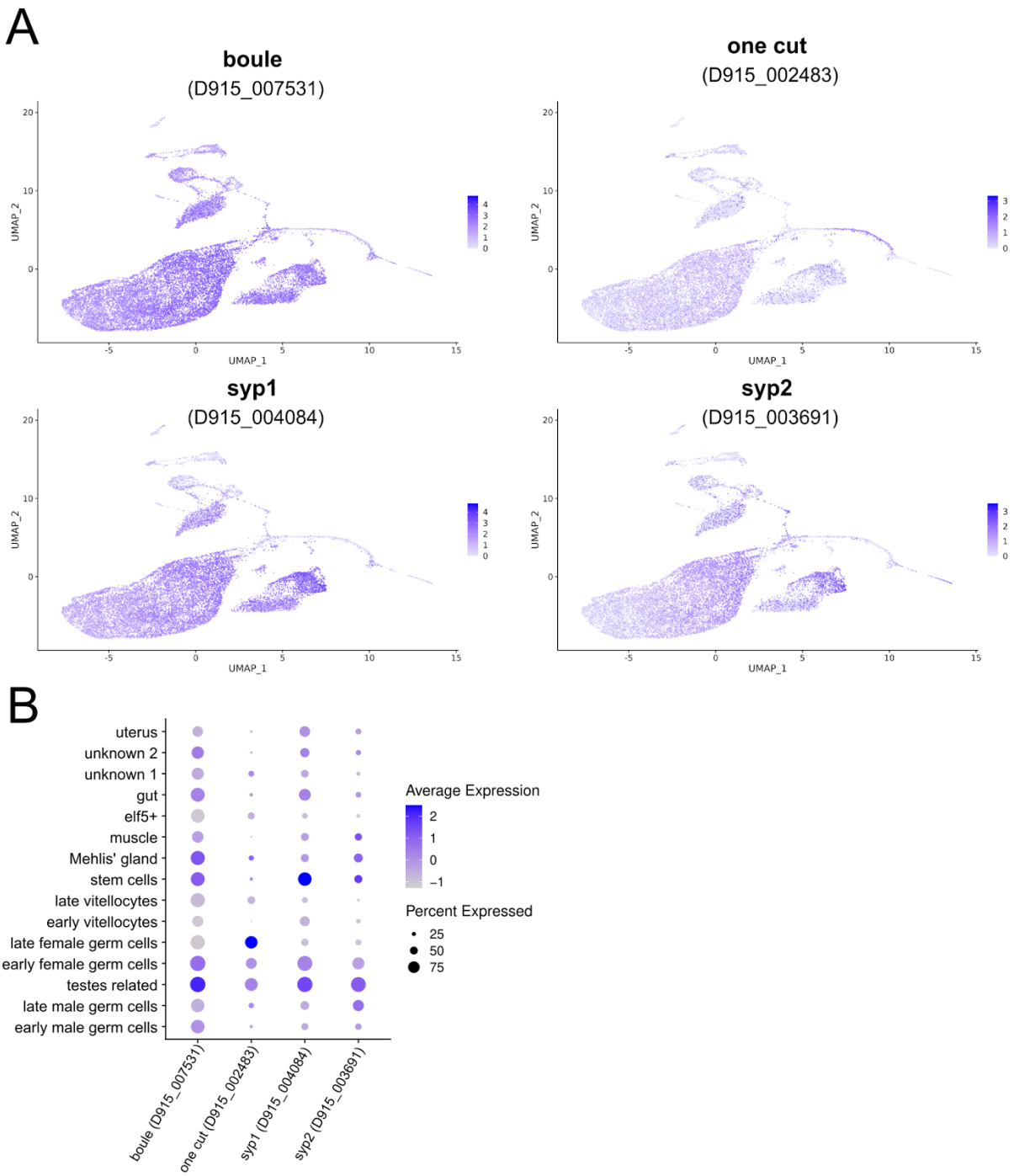


Figure 3.20 Known markers for the germline are expressed in the early germ cell clusters. A. UMAP plots colored by gene expression show the expression of selected marker genes. **B.** Profiles of gene expression over all clusters are illustrated as dotplot. The expression level is indicated by color, from blue (high expression) to lavender (low expression). The percentage with which the cluster cells express the given gene is represented by the size of the circles

Using RNA CISH, *meiob* expression was validated in the testes (Fig 3.21 B). The late male germ cell cluster showed high expression of genes encoding structural components, including tektins and tubulins, which are known parts of the sperm flagella (Fig 3.19). The gene encoding *tektin4* (D915_008401) was validated to be strongly expressed in the testes of the worm (Fig 3.21 A). Based on these observations, the early germ cell marker was suspected to contain both spermatogonia and spermatocytes. In contrast, the male germ cell cluster probably consists of the more mature spermatids and sperm.

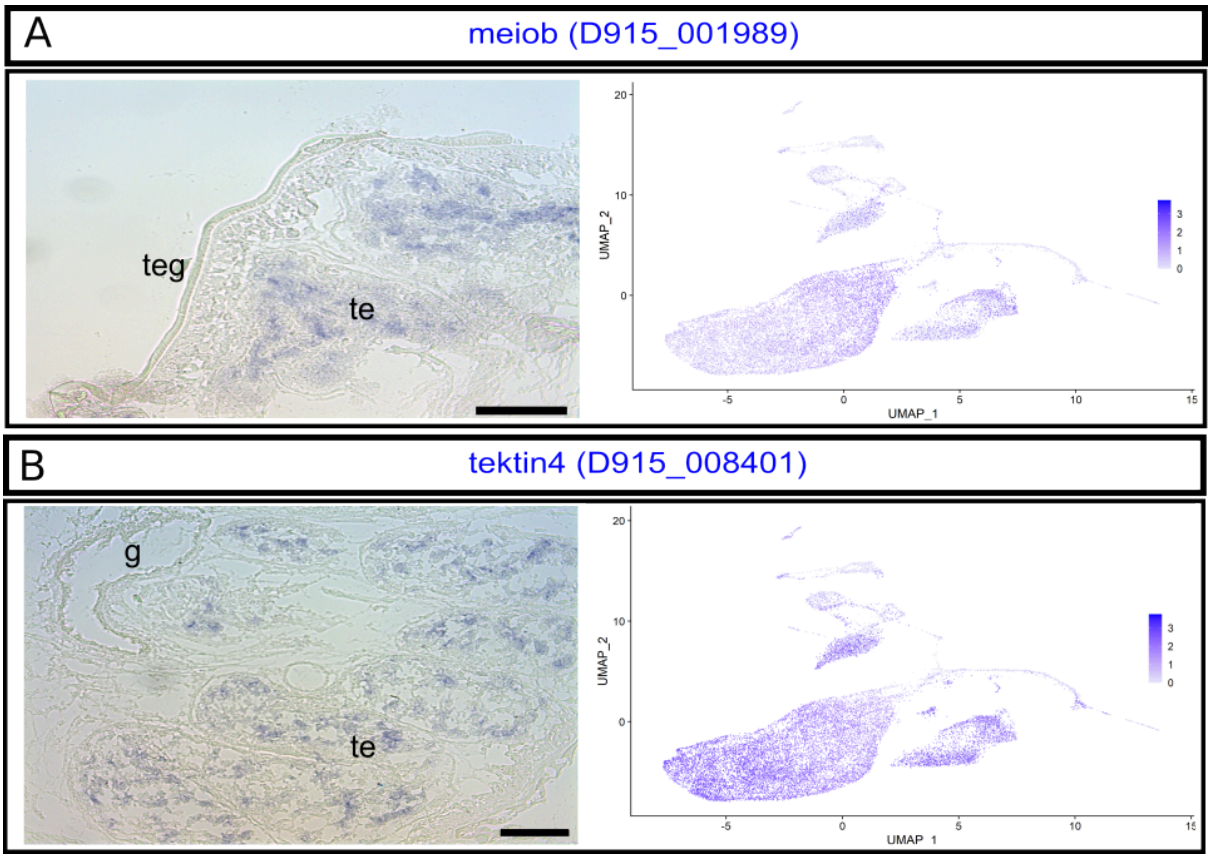


Figure 3.21 Predicted markers for the male germline can be validated *in situ*. A. ISH staining for transcripts of *tektin4* in blue and corresponding UMAP plots of gene expression. **B.** ISH staining of *meiob* transcripts, an early male germline marker in blue and corresponding UMAP plots of gene expression. Scale bar: 100 µm. Legend: gastrodermis (g), tegument (teg), testes (te). Sense controls were found to be negative (supplementary Fig S7).

Further evidence for this observation was gained from mapping the scRNA-seq data back on the spatial transcriptome of *F. hepatica* obtained by our lab [169]. This dataset utilized the 10x Genomics Visum workflow to interrogate the transcriptional architecture in cross sections of *F. hepatica* (Fig 3.22 A). By using a label transfer method, the cellular identity of spots in the spatial data can be predicted based on the single-cell data. Thereby a prediction score is computed, which is higher the more similar the gene expression of a spot is to a given cluster from the single-cell data. Using these predictions, it was evident that cells of the early male germ cell cluster and stem cells were located more on the periphery of the

testes, while the late male germ cell cluster was located more in the middle of the testicular lobe (Fig 3.22 C). This inward location of differentiating cells could also be visualized by immunolocalization using an anti-phosphorylated histone 3 (H3P) antibody. This modification is present during the transition from G2 to M phase and is used to label mitotically active cells [186,198]. Combining this staining with EdU showed a clear location of the mitotically active cells towards the inner layer of the reproductive organs (Fig 3.22 B).

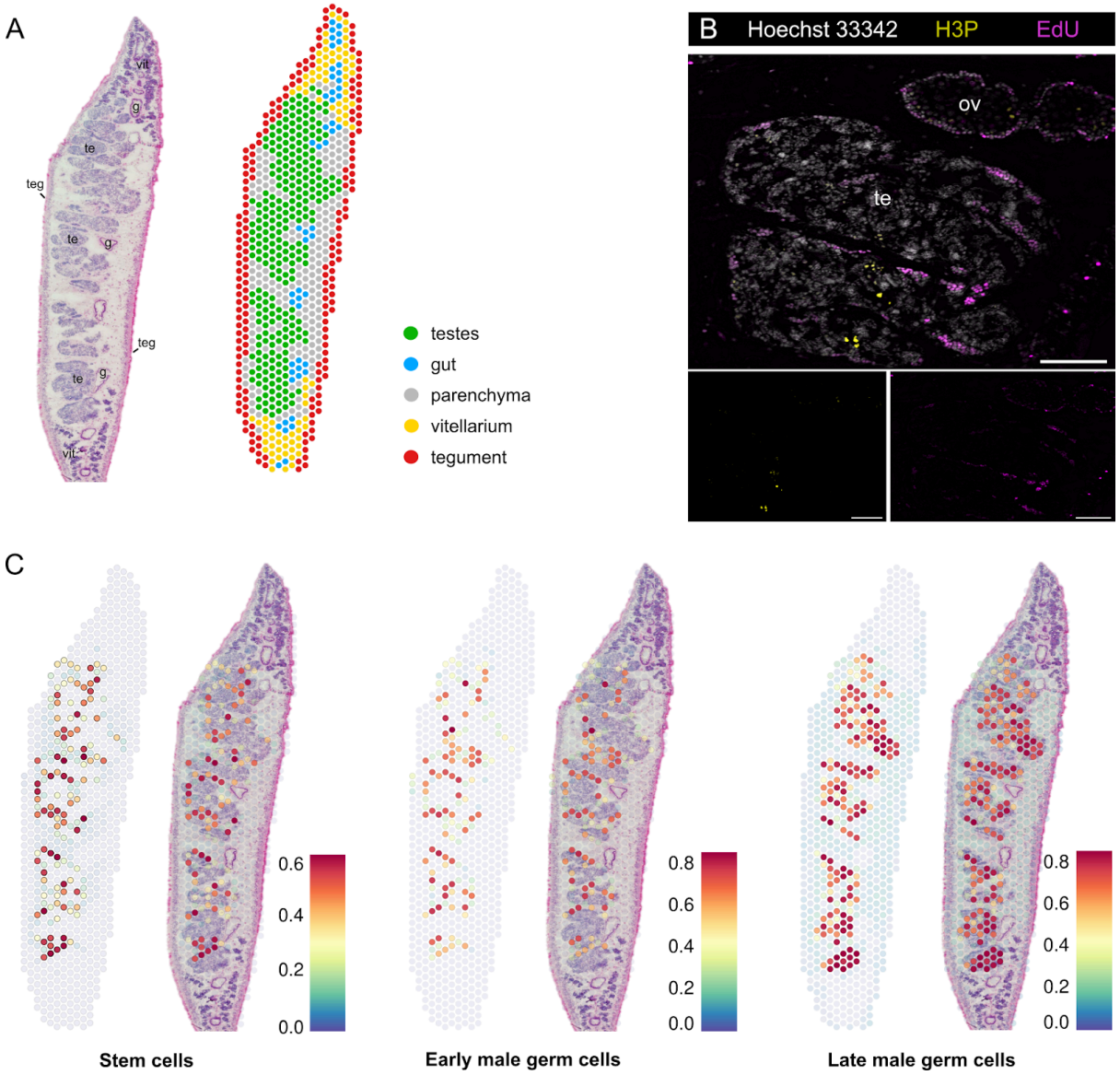


Figure 3.22 Male germline cells are located in spatially distinct locations. A. H&E stained section (left) and spatial projection of spots covering the tissue modified after [169]. **B** Confocal image of a combination of IHC for H3P (magenta), EdU staining (yellow) and nuclei stained with Hoechst (grey) on a transversal section. Scale bar: 100 µm. **C** Mapping of clusters onto the spatial transcriptome data generated by Gramberg et al. The color of the spots corresponds to the associated prediction scores for the clusters.

The GO term analysis of the early male germ cell cluster (Fig 3.23) shows enriched GO terms, including RNA binding and nucleotide binding, reminiscent of the stem cell cluster.

Meanwhile, the late male germ cell cluster had terms reflecting a differentiated cell state associated with the cytoskeleton, cilium assembly, and axoneme assembly [184].

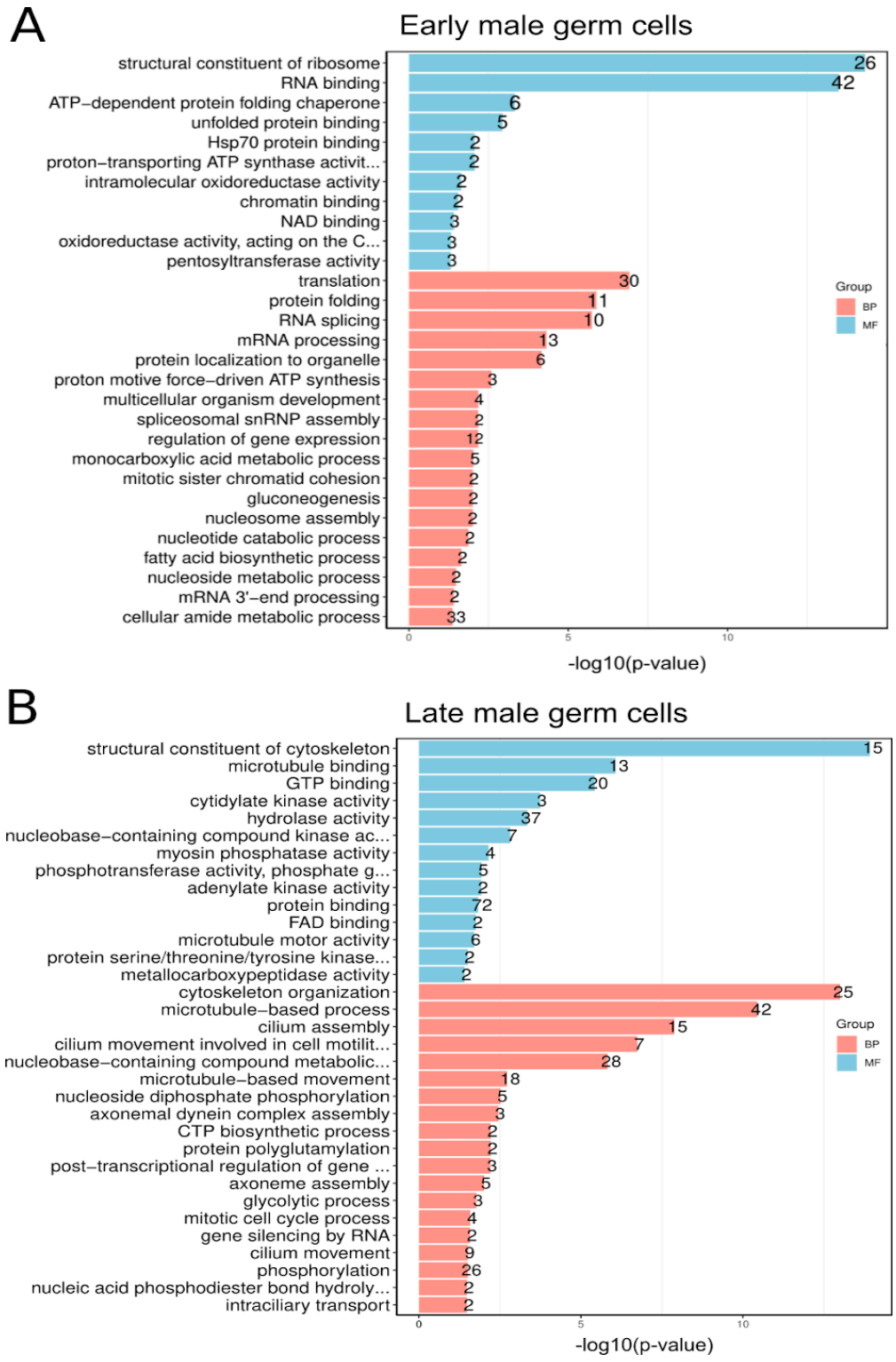


Figure 3.23 Male germline cells show GO terms reflecting their differentiation states. GO enrichment for biological processes (BP) and molecular function (MF) is shown. Marker genes of the early male germ cell cluster **(A)** or late male germ cell cluster **(B)** were used as an input. For each term the false discovery rate is plotted on the x-axis. A number of genes involved with each term is indicated.

The ovary of the liver fluke is a branched organ similar to the testes. Studies using microscopic methods identified three cell types: germline stem cells and cells labeled as oocytes 1 and 2 [193], which could be linked to the early and late female germcell clusters identified here. Initial identification of clusters associated with the ovary was achieved by detecting transcripts of the *bone marrow proteoglycan* or short *bmpg* (D915_005862). The transcripts of this gene were validated to be located in early and, to a lesser extent, late oocytes (Fig 3.19, 3.24), as was shown for schistosomes [111]. The noted expression in the uterus cluster could be derived from the oocytes within this cluster.

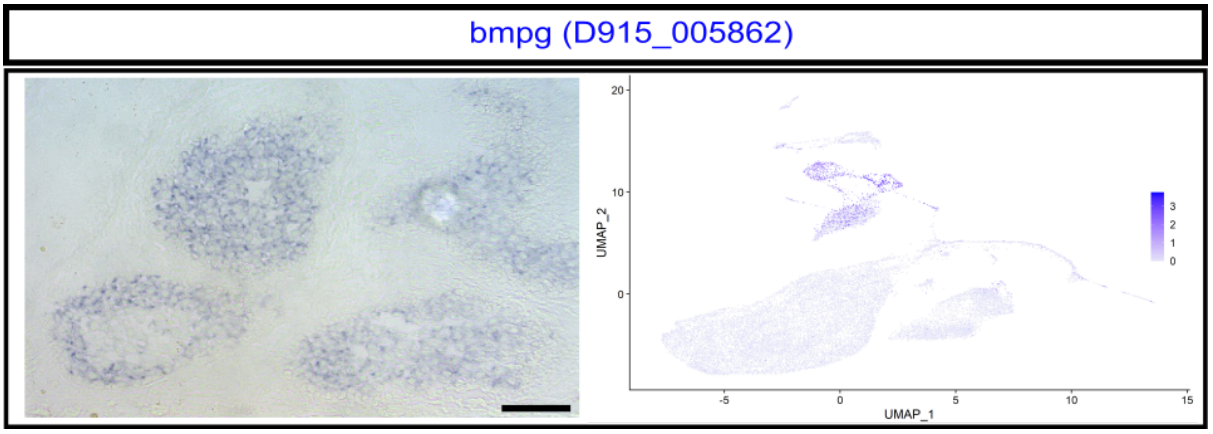


Figure 3.24 Validation of ovarian expression of *bmpg* by RNA *in situ* hybridization. A. ISH staining for transcripts of *bmpg* and corresponding UMAP plots of gene expression. Scale bar: 100 µm. Sense controls were found to be negative (supplementary Fig S7).

As for the male germ cells, the distinction between the periphery and more mature cells in the middle of the worm was envisaged by mapping the cells on the spatial data (Fig 3.25). Though this analysis failed to show clear spatial patterns for late female germ cells. Additionally prediction scores for the late female germ cell cluster were low (maximum of 0.02).

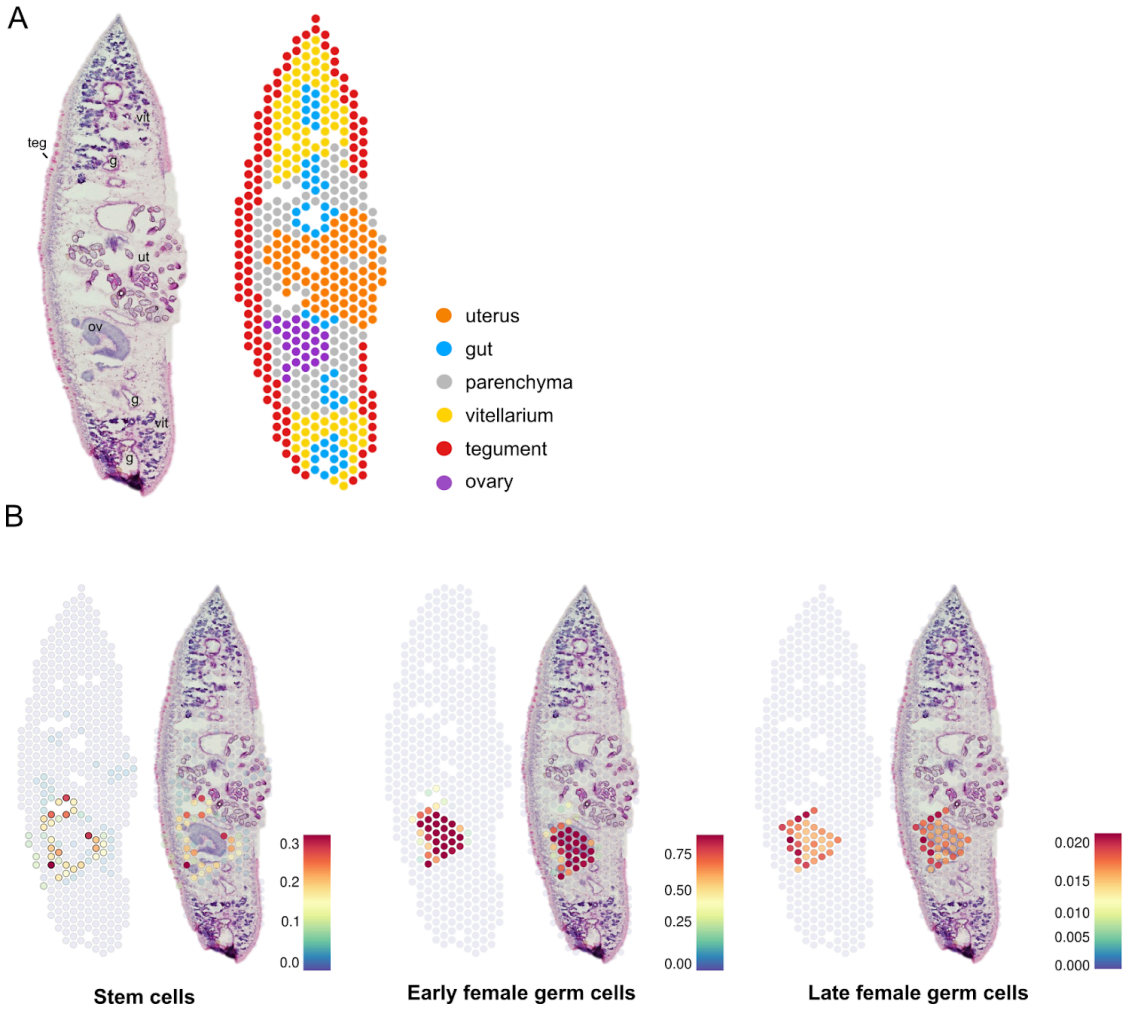


Figure 3.25 Cells of the female germline are correctly localized in the spatial transcriptome. A. H&E stained section (left) and spatial projection of spots covering the tissue modified after [169]. **B** Mapping of stem cell and early female germ cell clusters onto the spatial transcriptome data. The color of the spots corresponds to the associated prediction scores for the clusters.

A trend similar to the early male germ cells was observed for GO term analysis related to the female germ cell clusters (Fig 3.26). Early cells showed DNA binding and terms related to replication and metabolic needs. In contrast, the late female germ cell cluster was enriched for terms like signaling, phosphorylation, or nuclear receptor activity.

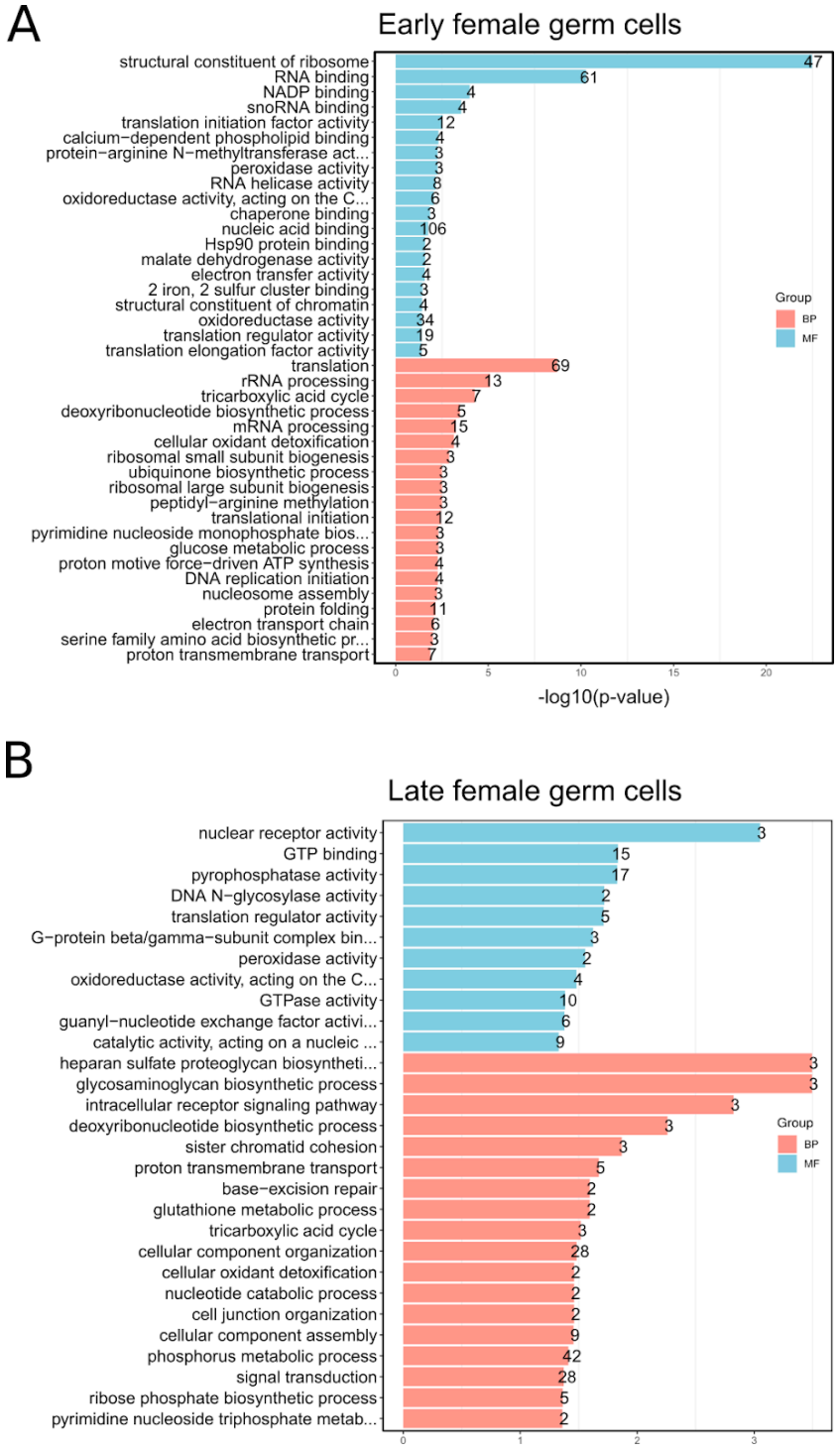


Figure 3.26 GO terms for the female germline show enrichment for signal transduction pathways. GO enrichment for biological processes (BP) and molecular function (MF). Marker genes of the early female germ cells cluster **(A)** or late female germ cells cluster **(B)** were used as an input. For each term the false discovery rate is plotted on the x-axis. Number of genes involved with each term are indicated.

Upon closer inspection of genes included in the GO term of signal transduction, several were involved in prominent signaling pathways. The study of signal transduction pathways, especially within reproductive biology, has a long history in schistosome research [199]. However, it was not yet in the focus for the liver fluke. Interestingly, core members of the

Wingless-related integration site (Wnt) signaling pathway were enriched in female germline cells. The Wnt pathway involves diverse processes like body axis patterning, cell fate specification, cell proliferation, and cell migration. It is essential for germline function in fruit flies [200]. Wnt members with enriched expression within the female germline clusters included the gene coding for the *adenomatous polyposis coli protein* (*apc*, D915_001684). Other genes involved with the Wnt pathway that were among the GO term of signal transduction were the *GSK3-beta interaction protein* (D915_010499) and *Ubiquitin thioesterase trabid* (*trabid*, D915_006722). To confirm the functional expression of the wnt pathway, all core members of the wnt pathway were identified in the *Fasciola* genome (Fig 3.27). This core group of proteins consists of a member of the frizzled receptor family, which acts through the protein Disheveled (dsh, D915_005287) upon binding its ligand, usually a wnt protein . This ligand binding usually requires a co-receptor, a role filled mainly by a Lipoprotein receptor-related protein (D915_010101). Upon activation, dsh disrupts the function of a destruction complex consisting of axin, the already mentioned APC, and the Glycogen synthase kinase 3 (D915_000845). Usually, this destruction complex degrades ß-catenin by recruiting Trabid, which marks it for the proteasome. Finally, ß-catenin acts as a transcriptional coregulator and regulates gene expression [201]. All the pathway members could be identified (3.27, supplementary table S2) by blasting amino acid sequences of known wnt members from model organisms against the liver flukes genome.

The identification of the related frizzled receptor was complicated by its genomic annotation. A blast search for previously identified receptors within this family [202] identified a gene annotated as a secreted protein (D915_010421) most similar to the query. The coding sequence of this gene appeared to be short, with an unusually long 5'UTR. To validate the open reading frame suggested, an *in silico* translation of the unsliced transcript of D915_010421 was performed. This suggested an open reading frame coding for 425 amino acids, compared to the initially expected 92. The new amino acid sequence contained a frizzled type domain typical for this receptor family. As all the components of the canonical members of wnt pathway components had notable expression within either one or both of the female germline clusters, this pathway is expected to play an essential role in oocyte maturation.

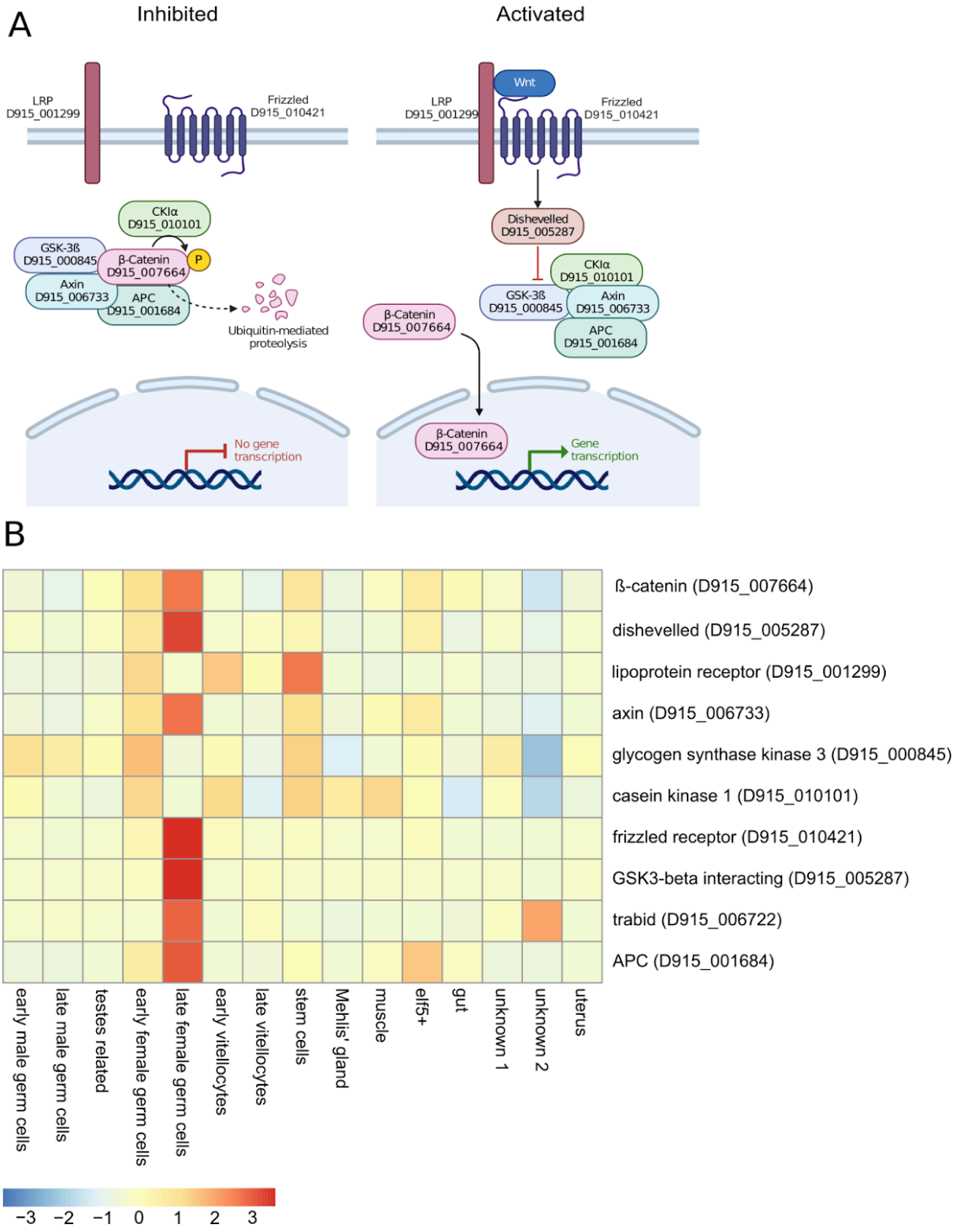


Figure 3.27 Important components of the canonical Wnt signaling pathway are present in oocytes. A. Schematic showing the canonical Wnt signaling pathway in its inhibited (left) and activated (right) state. Gene IDs of identified liver fluke orthologs are added. Figure was prepared using Biorender **B.** Heatmap of averaged expression values of these genes in the scRNA-seq data; the values are scaled by gene and centered.

Other transcripts associated with signaling pathway components were detected especially within the late female germline clusters, including a mitogen-activated kinase (Fhmp38, D915_005042), which was predicted to be the orthologue to the schistosomal mitogen-activated kinase 38 (Smp38). Smp38 was important for oviposition, and knockdown of this gene led to a decrease in ovary size [203]. Additionally, the Smp38 pathway was shown to be activated following exposure of immature females to male secretory products, additionally implicating its role in sexual maturation. Kinase pathways are complex and can be regulated by diverse stimuli. To inform the network in which Fhp38 might act in the oocytes, an interaction network was constructed using the GO term signal transduction from the previous GO term analysis. Here another kinase was shown to be part of the network, namely the Dual specificity mitogen-activated protein kinase kinase 4 (D915_010400) which in humans is known to activate p38 by phosphorylation [204]. The STRING network also showed a functional enrichment for genes associated with Rho, CDC42, and Rac signaling (Fig 3.28). These so-called G proteins are important signal-transmission molecules and interact with G protein-coupled receptors (GPCRs). Besides the already mentioned frizzled receptor, two more GPCRs were predicted to be markers for the oocyte clusters (D915_001255, D915_000428, Fig 3.29) but were not part of the Fhmp38 STRING network, as the lacked the necessary GO annotation. The elucidation of detailed mechanisms at play here lies beyond the scope of this study. Still, based on the evidence, it can be implicated that there might be a GPCR-mediated activation of the Fhp38 signaling pathway involved in oogenesis, as has been hinted at for schistosomes.

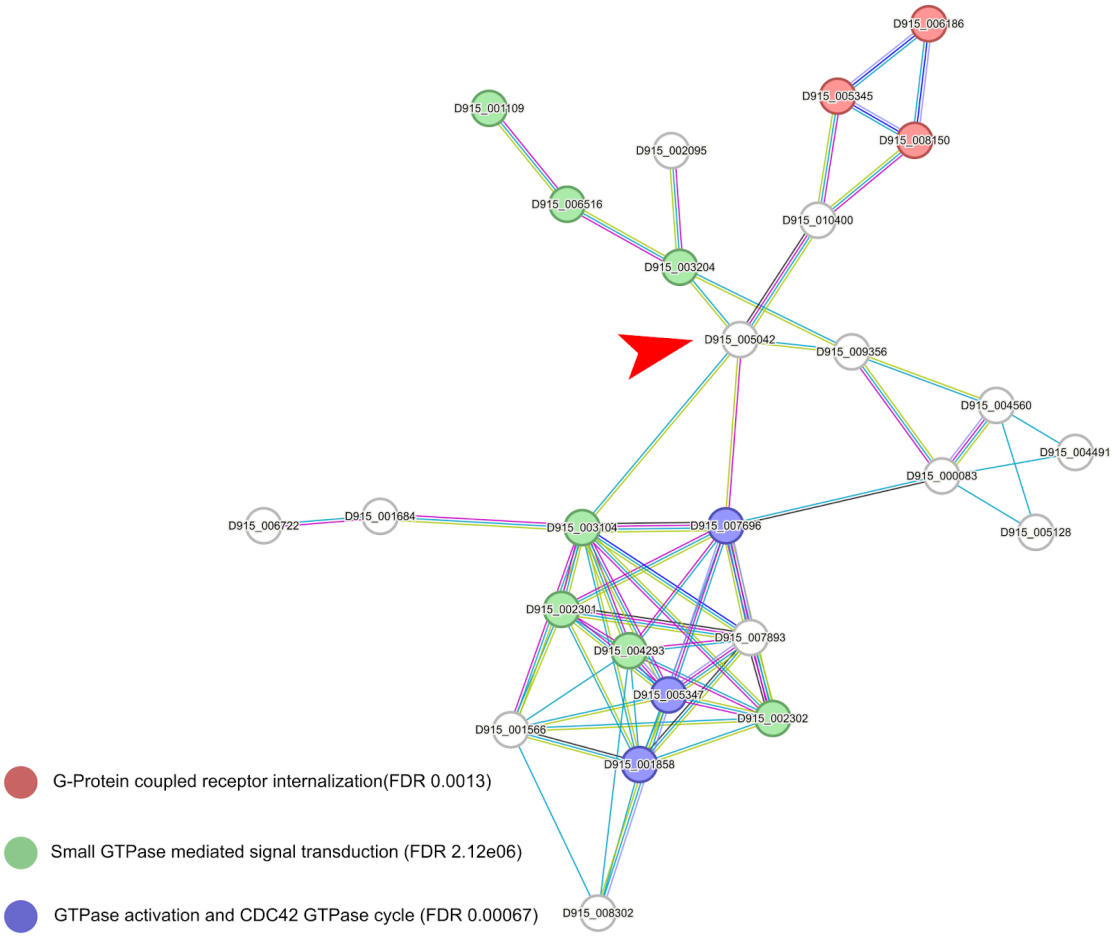


Figure 3.28 STRING analysis highlights signal transduction pathways enriched in the female germline. STRING analysis of marker genes associated with signal transduction within the female germ cell clusters. Connected nodes are based on evidence from homologs as predicted by STRING. Functional enrichment (*FDR*) is provided by STRING. A red arrow marks the p38 kinase.

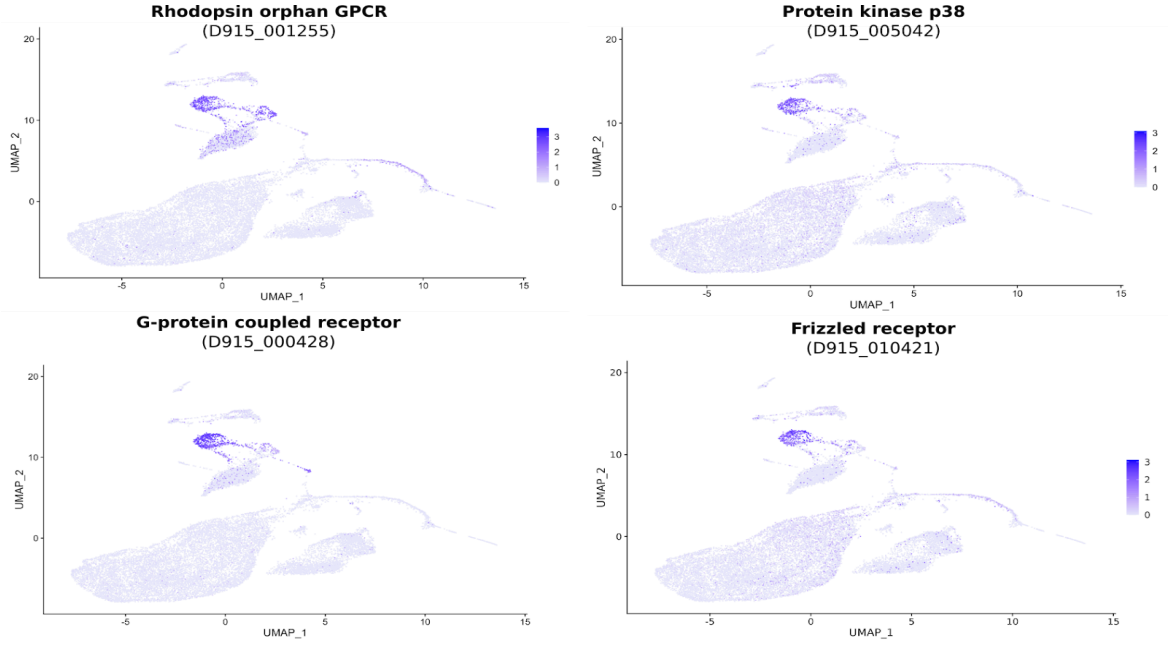


Figure 3.29 Potential components of G-protein signaling expressed in the female germline.
UMAP plots colored by gene expression show the expression of three selected GPCRs and *fhp38*.

As a final note on the extensive role of signaling in this cluster, there were several nuclear receptors (NRs) with preferential expression in the oocyte cluster. This family of receptors covers transcription factors that regulate essential biological processes. Structurally, they are divided into a highly variable region, a DNA-binding domain, a hinge region, and a ligand-binding domain. A recent study [205] identified and classified the NRs of 33 Platyhelminthes species, including *F. hepatica,* based on their DNA-binding domain. Based on this data, the NRs present within the markers in the oocyte clusters could be classified as thyroid receptor alpha (D915_005878), previously annotated as vitamin D3 receptor; thyroid hormone receptor beta (D915_003529) and the FTZF-1 alpha NR (D915_002517). While the *thyroid hormone receptor alpha* and *ftzf-1* were primarily transcribed within the female germline, the *thyroid hormone receptor beta* showed a more dispersed expression pattern (Fig 3.30).

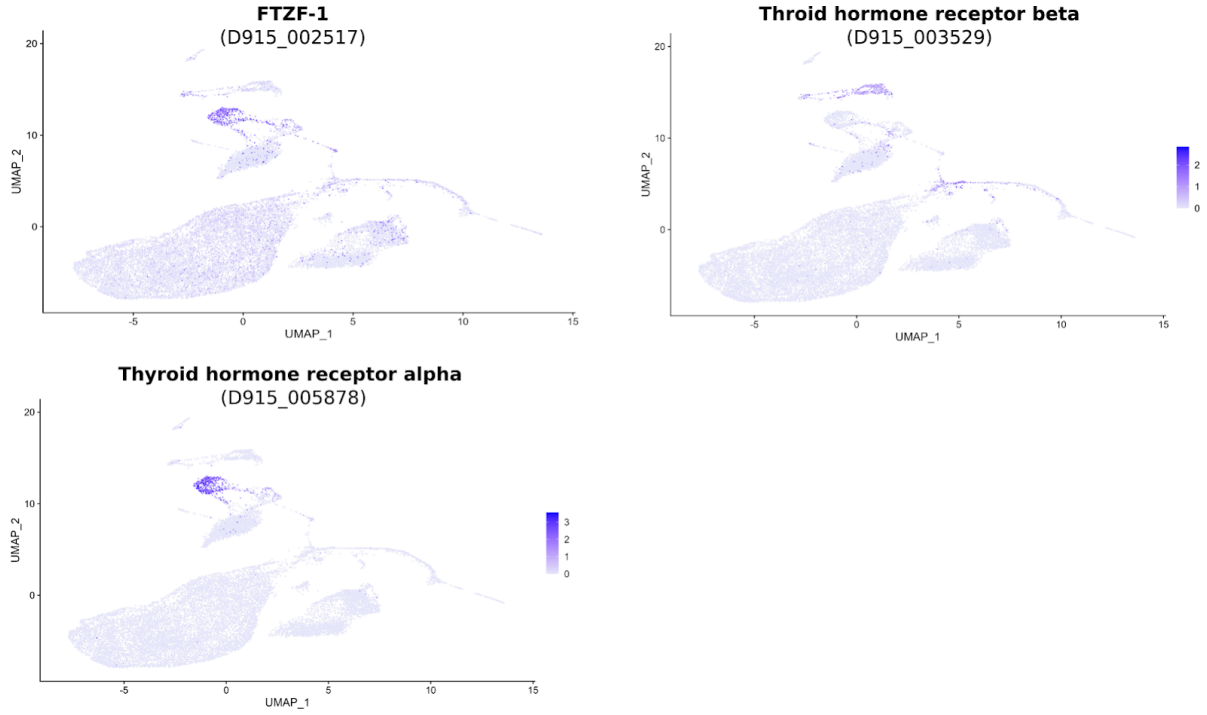


Figure 3.30 Selected nuclear receptors have marked expression in the female germline. UMAP plots colored by gene expression show the expression of selected marker genes.

3.3.3 Tracing conserved differentiation dynamics across a vitellocyte lineage

The vitellarium is unique to flatworms and is the source of vitellocytes, also called vitelline cells. These cells serve as the primary source of eggshell proteins and contain nutrients that later serve the developing embryo [181,182]. Several vitellocytes are next to the oocyte in the final egg, which demands a high output of vitelline cells to meet the required number for egg formation. Vitelline cells have been somewhat characterized in schistosome research, where cell markers have been identified for different differentiation stages. In this data, two distinct clusters could be identified, which were termed early and late vitellocytes due to the marked differential expression of tyrosinase (D915_002179) known to be a marker for early vitellocytes stage 2 to 3 (Fig 3.31) [206].

Visualizing the transcripts for both the mature vitellocyte marker *vitelline protein b1,* D915_010963 (*vb1*) and *tyrosinase* by CISH confirmed their presence in the vitellarium, though the distinction of cells proved difficult due to the naturally dark color of the vitellaria (Fig 3.32). The gene *vb1* is one of the most transcribed genes in adult worms [91], and the color reaction developed rapidly, explaining the darker stain.

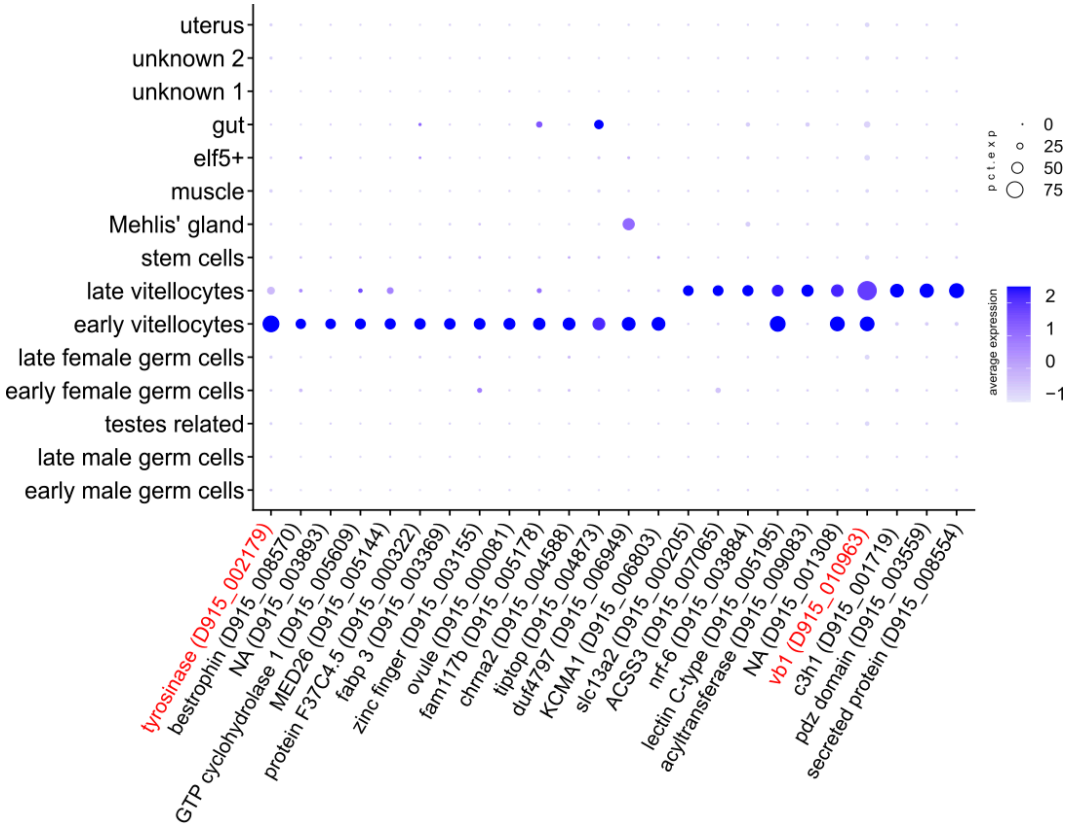


Figure 3.31 Two vitelline cell clusters had clearly defined marker genes. Dotplot shows marker gene expression over all clusters. The expression level is indicated by color, from blue (high expression) to lavender (low expression). The percentage with which the cluster cells express the given gene is represented by the size of the circles. Genes marked by red color were visualized by CISH.

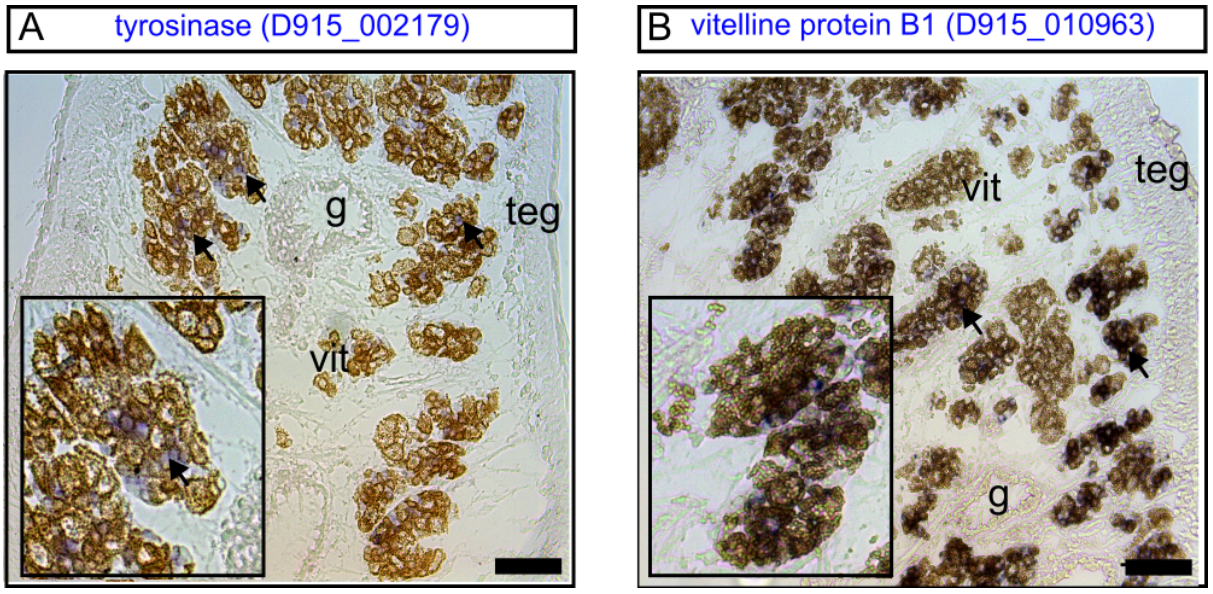


Figure 3.32 Transcripts of conserved marker genes localize to the vitelline follicles. Transversal sections were used for RNA *in situ* hybridization. ISH staining (blue) for *tyrosinase* transcripts **(A)** and *vb1* mRNA **(B)**. Scale bar: 100 µm. Legend: gastrodermis (g), tegument (teg), vitellarium (vit). Sense controls were found to be negative (supplementary Fig S7).

When mapping the clusters to the already described spatial dataset, as expected, the early vitellocytes cells were located exclusively in the spots associated with the vitellarium. The late vitellocytes cluster, besides being mapped to the vitelline tissue spots, also mapped to uterus spots, which might be related to the presence of mature vitellocytes within eggs (Fig 3.33).

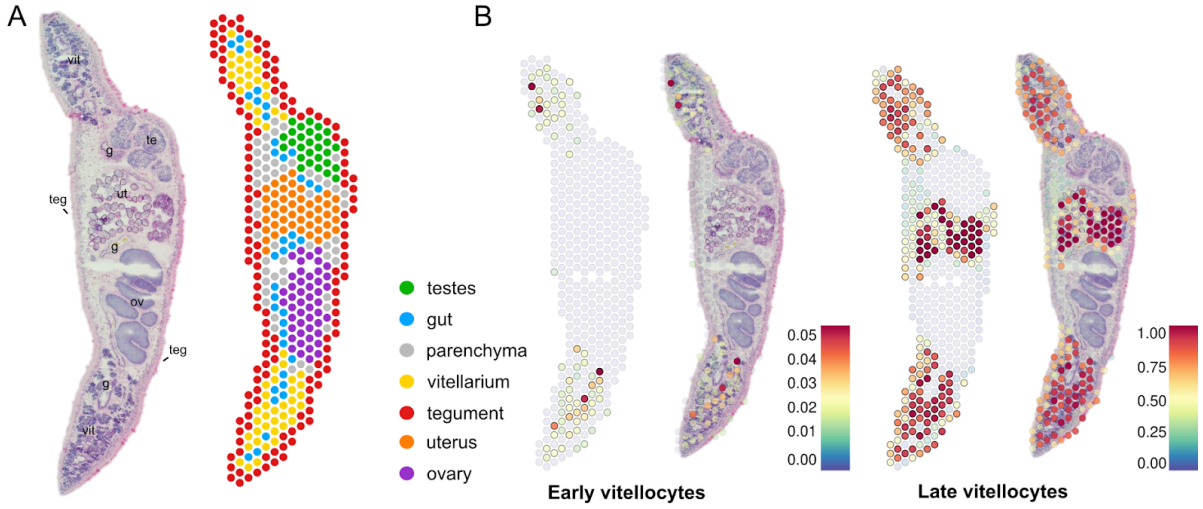


Figure 3.33 Vitellocyte cluster localization differs according to their differentiation state. A. H&E stained section (left) and spatial projection of spots covering the tissue modified after [169]. **B** Mapping of clusters onto the spatial transcriptome data. The color of the spots corresponds to the associated prediction scores for the clusters.

As described, the vitelline cell cluster was annotated in large parts by expression of conserved marker genes taken from the schistosome vitelline lineage [206]. A closer look at the genes characterizing the earlier stages of vitelline cells was taken to gain a better understanding of their differentiation. A nuclear hormone receptor termed vitellogenic factor 1 was shown to be expressed in early S1 cells, where it is a primary driver of differentiation [207]. The orthologue of said receptor in the liver fluke was identified annotated as nuclear hormone receptor HR96 (D915_001975) thereafter referred to as Vf1. As expected, it was expressed in a fraction of cells in the early vitellocytes cluster, partially coexpressing the proliferative marker *h2b* (Fig 3.34). Following the proposed lineage of cells, the expression patterns of *tyrosinase* and *vb1* that had already been described appeared to be partially overlapping.

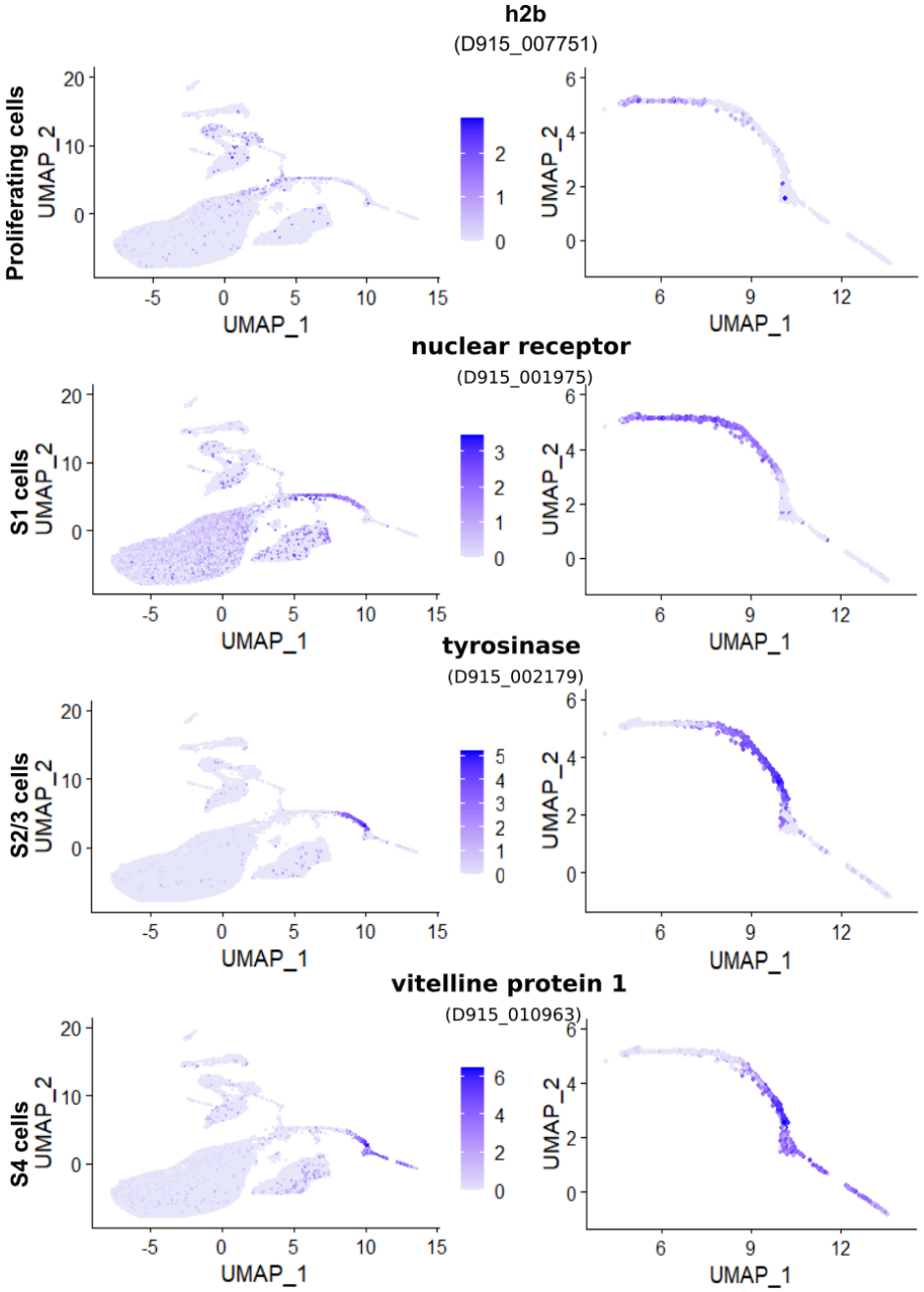


Figure 3.34 Transcription of selected genes recapitulates vitellocyte differentiation stages. UMAP feature plots showing the expression of conserved vitelline-cell marker genes shared with vitelline-cell markers of *S. mansoni*. An overview of all clusters and a close-up of the early and late vitellocytes clusters are shown.

To verify the proposed lineage and potentially identify driver genes of the differentiation in the vitelline lineage, RNA velocity was used on a subcluster set of the data [135]. The vitellocyte clusters were divided to a finer degree into five subclusters to predict driver genes in transition stages accurately. The analysis behind RNA velocity is based on slicing kinetics, these are computed based on the raw sequencing reads. The analysis divides reads from spliced and unspliced transcripts of genes. From this, latent time is inferred, meaning an approximation of the state of progression within the lineage. This latent time was the lowest

within cluster 0, which predicted these cells as the earliest point within the lineage. Rooting in cluster 0 the analysis implicated two possible directions of differentiation (Fig 3.35). One of the inferred directions ranged from cluster 0 to cluster 2, representing the most mature vitellocytes. The other lineage suggested a direction to the more stem cell-like cells, as indicated by their expression of *h2b* (Fig 3.35). The dualism within this lineage may suggest that one population of cells goes for differentiation while others represent self-renewing S1 cells. Looking at the velocity dynamics of single genes further validated the prior assumptions on the distinct expression of marker genes. The gene *vf1* showed high expression in the assumed early S1 cells and a sharp reduction represented by the negative velocity within the S2/3 cells. Overlapping with this reduction of *vf1* velocity was a marked induction of expression of both tyrosinase coding genes D915_002718 and D915_007816. A interesting observation was, that for genes showing a high expression in late vitelline cells unsliced reads were detectable in earlier cell types. It needs to be noted that most of the characteristic eggshell genes of late vitelline cells, including *vb1,* have no introns, making them inapplicable for this velocity analysis.

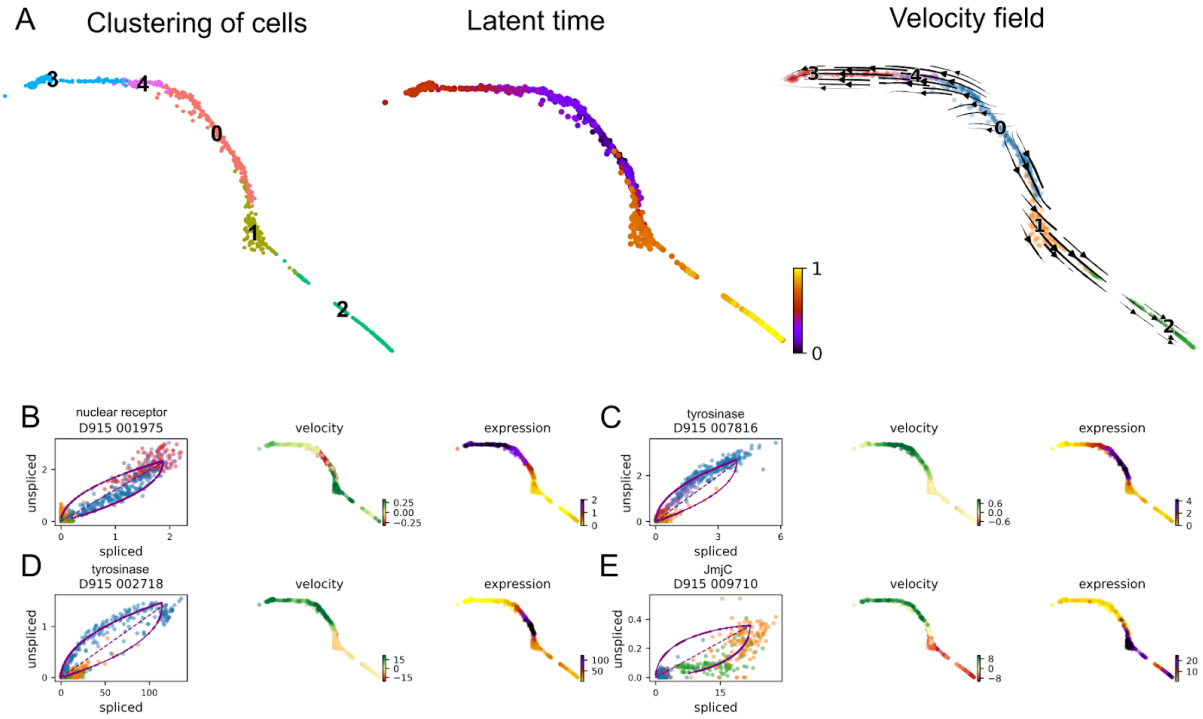


Figure 3.35 RNA velocity reveals gene dynamics over the vitelline lineage. A. UMAP showing the subclustering based on Seurat analysis (left) the temporal relationship between cells as estimated by latent time (middle) and the flow field indicating the direction of differentiation (right). Phase plots, as well as plots showing velocity and expression, were prepared for the nuclear receptor **(B),** two tyrosinases **(C** and **D),** and JmjC **(E)**.

The S1 vitellocytes are known to be capable of self-renewal, and there might be the possibility of two directions a cell might take, one being differentiation and the other maintaining proliferative capacity as a new S1 cell. In order to test this hypothesis the cells were scored based on their cell cycle state. To this end the gene expression of cell cycle marker genes was averaged and a score for each cell cycle phase was computed. These scores were then used to classify the cell cycle phase of each cell. The clusters had significant differences in cell state (Fig 3.36). For cluster 3, 75% of cells were expected to be in S-phase, while 4 and 1 had a more mixed profile. The clusters 0 and 2 had the highest scores for the G1 phase.

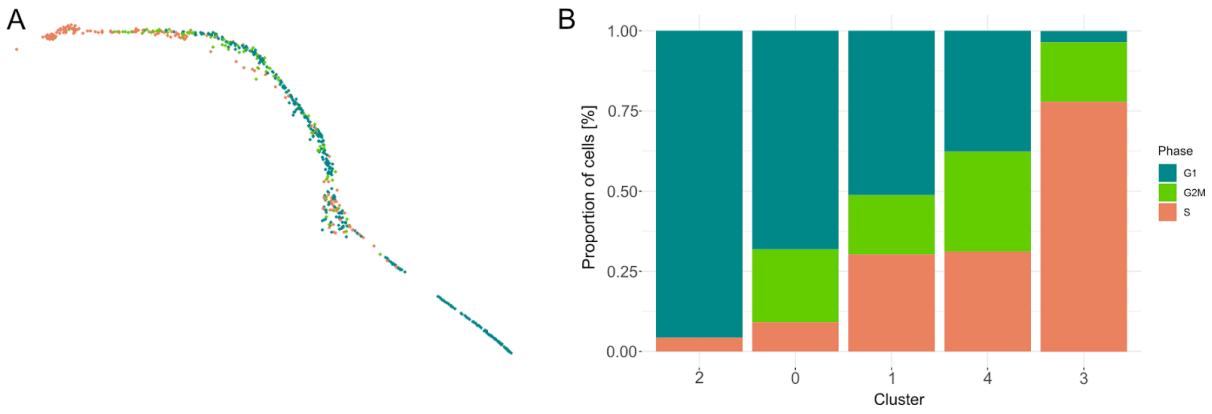


Figure 3.36 Subclusters within the vitelline lineage differed in cell cycle scores. A. UMAP plot using cell cycle scorings to color cells. **B.** Stacked bar plot displays the proportions of cell cycle stages per cluster.

3.3.4 Cells within the intestinal tract express genes involved in lipid metabolism

The liver fluke's gastrointestinal tract is a bifurcated tissue that branches through the whole body. These cells release numerous degrading enzymes, mainly cathepsins, to digest the ingested host blood and absorb the nutrients [208,209]. Cells of the gut were detected and classified based on the expression of selected cathepsins and showed a distinct transcriptional profile compared to the other clusters (Fig 3.37). Additionally, the intestinal expression of one of the cathepsins (D915_011077) was verified by CISH (Fig 3.38). Besides the high enrichment of GO terms related to proteolytic activity, there was an additional enrichment for genes in lipid binding.

Further investigating genes related to lipid metabolism, a gene coding for a phospholipase B-like protein (D915_003832) was explicitly expressed in the gut cells in the data set and *in situ* (Fig 3.37). Four other genes related to lipid metabolism were identified to be highly enriched within the gut cluster (D915_009347, D915_009349, D915_010479, D915_000724), all harboring an MD-2-related lipid recognition domain (ML domain) (Fig 3.37). These domains are implicated in functioning in human lipid transport [210]. Notably, all

these genes related to lipid metabolism were expressed either exclusively or preferentially in the adult or immature lifestage (supplementary Fig S2, [91])

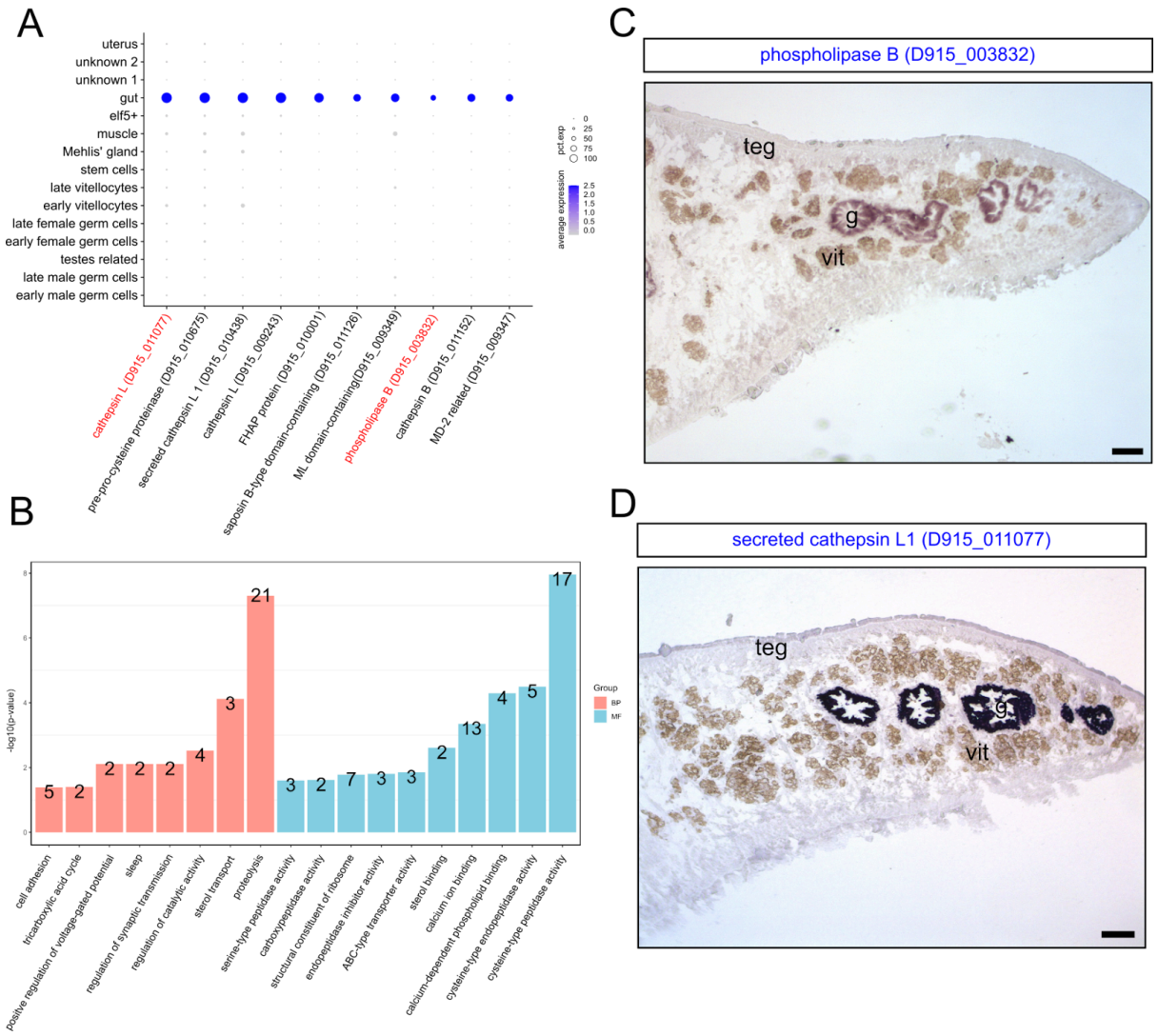


Figure 3.37 Identified gut cluster had distinct marker profile . A. Profiles of gene expression over all clusters illustrated as dotplot. The expression of predicted marker genes ofthe gut cluster over all clusters is shown. The expression level is indicated by color, from blue (high expression) to lavender (low expression). The percentage with which the cluster cells express the given gene is represented by the size of the circles. **B.** GO enrichment for biological processes (BP) and molecular function (MF) in the marker genes of the gut cluster. The number of genes involved with each process is indicated. **C**. Transversal sections of adult *F. hepatica* stained for transcripts of *phospholipase B*. **D.** Transversal sections of adult *F. hepatica* stained for transcripts of *cathepsin L1.* Scale bar: 100 µm. Legend: gastrodermis (g), tegument (teg), vitellarium (vit). Sense controls were found to be negative (supplementary Fig S7).

3.3.5 Transcriptional profiles of muscle cells are marked by protease inhibitors and signaling proteins

The muscular system of the liver fluke is essential for its survival; mirroring this, all major tissue types of the worm are lined with muscle, including the sucker, gut as well reproductive organs. The muscle of the parasite is of the invertebrate smooth muscle type, where the myocytes are connected with the muscle fibers via cytoplasmic connections [211,212]. Due to this, the expected distribution of myocytes might be diffuse and difficult to confirm by RNA *in situ* hybridization alone. To trace back potential muscle marker expression from scRNA-seq data to its associated fibers *in situ,* a muscle-specific antibody from the planaria *S. mediterranea [213]* was validated for immunohistochemistry. The antibody G610 stained the muscular system over the whole fluke's body (Fig 3.38 A). Muscle fibers were stained in proximity to the gut (Fig 3.38 E), the reproductive organs (Fig 3.38 C, D), and body musculature beneath the tegument (Fig 3.38 F).

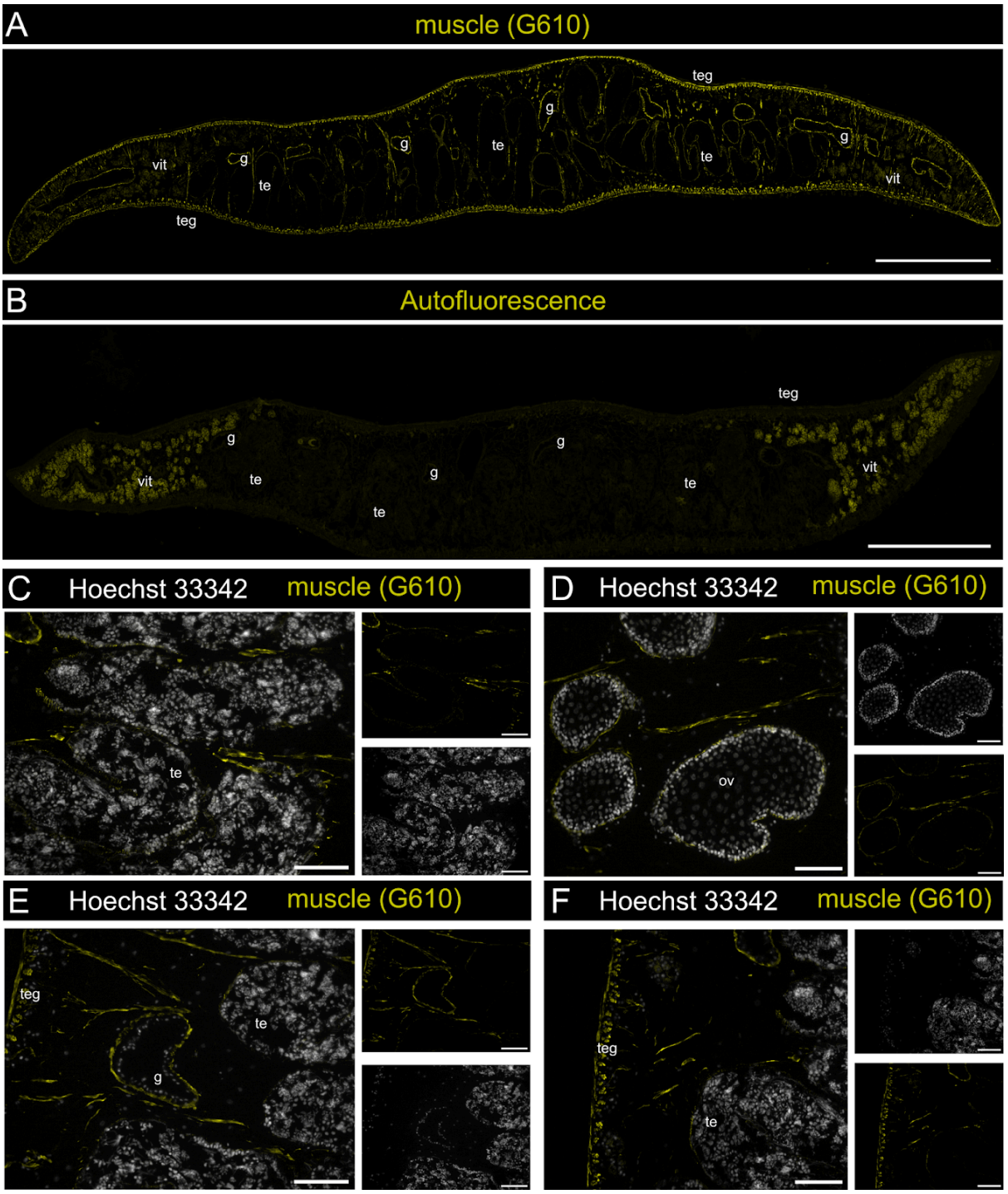


Figure 3.38 Staining using the antibody G610 marks all major musculature over the fluke's body. Staining was performed on transversal sections. **A.** Overview of the complete section showing staining of muscle fibers. **B.** Unstained control section showing autofluorescence in the vitellarium. Detailed view of staining in the testes **(C)**, ovary **(D),** intestine **(E),** and close to the tegument **(F)**. Scale bar: A-B 500 µm; C-F 100 µm. Legend: gastrodermis (g), tegument (teg), testes (te), ovary (ov), vitellarium (vit).

The muscle cluster expressed *collagen* (D915_008507) and *myosin* (D915_003945), which are known markers for musculature in related species [107,111]. These two markers showed a wide distribution of positive cells within the fluke's body, and these positive cells were associated with muscle fibers (Fig 3.39).

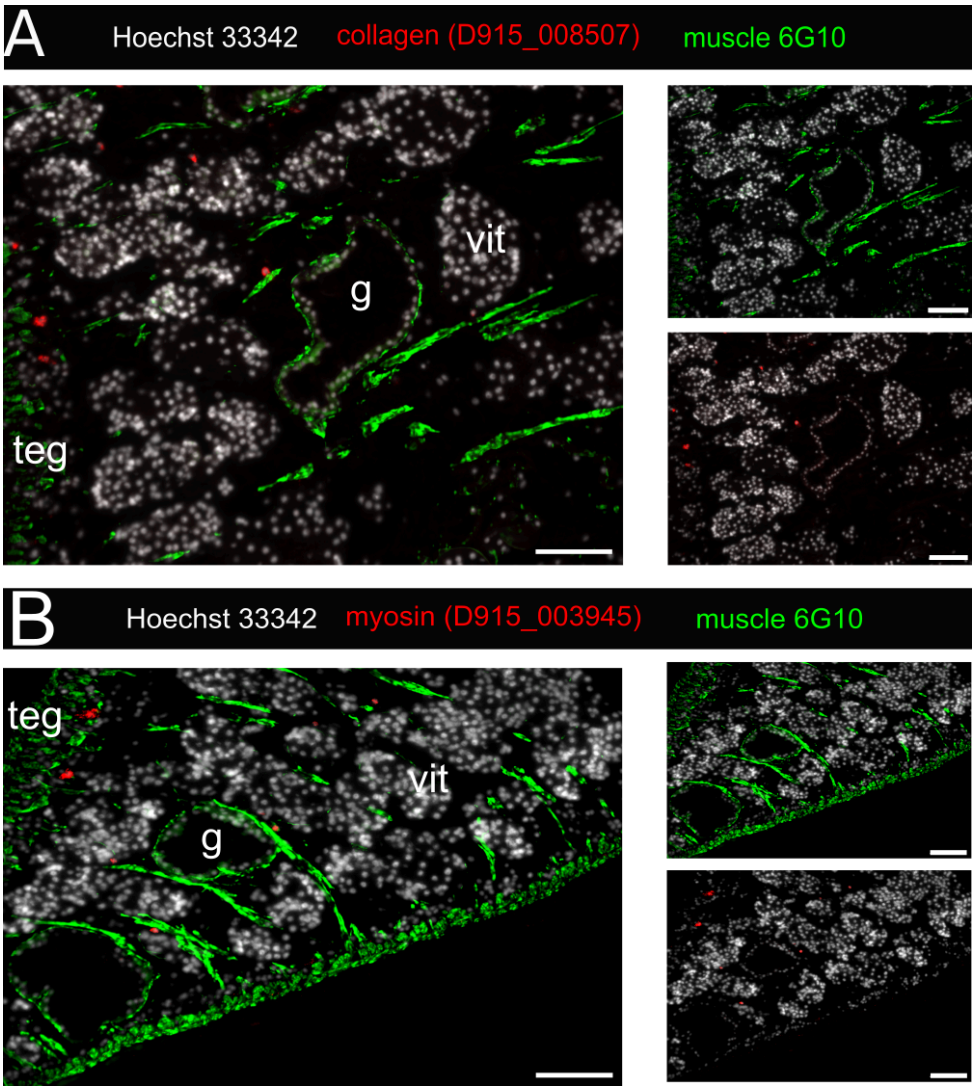


Figure 3.39 Marker genes for the muscle cluster associated with muscle fibers. Detailed view of FISH for collagen **(A)** and myosin **(B)** combined with immunolocalization of muscle fiber proteins (green). Scale bar: 100 µm. Legend: gastrodermis (g), tegument (teg), vitellarium (vit). Sense controls were found to be negative supplementary Fig S8.

Surprisingly, the data also showed that *stefin-2* (D915_009335) transcripts were highly abundant within the muscle cluster. Previous studies validated the presence of these types of proteins in gastrodermal cells and in the tegumental area [214]. Further looking into these genes, *stefin-1* (D915_009335) showed high transcription in the reproductive tract, while *stefin-2* and *stefin-3* (D915_001085) were present in various clusters, including muscle cells (Fig 3.40). Additionally, central regulators of signal transduction were found to be present as marker genes for the muscle cluster, among them proteinase kinase C (D915_006901) and G-protein coupled receptors (D915_002827, D915_003979), which are thought to be involved in FMRFamides related activation in the flukes muscle fibers [215] (Fig 3.41).

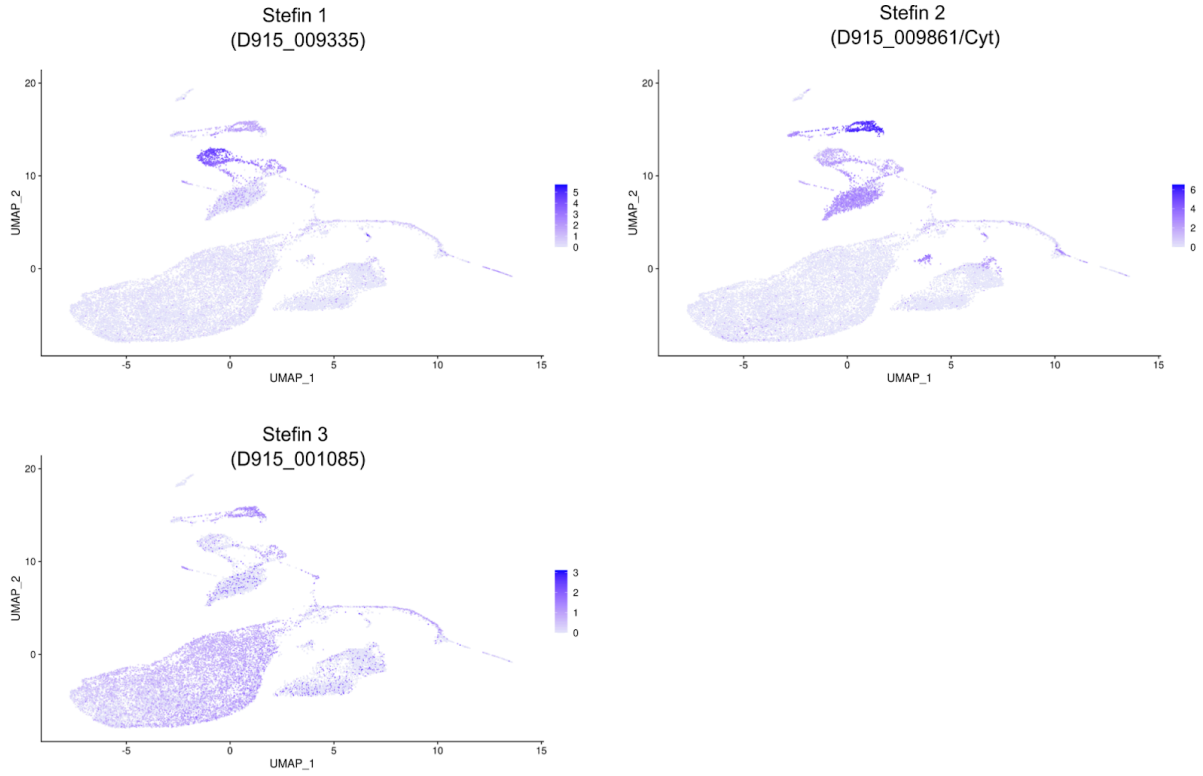


Figure 3.40 Protease inhibitors have unexpected expression profiles. UMAP feature plots showing the expression of three protease inhibitors.

3.3.6 The elf5+ cluster is associated with muscle function, protein kinase signaling, and cell adhesion

A cell cluster had a noticeable shared expression of marker genes with the muscle cluster (Fig 3.41) while having numerous distinct marker genes. The cluster was named elf5+ based on its expression of *elf5*, an ets-type transcription factor that has been shown to play a role in extracellular matrix composition in planarians [216].

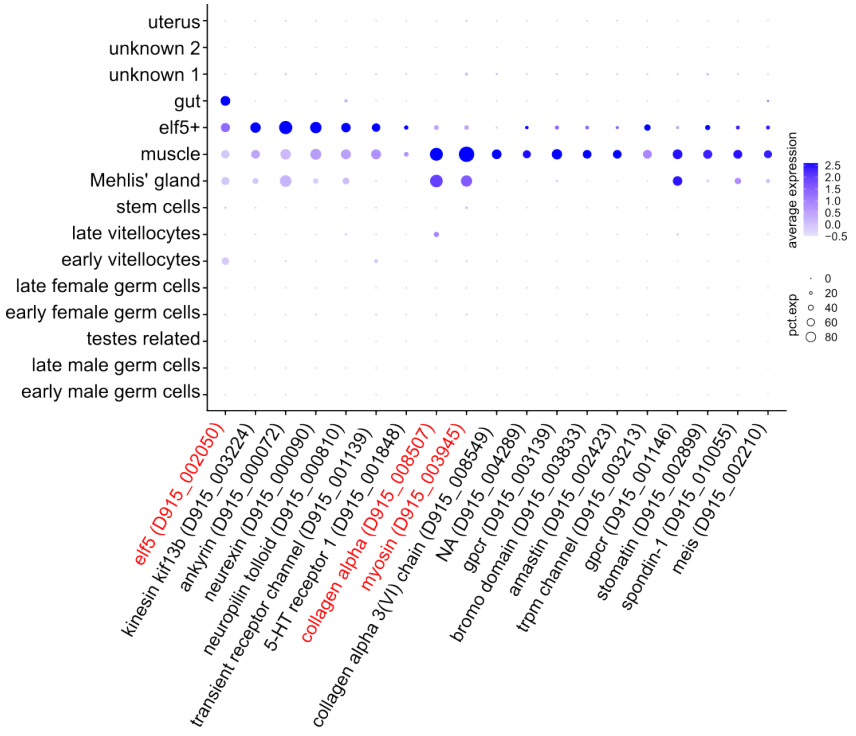


Figure 3.41 Marker genes of elf5+ cluster have some overlap with the muscle cluster. Profiles of gene expression over all clusters are illustrated as a dotplot. The expression of marker genes for muscle and elf5+ cluster is shown. The expression level is indicated by color, from blue (high expression) to lavender (low expression). The size of the circles represents the percentage with which the cells in a cluster express the given gene. Genes colored in red were used in ISH experiments.

A FISH based staining of *elf5* transcripts localized these cells within proximity to muscle fibers, additionally to expression within the intestine (Fig 3.42 A). As the marker genes for this cluster overlapped with the muscle cluster, and *elf5* transcripts were present in the muscle and gut cluster, a double FISH was performed to validate the annotation as a separate cell type. Combined FISH for transcripts of *collagen* and *elf5* together with staining of muscle fibers confirmed that *elf5*-positive cells did not stain positive for muscle transcripts but localized to muscle fibers, implicating a role in muscle function for this type of cells. Additionally, to gain more insight into the nature of these cells, *elf5* transcripts were co-stained together with an IHC for synapsin, which was used to stain neurons [213]. Still, no clear association with nerve cells could be seen for *elf5*-positive cells. However, *elf5*-positive

cells appeared close to the cerebral ganglion, with synapsin-negative structures potentially connecting to the structure.

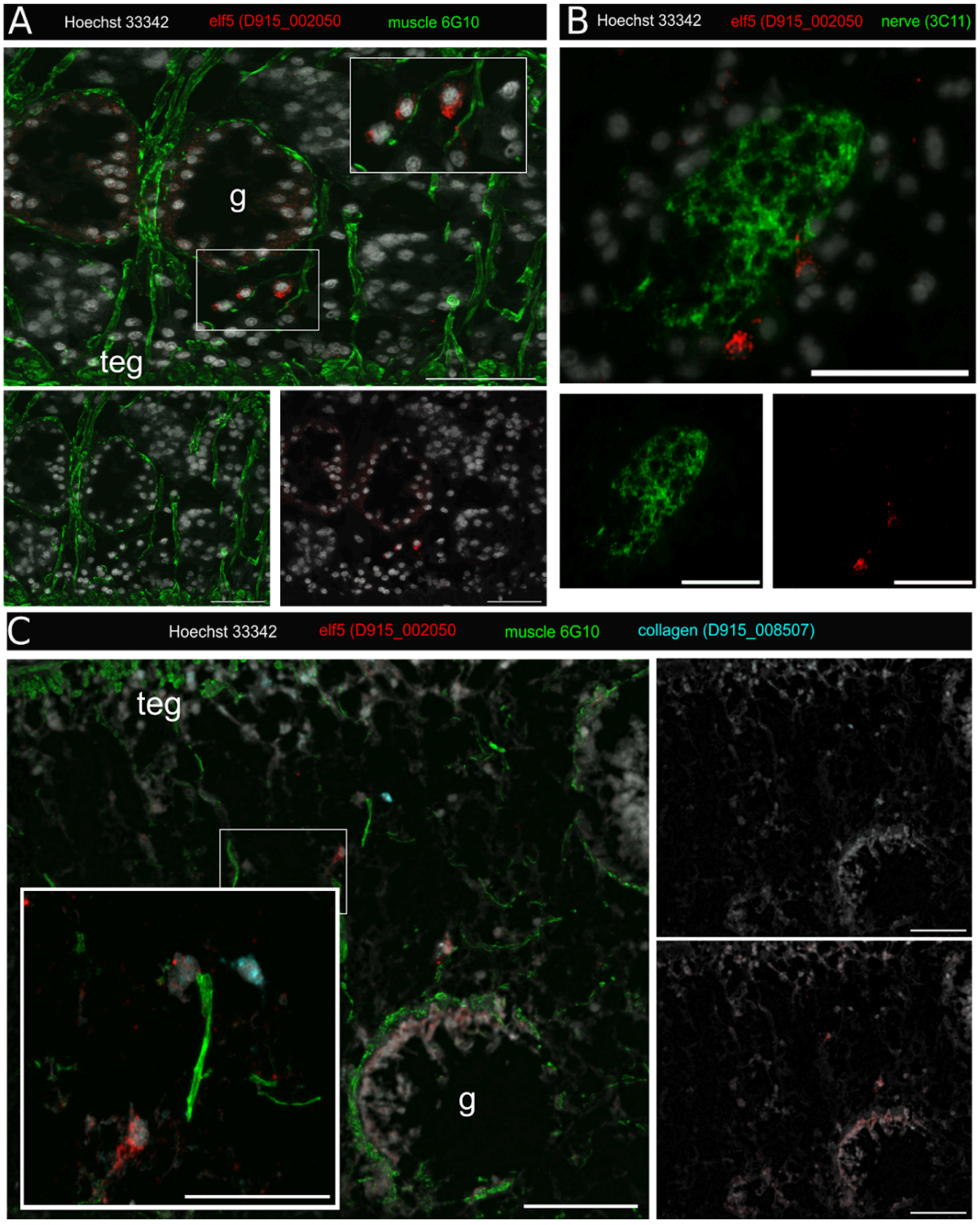


Figure 3.42 Cells positive for *elf5* transcripts did not stain positive as muscle or nerve cells. All images shown used transversal sections for staining. **A.** FISH for *elf5* transcripts combined with immunolocalization of muscle fiber proteins. Scale bar: 100 µm. **B.** FISH for *elf5* transcripts combined with immunolocalization of synapsin. Scale bar: 100 µm. **C** FISH for *elf5* and *collagen* transcripts combined with immunolocalization of muscle fiber proteins. Scale bar: 100 µm, insert scale bar 50 µm. Legend: gastrodermis (g), tegument (teg). Sense controls were found to be negative supplementary Fig S8.

One of the prominent genes among the specific marker genes for this cluster was a 5-HT receptor (D915_001848), which was implicated to play a role in serotonin-dependent activation of muscle function in flatworms [217,218]. Contrary to the muscle cluster with GO terms like transport, metabolism, and translation, the marker genes of th elf5+ cluster were enriched for cell signaling and signal transduction. Further analyzing the signal transduction terms brought up some interesting genes. These genes were, for example, the neuropilin and tolloid protein (D915_000810), which is known to be expressed in smooth muscle as well as neurons in humans [219] and has been proved to have functions at the neuromuscular junction in *Drosophila melanogaster* [220]. The binding partner's ankyrin (D915_000071, D915_002954) and spectrin beta (D915_003953) were preferentially expressed in the elf5+ cluster. These structural proteins are important in the linking of proteins such as membrane channels, receptors, and transporters to the cytoskeleton, thereby providing structural stability in response to muscle contraction [221–224]. In addition to these genes encoding ankyrin repeat domains among marker genes for this cluster, human orthologues such as ankyrin repeat domain 2 (ANKRD2) are involved in the stretch response of muscle cells.

Another prominent marker in the elf5 cluster was the gene encoding the transient receptor potential cation channel $TRPM_{PZQ}$ (D915_003213, Fig 3.41), with 30% of all elf5+ cluster cells showing transcripts for this gene. This calcium channel was shown to be a pan-trematode drug target, though its exact function is still unclear [76]. First results implicate its activation upon membrane stretch [225] in a ligand-dependent manner.

As a final step to further characterize the transcripts within the elf5+ cluster, a STRING analysis was performed using marker genes of this cluster (Fig 3.43). Enrichment for networks involved in focal adhesion and focal adhesion signaling were revealed in this analysis. Focal adhesions are adaptors between the plasma membrane and the cytoskeleton or anchor a cell to the extracellular matrix components. The network comprised genes encoding paxilin, talin as well as guanine nucleotide exchange factors and Rho GAP proteins (D915_006875, D915_001787, D915_001504, D915_000437, D915_000158). Interaction networks for these genes have been described before [226]. Centering the network was the focal adhesion kinase (D915_002353), which has a central function in focal adhesion, including turnover and response to changes in the extracellular matrix [227].

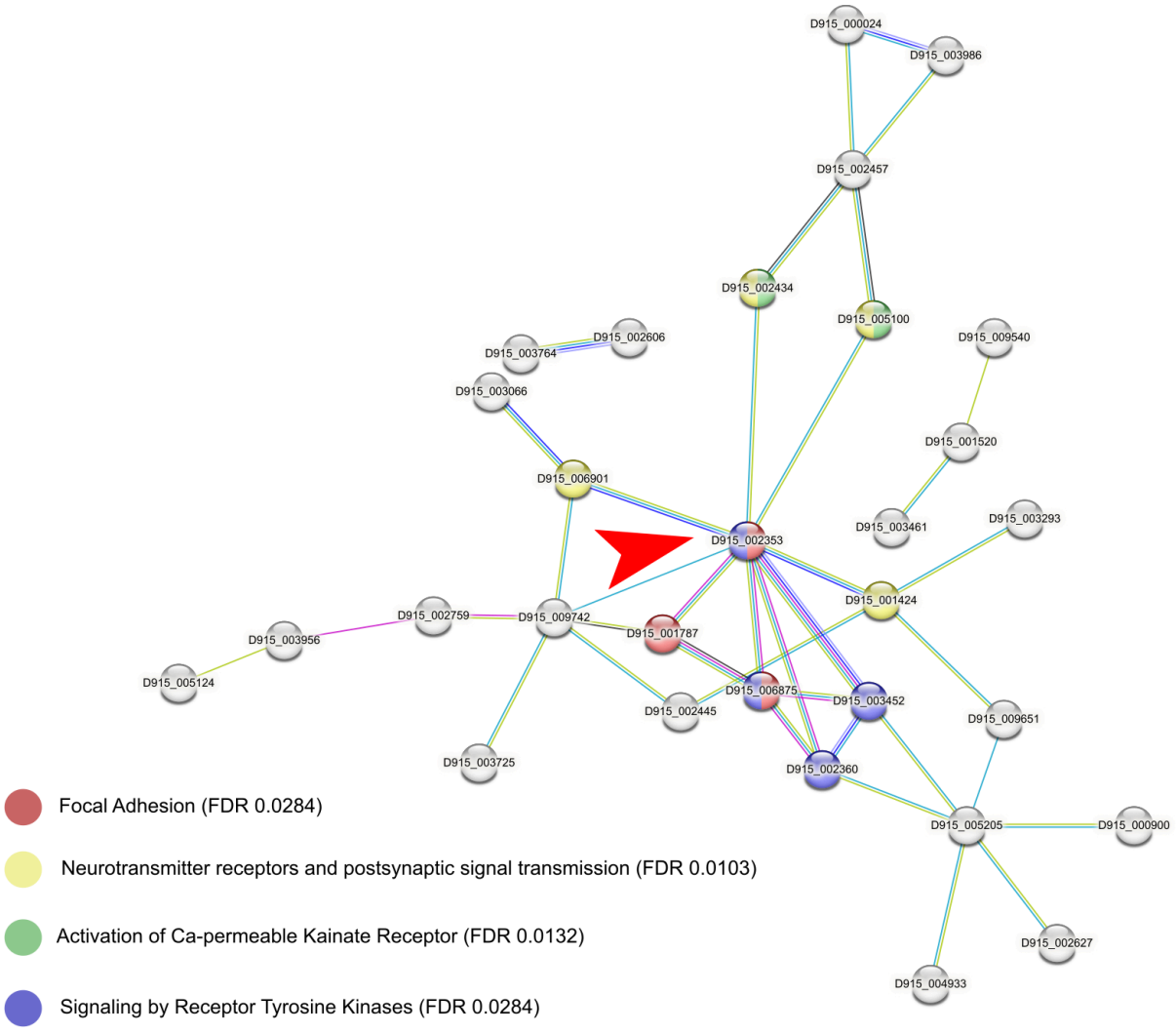


Figure 3.43 Identified markers show enrichment for focal adhesion network. STRING analysis of the top 100 marker genes for the elf5+ cluster connecting nodes are based on evidence of the function of homologs. Functional enrichment (*FDR*) is provided by STRING. A red arrow marks the focal adhesion kinase.

Besides the focal adhesion kinase, the elf5+ cluster expressed numerous other kinases as markers, indicating an important role of these cells in signal transduction and regulation of cellular processes. A total of 16 protein kinases (PKs) were identified in this cluster of cells, most of them at a highly enriched level compared to the other clusters (Fig 3.44). These PKs included a protein kinase C (PKC) D915_006901 and D915_003066 coding for protein kinase D. An overview of the proposed function for the found kinases can be found in the supplementary table S3. Most of those PKs are involved in extracellular signaling, like the mitogen-activated protein kinases (D915_006871, D915_003405) or protein kinase C/D. Additionally, several kinases were identified that are known to be sensitive to calcium or in other ways involved in calcium signaling, for example, the protein kinase fyna (D915_004177) or protein-tyrosine kinase 2 (D915_002353).

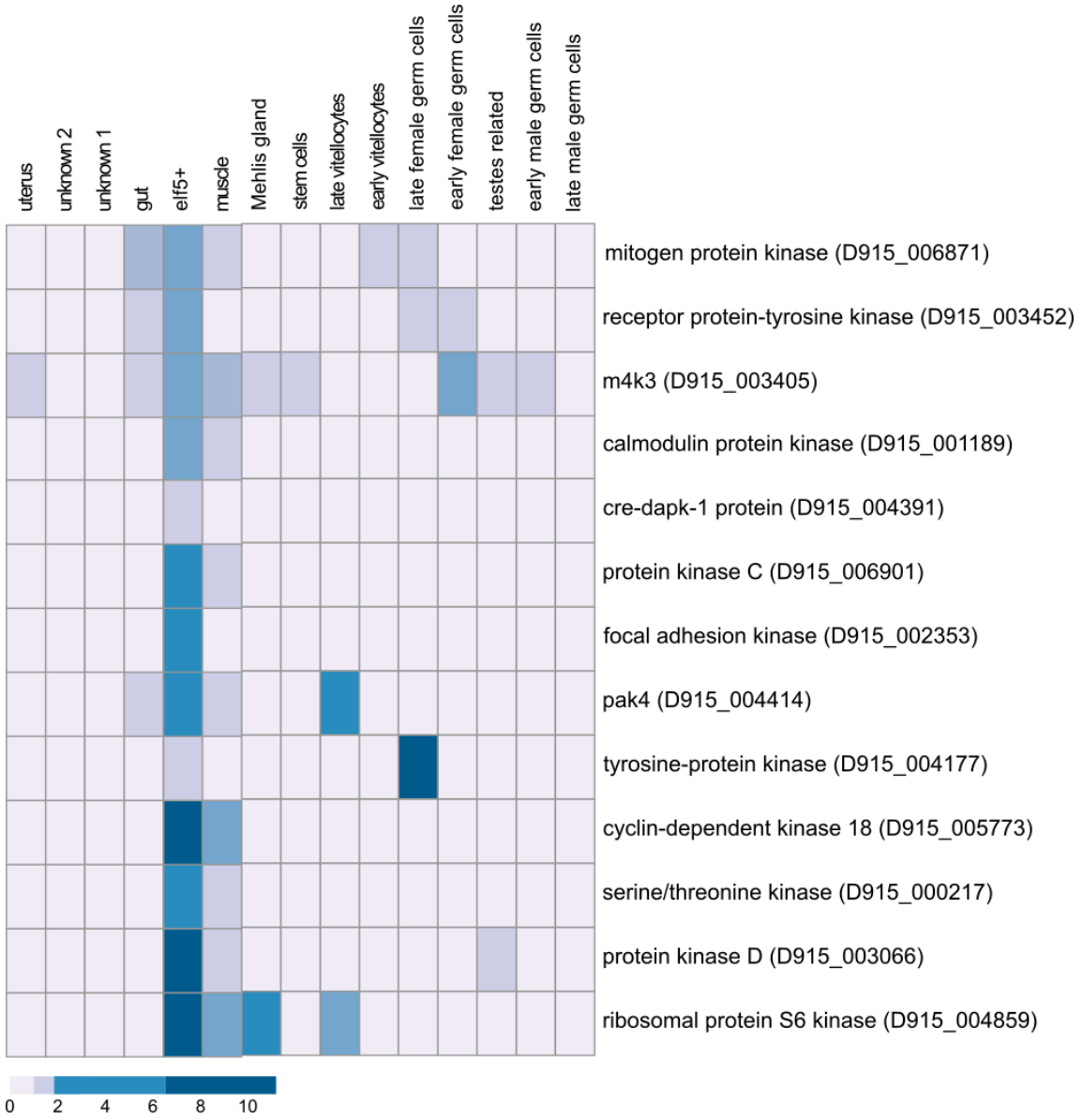


Figure 3.44 Various kinases show highly enriched expression in the elf5+ cluster. The average expression of protein kinase marker genes per cluster is displayed as a heat map. Expression values were individually centered and scaled for each row (each gene).

3.4 Improvement of the cell atlas by cell fixation

An improved way to enrich liver fluke cells was employed to improve the cell atlas project with a second refined scRNA-seq experiment. The experimental parts of this work were performed at the Stowers Institute for Medical Research in Kansas City. Hence, the cells had to be fixed and a fixation protocol was established prior.

3.4.1 Fixation and freezing for single cells

As the aim of a cell atlas using scRNA-seq is to profile the native transcriptional state of the cells, obtaining single cells has to be largely optimized and, if possible, should happen on the same day as library preparation. This process is time consuming and might benefit from fixation protocols that could decouple the sample preparation from the scRNA-seq experiment. An established fixation protocol would also be beneficial as parasite material is not readily available, and cells might need to be collected over several time points to generate enough samples. Another advantage of fixation is the possibility of transport. Several fixation protocols exist for other cell types but were not evaluated for the liver fluke. The fixation protocol presented in this work aimed to preserve cell morphology and protect RNA integrity. For the free-living relatives of the liver fluke, the planarian *S. mediterranea*, the already mentioned ACME dissociation protocol reported the simultaneous dissociation and fixation of cells using acetic methanol [158].

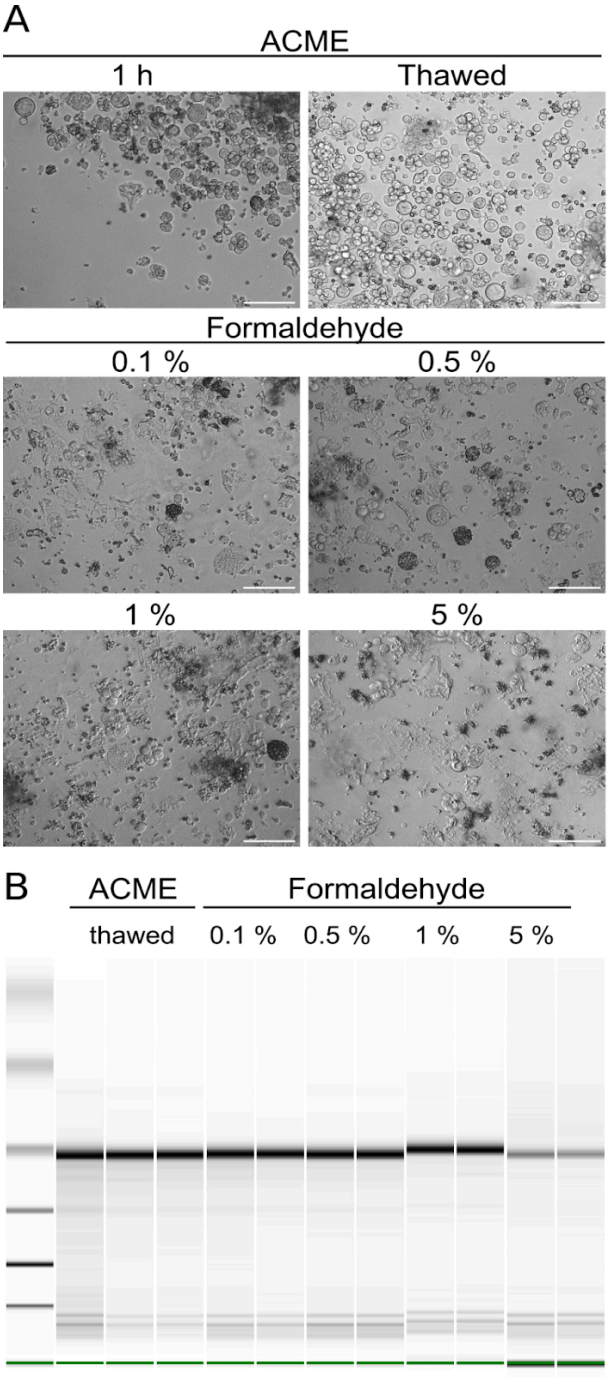


Figure 3.45 Fixation with ACME preserves cell morphology and RNA quality.
A. Brightfield images of cells fixed with ACME (top) after 1h or after freezing for 24h and varying formaldehyde concentrations (bottom). **B.** Bioanalyzer profiles for RNA of selected fixation protocols, showing the prominent 18S rRNA band.

While it could not be shown that ACME is beneficial for the dissociation of liver fluke tissue, ACME can also be used to fix cells obtained by enzymatic dissociation. To evaluate this fixation for liver fluke cells, they were fixed with ACME and frozen in PBS with 10% DMSO.
As a comparison, cells were fixed in varying concentrations of formaldehyde, which has been used in some scRNA-seq studies [228]. The cells fixed using ACME had a well preserved morphology comparable to viable cells (Fig 3.45 A) even after being frozen for one day. There was a notable reduction of lipid droplets from the cell suspension after fixation in ACME, which might be due to the washing out of lipids [229]. Fixation with formaldehyde caused the formation of aggregates with increasing concentrations (Fig 3.45 A). Evaluating the RNA integrity of fixed cells did prove the suitability of ACME fixation to preserve RNA quality (Fig 3.45 B). Likewise, no impact on RNA quality was detectable with formaldehyde in concentration between 0.1% to 1%, but beginning degradation could be seen with 5% (Fig 3.45 B). As ACME fixation seemed to have no impact on cell morphology and did not impact RNA integrity, it was deemed the suitable protocol for cell fixation.

3.4.2 Enrichment of somatic cells by FACS

As cell-specific markers like antibodies for enrichment via FACS are unavailable for the liver fluke, it was decided that the initial workflow should be modified by subdividing posterior and anterior regions by their DNA content (2n and 4n). Initial analysis of these parameters via cytometry showed cell populations that could be distinguished based on their staining for DRAQ5 (DNA content) and Concavalin A conjugated to Alexa488 (cytoplasm). An overrepresentation of 4n cells was observed, clearly illustrating the observed limitation of the first dataset (Fig 3.46). Additionally, populations of cells above and below the 4n and 1n cells were detected. The populations of cells above 4n are probably spermatocytes, which can be fused in rosettes of 8 to 32 nuclei depending on their differentiation state [230]. The population below 1n did not stain positive for DRAQ5 compared to an unstained control and represents cellular debris.
These samples were divided first into samples of either posterior and anterior samples which were further divided into samples of 2n and 4n cells. This resulted in sixteen libraries which were prepared using the 10x Genomics Chromium platform and analyzed with help from Carolyn Brewster at the Stowers Institute. Using this approach, considering the low number of 2n cells shown by FACS, these cells were enriched compared to the original dataset.
At the time of writing, the analysis of this vast dataset was not concluded, but initial results regarding improving the initial workflow can already be presented. A total of 66,319 cells was brought into the analysis, and a first clustering of those cells identified 18 distinct clusters (Fig 47 A) with a more heterogeneous distribution of cells between clusters than observed

before (supplementary table S4). Comparing marker genes between both datasets showed the expression of most identified markers (Fig 3.47 B). While some clusters had direct representation in the new data, others were completely new or were vastly improved by a higher cell number. For example, the stem cell cluster, with around 400 cells in the initial data, was increased to a size of 17,342 cells. Novel features in this data included, for example, the now clear representation of the vitelline lineage over S1, S2/3 to S4 cells by default clustering, a cluster expressing *p53* (D915_001973) marking tegumental progenitors in schistosomes and planarians [118,231], as well as a parenchymal cluster identified by expression of fatty acid binding proteins and the parenchymal marker *mf6* [232,233]. Of note was the observed high proportion of cells from the 2n samples in unidentified or underrepresented clusters (91% for the tegument cluster, 83% for intestinal cells, and 80% for S1 cells).

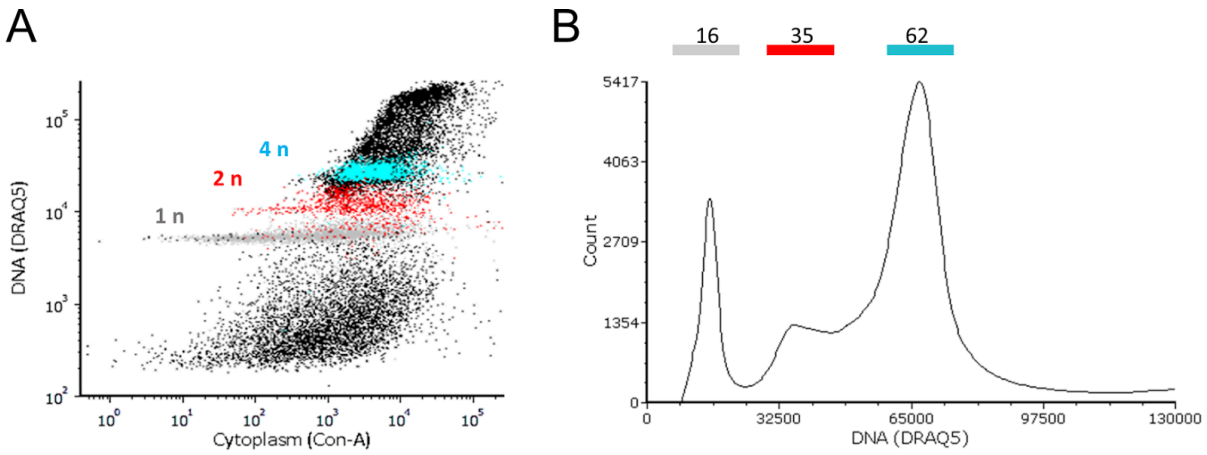


Figure 3.46 Identified clusters had distinct marker profiles. A. Flow cytometry profile of *F. hepatica* cells stained with DRAQ5 (DNA) and Concanavalin-A (cytoplasm). Cells are colored based on DNA content. **B**. Histogram of cells based on their DNA staining.

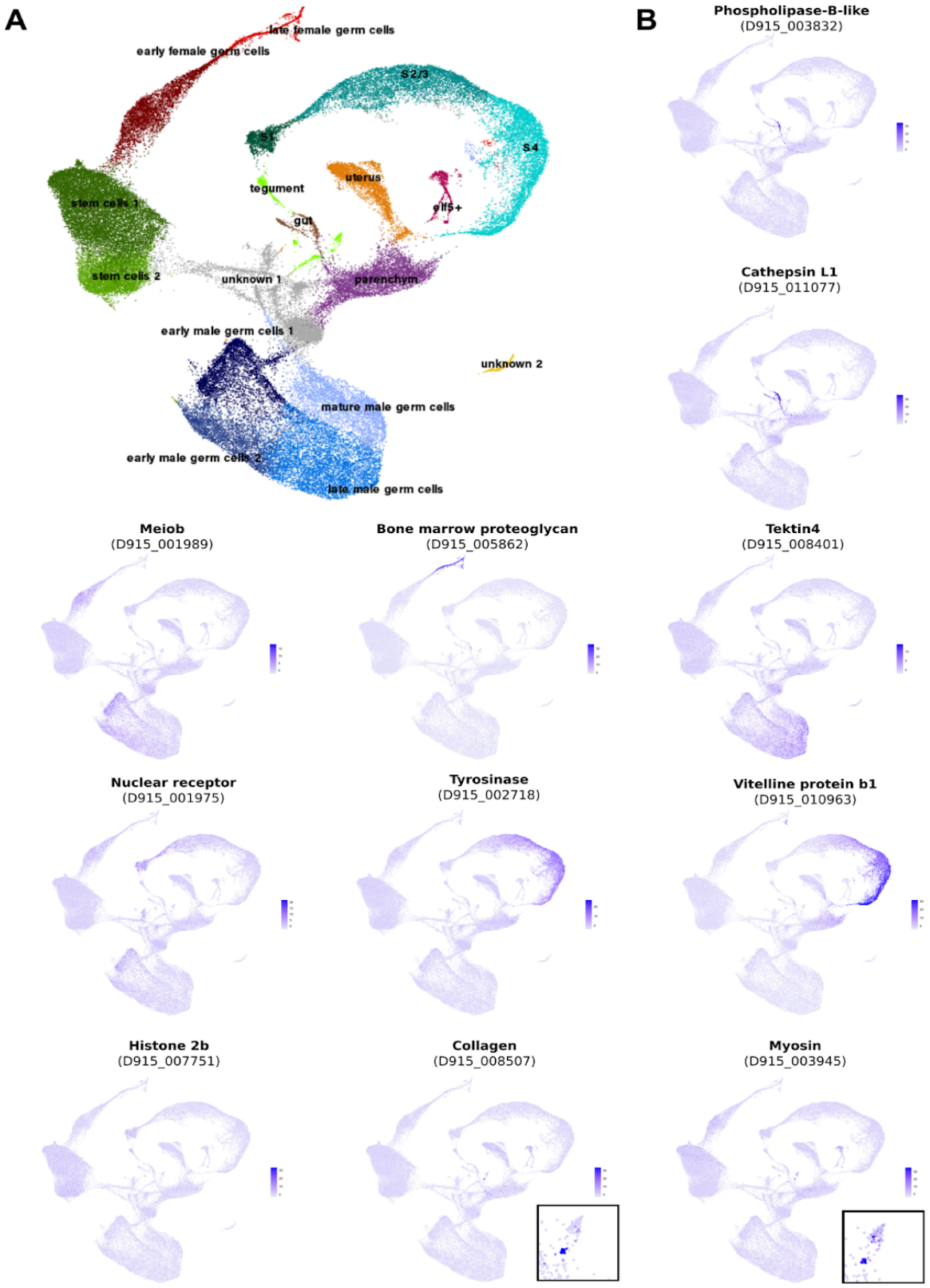


Figure 3.47 Identified clusters had distinct marker profiles. A. UMAP of 66,319 cells. Clusters are colored, and labels are added. **B**. UMAP feature plots showing the expression of various marker genes used throughout this project.

Additionally, the first sub-clustering of clusters hinted at a more heterogeneous composition of selected clusters (supplementary Fig S3 - 4). For example, sub-clustering of the tegument cluster gave 8 clusters, one expressing the schistosomal pan-neuronal marker *7b2,* indicating neuronal origin. Additionally, the sub-clustering of the two stem cell clusters revealed a cluster with a specific expression of a krüpple-like factor, a transcription factor that was shown to be a primordial germ cell marker in planarians (supplementary Fig S5 [234]). Cell cycle scoring was performed as the basis of this new dataset, exploring the influence of DNA content on the data structure (Fig 3.48). As expected, cells in the stem cell cluster scored high for proposed S-phase markers, while the cells within the potentially dividing oocyte cluster were enriched for G2M markers (Fig 3.48). Cells for somatic tissues like parenchymal or intestinal cells were high in G1 markers. The vitelline lineage was also clearly ordered based on the S-phase score, with the S1 cluster having almost 75% of cells in S-phase, followed by S2/3 cells with 20% and S4 with nearly no cells scoring for S-phase (Fig 3.48).

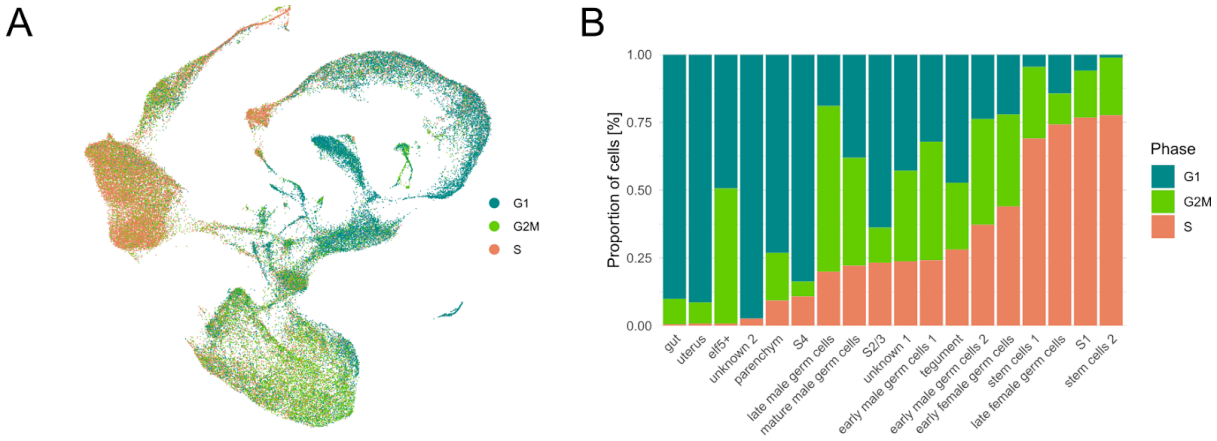


Figure 3.48 Identified clusters had distinct proportions of cell cycle stages. UMAP plot using cell cycle assignment to color cells. **B.** Stacked bar plot displays the proportions of cells within each cycle stages per cluster.

3.5 Leveraging the dataset for drug target discovery

In light of growing resistance to the drug of choice, triclabendazole, the identification of new drug targets for novel treatment strategies is a critical aspect of liver fluke research. The enrichment of the pan-trematode target $TRPM_{PZQ}$ within the elf5+ cluster to considerations regarding the druggability of this cell type. The family of PKs is known to contain many druggable targets in helminths [235]. Additionally, their importance in the progression and development of cancers has driven the development of various small-molecule inhibitors, which might have value in repurposing efforts for treating parasitic worms [77].

The kinase annotated as PAK4 (D915_004414) was selected as the candidate study for this work to illustrate this dataset's utility for identifying druggable targets. The gene had highest relative expression in the elf5+ and the late vitellocyte clusters, and clearly weaker expression in muscle and gut clusters (Fig 3.49 B). The family of p21-activated kinases (PAK) is characterized by a p21-binding domain (PBD), sometimes also called the Cdc42/Rac interactive binding motif [236]. PAK family members have not been characterized in trematodes, while in humans, PAK4 is implicated in cell migration and cytoskeletal reorganization. The kinase is also involved in various cancer stages, particularly the formation of metastasis and invasion of tissue by cancer cells [237]. In other organisms, PAK family members function in developmental processes, like in *D. melanogaster [238]* or functions like axon guidance and mechanotransduction [239–241]. Ensuring correct classification of FhPAK4, all other PAK family members within the liver flukes genome were identified based on their characteristic PBD domain (Fig 3.50 A).

The phylogenetic relationship of the PAK family members was resolved by comparing their amino acid sequences with related species and model organisms (Fig 3.50). The PAK family is divided into two groups that differ in structure and function [236,242]. FhPAK4 was grouped within group 2 of PAK family members and had a 44.6% sequence similarity with human PAK4. Comparing their kinase domains showed a 64.3% sequence similarity. Of note was the apparent absence of a group 2 PAK in *S. mansoni,* while one seemed to be encoded within the planarian genome (SMESG000022413) (Fig 3.50). As the other liver fluke PAKs grouped within the group 1 family of kinases, they were termed accordingly FhPAK1 (D915_001478), FhPAK2 (D915_004654) and FhPAK3 (D915_006992). The latter could be found within a previously described group [240,243] with *D. melanogaster* PAK3 and *C. elegans* MAX-2. Additionally, both schistosomal and planarian PAK were located within this separate group.

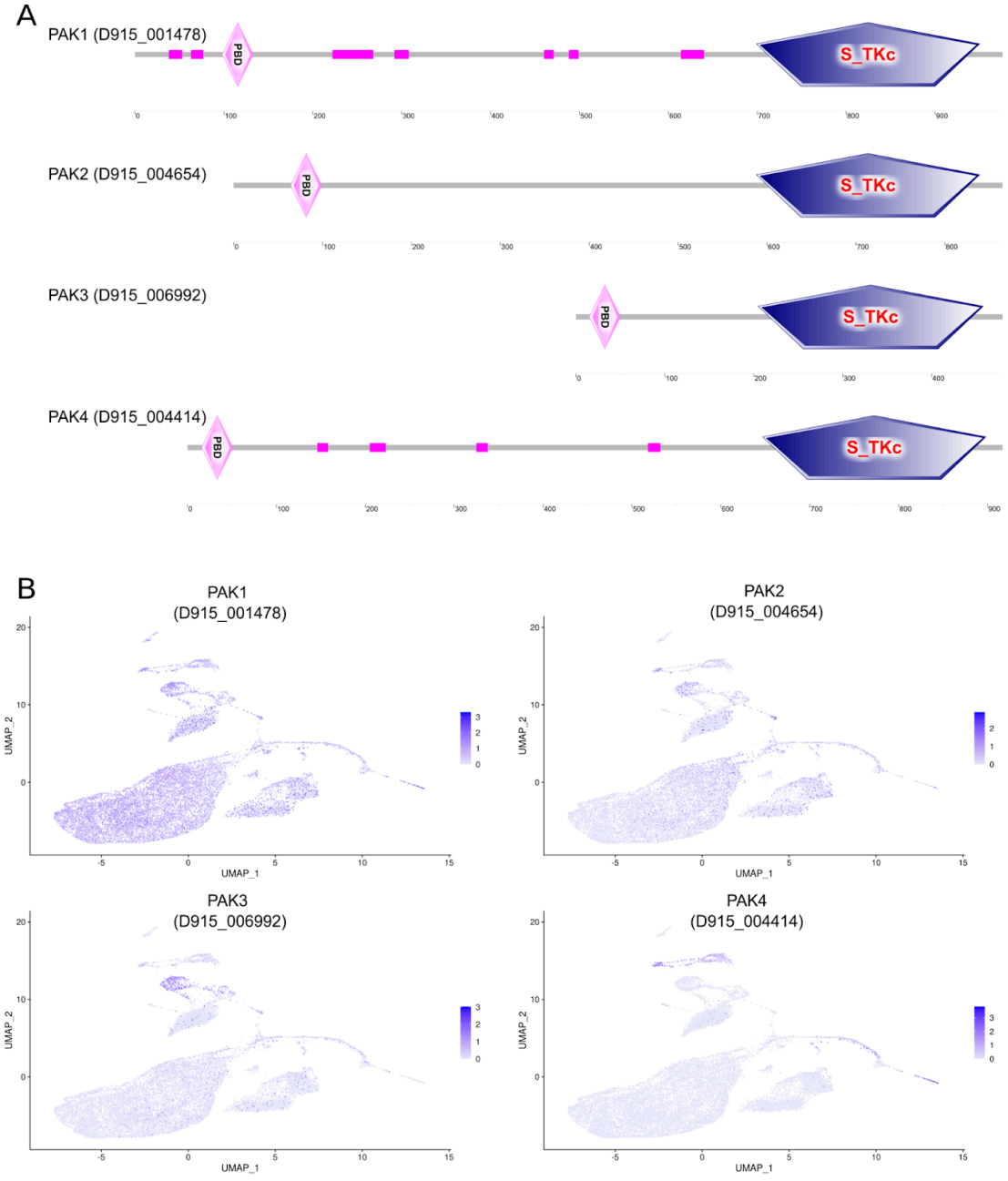


Figure 3.49 Domain structure and gene expression patterns of the four different PAKs. A SMART analysis confirmed the typical domain structure of *F. hepatica* PAK proteins, consisting of a serine/threonine kinase domain (S_TKc) and a p21-binding domain (PBD). **B** UMAP plots colored by gene expression show the expression of various PAK genes.

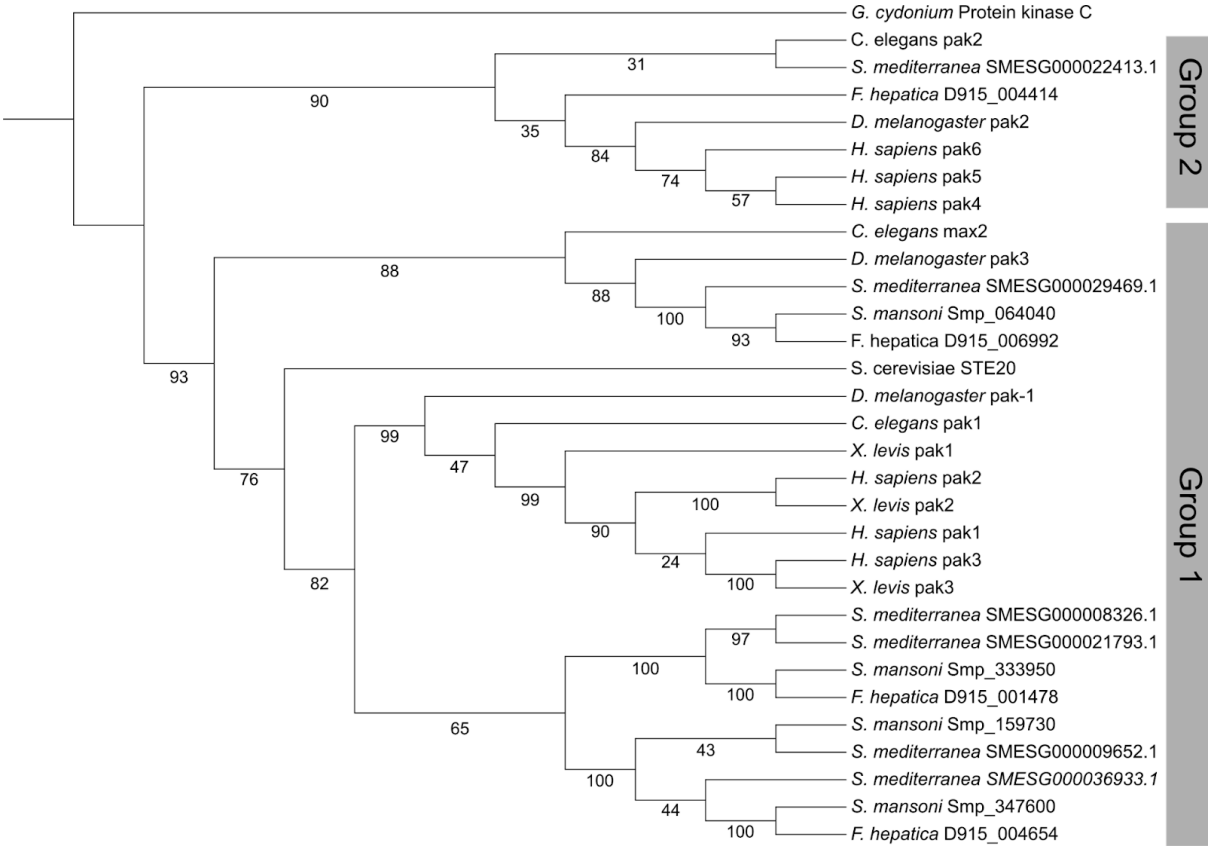


Figure 3.50 Phylogenetic analysis groups PAKs into their associated family groups. Phylogenetic tree of PAK orthologs of *F. hepatica* and other species (accession numbers are provided in supplementary table S6). p21 families are indicated, and protein kinase C from *Geodia cydonium* served as outgroup. Gene IDs in WormBase ParaSite (*S. mansoni* and *F. hepatica*) or Planosphere for *S. mediterranea* are indicated for previously unrecognized proteins.

Based on the scRNA-seq data, transcripts for *fhpak1* and *fhpak2* were transcribed in several tissues, while *fhpak3* was expressed in mature oocytes (Fig 3.49). All PAKs of the parasite appeared to be equally transcribed over the available bulk RNAseq data (supplementary Fig S6). Localizing *fhpak4* transcripts *in situ* showed positive cells beneath the tegumental area and weak staining in the gastrodermis (Fig 3.51). Unexpectedly, transcripts were also detected within the nucleus of oocytes (Fig 3.51).

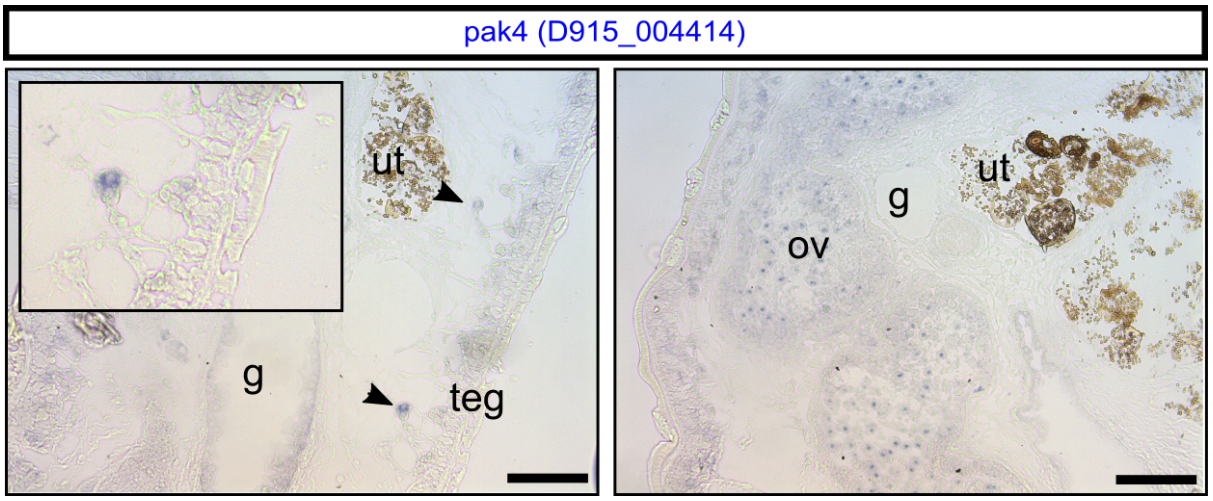


Figure 3.51 *Fhpak4* transcripts localized in cells beneath the tegument and oocytes. Transversal sections of adult *F. hepatica* stained for transcripts of *fhpak*4. Scale bar: 100 µm. Sense controls were found to be negative (supplementary Fig S7).

As noted earlier, there has been a lot of progress in developing small-molecule inhibitors for kinases. Inhibitors for human PAK4 have already been described [244,245], including a commercially available ATP-competitive inhibitor called LCH-774994, which has effects on proliferation, migration, and invasion of gastric cancer cells [246]. As the binding mechanism was predicted to occur between the hinge and p-loop regions of the human protein version [247], the required sequence similarities were validated. A high amino acid sequence similarity was found between all liver fluke PAKs and the human PAK4 for the hinge region, though the p-loop region was more similar between human and liver fluke PAK4s (Fig 3.52).

```
                                10        20        30        40        50        60
D915_004414/1-252     1 LEELGPIGEGSTGIVCLMRYRSTNHYVAVKRMNIFKQQRRELLFNEVMIMRSYPHPNIVEMFASHLI  67
human_pak4/1-252      1 LDNFIKIGEGSTGIVCIATVRSSGKLVAVKKMDLRKQQRRELLFNEVVIMRDYQHENVVEMYNSYLV  67
D915_006992/1-254     1 YTLIEKVGSGASGTVQRSRDNRTRQMVAIKVMRLDKQPNRDLIISEIEVMKELRHESIVNYLESYLL  67
D915_001478/1-252     1 YKRQEVIGQGASGVVYGGLEVATGRRVAIKEMNLKQQPKKELIVNEIYVMKAKKHPNVVNYLDSYLV  67
D915_004654/1-252     1 YETLGRIGHGASGIVQIGRDLKTGQRVAIKQMNLRKQPKKELILNEILVMRAYRNPNIVNYLDSYLL  67

                                70        80        90       100       110       120       130
D915_004414/1-252    68 K--DELWVVMEYLEGGALTRIVARTLMSEQQMATVCRAVLKALAFLHDHGIIHRDVKSDSILLSITG 132
human_pak4/1-252     68 G--DELWVVMEFLEGGALTDIVTHTRMNEEQIAAVCLAVLQALSVLHAQGVIHRDIKSDSILLTHDG 132
D915_006992/1-254    68 RQENELWVVMEYLDGGALTEVVTETVMDVPTIAAVTRECVKALAFLHDKNIIHRDIKSDNVLLGRQG 134
D915_001478/1-252    68 G--DELWVVMEYLDGGSLTDVVTETYMEEGQIAAVCRECLQALDFLHKNQVIHRDIKSDNILLGLDG 132
D915_004654/1-252    68 G--DELWVVMEYLDGGSLTDILTETCMSEAHIATVCRETLQALEFLHSKHVIHRDIKSDNILLGLDG 132

                               140       150       160       170       180       190       200
D915_004414/1-252   133 RVKLSDFGFCAQITPDT-PRRRSLVGTPYWMSPEVISR-KPYGTAVDVWSMGVLLIEMVDGEPTYFN 197
human_pak4/1-252    133 RVKLSDFGFCAQVSKEV-PRRKSLVGTPYWMAPELISR-LPYGPEVDIWSLGIMVIEMVDGEPPYFN 197
D915_006992/1-254   135 QVKVTDFGFCAQLGNRQ-SKRDTMVGTPYWMAPEVVNKTVQYGQKIDIWSLGIMVIEMIDGEPPYLH 200
D915_001478/1-252   133 SVKLTDFGFCAQLSSEQ-TKRSTMVGTPYWMAPEVVTR-KQYGPKVDIWSLGIMAIEMLDGEPPYLN 197
D915_004654/1-252   133 AVKLTDFGFCAQLSSDQDVKRNTMVGTPYWMAPEVVSR-KQYGNKVDIWSLGIMTLEMIEGEPPYLS 198

                               210       220       230       240       250
D915_004414/1-252   198 EPPLRVMRRIQEEAVPHLANPHKSSRRLNSFLGLMLVREPSCRATAAQLLLHPFI          252
human_pak4/1-252    198 EPPLKAMKMIRDNLPPRLKNLHKVSPSLKGFLDRLLVRDPAQRATAAELLKHPFL          252
D915_006992/1-254   201 EQPLRAIMLIQANGKPHPK-TKHIDSSMQHFLDRCLIVNPDARASAKDLLQDPFL          254
D915_001478/1-252   198 ENPVRALYLIGTNGKPEIKEREKLSPEFQDFLDRCLEVKVDQRASAQELLQHPFI          252
D915_004654/1-252   199 ENPLKALYLIATNGKPHFR-KDHLSPELLDFLDCCLEVDVPKRATAASLLTHRFI          252
```

Figure 3.52 Sequence alignment shows conserved regions necessary for drug action. Alignment of the human PAK4 amino acid sequence with *F. hepatica* PAK sequences. The binding sites of LCH-7749944 are colored: P-loop in blue and hinge region in gray.

It was, therefore, decided to test the inhibitor's efficacy against various stages of the liver fluke *in vitro*. The inhibitor showed efficacy against all tested stages (Fig 3.53). Adult worms showed reduced motility with 25 µM and 72h exposure to the inhibitor, while a concentration of 50 µM killed worms. Immature worms were affected more severely by this inhibitor, with lethal effect after 48h at the highest concentration and substantial reduction (score 1) after 48h with 25 µM. In line with the notion that juvenile worms might be more affected by this inhibitor were the observations with NEJs, which died within hours of inhibitor exposure. In addition, the inhibitor caused tegumental lesions and generally impaired tissue integrity of NEJs.

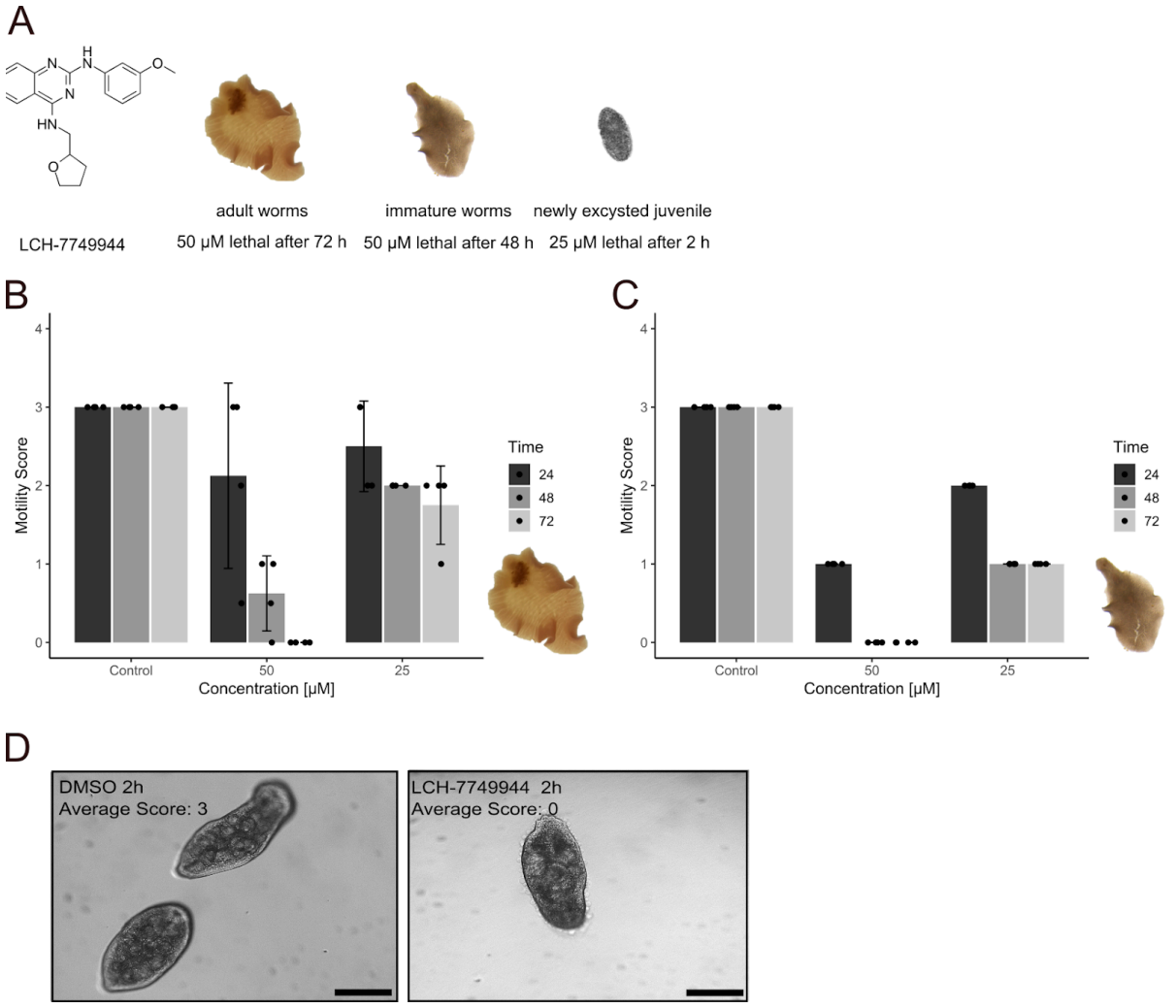


Figure 3.53 LCH-7749944 impacts all intra-mammalian stages *in vitro.* A. Motility scores of schistosome couples under varying concentrations of LCH-7749944 over 72h. Average scores in 10 worms per concentration are shown (n = 1). **B.** Representative image of schistosome worms treated with 50 µM inhibitor after 24 h. Scale bar: 500 µm **C**. Representative image of planarians treated with 12 µM inhibitor after 48 h showing severe lysis in treated worms. Scale bar: 200 µm.

To evaluate the utility of this inhibitor in related species, LCH-774994 was also tested against planarians and adult schistosomes. The inhibitor proved to be lethal against both worms, with observed lysis in planarians with inhibitor concentrations as low as 5 µM (Fig 3.54).

Schistosome adults were also impacted by inhibitor concentrations ranging from 50 µM to 12 µM. Treatment with 50 µM inhibitor *in vitro* was lethal after 24h, while lower concentrations affected motility between 30% - 60% compared to the control worms. Schistosome worms treated with the inhibitor showed bubbles in the tegument and tegumental shedding (Fig 3.55).

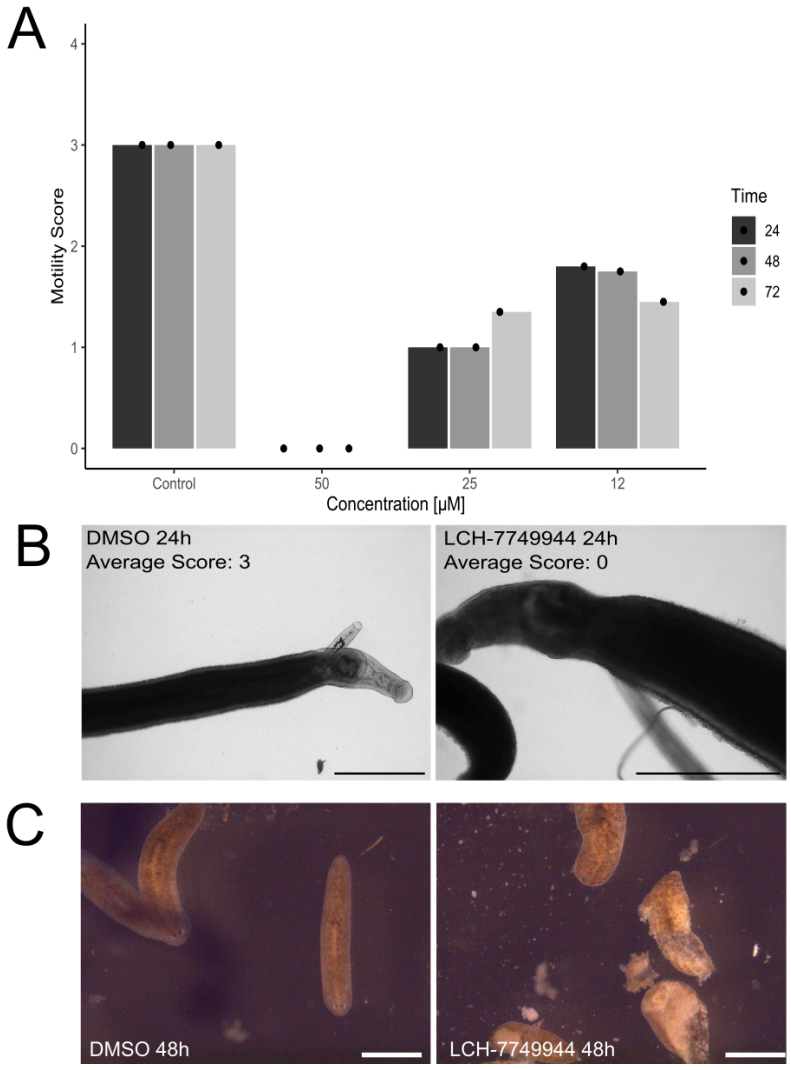


Figure 3.54 LCH-7749944 impacts other parasitic and free-living worms *in vitro.* A. Motility scores of schistosome couples under varying concentrations of LCH-7749944 over 72h. Average scores in 10 worms per concentration are shown (n = 1). **B.** Representative image of schistosome worms treated with 50 µM inhibitor after 24 h. Scale bar: 500 µm **C**. Representative image of schistosome worms treated with 12 µM inhibitor after 48 h showing severe lysis in treated worms. Scale bar: 200 µm.

As small molecule inhibitors might not just inhibit their target but also have off-target effects on other proteins, phenotypes must be evaluated using different methods. RNA interference (RNAi) has proven to be an effective method to silence the expression of selected target genes, thereby studying their function. Here, two long (400 - 500 bp) dsRNAs were designed against the mRNA of *fhpak4,* which were introduced to immature worms by soaking. Additionally, combined treatment of both dsRNAs in a final 50 ng/µl concentration was

included. Next to an untreated control, a non-helminth dsRNA designed against the bacterial resistance gene *neomycin* was included as another negative control. After cultivating the worms for 14 days *in vitro,* the knockdown level was evaluated using qRT-PCR. The untreated control resulted in a knockdown of 71% with dsRNA_1 and 75% with dsRNA_2. Combining dsRNAs resulted in an average knockdown of 62% (Fig 3.55).

Interestingly, compared to untreated worms, the non-helminth dsRNA showed a remarkable increase in transcript level, with an 86% increase (Fig 3.55). For that reason, the following analysis did not include *neomycin*. No change in motility was observed in the worms throughout the experiment.

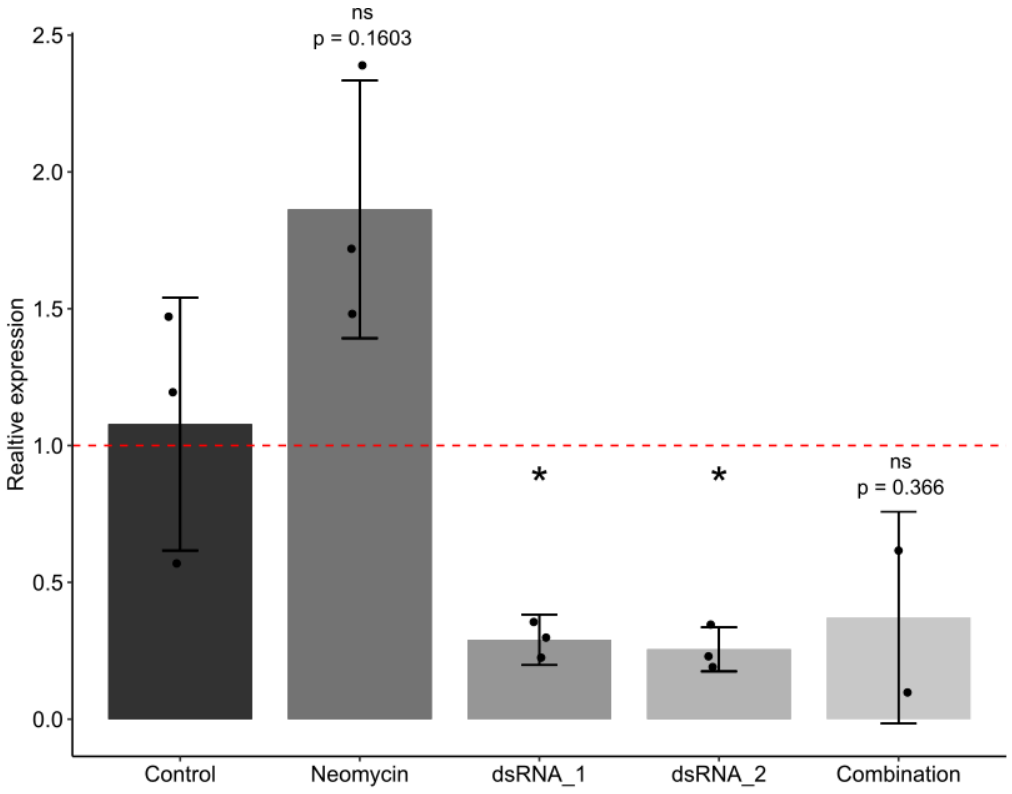


Figure 3.55 Knockdown of *Fhpak4* using dsRNA led to a reduction of transcript levels. Worms were treated with 50 ng/ml for 21 days. Columns represent the mean relative expression level compared to the control. The relative gene expression level of *fhpak4 is* compared to that of an untreated control. Each data point represents one qRT-PCR sample from a single worm; error bars represent standard deviation. Differences were determined to be statistically significant by using delta ct values and testing against the untreated control. Significant differences are indicated with * = p< 0.05 (Welch's t-test)

3.6 Mapping transcription factors in the liver fluke

Transcription factors (TFs) are diverse proteins that can influence gene expression by binding to DNA. This regulation of gene expression is essential to achieve complex biological processes like development, cell cycle progression, or stress response. As such, knowledge of TFs within organisms can greatly enhance understanding of biological processes. To date, no study has thoroughly mapped the transcription factor landscape within the liver fluke. Here, an approach from a recent planarian survey [172] was adapted to identify and group the TFs of *F. hepatica*. To this end, the liver fluke proteome was searched for the proteins containing a DNA-binding domain (DBD), resulting in a list of 645 proteins. As DNA binding domains are also present in some other proteins not meeting the criteria of a transcription factor, such as histones, a blast search was performed against known transcription factors in humans and planarians [127,172]. Applying suitable thresholds, 423 transcription factors were identified based on sequence similarities with other organisms, of which 97 genes were excluded from further analysis as they had matches for non-TFs. Finally, 117 genes were included as potential *F. hepatica* specific TFs as they had no match in either blast search, arriving at 599 DBD-containing proteins with potential transcription factor activity. Next, the obtained sequences were classified into their respective TF families. Of the identified TFs, a majority belonged to the family of zinc finger TFs, with 41% of the total TFs (Fig 3.56). Other genes identified were 9 GATA-like TFs, 5 DM domain-containing proteins, or two p53-like transcription factors.

As transcription factors are essential in coordinating gene expression during development, the expression of TFs was mapped across the life stages of the liver fluke using available data [91]. More than half of these factors were expressed in all screened stages (58%), though some were unique for each stage (Fig 3.57). For example, 5 TFs were unique for the adult stage. These genes encode for two SMAD orthologues (D915_009314, D915_009542), a basic helix loop helix TF (D915_002258), a zinc finger TF which had high similarities to the TF Escargot from *D. melanogaster* (D915_004057), and a potentially fragmented gene (D915_003896) which might be part of the centrosomal protein D915_003897. These genes were expressed at a low level and were not detected in the single-cell data, complicating a functional annotation. Still, SMAD TFs are involved in the TGF-beta signaling pathway, which has been implicated in egg development in *S. mansoni, [248,249]* and Escargot functions in stem cell maintenance in fruit flies [250].

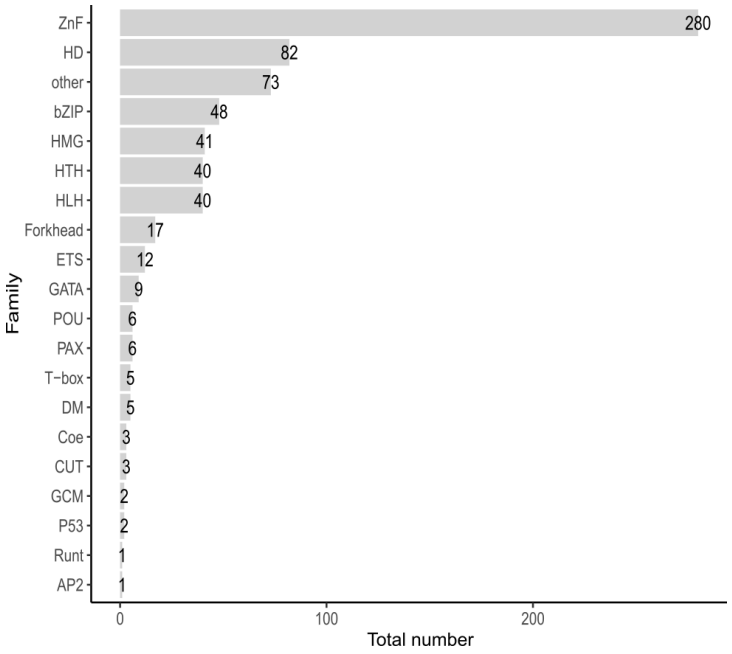


Figure 3.56 A diverse set of transcription factor families is encoded in the liver flukes genome. Categorization of genes within the identified TF by family. The identified number of genes per family is indicated for each bar.

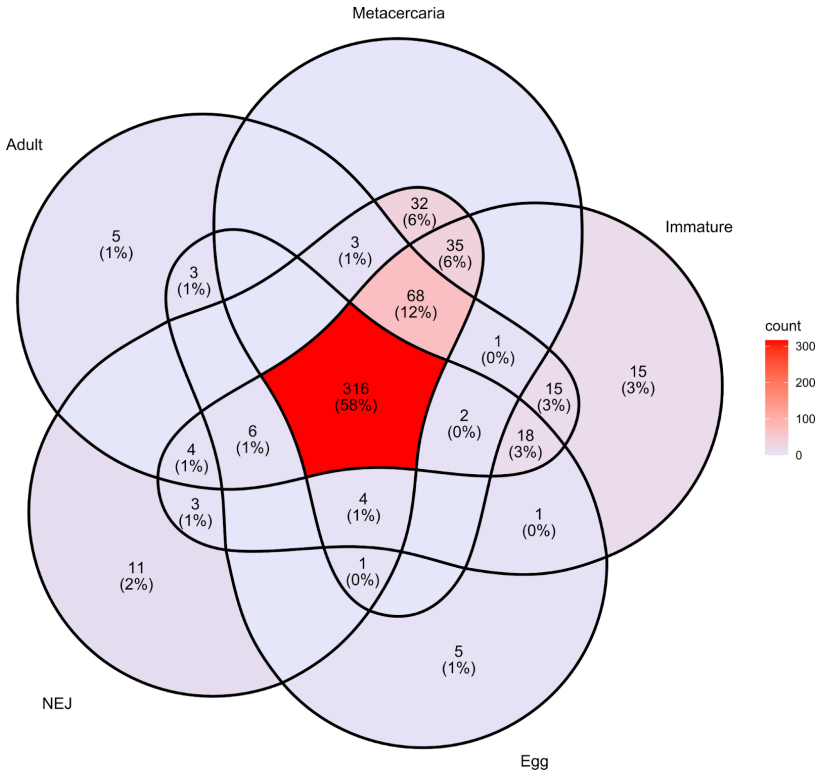


Figure 3.57 Selected transcription factors are uniquely expressed in different life stages. Venn diagram showing the expression of identified TFs over bulk RNAseq data in various life stages. The NEJ group consists of data summarized from NEJ samples after 1, 3, and 24 h.

Analyzing the expression of the 125 TFs that showed positive expression within the scRNA-seq data showed enrichment for particular transcription factors in cell types or

groups of cell types with related biological processes (Fig 3.58). Interestingly, a transcription factor encoding a fos transcription factor (D915-003851) was enriched in the elf5+ cluster, a known protein affected by MAPK signaling pathways, or transcripts for the gene *zeb-1* for which a role in cell migration could be shown in planarians [251]. The most distinct profile of TFs was demonstrated within the late female germ cell cluster, with more than 20 TFs showing strong enrichment. This included the already described thyroid hormone receptor (D915_005878), but also other transcription factors such as a GATA TFs with sequence similarity for GATA-2 [252], which is a maternal TFs important for development in frogs [252] or an HGM-box containing TF (D915_006158), determined to be an orthologue to the TCF family of TFs in planarians by BLAST. The TCF/LEF family is implicated in Wnt signaling, which was previously shown to be enriched in the late oocyte cluster [253].

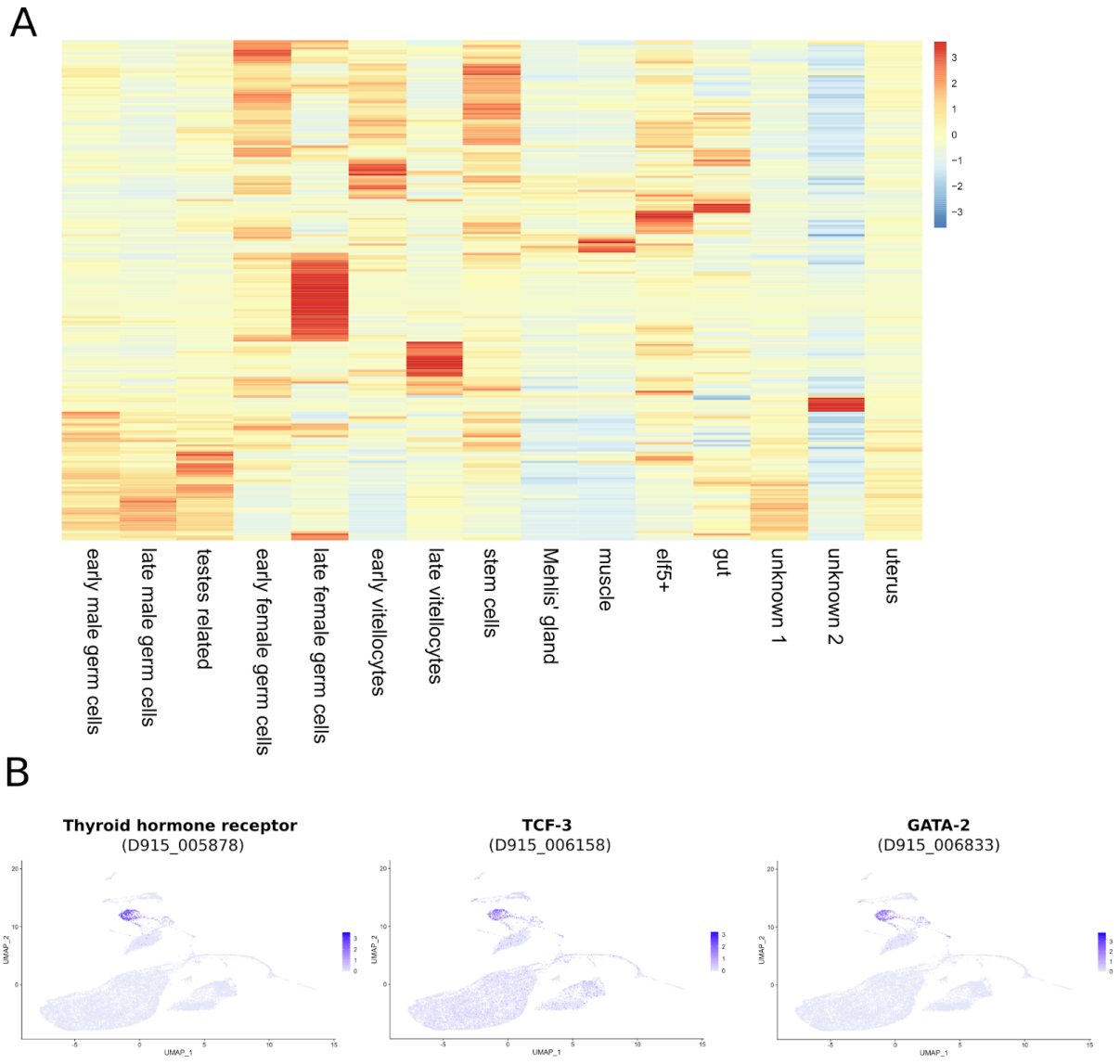


Figure 3.58 Various transcription factors have cell-specific expression profiles. A Heatmap of averaged expression values in the scRNA-seq data; the values are scaled by gene and centered. **B** UMAP feature plot of selected transcription factors.

4. Discussion

The common liver fluke *F. hepatica* is a globally relevant parasite of animals and humans. It causes major socioeconomic issues and is recognized by the WHO as an NTD [3]. The spread of resistance and limited knowledge of this parasite's biology highlights the need for more and better basic research to learn how this worm – and its cells - function on a molecular level. The work presented here aimed to generate a scRNA-seq atlas to further the understanding of the liver fluke's molecular biology.

4.1 Establishment and validation of a scRNA-seq workflow

An important prerequisite for the success of a scRNA-seq experiment is a sufficient quality of the input cells. Impaired viability and clumping of the cells can lead to low-quality data due to cell death or even complete failure of the experiment. An extensive number of different protocols were tested, varying proteases, temperature, and buffer conditions, from which optimized trypsin-based dissociation was deemed best suitable for scRNA-seq. Next, a FACS protocol for the enrichment of viable cells was established. Overall, the raw data looked promising with on average 30% of the 8716 total transcribed genes in adult worms detected per cell. These numbers are higher compared to the *S. mansoni* scRNA-seq data [111], where the 1600 median genes per cell represent 18% of the total genes transcribed (8890 genes; TPM >2) [131] across adult worms. Finally, 15 cell clusters were predicted, most of which could be directly linked to major tissue types of parasitic flaworms. Still, comparing the available bulk RNAseq datasets with the scRNA-seq data showed low correlation. A similar analysis for parasitic nematodes showed that RNAseq was able to recapitulate the cell-type distribution and transcriptional profile of the parasites [254]. The low correlation between the liver fluke datasets might be caused by the intentional enrichment of anterior cell types, which might lead to a skewed abundance of transcripts. Though several of the cell types of the worms were covered within this initial dataset, some were missing, such as cells of the nervous system or flame cells. Several possible reasons for this are related to the complex workflow of a scRNA-seq experiment. This section will provide a detailed discussion of the established method and possible improvements.

4.1.1 Evaluation of a trypsin-based dissociation protocol

As it is one of the essential steps in the scRNA-seq workflow, it was extensively improved and optimized. Especially the dissociation process of adult worms was subject to thorough validation. Prior to this work, no dissociation protocol for adult worms had been published.

While there have been reports of dissociation in the juvenile stages of the parasites using trypsin [85], this dissociation was carried out over 3 h, which could severely impact the cells' transcriptome. Another method described the enzymatic isolation of muscle fibers from the ventral sucker, using a combination of papain with trypsin [255]. This dissociation was effective at isolating fibers, but the authors noted a lack of nuclei, explained by the structure of the flatworm musculature [255], and the approach was unsuitable for scRNA-seq. While reviewing the literature, it was noticeable that most dissociations in flatworms utilized trypsin, including planarians [256,257] or *S. mansoni* adults [111] which could be due to a similar extracellular matrix composition of these flatworms. The best cell quality in this work was also achieved using trypsin at 37°C with HBSS devoid of calcium or magnesium ions. Other tested enzymes, like collagenases, showed little dissociation efficiency. As the worm secretes numerous protease inhibitors [258], it might be interesting to speculate that these might impact the efficacy of certain proteases used for dissociation.

Cells dissociated with trypsin showed good viability and high heterogeneity, with numerous morphologically distinct cells present. Testing for the presence of putative cell marker genes confirmed the presence of various cell types, including nerve cells. The apparent loss of these cells in the further process is discussed later (4.1.3 and 4.1.4). It must be noted that using a single protease for dissociation might cause a bias of specific cells or be a limiting factor in detecting others. Due to the complexity of the extracellular matrix of a whole organism, a single protease might not be sufficient in dissociating all cell types, and an approach using multiple proteases as that utilized for miracidia of *S. mansoni [259]* could lead to a better capture of these cell types.

The process of dissociation inevitably causes cell stress, which must be taken into account when interpreting the data obtained [260]. Additionally cell damage produced large amounts of debris, some of which were found to be lipid droplets from ruptured vitelline cells. The sensitive nature of these cells has already been described for vitelline cells of schistosomes [183]. Altogether, avoiding these cell ruptures proved impossible, but switching to wide-bore pipette tips for cell handling reduced it significantly. Even though the dissociation of primary tissue was carefully optimized, it still caused debris and some cell death, which is why FACS was performed using the viability dye calcein-AM, enriching viable cells.

4.1.2 Cell biases introduced by FACS and 10x technology

FACS is the standard method for the enrichment of various cells upstream of scRNA-seq. Even though it is routinely used to isolate viable cells, FACS is a harsh procedure where cells are stressed or even damaged by shear stress [261]. It has also been shown that the FACS step can be problematic for capturing trematode cells, as ciliary cells could only be

found in unsorted samples of the miracidial cell atlas [259]. Additionally, the cell atlas of schistosomal sporocysts failed to capture protonephridial cells and flame cells, which was potentially due to difficulties reported for calcein-based sorting strategies [113]. Considering the potential pitfalls of a FACS, the sorted cells within this project were carefully evaluated by checking RNA quality and by reevaluation of selected marker genes via RT-PCR. Most marker genes could be detected with no apparent difference in the cell suspension before and after FACS, but *h2b*-positive cells seemed to decrease after dissociation. An image-based approach using EdU staining of cells validated the presence of proliferating cells in the mixture, so this loss of cells was considered acceptable.

Similar to schistosomal datasets, cells of the protonephridial system and flame cells were missing from the dataset. It cannot be ruled out that the FACS approach used here did also affect the capture of these cells. Adding an unsorted sample to the 10x workflow might enhance the capture of these impacted cells, though the debris within the sample might hinder this approach. Another significant factor in a scRNA-seq method is the choice of technology. The method of choice was the 10x Chromium technology due to its outstanding throughput and broad accessibility. Chromium can capture a broad range of different cells but has a size limitation of a maximum of 30 µm in cell diameter [177], which has been shown to hinder the capture of penetration gland cells in miracidia [259]. This limitation in size also demands a filtering step for Chromium experiments. While it is not known that major cell types of the liver fluke exceed this cell size, other technologies might be needed to comprehensively detect these bigger cells. Methods like plate-based scRNA-seq have been successfully utilized to capture larger cells that have previously been missing from scRNA-seq data [259]. Additionally, split-pool barcoding has been successfully used in planarians [158,262], which uses combinatorial barcoding to label cells without the need for a microfluidic device like droplet-based methods.

4.1.3 Liver fluke specific challenges in scRNA-seq technology

Each organism has unique biological and cellular adaptations, for which a scRNA-seq workflow must account. In this work, one challenge working with the liver fluke was the abundance of cells relating to reproductive tissues. This was particularly apparent for the cells of the testes, which were found to contribute to most of the cells within the worm. This observation led to the division of the worm into an anterior and posterior part to enrich rare anterior cell types. The utility of this approach was confirmed as whole worm samples had a similar proportion of cells as the posterior worm samples. Similar problems have been shown in schistosomes, where cells of the vitellarium made up most of the sample in mature females, and an anterior sample was included to enrich rarer cell types [111]. Adding to this

problem is the liver fluke's cell number. As determined by this work, *F. hepatica* consists of around 10 million cells, while schistosome adults, for example, have some tens of thousands of cells [259]. This means that a much higher number of cells needs to be analyzed in case of liver flukes to achieve the same coverage of cells and cell types. Furthermore, an increase in size is not necessarily associated with an equal increase in the number of all cell types, which could complicate the analysis of rare cells from the nervous system [263]. The cell atlas project for planarians subdivided the worms into smaller pieces taken from various anatomical regions, thereby profiling almost all cell types within this worm [107]. Adopting a similar approach to the liver fluke might be tempting for more thorough profiling in the future.

A unique challenge in trematodes is the capture of tegument cells, as this tissue is a syncytium, showing no clear cell boundaries. A potential solution for capturing tegumental cells might be single-nuclei RNA sequencing (snRNA-seq) [264], which uses isolated nuclei instead of single cells. This has been described as having some major advantages to scRNA-seq, as it is suitable for frozen, fixed, or fresh samples, has a lower cell bias, and introduces lower dissociation bias to the cells [265]. One disadvantage of snRNA-seq is a reduced sensitivity for certain transcripts, as it only captures transcripts in the nucleus. Still, it is a suitable alternative.

4.1.4 Improvement of the scRNA-seq dataset

Fixation of cells after dissociation can benefit scRNA-seq projects as it allows for the decoupling of cell preparation from the capture of cells. This is necessary for rare samples that are not available at all times and need to be stored, or for the transport of cells processed at a different location. During the project, a research visit at the Sanchez-Alvarado lab (USA) necessitated the transport of fixed liver fluke cells. For that reason, fixation methods were evaluated to preserve the samples' cell morphology and RNA integrity. While both tested fixation methods (formaldehyde and ACME) preserved RNA integrity, ACME preserved the cells' morphology better. Recent comparisons contrasting different fixation protocols have found that both methanol and ACME do not significantly affect cell composition. Still, ACME could be shown to cause some extent of damage to RNA molecules, arguing for a switch to pure methanol fixation in future experiments [266,267]. Nevertheless, ACME fixation greatly benefited the present work in several ways. Surprisingly, the suspension contained fewer lipid droplets after fixation, which might be a secondary effect of methanol washing out the lipids. As the latter dataset contained a greater number of vitelline cells, the fixation might have had a stabilizing effect on these cells, thereby preventing their rupture. In collaboration with the Sanchez-Alvarado lab in the USA,

the initial workflow was modified to enhance the capture of missing cell types. This included splitting anterior and posterior samples further based on their DNA content. This greatly improved the capture of cells that were not part or underrepresented in the prior experiment, including parenchymal and gastrodermal cells. The cellular diversity was greatly enhanced and none of the original cell clusters were lost using this approach. Although the data set could not be analyzed conclusively at this point, it impressively illustrated the lessons learned in this project on how to approach the challenges posed by this parasite and can serve as a template on how to improve the cell atlas of *F. hepatica* in the future. Based on the lessons gained here, an optimal workflow would utilize fixed cells, including a unsorted samples. As this approach would hinder the enrichment of cells based on their DNA content a larger amount of cells need to be included using this approach.

4.2 Germline differentiation in the dataset

Most of the liver fluke's body consists of reproductive tissue, divided into the testes, ovary, and vitelline follicles. These tissues consume a lot of energy due to the worm's outstanding output of 25,000 eggs per day [24]. Due to their abundance within the liver fluke, it was possible to capture most of the cell types making up the reproductive tissues, involving early and late germ cells of both the male and female reproductive tracts and the differentiation lineage of vitelline cells. Ways to interrupt the reproductive cycle of the liver fluke have been discussed as a potential strategy to prevent further spread of the disease [52]. The search for such interventions could be greatly enhanced by the data presented here, as it serves as a gateway for an improved understanding of germline differentiation of the liver fluke. Most work considering germline differentiation in trematodes has been conducted in *S. mansoni,* mapping out germline-associated stem-cell populations and identifying the gene regulation pathways controlling differentiation by scRNA-seq [109,114,185]. As such in the following sections the liver fluke data is compared and set into context using the insights gained from these studies.

4.2.1 Implications of the cellular location of germline stem cells

The testes and ovary of schistosomes anatomically differ significantly from the liver fluke, as they are organized more spherically compared to the branched reproductive organs in *F. hepatica.* Despite this difference, there are some important similarities at the cellular level. For example, in both tissues, the stem cells are located at the periphery of the gonadal tissue, and thereby are spatially separated from differentiated cells [193,268]. By using EdU labeling and IHC for H3P, this relationship could be clearly shown for both the ovary and testes of the liver fluke. These two staining methods can distinguish between different cell-

cycle states [186,198], which could be a future tool for evaluating the effect of pharmacological intervention on the differentiation process at a more detailed level. The location of stem cells was also successfully illustrated by utilizing the spatial transcriptome of adult worms generated by our lab [169]. Specification of stem-cell fate, especially in the germline, relies on tight control through signal transduction pathways. The regulation of these pathways can depend on somatic supporting cells, such as the Sertoli cells in the mammalian male germline or distal tip and stromal cap cells in *C. elegans* and *D. melanogaster,* respectively [269–271]. For flatworms, research on the role of external cues for germline differentiation has gained attention due to the discovery of a tryptophan-derived dipeptide, ß-alanyl-tryptamine (BAAT), which was secreted from the male schistosome to trigger female maturation [272]. In sexually reproducing planarians, which are hermaphrodites, this peptide was produced by somatic niche cells, triggering germ cell development, which sparked the hypothesis that during pairing in schistosomes, the male mimics a quasi germ-cell niche [273]. While there are some similarities in germ-cell differentiation between planarians and schistosomes, important differences like the reproductive output have been noted [114]. The liver fluke shares the parasitic lifestyle and high reproductive capabilities with schistosomes, while still being a hermaphrodite. Studying the somatic niche in the hermaphrodite liver fluke, building up on the tools and data on their location and cellular relationships provided here, could yield insights into the development of separate sexes and might be a way to better understand how reproductive machinery of parasitic flatworms is maintained.

4.2.2 Conserved germline regulators in two parasitic flatworms

Understanding the regulatory system within germline stem cells and how these cells differ between schistosomes and liver flukes could enhance our understanding on how these worms maintain their life cycles. By analyzing gene expression of the early germ cells of both the male and female reproductive organs, orthologues of the key genes in the schistosome germ cells were shown to be transcribed in liver fluke cells. This included three *nanos* paralogs that were expressed in the stem-cell clusters and could be linked to their orthologue in schistosomes via phylogenetic analysis. While *nanos1* is a marker for germline stem cells in *S. mansoni [185],* this distinction could not be drawn here, which might be complicated by the shared set of genes between somatic and germline stem cells described before [109]. The use of other scRNA-seq methods, with greater sensitivity, has been shown to resolve stem-cell subpopulations [109,114]. Additionally, as RNA *in situ* hybridization for the *nanos* orthologues failed, other methods like hybridization chain reaction (HCR) should be considered to detect more challenging transcripts [274].

Other key regulators for which orthologs were identified were the RNA-binding protein Boule and a transcription factor classified as a Onecut family member. These two proteins inhibit proliferation in germline stem-cells and drive differentiation in the schistosome germline [114]. This is in contrast to the role of these regulators in planaria, where these regulators are required for stem cell maintenance [114,192]. Since their cellular expression did hint at a function in germline stem cells, it might be interesting to investigate the function of the Boule and Onecut also in the liver fluke.

Further intriguing aspects found by analyzing the data mainly concerned the female germline. One aspect was the specific expression of selected nuclear receptors within the late female germ-cell cluster. The expression of nuclear receptors in the germline was also found in *S. mansoni,* where orthologues of FTZF-1 and the thyroid hormone receptor were also found in the mature part of the ovary [275]. Their expression was mapped to the intermediate stage of schistosomal oocytes by scRNA-seq (Möscheid and Lu *et al*., unpublished data), which might hint at a mixed cell population in the present dataset. Nuclear receptors have been shown to be important for correct egg formation in trematodes [276]. An interesting feature is the apparent duplication of the thyroid hormone receptor after the evolutionary split from the free-living turbellaria [205,277]. While trematodes lack essential parts of the thyroid hormone synthesis pathway, like the thyroid peroxidase, it has been shown that host-derived thyroid hormones severely impact the growth of *S. mansoni [207,278]*. This implies that host-derived T3 or T4 might be the natural ligands of these parasite receptors.

Finally, all components of the canonical Wnt pathway could be identified, with several genes being preferentially transcribed within the late female germ cell cluster. While being mostly associated with development, the Wnt pathway is associated with numerous different functions. For example, it was shown to be essential for oocyte differentiation in mammalian reproduction [279]. Although a frizzled receptor (receptor for Wnt) could be localized in the reproductive organs of *S. japonicum* by IHC [280], no comprehensive study has been performed on its function in reproduction. The observation of expressed GPCRs within the female germ-cell clusters further reinforces the role of external stimuli, either worm or host derived as regulators for sexual reproduction. Another interesting observation on the Wnt pathway stems from research in the evolutionary basis of regeneration in several planarian species [281]. There the authors discovered an interesting dichotomy in the biological function of ß-catenin. This protein is central to the Wnt pathway, and upon activation it accumulates in the nucleus and influences gene expression [201]. In some planarians that lost regenerative ability, RNAi of ß-catenin can recover the ability of the worms to regenerate their heads. Though this was only obsered in assexual species, sexual worms depend on ß-catenin for the formation and maintenance of germ-cells. The discovery of a functioning

Wnt-pathway in the female germline of the liver fluke in this work could deliver a working hypothesis for the evolutionary loss of regeneration capacities in parasitic flatworms.

4.2.3 Analysis of the differentiation lineage in vitelline cells

Besides the major reproductive tissues, namely the ovary and testes, the vitellarium is represents another important structure of the flatworm reproductive system. This tissue is the source of vitellocytes, which provide a source of nutrients for the developing embryo and play a role in eggshell formation [282]. A unique feature of flatworms is the production of ectolecithal eggs, consisting of an oocyte and several vitellocytes [283]. In the liver fluke, one oocyte is surrounded by 30 vitelline cells [230], and considering the immense output of eggs [24], this results in a tremendous need for vitelline cells. The differentiation lineage of the vitellocytes is a continuous process involving four cell types, as has been determined by microscopical studies [181]. The described cell types were termed S for stem cell, intermediate cells 1 and 2, and mature cells, which mirrors the number and description of liver fluke vitellocytes in schistosomes where these cells are numbered S 1-4 cells [182]. This apparent similarity in cell types sparked interest in how these cells resemble each other on a molecular level. Initial analysis of the scRNA-seq data identified two clusters of vitellocytes termed early and late vitellocytes. Expression of schistosomal markers *vf1* and tyrosinases [206,207,284] showed a conserved expression of these markers within the lineage of these cells. The nuclear receptor Vf1 was initially identified in *S. mansoni* and might drive a pairing-induced gene regulatory program that drives the differentiation of vitellocytes. This points to the evolutionary conservation of the vitelline lineage and might have significance in the development of a pan-trematode therapeutic targeting the vitellarium to interrupt the parasite life cylce.

While the vitellocyte clusters were described as a lineage up to this point, scRNA-seq only captures snapshots of a cell at the time of cell lysis, and dynamical relationships are challenging to infer [285]. The method of RNA velocity tries to infer the trajectory of differential lineages by using spliced and unspliced counts of genes [134,135]. Here, this was applied to the proposed lineage of vitellocytes. Based on RNA velocity of the vitelline cells, there appeared to be two trajectories rooted within the subcluster 0, which had a marked expression of the nuclear receptor *vf1*. The global projection, the first trajectory appeared to be connected to the stem cell cluster, while the second trajectory was linked to more differentiating vitellocytes. Additional information on unspliced reads revealed a suppression of gene expression of the nuclear receptor precisely at the split of the two lineages. Considering the transcription of *vf1* in schistosomes, this receptor is also coexpressed with *h2b* [207]. Because RNAi of schistosomal *vf1* expression did not affect the

number of EdU-positive cells as a readout for cell proliferation, it was concluded on a role in differentiation. The potential expression within cycling cells within the liver fluke was also suggested by cell cycle scoring, leading to the following interpretation of the results: The two trajectories are rooted at the transitionary phase between differentiation and proliferation in S1 cells, one trajectory showing the final differentiation into mature S4 cells. In contrast, the other trajectory can be interpreted as stem cells undergoing self-renewal, starting with a cell cycle entry, shown by the high score for the S phase. The apparent lack of cells directing back through the transition phase could be due to a lack of sampling, which might be alleviated in the new dataset. Even though RNA velocity can be an interesting tool for validating potential differentiation trajectories or identifying novel driver genes, the method used here heavily relies on the model assumption that developmental progress is comparable to the half-life of RNA molecules, and violated model assumptions might cause results that fail to capture the biological truth [167,286]. Updated tools like cellrank2 [287] could be applied to the data to confirm the results, elucidate vitellocyte differentiation, or use these tools for other cells within the dataset.

Taken together there appears to be a marked conservation of regulators in the germline between schistosomes and the liver fluke. Shown by cell type specific expression of key TFs like Onecut or nuclear receptors including Vf1. Compared to the divergence of germline regulation in free-living flatworms [114], this could have implications considering the adaptation of reproduction to a parasitic lifestyle in trematodes.

4.3 Application for drug discovery

The scRNA-seq data presented here can aid with drug target discovery. With knowledge of cell type-specific transcription available to researchers, it is possible to identify essential targets, link them to biological processes, and thereby identify potential drug targets for further evaluation. As it is beyond the scope of this discussion to evaluate every potential target presented in this work, specific trends in the data are pointed out and analyzed in further detail. The expression of protein kinases within the dataset was particularly interesting, as these have been the focus of drug development for other diseases, and several small-molecule inhibitors were already developed for their inhibition [77]. Additionally, our group has shown the utility of kinases as a potential drug target in the liver fluke, and knowledge of the cell-specific expression of these proteins might greatly facilitate the discovery of suitable drug target candidates [80,121].

4.3.1 Potential drug targets in reproductive organs

The reproductive system of the liver fluke is a highly energy-consuming machinery producing an enormous number of eggs. The production of eggs and their secretion into the environment is essential for the parasite's life cycle. As such, targeting the reproductive organs could be a credible treatment strategy for liver fluke control, especially to prevent animal reinfection and control disease epidemiology. This work thoroughly captured several cells of the reproductive tissue and made comparisons to the schistosomal cells of the same tissues. Potential homologies were identified that could indicate the cross-species efficacy of treatments targeting these tissues. Notably the main flukicide TCBZ causes damage to reproductive tissues, which included a reduced number of cells and apoptosis within the tissue [288].

Here, a p38 kinase was described as preferentially expressed in the oocyte cells. Together with a functional network predicted by STRING analysis, this strongly suggests the role of the p38 pathway in female germ-cell development. The kinase p38 is a member of the MAP kinase family, a family of kinases involved with diverse biological processes like cell proliferation, differentiation, or stress response [289]. A main feature of MAP kinases is their involvement in pathways processing extracellular signals [289]. This central role of these kinases has sparked interest in their potential as drug targets in parasitic trematodes. In schistosomes, p38 was implicated in sexual maturation of female worms [290], oviposition, and embryonic development [203,291]. Due to its role in inflammatory diseases, several inhibitors have been developed for human p38 [292], and it might be promising to validate their efficacy against schistosomes as well as *F. hepatica in vitro.*

Another interesting set of proteins that were identified to be preferentially expressed within the germline were nuclear receptors, which were adresses already above. The role of nuclear receptors in several biological processes might make them a valuable drug target [207,278]. Despite these potential benefits, there are several challenges in targeting them for drug development. One of these challenges is the identification of the receptor ligands, which is further complicated by the formation of heterodimers of several nuclear receptors [293]. Knowledge of the exact ligands bound by nuclear receptors is critical for developing potential therapeutics, as antagonistic molecules have the most potential to be effective, as highlighted by already described molecules in use [294]. Further difficulties might arise due to homologies between human and parasite nuclear receptors [205], which might lead to side effects and must be thoroughly explored. A possibility might arise as trematodes are the only known organisms to encode two thyroid hormone receptors, that might be structurally distinct to mammalian receptors and could be exploited for drug target discovery. While the

role of nuclear receptors in liver fluke biology is an interesting line of work in itself, their role as drug targets is at best, hypothetical and needs to be carefully evaluated.

4.3.2 Transcriptional profile of gut cells reveals insights into lipid uptake and metabolism

Upon inspection of gastrodermal cells, gene related to lipid metabolism, and nutrient uptake were found to be explicitly expressed in these cells. The metabolism of lipids is particularly interesting, as the parasite does not synthesize its lipids *de novo* but takes them from the host, thereby altering the host's metabolism [295]. However, the detection of parasite-specific lipids indicated that the worm might modify host-derived lipids [296]. Nevertheless, the exact cellular mechanisms of these modifications are not clear. The expression of a phospholipase b-like protein in the liver fluke's gut cells might be important in the metabolism of host-derived lipids. Indeed, there have been descriptions of the production of fatty acids, especially parasite-derived eicosanoids, during infection [297]. Interestingly, an increase in thromboxane B2 to prostaglandin I2 ratio was noted between 8 to 11 weeks post-infection, which might be involved with the increased transcriptional expression in adult worms compared to younger ones [91,297]. Eicosanoids have been linked to immunomodulation and disease progression in helminth infections. The microfilaria of *Wuchereria bancrofti* and *Brugia malayi* have been shown to modulate the immune system by releasing prostaglandins, one group of eicosanoids [298,299]. Additional immunomodulatory effects were mediated by prostaglandin released from *S. mansoni* schistosomula during tissue migration [300]. The extent to which the identified phospholipase b participates in the production of immunomodulatory eicosanoids remains to be explored, but it might be a promising line for further drug discovery. The class of phospholipases has been researched in the context of inflammatory diseases or microbial infections. Several inhibitors have been developed, and some have also entered clinical trials [301]. It might be promising to investigate their efficacy in inhibiting the liver fluke phospholipases. This is further underlined by the expression of these enzymes within the gastrodermis of the worm, which is in direct contact with the host environment and might have promising accessibility for small molecules such as inhibitors.

Transcripts of genes encoding for a group of ML domain-containing proteins were also detected in the gastrodermal cells of the liver fluke. These domains have been implicated in cholesterol recognition and binding [302]. Additionally, these groups of proteins have been predicted to be part of the excretory/secretory product of *C. sinesis* [303]. Other lipid binding proteins like two major isoforms of fatty acid-binding proteins (FABP) are part of the excretory/secretory product of *F. hepatica* and have a role in nutrient acquisition. Additionally,

several FABP-mediated immunomodulatory effects were described, for example, by the suppression of cytokines in macrophages or promoting a Th2- immune response in dendritic cells [304–306]. Given the essential roles of lipid-binding proteins in the liver fluke, this newly recognized group of proteins might be an interesting avenue of research in host-parasite interaction. A downside that has to be noted is that the transcripts of all these genes are mostly limited to the immature and adult stages of the parasite, which might limit the utility of treatment strategies in targeting earlier stages.

4.3.3 Identification and evaluation of a highly druggable cell cluster

A previously unrecognized cell population was described, which was deemed highly valuable for drug target discovery. This discovery was initially driven by the preferential expression of $FhTRPM_{PZQ}$, which is the recently identified orthologue gene to the target of praziquantel in *S. mansoni [75]*. Characterizing these cells might deliver insights into the biological role of these cells and lead to the identification of promising drug targets. As the schistosomal $TRPM_{PZQ}$ is expressed in several cell types, including muscle and neurons, elf5 was used as a marker and for RNA *in situ* experiments. This identified elf5+ cells in proximity to both muscle and nerve cells, which suggested a neuromuscular role of these cells. Therefore, the function of these cells in critical biological processes like motility, attachment, and general interaction with the host environment could be expected, making it a promising target for drug target discovery. Further analysis of marker genes for this cluster underlined the potential role of these cells in response to external stimuli. A recent study also provided additional evidence for the proposed activation of $TRPM_{PZQ}$ by membrane stretch [225].

To gain further insight into the biological pathways underlining the function of these cells, a STRING network was constructed using genes enriched within this cluster and associated with signal transduction. Also 16 protein kinases expressed in this cluster, had several kinases which were found to have a role in calcium signaling. As mentioned earlier, protein kinases have been discussed as drug targets in helminths, including filaria, cestodes, and trematodes [77]. Their role in various biological processes, such as cell cycle, migration, and signal transduction, as well as the availability of small-molecule inhibitors from cancer research, make protein kinases a valuable target for drug repurposing [77]. The evaluation of kinases as a drug target in the liver fluke is in its infancy with only few published studies. An earlier study evaluated a phosphofructokinase as a potential drug target, but was not further followed due to the inefficiency of the used inhibitor *in vivo* [307–309]. Our lab has recently shown the variable transcription of several kinases across life stages, including *abl1* and *plk1* [121]. Furthermore, as a follow-up study, we evaluated the efficacy of two inhibitors of said kinases on the liver fluke [80]. The Abl1 inhibitor imatinib, which had promising

antischistosomal properties [310], was shown to be lethal for immature and adult *F. hepatica* [80], while the Plk1 inhibitor BI2536 was lethal for immature worms but only had effects on egg production in adult worms [80].

Here, several kinases that have not been the focus of drug target discovery before were identified and found to be expressed in the elf5+ cell cluster. This included a protein kinase C orthologue for which a role for body contraction and muscle function has been proposed for the liver fluke and schistosomes [215,311], and we showd recently that the PKCß inhibitor ruboxistaurin was lethal against adult *F. hepatica in vitro* [312]. Another example represents the focal adhesion kinase, which was found by STRING to be at the center of a network connected to kinase signaling mediated by receptor tyrosine kinases, but also processes involved with calcium permeable kainate receptors. The focal adhesion kinase is involved in calcium signaling in the human system [313]. Inhibitors of the focal adhesion kinase have been shown to have efficacies against several *Schistosoma* species, with comparable effects to PZQ after 24 h [314]. However, these tests were performed with high inhibitor concentrations and have to be carefully evaluated. Also, comparable changes in phosphorylation patterns have been notated between worms treated with a focal adhesion kinase inhibitor and PZQ, which could indicate a role of kinase activity in response to the calcium influx caused by PZQ [314]. This response of kinases to PZQ was also noted for the calcium/calmodulin-dependent protein kinase II [315]. The potential role of kinase signaling in response to PZQ exposure might be interesting for drug target development, further highlighting the importance of protein kinases in fluke biology.

4.3.4 PAK4 is a kinase with potential importance for fluke survival

As a way to investigate the druggability of the elf5+ cluster, the kinase PAK4 was evaluated as a potential drug target by testing the inhibitor LCH-7749944 *in vitro*. The family of PAKs has not been analyzed in parasitic flatworms to date. A total of 4 PAK kinases was identified based on the presence of a p21 binding site and a kinase domain. A phylogenetic analysis of the identified PAK sequences from model organisms, planarians, and *S. mansoni* showed that platyhelminths have the three typical group 1 PAKs. In contrast, the group 2 is only represented by one kinase. Insights into the role and function of PAKs are primarily based on human and mouse studies where group 1 PAKs are functionally associated with GTPase-mediated changes in cell morphology, motility, and cell-cycle progression [236,316,317]. Group 2 PAKs can be regulated by autophosphorylation and are not dependent on GTPase binding [316]. Functionally, they are associated with cytoskeletal reorganization, cell survival, and are important in tumor progression [236]. The finding that

flatworms just possess one group 2 PAK member could give insights into the evolutionary history of PAK, even though it remains to be elucidated if the apparent lack of group 2 PAK members is due to secondary loss of these genes or more group 2 PAKs have evolved in vertebrates. The four PAKs expressed in the liver fluke had some differences in their cellular expression patterns, for example, the transcripts of *pak3* were preferentially expressed in the oocyte cluster. This is especially interesting given the presence of GTPases in the ovary, which could be directly related to the function of PAK3. Focusing on *pak4*, RNA *in situ* hybridization showed transcripts in the gut, oocytes, and cells below the tegument, most likely cells of the elf5+ cluster. However, this remains to be proven by double FISH experiments. The expression in these cells suggested a vital role of this kinase in important tissues for the worms' survival and reproduction. Bulk-RNAseq data of liver flukes previously supported the transcription of *pak4* in eggs [91]. While unexpected at first sight, the detection of transcripts in the nucleus of oocytes, here by ISH, is also known from zebrafish [318].

Some progress has been made in developing small-molecule inhibitors against human PAK4 to infer with cancer progression. A commercially available inhibitor for PAK4 was selected for further evaluation of the druggability of liver fluke PAK4. The inhibitor LCH-7749944 had promising lethal effects against adults and immatures using 50 µM after 72 h and against NEJs after 2 h using the same concentrations. The short-acting time of the inhibitor against NEJs might indicate the efficacy of far lower concentrations against this stage. The used concentrations are comparable to the *in vitro* working concentrations of drugs currently used against the liver fluke, namely TCBZ or closantel [80]. Additionally, the inhibitor had substantial effects on schistosomes and planarians, which might indicate some promise in the pan-flatworm activity of this inhibitor. While the inhibitor has been shown to specifically inhibit PAK4 in humans cells at similar high concentrations as used against liver flukes [319–321], there might be the potential for effects caused by binding on proteins other than FhPAK4. The binding site of the inhibitor was largely conserved only between human and liver fluke PAK4 but not in other liver fluke PAKs [247]. However, the hinge region within the ATP-binding pocket was conserved between all analyzed sequences. Further analysis of the binding of the inhibitor is needed to be sure of its specificity, which could be enzyme activity assays of the liver fluke proteins exposed to LCH-7749944, western plots using treated worms that could detect the PAK phosphorylation state, or thermal shift assays between all liver fluke PAKs to show which enzymes are bound by the inhibitor. Furthermore, new molecules have been developed based on the structure of LCH-7749944, which could also be evaluated for their efficacy at lower concentrations [247] or chemically modified to fit the liver fluke enzyme.

Finally, to elucidate the essential role of FhPAK4 for the liver fluke, a knockdown was done in immature worms *in vitro,* which had no visible effect. The transcript levels were reduced by 70%, which could be increased by using other dsRNAs or increasing dsRNA concentration. Additionally, the reduction in protein level was not verified, and the remaining protein level could explain the absence of a phenotype. Another problem was illustrated by the apparent upregulation of *pak4* transcripts control worms treated with the non-helminth neomycin dsRNA. In a side-project during the period of this work, I could contribute to investigations on the use of non-helminth dsRNAs in *S. mansoni* [161]. These controls are essential in RNAi experiments to control for non-target-related effects caused by dsRNAs [322,323]. In this project, it was found that sequences homologous to *neomycin* caused changes in gene expression in schistosomes. Even though the liver fluke transcriptome presents another background, the results could be an argument for an optimized control study in *F. hepatica.*

Based on the promising results of LCH-7749944, PAK4 might be an interesting target for further drug discovery in liver fluke research, but further investigation is needed. The use of more optimized inhibitors, as well as further validation of the inhibitor's specificity, could be promising steps to develop this idea further. Additionally, optimized RNAi conditions could help elucidate the biological function of this kinase.

5. Outlook

The workflow and dataset presented here are the first descriptions of the cellular landscape of *F. hepatica* on a transcriptional level. This knowledge can be utilized in various ways to better understand the cellular biology of this parasitic flatworm and will serve as a future resource.

A potential application could build upon the prediction and prioritization of drug targets illustrated in this work for the protein kinase FhPAK4. While this project focused on just one family of proteins, the protein kinases, and within just one cell cluster, this approach could be explored on a bigger scale. Databases such as ChEMBL or DrugBank [324,325] contain information on inhibitors, drugs, and their targets. Based on a selection of potential druggable cell types, such as the elf5+ cluster but also muscle or intestinal cells, targets could be identified based on the availability of inhibitors stored in said databases and tested for efficacy *in vitro.*

Another potential application might be the comparison of cell types between different species, especially flatworms like schistosomes or planarians, for which scRNA-seq datasets are already available. This comparison could answer questions such as how parasites adapt to their unique environment on a cellular level. Of particular interest could be transcriptional differences in germ cells between parasitic flatworms and free-living species, which could reveal the adaptation of reproductive strategies.

The established workflow will also serve as template for further single-cell RNAseq-based research directions aiming at the generation of new biological insights. This could include elucidating the mode of action underlying the resistance to TCBZ. While bulk-RNAseq datasets comparing resistant and susceptible strains exist [326], there is no clear understanding of which cells are driving the resistance. Using scRNA-seq could shed light on how cells differ between resistant and susceptible parasites and further our understanding of TCBZ resistance, e.g. related to cell-type specific expression levels of drug efflux transporters. While the data presented focused on adult worms, the workflow could also be adapted to other lifestages like NEJs and immature worms. This would enable us to understand developmental processes on a cell scale, identifying how cell types adapt and differ during development.

Taken together, the applied methodology and obtained data can guide the development of novel research projects and serve as a tool to answer new and exciting questions.

6. Summary/Zusammenfassung

6.1 Summary

The liver fluke *Fasciola hepatica* is a widespread parasite affecting humans and animals. It causes considerable economic loss by infecting livestock. The World Health Organization also recognizes fasciolosis as a neglected tropical disease. Successful treatment of the disease heavily relies on the drug triclabendazole, and the widespread development of resistance makes the discovery of new treatment strategies essential.

Transcriptome analysis has the potential to reveal fundamental aspects of parasite biology and identify new drug targets. By design, classical RNAseq methods analyze gene expression of a mixture of various cells, which can thus only deliver transcriptome information of tissues or entire organisms without discrimination of different cell types. New technologies allow the analysis of the transcriptome of individual cells. This information gives scientists a new perspective for research on previously neglected organisms.
This work aimed to establish and characterize a single-cell transcriptome dataset for the liver fluke *F. hepatica*. For this purpose, a method for isolating and enriching cells from the adult parasite stage using the protease trypsin was established. Cellular composition, quality, and RNA integrity were optimized for transcriptome analysis. In addition, the fixation of cells with acidic methanol was identified as useful approach to preserve cells for transportation and later analysis.
Single-cell transcriptome analysis using the 10x Chromium technology was performed with separately dissociated anterior and posterior parts of the liver fluke to enrich for rare cells. A total of 15 cell clusters over 19,581 cells were identified and 8716 genes were expressed. Among the identified clusters were cells of the reproductive organs, including testes (4 clusters), ovary (2 clusters), vitellarium (2 clusters), and one cluster representing cells of the Mehlis‘ gland. Furthermore, stem cells, muscle cells, cells of the gastrodermis, and a previously unknown cell population named elf5+ due to the expression of an ETS transcription factor were identified. Marker genes were determined for the identified cells and verified for selected cell types using RNA *in situ* hybridization. These marker genes included a phospholipase-B-like gene previously unrecognized as a gene expressed in the gastrodermis, which might have implications for immunosuppression of the liver fluke in the final host. Classical markers like tektin for the male gem cells or the classical muscle markers myosin and collagen were verified. Gene ontology analysis with transcriptional markers of the clusters allowed a functional characterization of the clusters, allowing for a

better distinction between differentiated and proliferating cells. A joint approach of gene ontology analysis with STRING analysis was used to characterize signal transduction networks in both female germ cells and the elf5+ cluster. Leveraging this approach led to the identification of a signaling network, including the kinase p38, that might be involved in female germ-cell function or enrichment for focal adhesion signaling in the elf5+ cluster.
Furthermore, an RNA velocity analysis indicated a differentiation lineage of the vitellocytes. Two lineages were predicted over five subclusters. One lineage included clusters characterized by expression of the *histone 2b* and a nuclear receptor named Vf1 and was annotated as S1 vitelline cells reentering the cell cycle, while a second lineage over clusters with marked expression of two tyrosinases was assigned to maturing vitellocytes. Finally, the dataset's utility for discovering new treatment strategies was demonstrated. The newly described elf5+ cell cluster was identified as promising for drug discovery based on preferential transcription of an orthologue of the previously described drug target of praziquantel, $FhTRPM_{PZQ}$. Interest was focused on the specific expression of various protein kinases in this cell cluster. The p21-activated kinase PAK4 was tested for its potential as a drug target using the inhibitor LCH-7749944. The inhibitor significantly affected all life stages of the liver fluke that are found in the final host when exposed *in vitro* to concentrations of 25 to 50 µM. Preliminary data also suggested the activity of LCH-7749944 against related flatworm species. This confirms the utility of the *F. hepatica* single-cell data as a basis for drug discovery.

In summary, the first transcriptional characterization of single cells from liver flukes was achieved. A method for isolating liver fluke cells and adapting the technique to the worm's anatomical characteristics led to the first description of important cell types. This data set represents an important resource for the fundamental understanding of liver flukes and may facilitate the discovery of new active substances.

6.2 Zusammenfassung

Der Leberegel *Fasciola hepatica* ist ein weltweit verbreiteter Parasit, der sowohl Menschen als auch Tiere befällt. Durch die Infektion von Nutztieren verursacht er erheblichen wirtschaftlichen Schaden. Zudem ist die Erkrankung Fasciolose von der Weltgesundheitsorganisation als vernachlässigte Tropenerkrankung anerkannt. Eine erfolgreiche Behandlung der Erkrankung hängt in hohem Maße vom Medikament Triclabendazol ab, wobei die beginnende Entwicklung von Resistenzen die Entdeckung neuer Behandlungsstrategien essentiell macht.

Transkriptomanalysen besitzen ein hohes Potential für die Erforschung grundlegender Aspekte der Biologie von Parasiten und die Identifizierung neuer Wirkstoffziele. Klassische RNAseq-Methoden analysieren die Genexpression einer Mischung verschiedener Zellen, d.h. man erhält Transkriptominformation von Geweben oder ganzer Organismen ohne Differenzierung einzelner Zellpopulationen. Neue Technologien erlauben dagegen die Analyse des Transkriptoms einzelner Zellen. Informationen dieser Art geben Wissenschaftlern eine neue Perspektive für die Forschung an bisher vernachlässigten Organismen.
Ziel dieser Arbeit war die Etablierung und Beschreibung eines Einzelzell-Transkriptom-Datensatzes für den Leberegel *F. hepatica*. Dafür wurde eine Methode für die Vereinzelung und Anreicherung von Zellen mittels der Protease Trypsin etabliert. Zelluläre Komposition, Qualität und RNA-Integrität wurden für die Verwendung in einer Transkriptomanalyse optimiert. Zusätzlich wurde eine Fixierung von Zellen mit saurem Methanol als nützliche Methode identifiziert, um Zellen für den Transport und spätere Analyse aufzubewahren. Die Einzelzell-Transkriptomanalyse mit der 10x Chromium-Technologie wurde mit separat dissoziierten anterioren und posterioren Teilen des Leberegels durchgeführt, um auch seltene Zellen zu finden. Insgesamt wurden in den 19.581 analysierten Zellen 15 Cluster identifiziert und 8.716 Gene exprimiert. Zu den identifizierten Clustern gehörten Zellen der Fortpflanzungsorgane, darunter Testis (4 Cluster), Ovar (2 Cluster), Vitellarium (2 Cluster) und ein Cluster mit Zellen der Mehlis-Drüse. Darüber hinaus wurden Stammzellen, Muskelzellen, Zellen der Gastrodermis und eine bisher unbekannte Zellpopulation identifiziert, die aufgrund der Expression eines ETS-Transkriptionsfaktors mit elf5+ bezeichnet wurde. Für die identifizierten Zellen wurden Markergene bestimmt und für ausgewählte Zelltypen mittels RNA *in situ* Hybridisierung verifiziert. Zu diesen Markergenen gehörte ein Phospholipase-B-ähnliches Gen, das bisher nicht als ein in der Gastrodermis exprimiertes Gen bekannt war und das Auswirkungen auf die Immunsuppression des Leberegels im Endwirt haben könnte. Außerdem wurden klassische Marker wie Tektin für die

männlichen Geschlechtszellen oder die klassischen Muskelmarker Myosin und Kollagen nachgewiesen.

Eine Gen-Ontologie-Analyse mit Transkriptionsmarkern der Cluster ermöglichte eine funktionelle Charakterisierung der Cluster, die unter anderem eine bessere Unterscheidung zwischen differenzierten und proliferierenden Zellen erlaubte. Ein kombinierter Ansatz aus Gen-Ontologie-Analyse und STRING-Analyse wurde verwendet, um Signaltransduktions-netzwerke sowohl in weiblichen Keimzellen als auch im elf5+-Cluster zu charakterisieren. Die Nutzung dieses Ansatzes führte zur Identifizierung eines Signalnetzwerks, das die Kinase p38 einschließt, die möglicherweise an der Funktion der weiblichen Keimzellen beteiligt ist, sowie zur Identifikation des fokalen Adhäsionssignalwegs im elf5+-Cluster.

Darüber hinaus wies eine “RNA-Velocity”-Analyse auf zwei unterschiedliche Differenzierungslinien in den Vitellozyten-Clustern hin, die insgesamt fünf Subcluster umfassen. Eine Linie umfasst Cluster, die durch die Expression des *Histons 2b* und eines nukleären Rezeptors namens Vf1 gekennzeichnet sind und als S1 Vitellinzellen annotiert wurden, die in den Zellzyklus eintreten. Die zweite Linie beeinhaltet Cluster mit ausgeprägter Expression zweier Tyrosinasen und wurde reifenden Vitellinzellen zugeordnet.

Der neu beschriebene Cluster elf5+ wurde aufgrund der präferenziellen Transkription eines Orthologs des zuvor beschriebenen Wirkstoffziels von Praziquantel, $FhTRPM_{PZQ}$, als vielversprechendes Ziel für weitere Wirkstoffentwicklung identifiziert. Dabei fiel das Interesse auf die spezifische Expression einer Vielzahl von Proteinkinasen in diesem Zelltyp. Die p21-aktivierte Kinase PAK4 wurde mittels des Inhibitors LCH-7749944 hinsichtlich ihres Potentials als Wirkstoffziel getestet. Der Inhibitor schädigte alle Lebensstadien des Leberegels im Endwirt signifikant, wenn er *in vitro* in Konzentrationen von 25 bis 50 µM eingesetzt wurde. Vorläufige Daten deuteten außerdem auf die Aktivität von LCH-7749944 gegen verwandte Plattwurmarten hin. Dies bestätigt den Nutzen der Einzelzelldaten von *F. hepatica* als Grundlage für die Entdeckung von Wirkstoffen.

Zusammenfassend gelang erstmals eine transkriptionelle Analyse von Einzelzellen des Leberegels. Eine Methode zur Zellvereinzelung sowie die Anpassung der Methode an die anatomischen Eigenschaften des Wurms wurde erfolgreich etabliert und eine erste Beschreibung wichtiger Zelltypen erzielt. Dieser Datensatz stellt eine wichtige Ressource für die Grundlagenforschung zum Leberegel dar und kann in Zukunft die Entdeckung neuer Wirkstoffe erleichtern.

7. References

1. Araújo A, Reinhard K, Ferreira LF. Paleoparasitology – human parasites in ancient material. Advances in Parasitology Volume 90. London, Netherlands: Elsevier; 2015. pp. 349–387.

2. Hotez PJ, Brindley PJ, Bethony JM, King CH, Pearce EJ, Jacobson J. Helminth infections: the great neglected tropical diseases. J Clin Invest. 2008;118: 1311–1321.

3. Global report on neglected tropical diseases 2024. [cited 26 Aug 2024]. Available: https://www.who.int/teams/control-of-neglected-tropical-diseases/global-report-on-neglected-tropical-diseases-2024

4. Laumer CE, Hejnol A, Giribet G. Nuclear genomic signals of the “microturbellarian” roots of platyhelminth evolutionary innovation. Elife. 2015;4. doi:10.7554/eLife.05503

5. Cribb TH, Chisholm LA, Bray RA. Diversity in the Monogenea and Digenea: does lifestyle matter? Int J Parasitol. 2002;32: 321–328.

6. Littlewood DTJ, Rohde K, Bray RA, Herniou EA. Phylogeny of the Platyhelminthes and the evolution of parasitism. Biol J Linn Soc Lond. 1999;68: 257–287.

7. Koziol U. Evolutionary developmental biology (evo-devo) of cestodes. Exp Parasitol. 2017;180: 84–100.

8. Sotelo J, Del Brutto OH, Roman GC. Cysticercosis. Curr Clin Top Infect Dis. 1996;16: 240–259.

9. Eckert J, Deplazes P. Biological, epidemiological, and clinical aspects of echinococcosis, a zoonosis of increasing concern. Clin Microbiol Rev. 2004;17: 107–135.

10. Foodborne trematode infections. [cited 26 Aug 2024]. Available: https://www.who.int/news-room/fact-sheets/detail/foodborne-trematode-infections

11. Mas-Coma S. Epidemiology of fascioliasis in human endemic areas. J Helminthol. 2005;79: 207–216.

12. Calvani NED, Šlapeta J. *Fasciola* species introgression: Just a fluke or something more? Trends Parasitol. 2021;37: 25–34.

13. Cwiklinski K, O’Neill SM, Donnelly S, Dalton JP. A prospective view of animal and human Fasciolosis. Parasite Immunol. 2016;38: 558–568.

14. Santiago Mas-Coma, M. Adela Valero, and M. Dolores Bargues. Fascioliasis. In: Toledo R, Fried B, editors. Digenetic trematodes. Cham: Springer International Publishing; 2024. pp. 157–201.

15. Rowcliffe SA, Ollerenshaw CB. Observations on the bionomics of the egg of *Fasciola hepatica*. Ann Trop Med Parasitol. 1960;54: 172–181.

16. Smith G, Grenfell BT. The influence of water temperature and pH on the survival of *Fasciola hepatica* miracidia. Parasitology. 1984;88: 97–104.

17. Hussein A-NA, Hassan IM, Khalifa RMA. Development and hatching mechanism of *Fasciola* eggs, light and scanning electron microscopic studies. Saudi J Biol Sci. 2010;17: 247–251.

18. Anderson RM, Mercer JG, Wilson RA, Carter NP. Transmission of *Schistosoma mansoni* from man to snail: experimental studies of miracidial survival and infectivity in relation to larval age, water temperature, host size and host age. Parasitology. 1982;85: 339–360.

19. Correa AC, Escobar JS, Durand P, Renaud F, David P, Jarne P, et al. Bridging gaps in the molecular phylogeny of the Lymnaeidae (Gastropoda: Pulmonata), vectors of Fascioliasis. BMC Evol Biol. 2010;10: 381.

20. Mas-Coma S, Valero MA, Bargues MD. Fascioliasis. Adv Exp Med Biol. 2019;1154: 71–103.

21. Luzón-Peña M, Rojo-Vázquez FA, Gómez-Bautista M. The overwintering of eggs, intramolluscal stages and metacercariae of *Fasciola hepatica* under the temperatures of a Mediterranean area (Madrid, Spain). Vet Parasitol. 1994;55: 143–148.

22. Hodgkinson JE, Cwiklinski K, Beesley N, Hartley C, Allen K, Williams DJL. Clonal amplification of *Fasciola hepatica* in *Galba truncatula*: within and between isolate variation of triclabendazole-susceptible and -resistant clones. Parasit Vectors. 2018;11: 363.

23. Howell AK, Williams DJL. The epidemiology and control of liver flukes in cattle and sheep. Vet Clin North Am Food Anim Pract. 2020;36: 109–123.

24. Happich FA, Boray JC. Quantitative diagnosis of chronic fasciolosis The estimation of daily total egg production of *Fasciola hepatica* and the number of adult flukes in sheep by faecal egg counts. Aust Vet J. 1969;45: 329–331.

25. Mas-Coma S, Adela Valero M, Dolores Bargues M. *Fasciola* and *Fasciolopsis*. 1st Edition. Biology of Foodborne Parasites. 1st Edition. CRC Press; 2015. pp. 386–419.

26. Fascioliasis. 3 May 2019 [cited 3 Sep 2024]. Available: https://www.cdc.gov/dpdx/fascioliasis/index.html

27. Sangster N, Martínez-Moreno Á, Pérez J. Pathology, pathophysiology and clinical aspects. Dalton JP, editor. Fasciolosis. 2021; 145–146.

28. Webb CM, Cabada MM. Recent developments in the epidemiology, diagnosis, and treatment of *Fasciola* infection. Curr Opin Infect Dis. 2018;31: 409–414.

29. Rosas-Hostos Infantes LR, Paredes Yataco GA, Ortiz-Martínez Y, Mayer T, Terashima A, Franco-Paredes C, et al. The global prevalence of human fascioliasis: a systematic review and meta-analysis. Ther Adv Infect Dis. 2023;10: 20499361231185413.

30. Mas-Coma MS, Esteban JG, Bargues MD. Epidemiology of human fascioliasis: a review and proposed new classification. Bull World Health Organ. 1999;77: 340–346.

31. Raue K, Heuer L, Böhm C, Wolken S, Epe C, Strube C. 10-year parasitological examination results (2003 to 2012) of faecal samples from horses, ruminants, pigs, dogs, cats, rabbits and hedgehogs. Parasitol Res. 2017;116: 3315–3330.

32. *Fasciola hepatica* in Ruminants. In: MSD Veterinary Manual [Internet]. [cited 3 Sep 2024]. Available: https://www.msdvetmanual.com/digestive-system/fluke-infections-in-ruminants/fasciola-hepatica-in-ruminants

33. Clery D, Torgerson P, Mulcahy G. Immune responses of chronically infected adult cattle to *Fasciola hepatica*. Vet Parasitol. 1996;62: 71–82.

34. Doy TG, Hughes DL. *Fasciola hepatica*: site of resistance to reinfection in cattle. Exp Parasitol. 1984;57: 274–278.

35. Piedrafita D, Parsons JC, Sandeman RM, Wood PR, Estuningsih SE, Partoutomo S, et al. Antibody-dependent cell-mediated cytotoxicity to newly excysted juvenile *Fasciola hepatica* in vitro is mediated by reactive nitrogen intermediates: Nitric oxide mediated killing of juvenile liver fluke. Parasite Immunol. 2001;23: 473–482.

36. Phiri IK, Phiri AM, Harrison LJS. Serum antibody isotype responses of Fasciola-infected sheep and cattle to excretory and secretory products of Fasciola species. Vet Parasitol. 2006;141: 234–242.

37. Leiper JWG. The longevity of *Fasciola hepatica*. J Helminthol. 1938;16: 173–176.

38. Takeuchi-Storm N, Denwood M, Petersen HH, Enemark HL, Stensgaard A-S, Sengupta ME, et al. Patterns of *Fasciola hepatica* infection in Danish dairy cattle: implications for on-farm control of the parasite based on different diagnostic methods. Parasit Vectors. 2018;11: 674.

39. Gandhi P, Schmitt EK, Chen C-W, Samantray S, Venishetty VK, Hughes D. Triclabendazole in the treatment of human fascioliasis: a review. Trans R Soc Trop Med Hyg. 2019;113: 797–804.

40. Boray JC, Crowfoot PD, Strong MB, Allison JR, Schellenbaum M, Von Orelli M, et al. Treatment of immature and mature *Fasciola hepatica* infections in sheep with triclabendazole. Vet Rec. 1983;113: 315–317.

41. Smeal MG, Hall CA. The activity of triclabendazole against immature and adult *Fasciola hepatica* infections in sheep. Aust Vet J. 1983;60: 329–331.

42. Hanna REB, McMahon C, Ellison S, Edgar HW, Kajugu P-E, Gordon A, et al. *Fasciola hepatica*: a comparative survey of adult fluke resistance to triclabendazole, nitroxynil and closantel on selected upland and lowland sheep farms in Northern Ireland using faecal egg counting, coproantigen ELISA testing and fluke histology. Vet Parasitol. 2015;207: 34–43.

43. Toner E, Brennan GP, Hanna REB, Edgar HWJ, Fairweather I. Disruption of egg formation by *Fasciola hepatica* following treatment in vivo with triclabendazole in the sheep host. Vet Parasitol. 2011;177: 79–89.

44. Chambers E, Ryan LA, Hoey EM, Trudgett A, McFerran NV, Fairweather I, et al. Liver fluke β-tubulin isotype 2 binds albendazole and is thus a probable target of this drug. Parasitol Res. 2010;107: 1257–1264.

45. Robinson MW, McFerran N, Trudgett A, Hoey L, Fairweather I. A possible model of benzimidazole binding to beta-tubulin disclosed by invoking an inter-domain movement. J Mol Graph Model. 2004;23: 275–284.

46. Ranjan P, Kumar SP, Kari V, Jha PC. Exploration of interaction zones of β-tubulin colchicine binding domain of helminths and binding mechanism of anthelmintics. Comput Biol Chem. 2017;68: 78–91.

47. Lang BZ, Hall RF, Kirstein KW, Wescott RB. Reduced reproductive potential of *Fasciola hepatica* surviving in cattle treated with albendazole. Vet Med Small Anim Clin. 1980;75: 1853–1857.

48. Fairweather I, Boray JC. Fasciolicides: efficacy, actions, resistance and its management. Vet J. 1999;158: 81–112.

49. Meaney M, Allister J, McKinstry B, McLaughlin K, Brennan GP, Forbes AB, et al. *Fasciola hepatica*: morphological effects of a combination of triclabendazole and clorsulon against mature fluke. Parasitol Res. 2006;99: 609–621.

50. Meaney M, Allister J, McKinstry B, McLaughlin K, Brennan GP, Forbes AB, et al. *Fasciola hepatica*: ultrastructural effects of a combination of triclabendazole and clorsulon against mature fluke. Parasitol Res. 2007;100: 1091–1104.

51. Overend DJ, Bowen FL. Resistance of *Fasciola hepatica* to triclabendazole. Aust Vet J. 1995;72: 275–276.

52. Fairweather I, Brennan GP, Hanna REB, Robinson MW, Skuce PJ. Drug resistance in liver flukes. Int J Parasitol Drugs Drug Resist. 2020;12: 39–59.

53. Cabada MM, Lopez M, Cruz M, Delgado JR, Hill V, White AC Jr. Treatment failure after multiple courses of triclabendazole among patients with fascioliasis in Cusco, Peru: A case series. PLoS Negl Trop Dis. 2016;10: e0004361.

54. Ramadan HK-A, Hassan WA, Elossily NA, Ahmad AA, Mohamed AA, Abd-Elkader AS, et al. Evaluation of nitazoxanide treatment following triclabendazole failure in an outbreak of human fascioliasis in Upper Egypt. PLoS Negl Trop Dis. 2019;13: e0007779.

55. Kahl A, von Samson-Himmelstjerna G, Helm C, Hodgkinson J, Williams D, Weiher W, et al. Efficacy of flukicides against *Fasciola hepatica* and first report of triclabendazole resistance on German sheep farms. Int J Parasitol Drugs Drug Resist. 2023;23: 94–105.

56. Novobilský A, Averpil HB, Höglund J. The field evaluation of albendazole and triclabendazole efficacy against *Fasciola hepatica* by coproantigen ELISA in naturally infected sheep. Vet Parasitol. 2012;190: 272–276.

57. Novobilský A, Höglund J. First report of closantel treatment failure against *Fasciola hepatica* in cattle. Int J Parasitol Drugs Drug Resist. 2015;5: 172–177.

58. Robles-Pérez D, Martínez-Pérez JM, Rojo-Vázquez FA, Martínez-Valladares M. The diagnosis of fasciolosis in feces of sheep by means of a PCR and its application in the detection of anthelmintic resistance in sheep flocks naturally infected. Vet Parasitol. 2013;197: 277–282.

59. Alvarez LI, Lanusse CE, Williams DJL, Fairweather I, Hodgkinson JE. Flukicidal drugs:Pharmaco-therapeutics and drug resistance. In: Dalton JP, editor. Fasciolosis 2nd Edition. Wallingford, England: CABI Publishing; 2020. pp. 227–230.

60. Wolstenholme AJ, Fairweather I, Prichard R, von Samson-Himmelstjerna G, Sangster NC. Drug resistance in veterinary helminths. Trends Parasitol. 2004;20: 469–476.

61. Alvarez LI, Solana HD, Mottier ML, Virkel GL, Fairweather I, Lanusse CE. Altered drug influx/efflux and enhanced metabolic activity in triclabendazole-resistant liver flukes. Parasitology. 2005;131: 501–510.

62. Savage J, Meaney M, Brennan GP, Hoey E, Trudgett A, Fairweather I. Increased action of triclabendazole (TCBZ) in vitro against a TCBZ-resistant isolate of *Fasciola hepatica* following its co-incubation with the P-glycoprotein inhibitor, R(+)-verapamil. Exp Parasitol. 2013;135: 642–653.

63. Meaney M, Savage J, Brennan GP, Hoey E, Trudgett A, Fairweather I. Increased susceptibility of a triclabendazole (TCBZ)-resistant isolate of *Fasciola hepatica* to TCBZ following co-incubation in vitro with the P-glycoprotein inhibitor, R(+)-verapamil. Parasitology. 2013;140: 1287–1303.

64. Radio S, Fontenla S, Solana V, Matos Salim AC, Araújo FMG, Ortiz P, et al. Pleiotropic alterations in gene expression in Latin American *Fasciola hepatica* isolates with different susceptibility to drugs. Parasit Vectors. 2018;11: 56.

65. Robinson MW, Lawson J, Trudgett A, Hoey EM, Fairweather I. The comparative metabolism of triclabendazole sulphoxide by triclabendazole-susceptible and triclabendazole-resistant *Fasciola hepatica*. Parasitol Res. 2004;92: 205–210.

66. Fernández V, Estein S, Ortiz P, Luchessi P, Solana V, Solana H. A single amino acid substitution in isozyme GST mu in Triclabendazole resistant *Fasciola hepatica* (Sligo strain) can substantially influence the manifestation of anthelmintic resistance. Exp Parasitol. 2015;159: 274–279.

67. Beesley NJ, Cwiklinski K, Allen K, Hoyle RC, Spithill TW, La Course EJ, et al. A major locus confers triclabendazole resistance in *Fasciola hepatica* and shows dominant inheritance. PLoS Pathog. 2023;19: e1011081.

68. Molina-Hernández V, Mulcahy G, Pérez J, Martínez-Moreno Á, Donnelly S, O'Neill SM, et al. *Fasciola hepatica* vaccine: we may not be there yet but we're on the right road. Vet Parasitol. 2015;208: 101–111.

69. Ullah R, Rehman A, Zafeer MF, Rehman L, Khan YA, Khan MAH, et al. Anthelmintic potential of thymoquinone and curcumin on *Fasciola gigantica*. PLoS One. 2017;12: e0171267.

70. Alvarez-Mercado JM, Ibarra-Velarde F, Alonso-Díaz MÁ, Vera-Montenegro Y, Avila-Acevedo JG, García-Bores AM. *In vitro* antihelmintic effect of fifteen tropical plant extracts on excysted flukes of *Fasciola hepatica*. BMC Vet Res. 2015;11: 45.

71. Hernández Alvarez L, Naranjo Feliciano D, Hernández González JE, Soares RO, Barreto Gomes DE, Pascutti PG. Insights into the interactions of *Fasciola hepatica* cathepsin L3 with a substrate and potential novel inhibitors through in silico approaches. PLoS Negl Trop Dis. 2015;9: e0003759.

72. Valderas-García E, Castilla-Gómez de Agüero V, González Del Palacio L, Galli G, Escala N, Ruiz-Somacarrera M, et al. New benzimidazole derivative compounds with in vitro fasciolicidal properties. Parasit Vectors. 2024;17: 173.

73. McConville M, Hanna REB, Brennan GP, Edgar HWJ, McConnell S, McCoy M, et al. Impact of compound alpha treatment *in vivo* on egg production by the liver fluke, *Fasciola hepatica*. Vet Parasitol. 2012;187: 183–195.

74. Park S-K, Gunaratne GS, Chulkov EG, Moehring F, McCusker P, Dosa PI, et al. The anthelmintic drug praziquantel activates a schistosome transient receptor potential channel. J Biol Chem. 2019;294: 18873–18880.

75. Rohr CM, Sprague DJ, Park S-K, Malcolm NJ, Marchant JS. Natural variation in the binding pocket of a parasitic flatworm TRPM channel resolves the basis for praziquantel sensitivity. Proc Natl Acad Sci U S A. 2023;120: e2217732120.

76. Sprague DJ, Park S-K, Gramberg S, Bauer L, Rohr CM, Chulkov EG, et al. Target-based discovery of a broad spectrum flukicide. bioRxivorg. 2023. doi:10.1101/2023.09.22.559026

77. Pereira Moreira B, Weber MHW, Haeberlein S, Mokosch AS, Spengler B, Grevelding CG, et al. Drug repurposing and DE Novo drug discovery of protein kinase inhibitors as new drugs against schistosomiasis. Molecules. 2022;27: 1414.

78. Melnikova I, Golden J. Targeting protein kinases. Nat Rev Drug Discov. 2004;3: 993–994.

79. Roskoski R Jr. Properties of FDA-approved small molecule protein kinase inhibitors: A 2024 update. Pharmacol Res. 2024;200: 107059.

80. Morawietz CM, Houhou H, Puckelwaldt O, Hehr L, Dreisbach D, Mokosch A, et al. Targeting kinases in *Fasciola hepatica*: Anthelminthic effects and tissue distribution of selected kinase inhibitors. Front Vet Sci. 2020;7: 611270.

81. Flores-Velázquez LM, Ruiz-Campillo MT, Herrera-Torres G, Martínez-Moreno Á, Martínez-Moreno FJ, Zafra R, et al. Fasciolosis: pathogenesis, host-parasite interactions, and implication in vaccine development. Front Vet Sci. 2023;10: 1270064.

82. Wendt GR, Collins JJ 3rd. Schistosomiasis as a disease of stem cells. Curr Opin Genet Dev. 2016;40: 95–102.

83. McCusker P, Clarke NG, Gardiner E, Armstrong R, McCammick EM, McVeigh P, et al. Neoblast-like stem cells of *Fasciola hepatica*. PLoS Pathog. 2024;20: e1011903.

84. Robb E, McCammick EM, Wells D, McVeigh P, Gardiner E, Armstrong R, et al. Transcriptomic analysis supports a role for the nervous system in regulating growth and development of *Fasciola hepatica* juveniles. PLoS Negl Trop Dis. 2022;16: e0010854.

85. McCusker P, McVeigh P, Rathinasamy V, Toet H, McCammick E, O'Connor A, et al. Stimulating neoblast-like cell proliferation in juvenile *Fasciola hepatica* supports growth and progression towards the adult phenotype in vitro. PLoS Negl Trop Dis. 2016;10: e0004994.

86. Cwiklinski K, Dalton JP. Advances in *Fasciola hepatica* research using "omics" technologies. Int J Parasitol. 2018;48: 321–331.

87. Cwiklinski K, Jewhurst H, McVeigh P, Barbour T, Maule AG, Tort J, et al. Infection by the helminth parasite Fasciola hepatica requires rapid regulation of metabolic, virulence, and invasive factors to adjust to its mammalian host. Mol Cell Proteomics. 2018;17: 792–809.

88. Cwiklinski K, Dalton JP. Omics tools enabling vaccine discovery against fasciolosis. Trends Parasitol. 2022;38: 1068–1079.

89. Le TH, Blair D, McManus DP. Complete DNA sequence and gene organization of the mitochondrial genome of the liverfluke, *Fasciola hepatica* L. (Platyhelminthes; Trematoda). Parasitology. 2001;123: 609–621.

90. Young ND, Hall RS, Jex AR, Cantacessi C, Gasser RB. Elucidating the transcriptome of Fasciola hepatica - a key to fundamental and biotechnological discoveries for a neglected parasite. Biotechnol Adv. 2010;28: 222–231.

91. Cwiklinski K, Dalton JP, Dufresne PJ, La Course J, Williams DJ, Hodgkinson J, et al. The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. Genome Biol. 2015;16: 71.

92. McNulty SN, Tort JF, Rinaldi G, Fischer K, Rosa BA, Smircich P, et al. Genomes of *Fasciola hepatica* from the Americas reveal colonization with *Neorickettsia* endobacteria related to the agents of Potomac horse and human Sennetsu fevers. PLoS Genet. 2017;13: e1006537.

93. Choi Y-J, Fontenla S, Fischer PU, Le TH, Costábile A, Blair D, et al. Adaptive radiation of the flukes of the family Fasciolidae inferred from genome-wide comparisons of key species. Mol Biol Evol. 2020;37: 84–99.

94. Fuchs M-A, Ryan LA, Chambers EL, Moore CM, Fairweather I, Trudgett A, et al. Differential expression of liver fluke β-tubulin isotypes at selected life cycle stages. Int J Parasitol. 2013;43: 1133–1139.

95. Ryan LA, Hoey E, Trudgett A, Fairweather I, Fuchs M, Robinson MW, et al. *Fasciola hepatica* expresses multiple alpha- and beta-tubulin isotypes. Mol Biochem Parasitol. 2008;159: 73–78.

96. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009;6: 377–382.

97. Aldridge S, Teichmann SA. Single cell transcriptomics comes of age. Nat Commun. 2020;11: 4307.

98. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol Syst Biol. 2019;15: e8746.

99. Mereu E, Lafzi A, Moutinho C, Ziegenhain C, McCarthy DJ, Álvarez-Varela A, et al. Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. Nat Biotechnol. 2020;38: 747–755.

100. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015;161: 1202–1214.

101. Zilionis R, Nainys J, Veres A, Savova V, Zemmour D, Klein AM, et al. Single-cell barcoding and sequencing using droplet microfluidics. Nat Protoc. 2017;12: 44–73.

102. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8: 14049.

103. Zhang X, Li T, Liu F, Chen Y, Yao J, Li Z, et al. Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems. Mol Cell. 2019;73: 130–142.e5.

104. Farnsworth DR, Saunders LM, Miller AC. A single-cell transcriptome atlas for zebrafish development. Dev Biol. 2020;459: 100–108.

105. Liao Y, Ma L, Guo Q, E W, Fang X, Yang L, et al. Cell landscape of larval and adult *Xenopus laevis* at single-cell resolution. Nat Commun. 2022;13: 4306.

106. Sebé-Pedrós A, Saudemont B, Chomsky E, Plessier F, Mailhé M-P, Renno J, et al. Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-seq. Cell. 2018;173: 1520–1534.e20.

107. Fincher CT, Wurtzel O, de Hoog T, Kravarik KM, Reddien PW. Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*. Science. 2018;360. doi:10.1126/science.aaq1736

108. Plass M, Solana J, Wolf FA, Ayoub S, Misios A, Glažar P, et al. Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. Science. 2018;360. doi:10.1126/science.aaq1723

109. Wang B, Lee J, Li P, Saberi A, Yang H, Liu C, et al. Stem cell heterogeneity drives the parasitic life cycle of *Schistosoma mansoni*. Elife. 2018;7. doi:10.7554/eLife.35449

110. DeLaughter DM. The use of the Fluidigm C1 for RNA expression analyses of single cells. Curr Protoc Mol Biol. 2018;122: e55.

111. Wendt G, Zhao L, Chen R, Liu C, O'Donoghue AJ, Caffrey CR, et al. A single-cell RNA-seq atlas of *Schistosoma mansoni* identifies a key regulator of blood feeding. Science. 2020;369: 1644–1649.

112. Diaz Soria CL, Lee J, Chong T, Coghlan A, Tracey A, Young MD, et al. Single-cell atlas of the first intra-mammalian developmental stage of the human parasite *Schistosoma mansoni*. Nat Commun. 2020;11: 6411.

113. Diaz Soria CL, Attenborough T, Lu Z, Fontenla S, Graham J, Hall C, et al. Single-cell transcriptomics of the human parasite *Schistosoma mansoni* first intra-molluscan stage reveals tentative tegumental and stem-cell regulators. Sci Rep. 2024;14: 5974.

114. Li P, Nanes Sarfati D, Xue Y, Yu X, Tarashansky AJ, Quake SR, et al. Single-cell analysis of *Schistosoma mansoni* identifies a conserved genetic program controlling germline stem cell fate. Nat Commun. 2021;12: 485.

115. Attenborough T, Rawlinson KA, Soria CLD, Ambridge K, Sankaranarayanan G, Graham J, et al. A single-cell atlas of the miracidium larva of the human blood fluke *Schistosoma mansoni*: cell types, developmental pathways and tissue architecture. eLife. 2024. doi:10.7554/elife.95628.2

116. Wang J, Paz C, Padalino G, Coghlan A, Lu Z, Gradinaru I, et al. Large-scale RNAi screening uncovers therapeutic targets in the parasite *Schistosoma mansoni*. Science. 2020;369: 1649–1653.

117. Morales-Vicente DA, Zhao L, Silveira GO, Tahira AC, Amaral MS, Collins JJ 3rd, et al. Single-cell RNA-seq analyses show that long non-coding RNAs are conspicuously expressed in *Schistosoma mansoni* gamete and tegument progenitor cell populations. Front Genet. 2022;13: 924877.

118. Wendt GR, Shiroor DA, Adler CE, Collins JJ 3rd. Convergent evolution of a genotoxic stress response in a parasite-specific p53 homolog. Proc Natl Acad Sci U S A. 2022;119: e2205201119.

119. Collins JJ 3rd, Hou X, Romanova EV, Lambrus BG, Miller CM, Saberi A, et al. Genome-wide analyses reveal a role for peptide hormones in planarian germline development. PLoS Biol. 2010;8: e1000509.

120. Rice-Ficht AC, Dusek KA, Kochevar GJ, Waite JH. Eggshell precursor proteins of *Fasciola hepatica*, I. Structure and expression of vitelline protein B. Mol Biochem Parasitol. 1992;54: 129–141.

121. Houhou H, Puckelwaldt O, Strube C, Haeberlein S. Reference gene analysis and its use for kinase expression profiling in *Fasciola hepatica*. Sci Rep. 2019;9: 15867.

122. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49: W293–W296.

123. Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res. 2007;35: W43–6.

124. UniProt Consortium. UniProt: The universal protein knowledgebase in 2023. Nucleic Acids Res. 2023;51: D523–D531.

125. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. Nucleic Acids Res. 2021;49: D458–D460.

126. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30: 1236–1240.

127. Shen W-K, Chen S-Y, Gan Z-Q, Zhang Y-Z, Yue T, Chen M-M, et al. AnimalTFDB 4.0: a comprehensive animal transcription factor database updated with variation and expression annotations. Nucleic Acids Res. 2023;51: D39–D45.

128. Planosphere: Sánchez Lab planarian resources. [cited 24 Sep 2024]. Available: https://planosphere.stowers.org/

129. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite - a comprehensive resource for helminth genomics. Mol Biochem Parasitol. 2017;215: 2–10.

130. Lu Z, Sessler F, Holroyd N, Hahnel S, Quack T, Berriman M, et al. A gene expression atlas of adult *Schistosoma mansoni* and their gonads. Sci Data. 2017;4: 170118.

131. Lu Z, Sessler F, Holroyd N, Hahnel S, Quack T, Berriman M, et al. *Schistosome* sex matters: a deep view into gonad-specific and pairing-dependent transcriptomes reveals a complex gender interplay. Sci Rep. 2016;6: 31150.

132. Sievers F, Higgins DG. Clustal Omega, accurate alignment of very large numbers of sequences. Methods Mol Biol. 2014;1079: 105–116.

133. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10. doi:10.1093/gigascience/giab008

134. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. Nature. 2018;560: 494–498.

135. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol. 2020;38: 1408–1414.

136. Van Rossum G, Drake FL, Others. Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam; 1995.

137. Executable Books Community. Jupyter Book. Zenodo; 2020. doi:10.5281/ZENODO.4539666

138. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10: 421.

139. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 2011;39: W29–37.

140. Inkscape Project. Inkscape. 2020. Available: https://inkscape.org

141. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9: 676–682.

142. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32: 1792–1797.

143. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30: 772–780.

144. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61: 539–542.

145. R Core Team. The R Project for Statistical Computing. Vienna, Austria; 2023 [cited 3 Sep 2024]. Available: https://www.R-project.org/

146. RStudio Team. Posit. In: Posit [Internet]. Boston, MA; 9 Sep 2022 [cited 3 Sep 2024]. Available: http://www.rstudio.com/

147. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr. 2023 [cited 3 Sep 2024]. Available: https://dplyr.tidyverse.org

148. Wickham H. stringr. 2023 [cited 3 Sep 2024]. Available: https://stringr.tidyverse.org

149. Wickham H, Vaughan D, Girlich M. tidyr. 2024 [cited 3 Sep 2024]. Available: https://tidyr.tidyverse.org

150. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 2024;42: 293–304.

151. Wickham H. Ggplot2: Elegant graphics for data analysis. 2nd ed. Basel, Switzerland: Springer International Publishing; 2016. Available: https://ggplot2.tidyverse.org

152. Alexa A, Rahnenfuhrer J. topGO: Enrichment analysis for gene ontology. R package version. 2010.

153. Gu Z, Hübschmann D. Make interactive complex heatmaps in R. Bioinformatics. 2022;38: 1460–1462.

154. Gao C-H, Yu G, Cai P. GgVennDiagram: An intuitive, easy-to-use, and highly customizable R package to generate Venn diagram. Front Genet. 2021;12: 706907.

155. McVeigh P, McCammick EM, McCusker P, Morphew RM, Mousley A, Abidi A, et al. RNAi dynamics in Juvenile *Fasciola spp.* Liver flukes reveals the persistence of gene silencing in vitro. PLoS Negl Trop Dis. 2014;8: e3185.

156. McCusker P, Hussain W, McVeigh P, McCammick E, Clarke NG, Robb E, et al. RNA interference dynamics in juvenile *Fasciola hepatica* are altered during *in vitro* growth and development. Int J Parasitol Drugs Drug Resist. 2020;14: 46–55.

157. Ramirez B, Bickle Q, Yousif F, Fakorede F, Mouries M-A, Nwaka S. Schistosomes: challenges in compound screening. Expert Opin Drug Discov. 2007;2: S53–61.

158. García-Castro H, Kenny NJ, Iglesias M, Álvarez-Campos P, Mason V, Elek A, et al. ACME dissociation: a versatile cell fixation-dissociation method for single-cell transcriptomics. Genome Biol. 2021;22: 89.

159. Cancela M, Maggioli G. *Fasciola hepatica*: Methods and protocols. 1st ed. Cancela M, Maggioli G, editors. New York, NY: Springer; 2020.

160. Jensen E. Technical review: *In situ* hybridization: AR insights. Anat Rec (Hoboken). 2014;297: 1349–1353.

161. Moescheid MF, Puckelwaldt O, Beutler M, Haeberlein S, Grevelding CG. Defining an optimal control for RNAi experiments with adult *Schistosoma mansoni*. Sci Rep. 2023;13: 9766.

162. Lück S, Kreszies T, Strickert M, Schweizer P, Kuhlmann M, Douchkov D. SiRNA-finder (si-Fi) software for RNAi-target design and off-target prediction. Front Plant Sci. 2019;10: 1023.

163. Haçarız O, Sayers G. *Fasciola hepatica* - where is 28S ribosomal RNA? Exp Parasitol. 2013;135: 426–429.

164. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25: 402–408.

165. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience. 2018;7. doi:10.1093/gigascience/giy083

166. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184: 3573–3587.e29.

167. Barile M, Imaz-Rosshandler I, Inzani I, Ghazanfar S, Nichols J, Marioni JC, et al. Coordinated changes in gene expression kinetics underlie both mouse and human erythroid maturation. Genome Biol. 2021;22: 197.

168. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016;352: 189–196.

169. Gramberg S, Puckelwaldt O, Schmitt T, Lu Z, Haeberlein S. Spatial transcriptomics of a parasitic flatworm provides a molecular map of drug targets and drug resistance genes. Nat Commun. 2024;15: 1–19.

170. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive integration of single-cell data. Cell. 2019;177: 1888–1902.e21.

171. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51: D638–D646.

172. King HO, Owusu-Boaitey KE, Fincher CT, Reddien PW. A transcription factor atlas of stem cell fate in planarians. Cell Rep. 2024;43: 113843.

173. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14: 417–419.

174. 10x Genomics. Cell Preparation for Single Cell Protocols. 10x Genomics; 2023 Jun.

175. Santangelo C. Methods and Materials: Basic Primary Cell Isolation. In: Worthington Biochemical [Internet]. 2011 [cited 3 Sep 2024]. Available: http://www.worthington-biochem.com/tissuedissociation/basic.html

176. Reichard A, Asosingh K. Best practices for preparing a single cell suspension from solid tissues for flow cytometry: Single cell suspension preparation. Cytometry A. 2019;95: 219–226.

177. What is the range of compatible cell sizes? In: 10X Genomics [Internet]. [cited 9 Sep 2024]. Available: https://kb.10xgenomics.com/hc/en-us/articles/218170543-What-is-the-range-of-compatible-cell-sizes

178. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol. 2006;7: 3.

179. Peiris TH, García-Ojeda ME, Oviedo NJ. Alternative flow cytometry strategies to analyze stem cells and cell death in planarians. Regeneration (Oxf). 2016;3: 123–135.

180. Dalton JP, Neill SO, Stack C, Collins P, Walshe A, Sekiya M, et al. *Fasciola hepatica* cathepsin L-like proteases: biology, function, and potential in the development of first generation liver fluke vaccines. Int J Parasitol. 2003;33: 1173–1181.

181. Threadgold LT. *Fasciola hepatica*: stereological analysis of vitelline cell development. Exp Parasitol. 1982;54: 352–365.

182. Erasmus DA. *Schistosoma mansoni:* development of the vitelline cell, its role in drug sequestration, and changes induced by Astiban. Exp Parasitol. 1975;38: 240–256.

183. Lu Z, Quack T, Hahnel S, Gelmedin V, Pouokam E, Diener M, et al. Isolation, enrichment and primary characterisation of vitelline cells from *Schistosoma mansoni* obtained by the organ isolation method. Int J Parasitol. 2015;45: 663–672.

184. Stitt AW, Fairweather I. Spermatogenesis and the fine structure of the mature spermatozoon of the liver fluke, *Fasciola hepatica* (Trematoda: Digenea). Parasitology. 1990;101 Pt 3: 395–407.

185. Collins JJ 3rd, Wang B, Lambrus BG, Tharp ME, Iyer H, Newmark PA. Adult somatic stem cells in the human parasite *Schistosoma mansoni*. Nature. 2013;494: 476–479.

186. Newmark PA, Sánchez Alvarado A. Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. Dev Biol. 2000;220: 142–153.

187. De Keuckelaere E, Hulpiau P, Saeys Y, Berx G, van Roy F. Nanos genes and their role in development and beyond. Cell Mol Life Sci. 2018;75: 1929–1946.

188. Zhang Y-F, Wang Y-X, Zhang N-, Lin Z-H, Wang L-R, Feng Y, et al. Prognostic alternative splicing regulatory network of RBM25 in hepatocellular carcinoma. Bioengineered. 2021;12: 1202–1211.

189. Li D, Yu W, Lai M. Towards understandings of serine/arginine-rich splicing factors. Acta Pharm Sin B. 2023;13: 3181–3207.

190. Overall CM, Dean RA. Degradomics: systems biology of the protease web. Pleiotropic roles of MMPs in cancer. Cancer Metastasis Rev. 2006;25: 69–75.

191. Blobel CP. ADAMs: key components in EGFR signalling and development. Nat Rev Mol Cell Biol. 2005;6: 32–43.

192. Iyer H, Issigonis M, Sharma PP, Extavour CG, Newmark PA. A premeiotic function for boule in the planarian *Schmidtea mediterranea*. Proc Natl Acad Sci U S A. 2016;113: E3509–18.

193. Hanna REB, Moffett D, Forster FI, Trudgett AG, Brennan GP, Fairweather I. *Fasciola hepatica*: a light and electron microscope study of the ovary and of the development of oocytes within eggs in the uterus provides an insight into reproductive strategy. Vet Parasitol. 2016;221: 93–103.

194. Zheng YH, Rengaraj D, Choi JW, Park KJ, Lee SI, Han JY. Expression pattern of meiosis associated SYCP family members during germline development in chickens. J Reprod Fertil. 2009;138: 483–492.

195. Wang PJ, McCarrey JR, Yang F, Page DC. An abundance of X-linked genes expressed in spermatogonia. Nat Genet. 2001;27: 422–426.

196. Luo M, Yang F, Leu NA, Landaiche J, Handel MA, Benavente R, et al. MEIOB exhibits single-stranded DNA-binding and exonuclease activities and is essential for meiotic recombination. Nat Commun. 2013;4: 2788.

197. Souquet B, Abby E, Hervé R, Finsterbusch F, Tourpin S, Le Bouffant R, et al. MEIOB targets single-strand DNA and is necessary for meiotic recombination. PLoS Genet. 2013;9: e1003784.

198. Hendzel MJ, Wei Y, Mancini MA, Van Hooser A, Ranalli T, Brinkley BR, et al. Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. Chromosoma. 1997;106: 348–360.

199. Beckmann S, Quack T, Burmeister C, Buro C, Long T, Dissous C, et al. *Schistosoma mansoni*: signal transduction processes during the development of the reproductive organs. Parasitology. 2010;137: 497–520.

200. Waghmare I, Page-McCaw A. Wnt signaling in stem cell maintenance and differentiation in the *Drosophila* germarium. Genes (Basel). 2018;9. doi:10.3390/genes9030127

201. Liu J, Xiao Q, Xiao J, Niu C, Li Y, Zhang X, et al. Wnt/β-catenin signalling: function, biological mechanisms, and therapeutic opportunities. Signal Transduct Target Ther. 2022;7: 3.

202. McVeigh P, McCammick E, McCusker P, Wells D, Hodgkinson J, Paterson S, et al. Profiling G protein-coupled receptors of *Fasciola hepatica* identifies orphan rhodopsins unique to phylum Platyhelminthes. Int J Parasitol Drugs Drug Resist. 2018;8: 87–103.

203. Avelar L das GA, Gava SG, Neves RH, Silva MCS, Araújo N, Tavares NC, et al. Smp38 MAP kinase regulation in *Schistosoma mansoni*: Roles in survival, oviposition, and protection against oxidative stress. Front Immunol. 2019;10: 21.

204. Han J, Wu J, Silke J. An overview of mammalian p38 mitogen-activated protein kinases, central regulators of cell stress and receptor signaling. F1000Res. 2020;9: 653.

205. Wu W, LoVerde PT. Identification and evolution of nuclear receptors in Platyhelminths. PLoS One. 2021;16: e0250750.

206. Wang J, Collins JJ 3rd. Identification of new markers for the *Schistosoma mansoni* vitelline lineage. Int J Parasitol. 2016;46: 405–410.

207. Wang J, Chen R, Collins JJ 3rd. Systematically improved in vitro culture conditions reveal new insights into the reproductive biology of the human parasite *Schistosoma mansoni*. PLoS Biol. 2019;17: e3000254.

208. Bennett APS, de la Torre-Escudero E, Dermott SSE, Threadgold LT, Hanna REB, Robinson MW. *Fasciola hepatica* gastrodermal cells selectively release extracellular vesicles via a novel atypical secretory mechanism. Int J Mol Sci. 2022;23: 5525.

209. Collins PR, Stack CM, O'Neill SM, Doyle S, Ryan T, Brennan GP, et al. Cathepsin L1, the major protease involved in liver fluke (*Fasciola hepatica*) virulence: propetide cleavage sites and autoactivation of the zymogen secreted from gastrodermal cells. J Biol Chem. 2004;279: 17038–17046.

210. McCauliff LA, Langan A, Li R, Ilnytska O, Bose D, Waghalter M, et al. Intracellular cholesterol trafficking is dependent upon NPC2 interaction with lysobisphosphatidic acid. Elife. 2019;8. doi:10.7554/eLife.50832

211. Robinson MW, Hanna REB, Fairweather I. Development of *Fasciola hepatica* in the mammalian host. Fasciolosis. UK: CABI; 2021. pp. 65–111.

212. Pax RA, Day TA, Miller CL, Bennett JL. Neuromuscular physiology and pharmacology of parasitic flatworms. Parasitology. 1996;113 Suppl: S83–96.

213. Ross KG, Omuro KC, Taylor MR, Munday RK, Hubert A, King RS, et al. Novel monoclonal antibodies to study tissue regeneration in planarians. BMC Dev Biol. 2015;15: 2.

214. Cancela M, Corvo I, DA Silva E, Teichmann A, Roche L, Díaz A, et al. Functional characterization of single-domain cystatin-like cysteine proteinase inhibitors expressed by the trematode *Fasciola hepatica*. Parasitology. 2017;144: 1695–1707.

215. Graham MK, Fairweather I, McGeown JG. Second messengers mediating mechanical responses to the FARP GYIRFamide in the fluke *Fasciola hepatica*. Am J Physiol Regul Integr Comp Physiol. 2000;279: R2089–94.

216. Dubey VK, Sarkar SR, Lakshmanan V, Dalmeida R, Gulyani A, Palakodeti D. *S. mediterranea* ETS-1 regulates the function of cathepsin-positive cells and the epidermal lineage landscape via basement membrane remodeling. J Cell Sci. 2022;135. doi:10.1242/jcs.259900

217. Holmes SD, Fairweather I. *Fasciola hepatica*: the effects of neuropharmacological agents upon in vitro motility. Exp Parasitol. 1984;58: 194–208.

218. Kreshchenko N, Terenina N, Ermakov A. Serotonin signalling in flatworms: An immunocytochemical localisation of 5-HT7 type of serotonin receptors in *Opisthorchis felineus* and *Hymenolepis diminuta*. Biomolecules. 2021;11: 1212.

219. Rossignol M, Gagnon ML, Klagsbrun M. Genomic organization of human neuropilin-1 and neuropilin-2 genes: identification and distribution of splice variants and soluble isoforms. Genomics. 2000;70: 211–222.

220. Kim Y-J, Bao H, Bonanno L, Zhang B, Serpe M. *Drosophila* Neto is essential for clustering glutamate receptors at the neuromuscular junction. Genes Dev. 2012;26: 974–987.

221. Bennett V, Healy J. Membrane domains based on ankyrin and spectrin associated with cell-cell interactions. Cold Spring Harb Perspect Biol. 2009;1: a003012.

222. Mohler PJ, Yoon W, Bennett V. Ankyrin-B targets beta2-spectrin to an intracellular compartment in neonatal cardiomyocytes. J Biol Chem. 2004;279: 40185–40193.

223. York NS, Sanchez-Arias JC, McAdam ACH, Rivera JE, Arbour LT, Swayne LA. Mechanisms underlying the role of ankyrin-B in cardiac and neurological health and disease. Front Cardiovasc Med. 2022;9: 964675.

224. Boskovic S, Marín-Juez R, Jasnic J, Reischauer S, El Sammak H, Kojic A, et al. Characterization of zebrafish (*Danio rerio*) muscle ankyrin repeat proteins reveals their conserved response to endurance exercise. PLoS One. 2018;13: e0204312.

225. Chulkov EG, Isaeva E, Stucky CL, Marchant JS. Use the force, fluke: Ligand-independent gating of *Schistosoma mansoni* ion channel TRPMPZQ. Int J Parasitol. 2023;53: 427–434.

226. Wells CM, Whale AD, Parsons M, Masters JRW, Jones GE. PAK4: a pluripotent kinase that regulates prostate cancer cell adhesion. J Cell Sci. 2010;123: 1663–1673.

227. Tan X, Yan Y, Song B, Zhu S, Mei Q, Wu K. Focal adhesion kinase: from biological functions to therapeutic strategies. Exp Hematol Oncol. 2023;12: 83.

228. Russell JN, Clements JE, Gama L. Quantitation of gene expression in formaldehyde-fixed and fluorescence-activated sorted cells. PLoS One. 2013;8: e73849.

229. DiDonato D, Brasaemle DL. Fixation methods for the study of lipid droplets by immunofluorescence microscopy. J Histochem Cytochem. 2003;51: 773–780.

230. Hanna REB, Fairweather I, Robinson MW. The reproductive system of *Fasciola hepatica*. In: Dalton JP, editor. Fascioliasis 2nd Edition. Wallingford, England: CABI Publishing; 2022. pp. 112–141.

231. van Wolfswinkel JC, Wagner DE, Reddien PW. Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. Cell Stem Cell. 2014;15: 326–339.

232. Martínez-Sernández V, Mezo M, González-Warleta M, Perteguer MJ, Muiño L, Guitián E, et al. The MF6p/FhHDM-1 major antigen secreted by the trematode parasite *Fasciola hepatica* is a heme-binding protein. J Biol Chem. 2014;289: 1441–1456.

233. Pankao V, Sirisriro A, Grams R, Vichasri-Grams S, Meepool A, Kangwanrangsan N, et al. Classification of the parenchymal cells in *Fasciola gigantica* based on ultrastructure and their expression of fatty acid binding proteins (FABPs). Vet Parasitol. 2006;142: 281–292.

234. Issigonis M, Redkar AB, Rozario T, Khan UW, Mejia-Sanchez R, Lapan SW, et al. A Krüppel-like factor is required for development and regeneration of germline and yolk cells from somatic stem cells in planarians. PLoS Biol. 2022;20: e3001472.

235. Doerig C. Protein kinases as targets for anti-parasitic chemotherapy. Biochim Biophys Acta. 2004;1697: 155–168.

236. Zhao Z-S, Manser E. PAK family kinases: Physiological roles and regulation: Physiological roles and regulation. Cell Logist. 2012;2: 59–68.

237. Won S-Y, Park J-J, Shin E-Y, Kim E-G. PAK4 signaling in health and disease: defining the PAK4-CREB axis. Exp Mol Med. 2019;51: 1–9.

238. Melzig J, Rein KH, Schäfer U, Pfister H, Jäckle H, Heisenberg M, et al. A protein related to p21-activated kinase (PAK) that is involved in neurogenesis in the *Drosophila* adult central nervous system. Curr Biol. 1998;8: 1223–1226.

239. Zhang H, Landmann F, Zahreddine H, Rodriguez D, Koch M, Labouesse M. A tension-induced mechanotransduction pathway promotes epithelial morphogenesis. Nature. 2011;471: 99–103.

240. Lucanic M, Kiley M, Ashcroft N, L'etoile N, Cheng H-J. The *Caenorhabditis elegans* P21-activated kinases are differentially required for UNC-6/netrin-mediated commissural motor axon guidance. Development. 2006;133: 4549–4559.

241. Peters EC, Gossett AJ, Goldstein B, Der CJ, Reiner DJ. Redundant canonical and noncanonical *Caenorhabditis elegans* p21-activated kinase signaling governs distal tip cell migrations. G3 (Bethesda). 2013;3: 181–195.

242. Ha BH, Morse EM, Turk BE, Boggon TJ. Signaling, regulation, and specificity of the type II p21-activated kinases. J Biol Chem. 2015;290: 12975–12983.

243. Mentzel B, Raabe T. Phylogenetic and structural analysis of the *Drosophila melanogaster* p21-activated kinase DmPAK3. Gene. 2005;349: 25–33.

244. Dummler B, Ohshiro K, Kumar R, Field J. Pak protein kinases and their role in cancer. Cancer Metastasis Rev. 2009;28: 51–63.

245. Li Y, Lu Q, Xie C, Yu Y, Zhang A. Recent advances on development of p21-activated kinase 4 inhibitors as anti-tumor agents. Front Pharmacol. 2022;13: 956220.

246. Zhang J, Wang J, Guo Q, Wang Y, Zhou Y, Peng H, et al. LCH-7749944, a novel and potent p21-activated kinase 4 inhibitor, suppresses proliferation and invasion in human gastric cancer cells. Cancer Lett. 2012;317: 24–32.

247. Wu T, Pang Y, Guo J, Yin W, Zhu M, Hao C, et al. Discovery of 2-(4-Substituted-piperidin/piperazine-1-yl)-N-(5-cyclopropyl-1H-pyrazol-3-yl)-quinazoline -2,4-diamines as PAK4 Inhibitors with Potent A549 Cell Proliferation, Migration, and Invasion Inhibition Activity. Molecules. 2018;23: 417.

248. Loverde PT, Osman A, Hinck A. *Schistosoma mansoni*: TGF-beta signaling pathways. Exp Parasitol. 2007;117: 304–317.

249. Freitas TC, Jung E, Pearce EJ. TGF-beta signaling controls embryo development in the parasitic flatworm *Schistosoma mansoni*. PLoS Pathog. 2007;3: e52.

250. Korzelius J, Naumann SK, Loza-Coll MA, Chan JS, Dutta D, Oberheim J, et al. Escargot maintains stemness and suppresses differentiation in *Drosophila* intestinal stem cells. EMBO J. 2014;33: 2967–2982.

251. Abnave P, Aboukhatwa E, Kosaka N, Thompson J, Hill MA, Aboobaker AA. Epithelial-mesenchymal transition transcription factors control pluripotent adult stem cell migration in vivo in planarians. Development. 2017;144: 3440–3453.

252. Partington GA, Bertwistle D, Nicolas RH, Kee WJ, Pizzey JA, Patient RK. GATA-2 is a maternal transcription factor present in *Xenopus* oocytes as a nuclear complex which is maintained throughout early development. Dev Biol. 1997;181: 144–155.

253. Liu B, He Y, Wu X, Lin Z, Ma J, Qiu Y, et al. Mapping putative enhancers in mouse oocytes and early embryos reveals TCF3/12 as key folliculogenesis regulators. Nat Cell Biol. 2024;26: 962–974.

254. Henthorn CR, Airs PM, Neumann EK, Zamanian M. Resolving the origins of secretory products and anthelmintic responses in a human parasitic nematode at single-cell resolution. Elife. 2023;12. doi:10.7554/eLife.83100

255. Kumar D, McGeown JG, Reynoso-Ducoing O, Ambrosio JR, Fairweather I. Observations on the musculature and isolated muscle fibres of the liver fluke, *Fasciola hepatica*. Parasitology. 2003;127: 457–473.

256. García-Castro H, Solana J. Single-cell transcriptomics in planaria: new tools allow new insights into cellular and evolutionary features. Biochem Soc Trans. 2022;50: 1237–1246.

257. Hayashi T, Asami M, Higuchi S, Shibata N, Agata K. Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. Dev Growth Differ. 2006;48: 371–380.

258. Cwiklinski K, Dalton JP. Exploiting comparative omics to understand the pathogenic and virulence-associated protease: Anti-protease relationships in the zoonotic parasites *Fasciola hepatica* and *Fasciola gigantica*. Genes (Basel). 2022;13: 1854.

259. Attenborough T, Rawlinson KA, Diaz Soria CL, Ambridge K, Sankaranarayanan G, Graham J, et al. A single-cell atlas of the miracidium larva of *Schistosoma mansoni* reveals cell types, developmental pathways, and tissue architecture. Elife. 2024;13. doi:10.7554/eLife.95628

260. Denisenko E, Guo BB, Jones M, Hou R, de Kock L, Lassmann T, et al. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. Genome Biol. 2020;21: 130.

261. Telford WG. Flow cytometry and cell sorting. Front Med (Lausanne). 2023;10: 1287884.

262. Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science. 2018;360: 176–182.

263. Emili E, Pérez-Posada A, Christodoulou MD, Solana J. Allometry of cell types in planarians by single cell transcriptomics. bioRxiv. 2023. p. 2023.11.01.565140. doi:10.1101/2023.11.01.565140

264. Grindberg RV, Yee-Greenbaum JL, McConnell MJ, Novotny M, O'Shaughnessy AL, Lambert GM, et al. RNA-sequencing from single nuclei. Proc Natl Acad Sci U S A. 2013;110: 19802–19807.

265. Wu H, Kirita Y, Donnelly EL, Humphreys BD. Advantages of single-nucleus over single-cell RNA sequencing of adult kidney: Rare cell types and novel cell states revealed in fibrosis. J Am Soc Nephrol. 2019;30: 23–32.

266. Wang X, Yu L, Wu AR. The effect of methanol fixation on single-cell RNA sequencing data. BMC Genomics. 2021;22: 420.

267. Gutiérrez-Franco A, Ake F, Hassan MN, Cayuela NC, Mularoni L, Plass M. Methanol fixation is the method of choice for droplet-based single-cell transcriptomics of neural cells. Commun Biol. 2023;6: 522.

268. Lee J. Planarians to schistosomes: an overview of flatworm cell-types and regulators. J Helminthol. 2023;97: e7.

269. Cantú AV, Altshuler-Keylin S, Laird DJ. Discrete somatic niches coordinate proliferation and migration of primordial germ cells via Wnt signaling. J Cell Biol. 2016;214: 215–229.

270. Xie T, Spradling AC. A niche maintaining germ line stem cells in the *Drosophila* ovary. Science. 2000;290: 328–330.

271. Lehmann R. Germline stem cells: origin and destiny. Cell Stem Cell. 2012;10: 729–739.

272. Chen R, Wang J, Gradinaru I, Vu HS, Geboers S, Naidoo J, et al. A male-derived nonribosomal peptide pheromone controls female schistosome development. Cell. 2022;185: 1506–1520.e17.

273. Issigonis M, Browder KL, Chen R, Collins JJ 3rd, Newmark PA. A niche-derived nonribosomal peptide triggers planarian sexual development. Proc Natl Acad Sci U S A. 2024;121: e2321349121.

274. Choi HMT, Schwarzkopf M, Fornace ME, Acharya A, Artavanis G, Stegmaier J, et al. Third-generation *in situ* hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. Development. 2018;145. doi:10.1242/dev.165753

275. Romero AA, Cobb SA, Collins JNR, Kliewer SA, Mangelsdorf DJ, Collins JJ 3rd. The *Schistosoma mansoni* nuclear receptor FTZ-F1 maintains esophageal gland function via transcriptional regulation of meg-8.3. PLoS Pathog. 2021;17: e1010140.

276. Carneiro VC, de Abreu da Silva IC, Torres EJL, Caby S, Lancelot J, Vanderstraete M, et al. Epigenetic changes modulate schistosome egg formation and are a novel target for reducing transmission of schistosomiasis. PLoS Pathog. 2014;10: e1004116.

277. Wu W, Niles EG, LoVerde PT. Thyroid hormone receptor orthologues from invertebrate species with emphasis on *Schistosoma mansoni*. BMC Evol Biol. 2007;7: 150.

278. Saule P, Adriaenssens E, Delacre M, Chassande O, Bossu M, Auriault C, et al. Early variations of host thyroxine and interleukin-7 favor *Schistosoma mansoni* development. J Parasitol. 2002;88: 849–855.

279. Habara O, Logan CY, Kanai-Azuma M, Nusse R, Takase HM. WNT signaling in pre-granulosa cells is required for ovarian folliculogenesis and female fertility. Development. 2021;148. doi:10.1242/dev.198846

280. Ye Z, Xu J, Feng X, Jia Y, Fu Z, Hong Y, et al. Spatiotemporal expression pattern of Sjfz7 and its expression comparison with other frizzled family genes in developmental stages of *Schistosoma japonicum*. Gene Expr Patterns. 2019;32: 44–52.

281. Vila-Farré M, Rozanski A, Ivanković M, Cleland J, Brand JN, Thalen F, et al. Evolutionary dynamics of whole-body regeneration across planarian flatworms. Nat Ecol Evol. 2023;7: 2108–2124.

282. Shinn GL. Formation of egg capsules by flatworms (phylum Platyhelminthes). Trans Am Microsc Soc. 1993;112: 18.

283. Collins JJ 3rd. Platyhelminthes. Curr Biol. 2017;27: R252–R256.

284. Fitzpatrick JM, Hirai Y, Hirai H, Hoffmann KF. Schistosome egg production is dependent upon the activities of two developmentally regulated tyrosinases. FASEB J. 2007;21: 823–835.

285. Gorin G, Fang M, Chari T, Pachter L. RNA velocity unraveled. PLoS Comput Biol. 2022;18: e1010492.

286. Bergen V, Soldatov RA, Kharchenko PV, Theis FJ. RNA velocity-current challenges and future perspectives. Mol Syst Biol. 2021;17: e10282.

287. Weiler P, Lange M, Klein M, Pe'er D, Theis F. CellRank 2: unified fate mapping in multiview single-cell data. Nat Methods. 2024;21: 1196–1205.

288. Hanna R. *Fasciola hepatica*: Histology of the reproductive organs and differential effects of triclabendazole on drug-sensitive and drug-resistant fluke isolates and on flukes from selected field cases. Pathogens. 2015;4: 431–456.

289. Chen RE, Thorner J. Function and regulation in MAPK signaling pathways: lessons learned from the yeast *Saccharomyces cerevisiae*. Biochim Biophys Acta. 2007;1773: 1311–1340.

290. Shakir EMN, Rinaldi G, Kirk RS, Walker AJ. *Schistosoma mansoni* excretory-secretory products induce protein kinase signalling, hyperkinesia, and stem cell proliferation in the opposite sex. Commun Biol. 2023;6: 985.

291. Ressurreição M, Rollinson D, Emery AM, Walker AJ. A role for p38 mitogen-activated protein kinase in early post-embryonic development of *Schistosoma mansoni*. Mol Biochem Parasitol. 2011;180: 51–55.

292. Hammaker D, Firestein GS. “Go upstream, young man”: lessons learned from the p38 saga. Ann Rheum Dis. 2010;69 Suppl 1: i77–82.

293. Evans RM, Mangelsdorf DJ. Nuclear receptors, RXR, and the Big Bang. Cell. 2014;157: 255–266.

294. Schulman IG. Nuclear receptors as drug targets for metabolic disease. Adv Drug Deliv Rev. 2010;62: 1307–1315.

295. Alvarez Rojas CA, Ansell BRE, Hall RS, Gasser RB, Young ND, Jex AR, et al. Transcriptional analysis identifies key genes involved in metabolism, fibrosis/tissue repair and the immune response against *Fasciola hepatica* in sheep liver. Parasit Vectors. 2015;8: 124.

296. Oldenborg V, Van Vugt F, Van Golde LM. Composition and metabolism of phospholipids of *Fasciola hepatics*, the common liver fluke. Biochim Biophys Acta. 1975;398: 101–110.

297. Ali SF, Joachim A, Daugschies A. Eicosanoid production by adult *Fasciola hepatica* and plasma eicosanoid patterns during fasciolosis in sheep. Int J Parasitol. 1999;29: 743–748.

298. Liu LX, Buhlmann JE, Weller PF. Release of prostaglandin E2 by microfilariae of *Wuchereria bancrofti* and *Brugia malayi*. Am J Trop Med Hyg. 1992;46: 520–523.

299. Liu LX, Weller PF. Intravascular filarial parasites inhibit platelet aggregation. Role of parasite-derived prostanoids. J Clin Invest. 1992;89: 1113–1120.

300. Angeli V, Faveeuw C, Roye O, Fontaine J, Teissier E, Capron A, et al. Role of the parasite-derived prostaglandin D2 in the inhibition of epidermal Langerhans cell migration during schistosomiasis infection. J Exp Med. 2001;193: 1135–1147.

301. Nikolaou A, Kokotou MG, Vasilakaki S, Kokotos G. Small-molecule inhibitors as potential therapeutics and as tools to understand the role of phospholipases A2. Biochim Biophys Acta Mol Cell Biol Lipids. 2019;1864: 941–956.

302. Infante R, Brown M, Goldstein J. NPC2 facilitates bidirectional transfer of cholesterol between NPC1 and lipid bilayers, a potential step in cholesterol egress from lysosomes. FASEB J. 2009;23. doi:10.1096/fasebj.23.1_supplement.521.5

303. Huang Y, Chen W, Wang X, Liu H, Chen Y, Guo L, et al. The carcinogenic liver fluke, *Clonorchis sinensis*: new assembly, reannotation and analysis of the genome and characterization of tissue transcriptomes. PLoS One. 2013;8: e54732.

304. Zawistowska-Deniziak A, Lambooij JM, Kalinowska A, Patente TA, Łapiński M, van der Zande HJP, et al. *Fasciola hepatica* fatty acid binding protein 1 modulates T cell polarization by promoting dendritic cell thrombospondin-1 secretion without affecting metabolic homeostasis in obese mice. Front Immunol. 2022;13: 884663.

305. Figueroa-Santiago O, Espino AM. *Fasciola hepatica* fatty acid binding protein induces the alternative activation of human macrophages. Infect Immun. 2014;82: 5005–5012.

306. Ramos-Benítez MJ, Ruiz-Jiménez C, Aguayo V, Espino AM. Recombinant *Fasciola hepatica* fatty acid binding protein suppresses toll-like receptor stimulation in response to multiple bacterial ligands. Sci Rep. 2017;7: 5455.

307. Mansour TE, Mansour JM. Effects of serotonin (5-hydroxytryptamine) and adenosine 3‘,5‘-phosphate on phosphofructokinase from the liver fluke *Fasciola hepatica*. J Biol Chem. 1962;237: 629–634.

308. Schulman MD, Valentino D, Cifelli S, Ostlind DA. Dose-dependent pharmacokinetics and efficacy of MK-401 against old, and young-mature infections of *Fasciola hepatica* in the rat. J Parasitol. 1982;68: 603–608.

309. Fairweather I, Holmes SD, Threadgold LT. *Fasciola hepatica*: motility response to fasciolicides *in vitro*. Exp Parasitol. 1984;57: 209–224.

310. Beckmann S, Grevelding CG. Imatinib has a fatal impact on morphology, pairing stability and survival of adult *Schistosoma mansoni in vitro*. Int J Parasitol. 2010;40: 521–526.

311. Ressurreição M, De Saram P, Kirk RS, Rollinson D, Emery AM, Page NM, et al. Protein kinase C and extracellular signal-regulated kinase regulate movement, attachment, pairing and egg release in *Schistosoma mansoni*. PLoS Negl Trop Dis. 2014;8: e2924.

312. Gramberg S, Puckelwaldt O, Schmitt T, Lu Z, Haeberlein S. Spatial transcriptomics of a parasitic flatworm provides a molecular map of vaccine candidates, drug targets and drug resistance genes. bioRxiv. 2023. p. 2023.12.11.571084. doi:10.1101/2023.12.11.571084

313. Gui P, Wu X, Ling S, Stotz SC, Winkfein RJ, Wilson E, et al. Integrin receptor activation triggers converging regulation of Cav1.2 calcium channels by c-Src and protein kinase A pathways. J Biol Chem. 2006;281: 14015–14025.

314. Thawornkuno C, Srisuksai K, Simanon N, Adisakwattana P, Ampawong S, Boonyuen U, et al. A reanalysis and integration of transcriptomics and proteomics datasets unveil novel drug targets for Mekong schistosomiasis. Sci Rep. 2024;14: 12969.

315. Hirst NL, Lawton SP, Walker AJ. CaMKII regulates neuromuscular activity and survival of the human blood fluke *Schistosoma mansoni*. Sci Rep. 2022;12: 19831.

316. Bokoch GM. Biology of the p21-activated kinases. Annu Rev Biochem. 2003;72: 743–781.

317. Eswaran J, Soundararajan M, Kumar R, Knapp S. UnPAKing the class differences among p21-activated kinases. Trends Biochem Sci. 2008;33: 394–403.

318. Law SHW, Sargent TD. Maternal pak4 expression is required for primitive myelopoiesis in zebrafish. Mech Dev. 2013;130: 181–194.

319. Guo Q, Su N, Zhang J, Li X, Miao Z, Wang G, et al. PAK4 kinase-mediated SCG10 phosphorylation involved in gastric cancer metastasis. Oncogene. 2014;33: 3277–3287.

320. Ramos-Alvarez I, Lee L, Jensen RT. Cyclic AMP-dependent protein kinase A and EPAC mediate VIP and secretin stimulation of PAK4 and activation of Na+,K+-ATPase in pancreatic acinar cells. Am J Physiol Gastrointest Liver Physiol. 2019;316: G263–G277.

321. Ramos-Alvarez I, Jensen RT. P21-activated kinase 4 in pancreatic acinar cells is activated by numerous gastrointestinal hormones/neurotransmitters and growth factors by novel signaling, and its activation stimulates secretory/growth cascades. Am J Physiol Gastrointest Liver Physiol. 2018;315: G302–G317.

322. Stefanić S, Dvořák J, Horn M, Braschi S, Sojka D, Ruelas DS, et al. RNA interference in *Schistosoma mansoni* schistosomula: selectivity, sensitivity and operation for larger-scale screening. PLoS Negl Trop Dis. 2010;4: e850.

323. Gava SG, Tavares NC, Salim AC de M, Araújo FMG de, Oliveira G, Mourão MM. *Schistosoma mansoni:* Off-target analyses using nonspecific double-stranded RNAs as control for RNAi experiments in schistosomula. Exp Parasitol. 2017;177: 98–103.

324. Zdrazil B, Felix E, Hunter F, Manners EJ, Blackshaw J, Corbett S, et al. The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. Nucleic Acids Res. 2024;52: D1180–D1192.

325. Knox C, Wilson M, Klinger CM, Franklin M, Oler E, Wilson A, et al. DrugBank 6.0: The DrugBank knowledgebase for 2024. Nucleic Acids Res. 2024;52: D1265–D1275.

326. Miranda-Miranda E, Cossio-Bayugar R, Aguilar-Díaz H, Narváez-Padilla V, Sachman-Ruíz B, Reynaud E. Transcriptome assembly dataset of anthelmintic response in *Fasciola hepatica*. Data Brief. 2021;35: 106808.

327. Cwiklinski K, Robinson MW, Donnelly S, Dalton JP. Complementary transcriptomic and proteomic analyses reveal the cellular and molecular processes that drive growth and development of *Fasciola hepatica* in the host liver. BMC Genomics. 2021;22: 46.

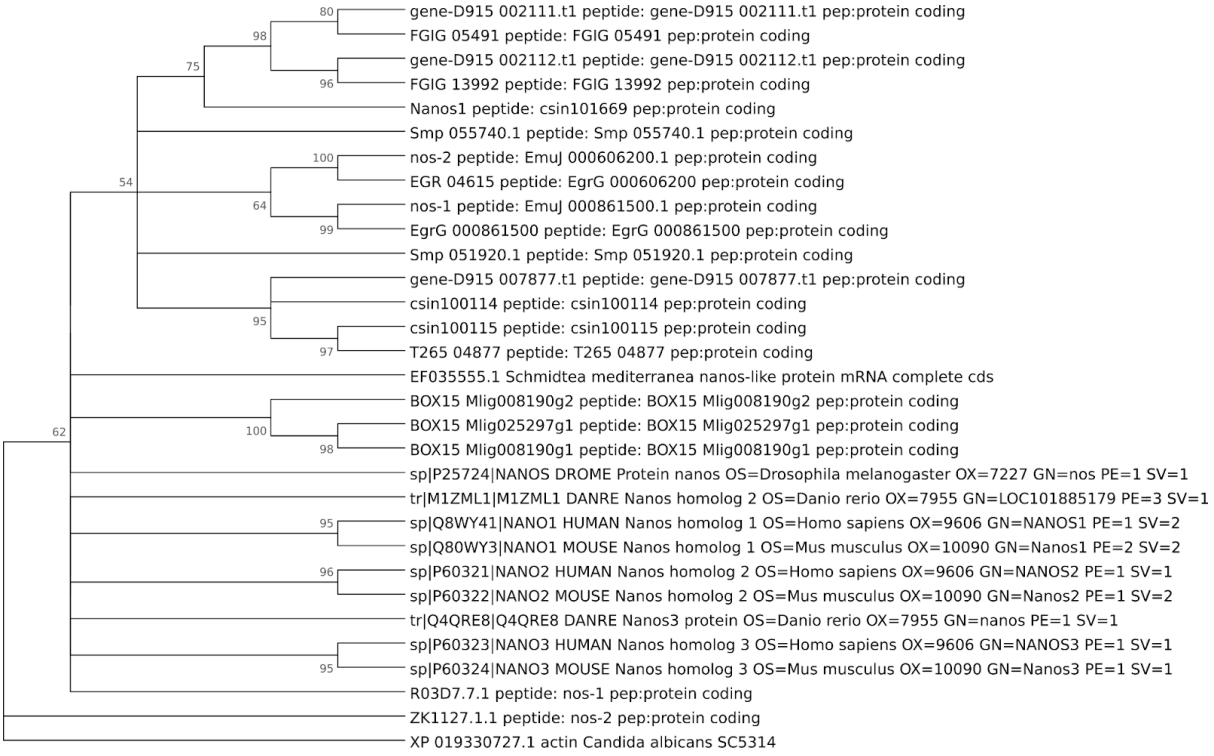
328. Robinson MW, Menon R, Donnelly SM, Dalton JP, Ranganathan S. An integrated transcriptomics and proteomics analysis of the secretome of the helminth pathogen *Fasciola hepatica*: proteins associated with invasion and infection of the mammalian host. Mol Cell Proteomics. 2009;8: 1891–1907.

329. Wilson RA, Wright JM, de Castro-Borges W, Parker-Manuel SJ, Dowle AA, Ashton PD, et al. Exploring the *Fasciola hepatica* tegument proteome. Int J Parasitol. 2011;41: 1347–1359.

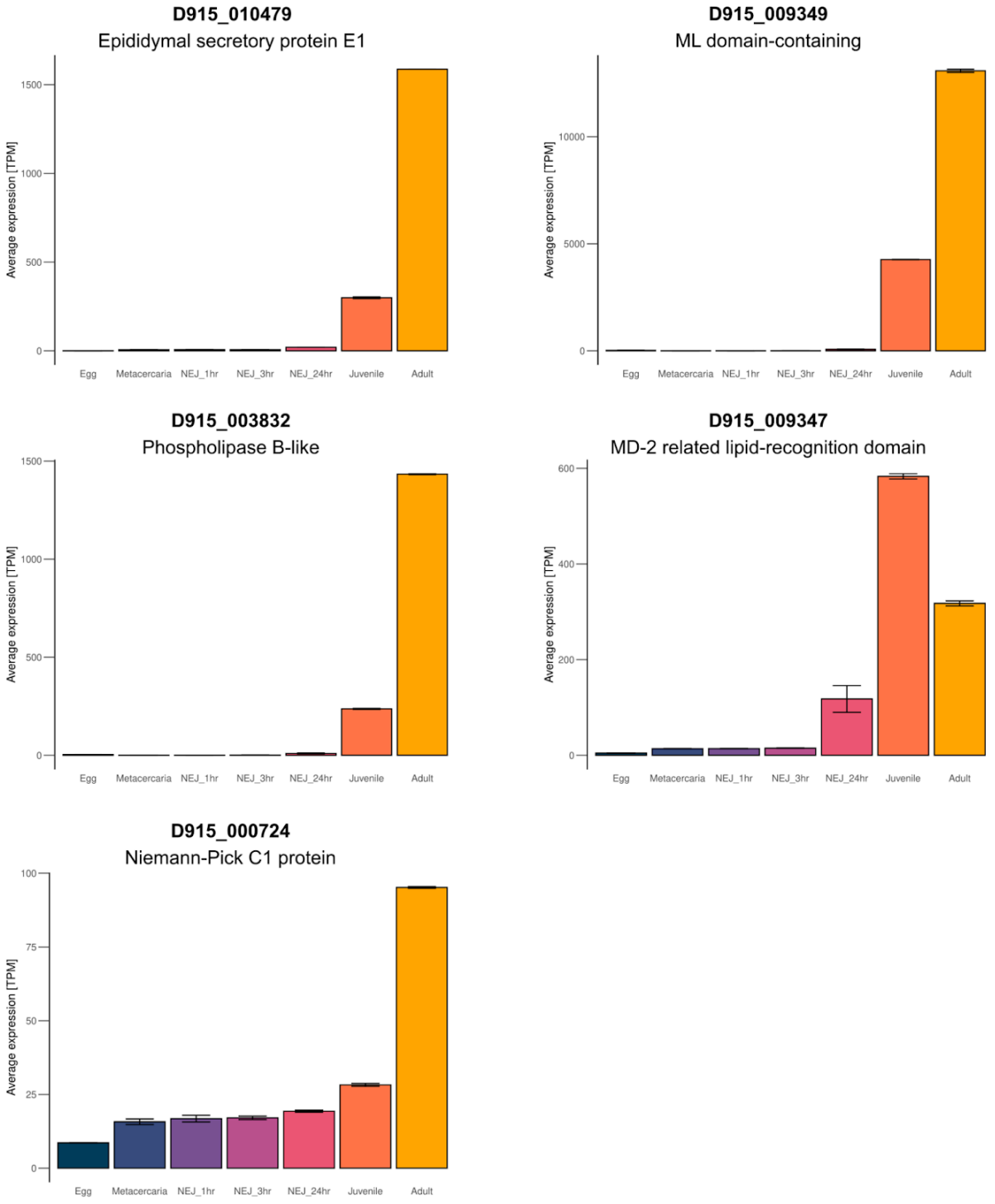
330. Di Maggio LS, Tirloni L, Pinto AFM, Diedrich JK, Yates JR Iii, Benavides U, et al. Across intra-mammalian stages of the liver fluke *Fasciola hepatica*: a proteomic study. Sci Rep. 2016;6: 32796.

8. Appendix

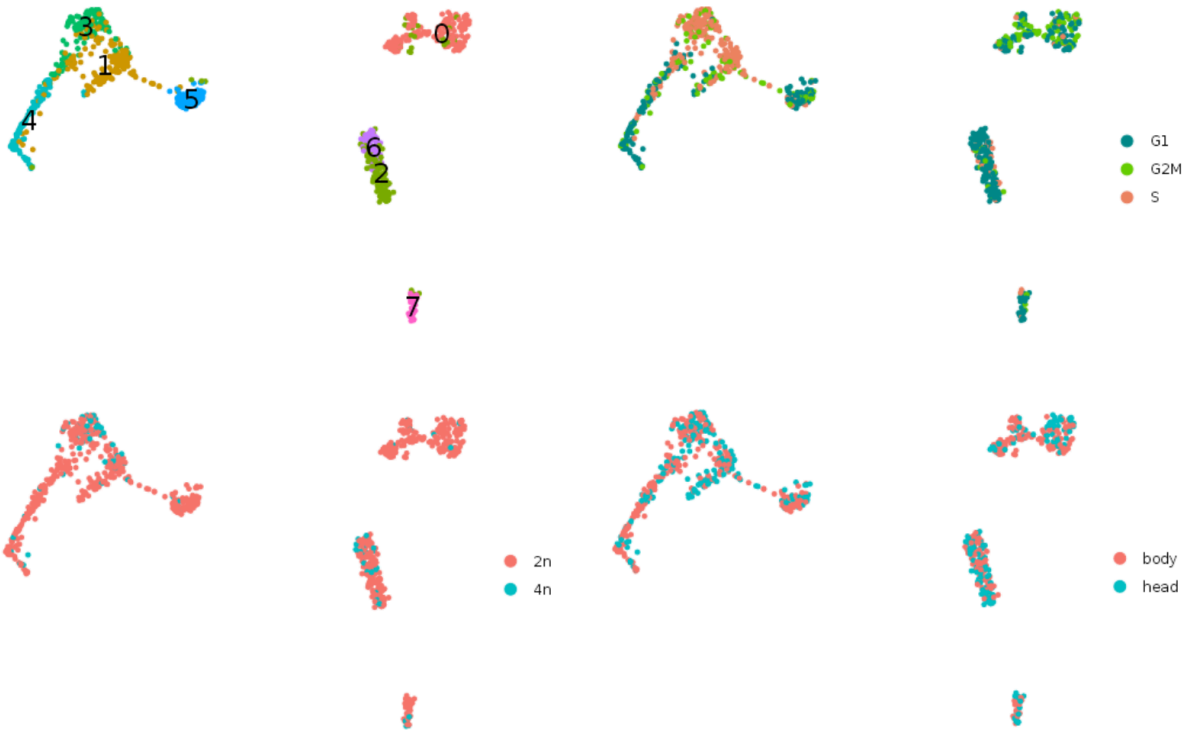
8.1 Supplementary figures



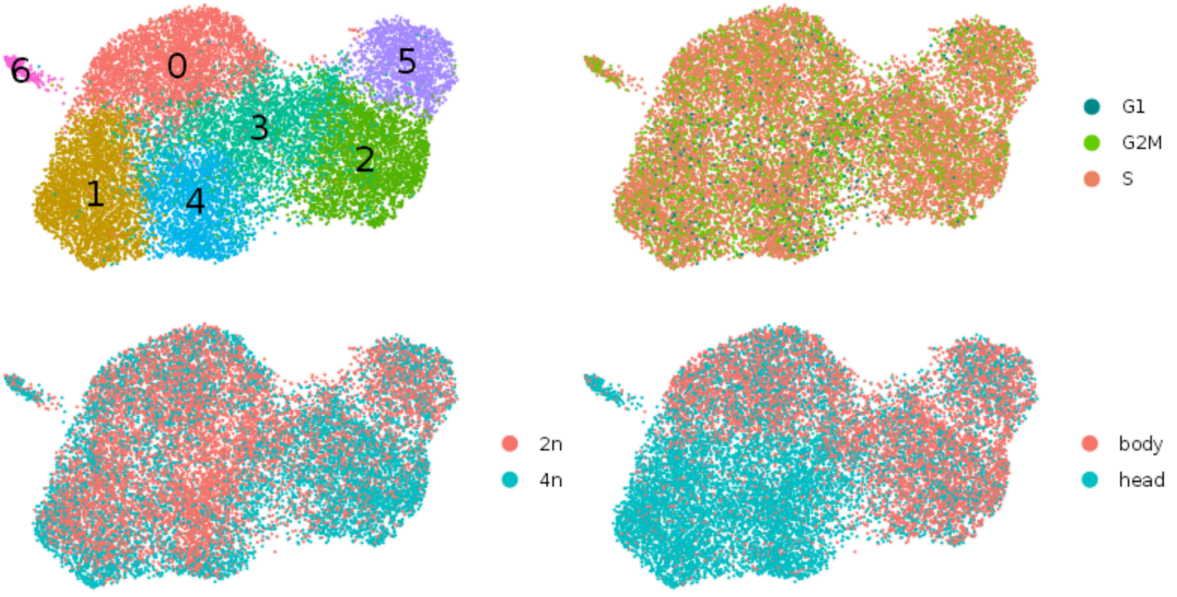
Suppl. Figure S1 Phylogenetic analysis of selected PAKs. Phyllogenetic tree of PAK orthologs of *F. hepatica* and other species. The protein kinase C from *Geodia cydonium* served as outgroup. Gene IDs are indicated.



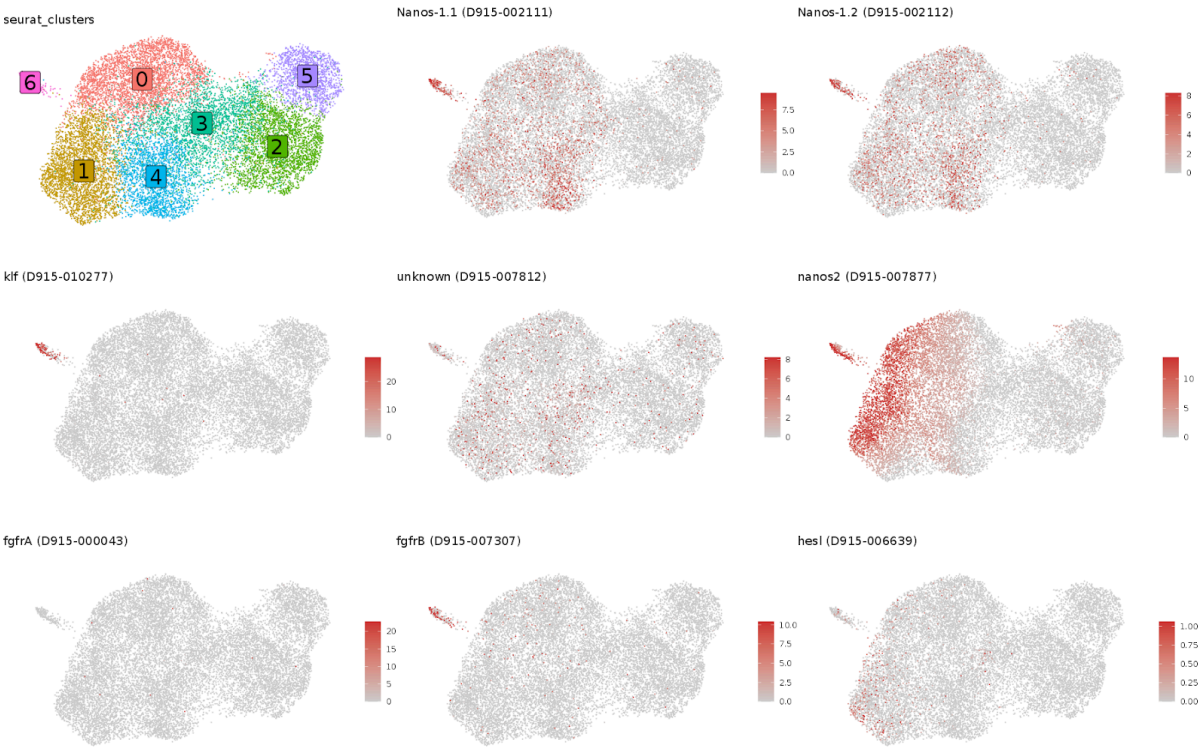
Suppl. Figure S2 Lifestage RNAseq data show high expression of lipid metabolism related genes in mature stages. Expression plots of selected lipid metabolism related genes. Raw reads as produced by [91] were remapped to the genome used in this work [92].



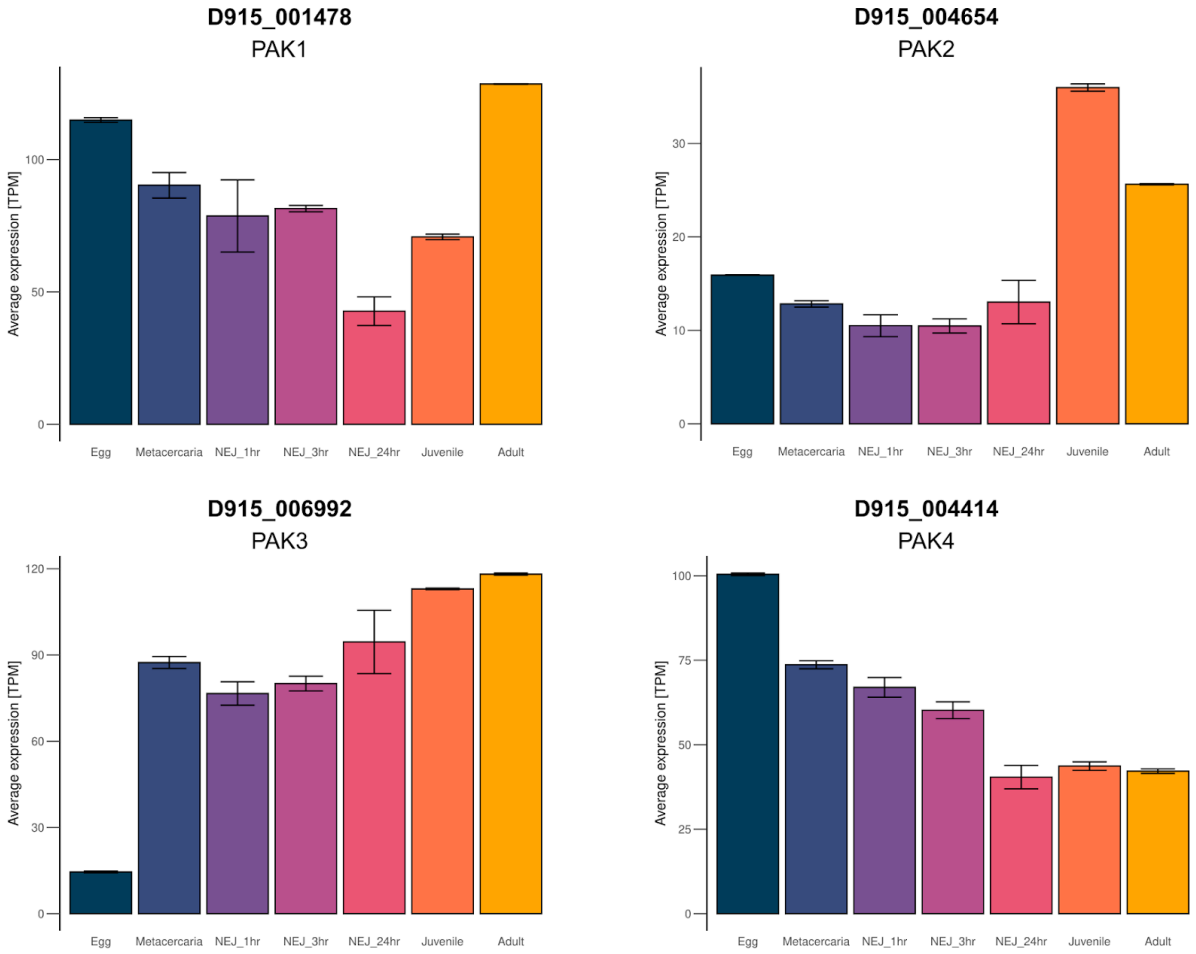
Suppl. Figure S3 Subclustering of the tegumental cluster produces distinct clusters. UMAP plots colored by cluster (top left), cell cycle stage (top right), 2n or 4n population (bottom right) or body part (bottom left).



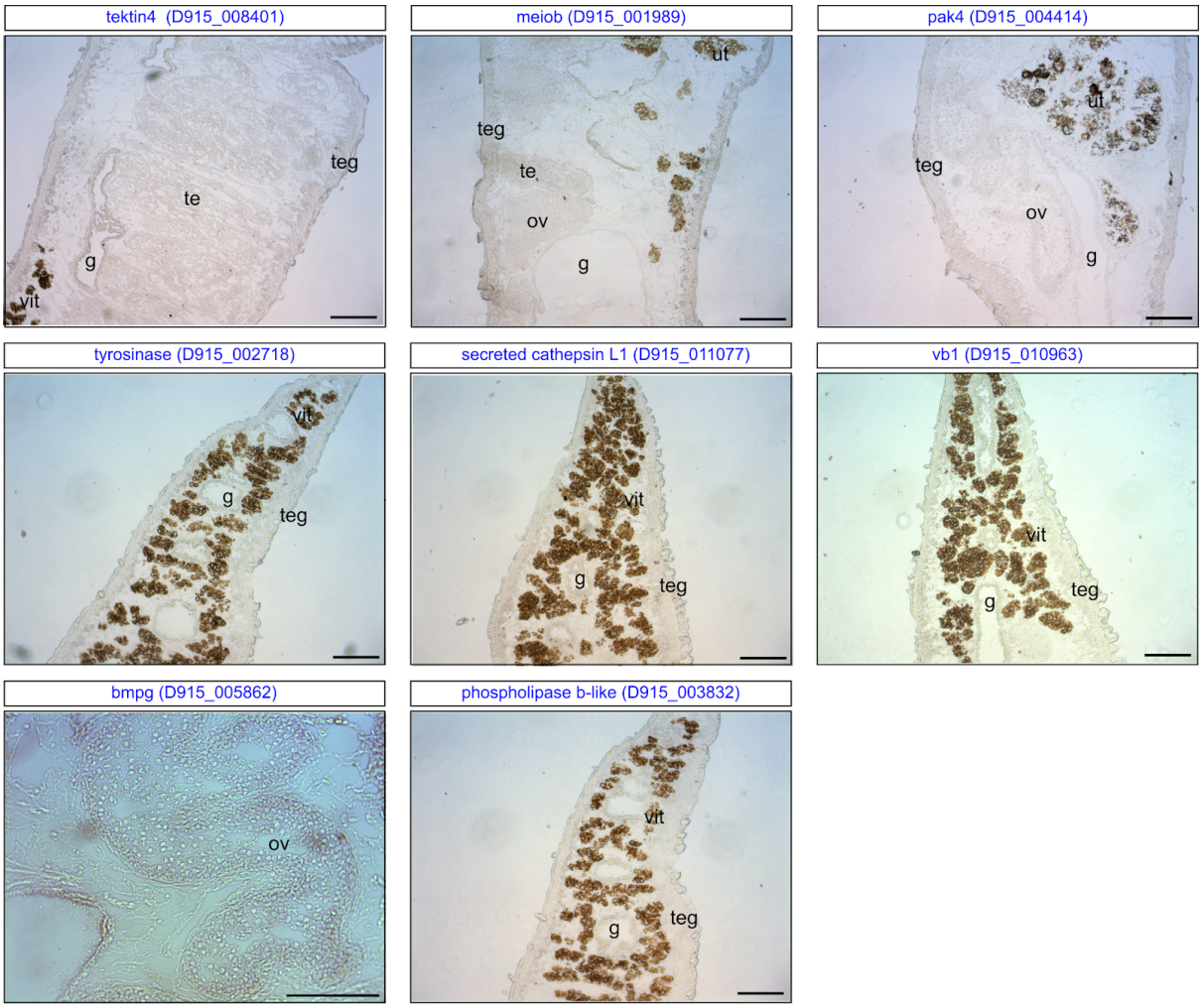
Suppl. Figure S4 Subclustering of stem cell cluster 1 and 2 did not clearly produce expected subclusters. UMAP plots colored by cluster (top left), cell cycle stage (top right), 2n or 4n population (bottom right) or body part (bottom left).



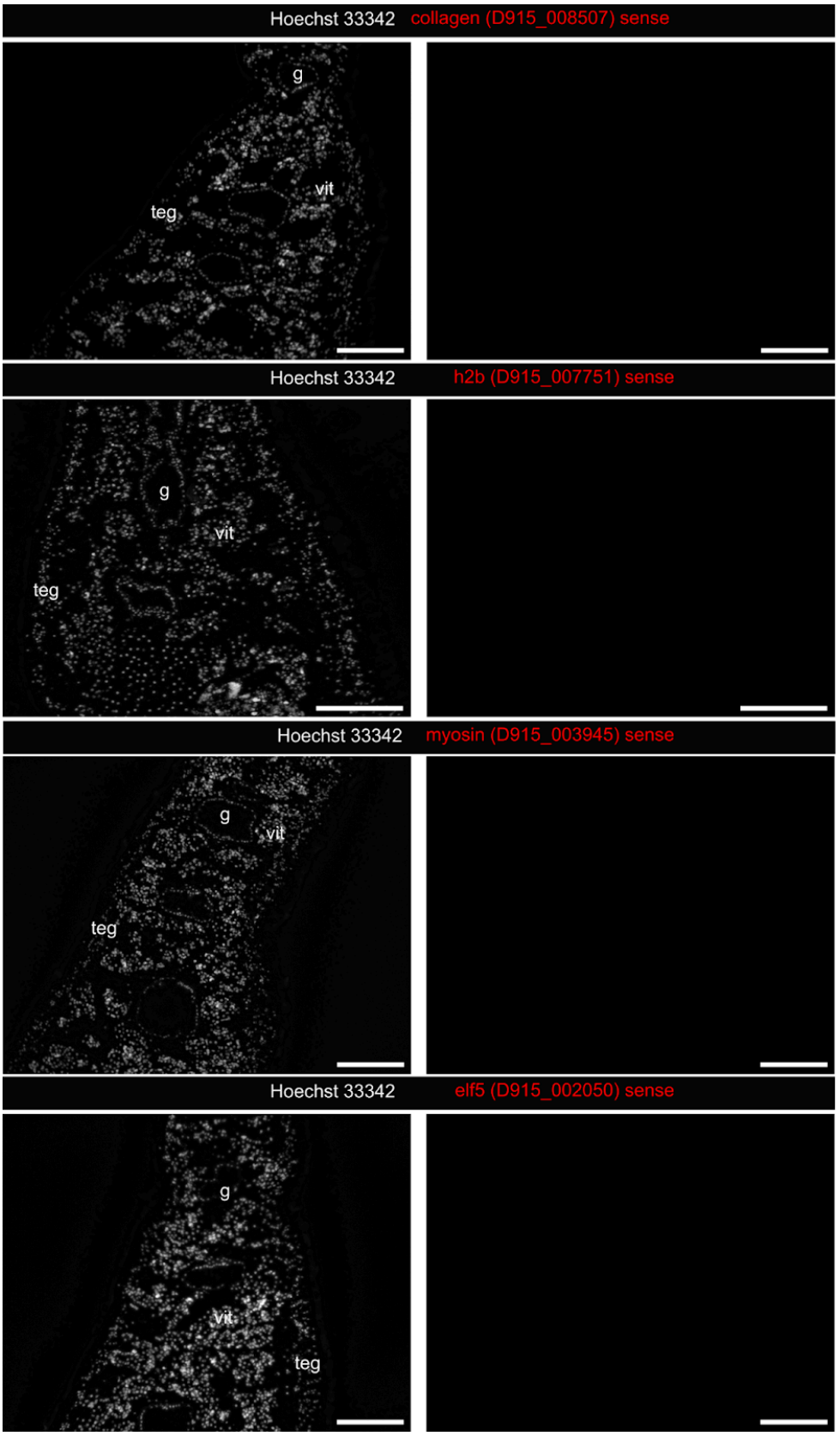
Suppl. Figure S5 Subclustering of stem cell and expression of subcluster markers. UMAP plots colored by cluster (top left) and UMAPs colored by the expression of selected stem cell marker genes.



Suppl. Figure S6 Lifestage RNAseq of four PAKs encoded in the *F. hepatica* genome. Expression plots of *pak* 1-4. Raw reads as produced by [91] were remapped to the genome used in this work [92].



Suppl. Figure S7 Sense probes produce no signal in colorimetric RNA *in situ* hybridization. Transversal section used for control with sense probes for transcripts by CISH. Scale bar: 200 µm. Legend: gasterodermis (g), tegument (teg), testes (te), ovary (ov), vitellarium (vit).



Suppl. Figure S8 Sense probes produce no signal in fluorescent RNA in situ hybridization. Transversal section used for control with sense probes for transcripts by FISH. Scale bar: 200 µm. Legend: gasterodermis (g), tegument (teg), vitellarium (vit).

8.2 Supplementary tables

Supplementary table S1 Blast results of cell cycle markers Human gene ids as derived from Tirosh et al 2016 [168], sequences were blasted against the *F. hepatica* genome

Human_ID	Fasciola_ID	percent_ID	lenght	mismatch	evalue	bitscore	phase
ENSG00000010292	D915_008067	37.312	796	409	1.81e-149	492	G2/M
ENSG00000011426	D915_007052	31.915	188	97	1.1e-23	107	G2/M
ENSG00000012963	D915_009724	33.573	417	233	5.55e-69	223	S
ENSG00000013810	D915_000754	41.791	134	78	2.89e-27	107	G2/M
ENSG00000049541	D915_004377	71.384	318	91	2.84e-177	494	S
ENSG00000051180	D915_003490	83.041	342	54	0	580	S
ENSG00000072571	D915_006444	20.938	320	209	13	44.3	G2/M
ENSG00000073111	D915_006290	69.619	734	213	0	1075	S
ENSG00000075218	D915_007410	32.174	115	67	1.5	31.2	G2/M
ENSG00000076003	D915_006290	30.988	668	369	8.59e-98	319	S
ENSG00000076248	D915_002571	62.766	188	66	7.68e-85	253	S
ENSG00000077514	D915_001901	29.167	120	76	1.38e-07	52.8	S
ENSG00000080986	D915_010238	41.121	107	61	3.69e-17	80.9	G2/M
ENSG00000087586	D915_007263	55.849	265	117	7.36e-115	335	G2/M
ENSG00000088325	D915_005221	31.507	219	122	2.67e-11	64.3	G2/M
ENSG00000089685	D915_003400	43.802	121	63	1.77e-30	107	G2/M
ENSG00000092140	D915_004019	38.525	122	68	3.59e-18	78.2	G2/M
ENSG00000092470	D915_006554	24.359	156	99	4.43e-06	48.9	S
ENSG00000092853	D915_007615	32.353	170	99	3.55e-12	70.1	S
ENSG00000093009	D915_009211	35.569	343	182	3.52e-63	205	S
ENSG00000094804	D915_009687	33.793	435	236	4.72e-65	219	S
ENSG00000094916	D915_003214	46.377	207	64	2.83e-52	166	G2/M
ENSG00000095002	D915_006999	48.598	535	253	6.64e-173	513	S
ENSG00000100297	D915_009119	53.907	755	257	0	808	S
ENSG00000100401	D915_002760	40.811	370	209	3.44e-78	261	G2/M
ENSG00000101868	D915_001192	38.497	1317	538	0	818	S
ENSG00000102974	D915_005692	38.75	160	97	5.95e-31	127	G2/M
ENSG00000104738	D915_009918	55.341	777	310	0	876	S
ENSG00000112742	D915_008029	35.88	301	166	3.13e-51	191	G2/M
ENSG00000113810	D915_006387	41.316	1307	660	0	950	G2/M
ENSG00000114346	D915_003446	31.746	315	196	5,00E-41	163	G2/M
ENSG00000115163	D915_010891	51.515	132	55	4.57e-33	112	G2/M

ENSG00000117399	D915_006257	41.558	462	244	5.61e-125	374	G2/M
ENSG00000117650	D915_006187	41.887	265	145	2.42e-58	202	G2/M
ENSG00000117724	D915_001775	24.862	362	238	5.6e-05	48.1	G2/M
ENSG00000117748	D915_007076	37.5	264	159	4.63e-59	187	S
ENSG00000119969	D915_000815	40.604	596	268	3.45e-135	426	S
ENSG00000126787	D915_001129	29.448	163	73	2.61e-10	63.2	G2/M
ENSG00000131153	D915_009013	43.275	171	97	1.07e-52	166	S
ENSG00000131747	D915_009008	62.951	1247	442	0	1639	G2/M
ENSG00000132646	D915_011058	64.341	258	92	3.41e-128	363	S
ENSG00000132780	D915_005194	29.412	221	128	1.04e-22	100	S
ENSG00000136492	D915_007941	35.424	813	420	8.81e-163	498	S
ENSG00000136982	D915_004955	24.272	412	238	5.34e-21	93.2	S
ENSG00000137804	D915_007767	28.571	105	60	0.34	32	G2/M
ENSG00000137807	D915_004538	32.127	884	452	2.37e-112	365	G2/M
ENSG00000138160	D915_002162	52.015	521	233	0	542	G2/M
ENSG00000138182	D915_003878	29.155	710	395	4.46e-79	288	G2/M
ENSG00000138778	D915_007155	44.565	368	186	5.02e-78	273	G2/M
ENSG00000139354	D915_000920	32.796	186	100	2.39e-22	101	G2/M
ENSG00000142945	D915_000033	48.148	378	149	5.28e-106	341	G2/M
ENSG00000143228	D915_006225	23.707	232	145	2	39.3	G2/M
ENSG00000143401	D915_008153	44.509	173	90	2.39e-33	122	G2/M
ENSG00000143476	D915_007084	34.65	329	189	1.05e-47	176	S
ENSG00000143815	D915_009518	33.952	377	224	4.9e-49	179	G2/M
ENSG00000151725	D915_006552	22.727	110	78	0.15	33.5	S
ENSG00000156802	D915_010718	82.946	129	22	3.24e-70	233	S
ENSG00000157456	D915_007018	34.392	189	116	1.15e-25	108	G2/M
ENSG00000158402	D915_003575	42.405	158	65	2.19e-29	112	G2/M
ENSG00000159259	D915_008281	33.094	417	193	2.43e-55	198	S
ENSG00000162607	D915_006266	27.119	177	103	1.52e-15	80.1	S
ENSG00000164104	D915_000513	51.445	173	69	2.43e-54	172	G2/M
ENSG00000167325	D915_005630	65.69	819	252	0	1160	S
ENSG00000168496	D915_009490	71.5	200	57	3.07e-105	307	S
ENSG00000169607	D915_003293	25.676	148	91	0.23	30.8	G2/M
ENSG00000169679	D915_004004	32.031	128	84	4.47e-17	75.5	G2/M
ENSG00000170312	D915_002458	61.486	296	109	2.09e-133	383	G2/M
ENSG00000171848	D915_007359	71.385	325	90	5.95e-176	493	S
ENSG00000174371	D915_007644	42.365	406	172	1.03e-92	308	S

ENSG00000175063	D915_007715	58.955	134	55	1.21e-56	174	G2/M
ENSG00000175216	D915_006461	36.033	1457	846	0	852	G2/M
ENSG00000176890	D915_003456	66.667	291	96	7.91e-145	409	S
ENSG00000178999	D915_007263	57.197	264	113	2.46e-114	332	G2/M
ENSG00000188229	D915_002311	97.964	442	9	0	922	G2/M
ENSG00000197299	D915_003541	45.525	771	366	0	632	S
ENSG00000198056	D915_009935	56.667	180	70	9.74e-68	213	S

Supplementary table S2 Accession numbers of sequences used for nanos phylogenetic analysis

ID	Species	Database
Smp_055740	*Schistosoma mansoni*	WormBase ParaSite Version 18
Smp_051920	*Schistosoma mansoni*	WormBase ParaSite Version 18
EF035555	*Schmidtea mediterranea*	NIH GeneBank
P25724	*Drosophila melanogaster*	NIH GeneBank
Q8WY41	*Homo sapiens*	NIH GeneBank
P60321	*Homo sapiens*	NIH GeneBank
P60323	*Homo sapiens*	NIH GeneBank
Q80WY3	*Mus musculus*	NIH GeneBank
P60324	*Mus musculus*	NIH GeneBank
P60322	*Mus musculus*	NIH GeneBank
Q4QRE8	*Danio rerio*	NIH GeneBank
M1ZML1	*Danio rerio*	NIH GeneBank
D915_002111	*Fasciola hepatica*	WormBase ParaSite Version 18
D915_002112	*Fasciola hepatica*	WormBase ParaSite Version 18
D915_007877	*Fasciola hepatica*	WormBase ParaSite Version 18
EmuJ_000606200	*Echinococcus multilocularis*	WormBase ParaSite Version 18
EmuJ_000861500	*Echinococcus multilocularis*	WormBase ParaSite Version 18
EgrG_000606200	*Echinococcus granulosus*	WormBase ParaSite Version 18
EgrG_000861500	*Echinococcus granulosus*	WormBase ParaSite Version 18
csin100115	*Clonorchis sinensis*	WormBase ParaSite Version 18

csin101669	*Clonorchis sinensis*	WormBase ParaSite Version 18
R03D7.7	*Caenorhabditis elegans*	Wormbase Version 290
ZK1127	*Caenorhabditis elegans*	Wormbase Version 290
FGIG_13992	*Fasciola gigantica*	WormBase ParaSite Version 18
FGIG_05491	*Fasciola gigantica*	WormBase ParaSite Version 18
BOX15_Mlig025297g1	*Macrostomum lignano*	WormBase ParaSite Version 18
BOX15_Mlig008190g2	*Macrostomum lignano*	WormBase ParaSite Version 18
BOX15_Mlig008190g1	*Macrostomum lignano*	WormBase ParaSite Version 18
T265_04877	*Opisthorchis viverrini*	WormBase ParaSite Version 18
XP_019330727.1	*Candida albicans*	NIH GeneBank

Supplementary table S3 Blast results for wnt members Original sequences were derived from NIH GeneBank and Blast was performed against the *F. hepatica* genome at WormBase ParaSite [129]

Name	*F. hepatica* ID	Query ID	E-value
ß-catenin	D915_007664	P35222	1.1e-60
Dishevelled	D915_005287	O14641	1.5e-41
Lipoprotein receptor	D915_001299	P01130	3.1e-65
Axin	D915_006733	Q9V407	1.4e-08
Glycogen synthase kinase 3	D915_000845	P49841	1.8e-84
Casein kinase 1	D915_010101	P48729	1.3e-141
Frizzled receptor	D915_010421	BN1106_s905B000183	0.0

Supplementary table S4 Kinases enriched in the elf5+ cluster and their involved pathways Kinase amino acid sequences were blasted against the landmark database usinge SmartBLAST. Shared pathways were identified by using the orthologue sequences and identifying shared KEGG pathways

F. hepatica ID	Target ID	Description Target	E-value	Pathway
D915_006871	NP_958915.1	mitogen-activated protein kinase 3	0,00E+00	MAPK signaling pathway
D915_003452	NP_956413.2	erbB-2	2,00E-72	MAPK signaling pathway,ErbB signaling pathway,Calcium signaling pathway
D915_003405	XP_009291340.1	mitogen-activated protein kinase kinase kinase kinase 3	4,00E-116	MAPK signaling pathway
D915_001189	NP_001014696.1	calcium/calmodulin-dependent protein kinase	0,00E+00	Calcium signaling pathway
D915_004391	XP_047278844.1	death-associated protein kinase 1	3,00E-53	
D915_006901	XP_005164062.1	protein kinase C beta	0,00E+00	MAPK signaling pathway,ErbB signaling pathway,Calcium signaling pathway
D915_002353	XP_005158636.1	protein-tyrosine kinase 2/focal adhesion kinase	6,00E-11	ErbB signaling pathway
D915_004414	XP_017173708.1	serine/threonine protein kinase PAK4	5,00E-128	ErbB signaling pathway
D915_004177	NP_001315092.1	tyrosine-protein kinase fyna	0,00E+00	
D915_005773	NP_148978.2	cyclin-dependent kinase 16	2,00E+115	
D915_000217	NP_001259779.1	ribosomal protein S6 kinase ß	0,00E+00	ErbB signaling pathway
D915_003066	XP_047287546.1	serine/threonin-protein kinase D	2,00E-149	
D915_004859	XP_005165175.1	ribosomal protein S6 kinase a	0	MAPK signaling pathway

Supplementary table S5 Cell number distributed over the various clusters

Cluster	Number of cells
stem cells 1	12399
late male germ cells	5844
S2/3	5384
unknown 1	5279
S4	5065
parenchym	4943
stem cells 2	4263
mature male germ cells	4262
early male germ cells 1	4087
early female germ cells	3730
early male germ cells 2	3172
uterus	3141
S1	1302
late female germ cells	1021
tegument	848
elf5+	739
gut	575
unknown 2	265

Supplementary table S6 Accession numbers of sequences used for PAK phylogenetic analysis

ID	Species	Database
WBGene00003911	*Caenorhabditis elegans*	Wormbase Version 290
WBGene00003144	*Caenorhabditis elegans*	Wormbase Version 290
WBGene00006443	*Caenorhabditis elegans*	Wormbase Version 290
NP_001262313.1	*Drosophila melanogaster*	NIH GeneBank
NP_001262638.1	*Drosophila melanogaster*	NIH GeneBank
NP_523375.2	*Drosophila melanogaster*	NIH GeneBank
ENSXETP00000025827	*Xenopus laevis*	Ensembl Release 110
ENSXETP00000092599	*Xenopus laevis*	Ensembl Release 110

ENSXETP00000025828	*Xenopus laevis*	Ensembl Release 110
NP_001363202	*Homo sapiens*	NIH GeneBank
NP_002568.2	*Homo sapiens*	NIH GeneBank
NP_001121638.1	*Homo sapiens*	NIH GeneBank
NP_001014831.1	*Homo sapiens*	NIH GeneBank
NP_065074.1	*Homo sapiens*	NIH GeneBank
NP_001122100.1	*Homo sapiens*	NIH GeneBank
S000000999	*Saccharomyces cerevisiae*	*Saccharomyces* Genome Database
D915_004414	*Fasciola hepatica*	WormBase ParaSite Version 18
D915_004654	*Fasciola hepatica*	WormBase ParaSite Version 18
D915_001478	*Fasciola hepatica*	WormBase ParaSite Version 18
D915_006992	*Fasciola hepatica*	WormBase ParaSite Version 18
O96997	*Geodia cydonium*	UniProt
SMESG000022413	*Schmidtea mediterranea*	PlanMine v3.0
SMESG000029469	*Schmidtea mediterranea*	PlanMine v3.0
SMESG000008329	*Schmidtea mediterranea*	PlanMine v3.0
SMESG000021793	*Schmidtea mediterranea*	PlanMine v3.0
SMESG000009652	*Schmidtea mediterranea*	PlanMine v3.0
SMESG000036933	*Schmidtea mediterranea*	PlanMine v3.0
Smp_064040	*Schistosoma mansoni*	WormBase ParaSite Version 18
Smp_333950	*Schistosoma mansoni*	WormBase ParaSite Version 18
Smp_159730	*Schistosoma mansoni*	WormBase ParaSite Version 18
Smp_347600	*Schistosoma mansoni*	WormBase ParaSite Version 18

9. Acknowledgments

At the end of this project, I would like to express my sincerest gratitude to all those who contributed to and helped during it. Without them, this would not have been possible.

First, I must sincerely thank my supervisor, Simone Häberlein, who ignited my passion for parasites that have yet to be extinguished. Thanks for giving me the time and freedom to make mistakes and providing a trusting and encouraging work environment.

Thanks to Nikola-Michael Prpic-Schäper for acting as my supervisor in the biology department.

I would also like to thank Christoph Grevelding for his valuable advice and help in navigating the university's administrative struggles.

A warm thank you to Alejandro Sanchez-Alvarado who welcomed me into his lab with open arms. Thanks also to Robb, Ekasit, Biff, Dung, Carolyn and Carlos for making my stay in the USA a truly memorable experience.

I owe a special thanks to Max Möscheid. For numerous conversations and encouragement after failed experiments. Thank you for joint training sessions, coffees and suffering through a mor than a few nights together. Without you, this work would have been much harder and way less exciting.

Also to my partner in *Fasciola* crime Svenja Gramberg. Your calm and determined manner was a real inspiration and without your knack for cryosectioning the images in this work would have been a nightmare.

Thanks also to Tobias Schmitt, not only for his excellent technical assistance, but also for his pleasant habit of always having a cheeky smirk on his lips

To all my colleagues Pia, Sophie, Rashmi, Sagar, Hicham, Mudassar, Monique, Gabo, Lisa, Alena, Xuesong, Olli, Martin and Thomas, Thank you for professional support and a friendly environment that made it hard to stay frustrated.

To my “outside” friends, thank you for keeping me grounded and for your unconditional support.

To my parents. Thank you for your support and for always giving me the feeling that I was safe and supported no matter the choices I made in life.

Last but not least Nici, thank you for being my light in every dark hour, without you none of this would have been achievable.

10. Contributions

10.1 Publications

Puckelwaldt, O., Gramberg, S., Ajmera, S., Koepke, J., Samakovlis, C., & Haeberlein, S. (2024). Single-cell transcriptomics identifies a p21-activated kinase important for survival of the zoonotic parasite *Fasciola hepatica*. In *bioRxiv*. https://doi.org/10.1101/2024.03.26.586785

Moescheid, MF**.,** Lu, Z**.,** Diaz Soria, C**.,** Quack, T**.,** Holroyd, N**.,** Holzaepfel, P**., Puckelwaldt, O.,** Haeberlein, S**.,** Rinaldi, G**.,** Berriman, M**.,** Grevelding, CG. The retinoic acid family-like nuclear receptor SmRAR identified by single-cell transcriptomics of ovarian cells controls oocyte differentiation in *Schistosoma mansoni. Nucleic Acids Res.* In revision. 2024.

Gramberg, S., **Puckelwaldt, O**., Schmitt, T., Lu, Z., & Haeberlein, S. (2024). Spatial transcriptomics of a parasitic flatworm provides a molecular map of drug targets and drug resistance genes. *Nature Communications*, *15*(1), 1–19.

Moescheid, M. F*., **Puckelwaldt, O***., Beutler, M., Haeberlein, S., & Grevelding, C. G. (2023). Defining an optimal control for RNAi experiments with adult *Schistosoma mansoni*. *Scientific Reports*, *13*(1), 9766.

**The authors contributed equally to this work*

Morawietz, C. M., Houhou, H., **Puckelwaldt, O**., Hehr, L., Dreisbach, D., Mokosch, A., Roeb, E., Roderfeld, M., Spengler, B., & Haeberlein, S. (2020). Targeting kinases in *Fasciola hepatica*: Anthelminthic effects and tissue distribution of selected kinase inhibitors. *Frontiers in Veterinary Science*, *7*, 611270.

10.2 Conferences and Talks

28th Helmithological Days of the Czech Society for Parasitology, 06-10 May 2024, Šumava, Czech Republik
Talk: **Puckelwaldt, O.,** Brewster, CE**.,** Mann, F**.,** Gramberg, S**.,** Schmitt, T**.,** Malloy, S**.,** Hall, KE., Perera, A**.,** Sanchez-Alvarado, A**.,** Häberlein, S. Mapping the gene-expression landscape in *F. hepatica* by spatial transcriptomics

Open Mic Science Seminar at the Stowers Institute for Medical Research 30 June 2023
Invited Talk: **Puckelwaldt, O.**, Gramberg, S., Brewster, CE., Mann, F., Häberlein, S. Utilizing NGS technology to understand a neglected flatworm

Molecular Helminthology Meeting 2023, 26-29 June 2023, Madison, Wisconsin
Talk: **Puckelwaldt, O.**, Gramberg, S., Kopke, J., Samakovlis, C., Häberlein, S. Uncovering the cellular diversity of *Fasciola hepatica* utilizing single-cell transcriptomics

27th Helmithological Days 2023 of the Czech Society for Parasitology, 08-12 May 2023, Hradec Králové, Czech Republik
Invited Talk: **Puckelwaldt, O.**, Gramberg, S., Schmitt, T., Häberlein, S. A gene expression atlas by single-cell and spatial transcriptomics as valuable resource in liver fluke research

30th Annual meeting of the German Society for Parasitology 2023, 15-17 March 2023, Giessen, Germany.
Talk: **Puckelwaldt, O.**, Hoek, A., Kopke, J., Samakovlis, C., Goesmann, A., Häberlein, S. Towards a cell atlas for the liver fluke *Fasciola hepatica.*

15th Annual GGL Conference 2022, 14-15 September 2022, Giessen, Germany.
Talk: **Puckelwaldt, O.**, Hoek, A., Kopke, J., Samakovlis, C., Goesmann, A., Grevelding, CG., Häberlein, S. High-resolution transcriptomics of the liver fluke *Fasciola hepatica.*

15th International Conference of Parasitology (ICOPA) 2022, 21-26 August 2022, Copenhagen, Denmark.
Talk: **Puckelwaldt, O.**, Hoek, A., Kopke, J., Samakovlis, C., Goesmann, A., Grevelding, CG., Häberlein, S. Toward a cell atlas of *Fasciola hepatica* utilizing single cell transcriptomics.

14th Annual GGL Conference 2021, 29-30 September 2021, Giessen, Germany.
Talk: **Puckelwaldt, O.**, Hoek, A., Kopke, J., Samakovlis, C., Goesmann, A., Grevelding, CG., Häberlein, S. High-resolution transcriptomics of the liver fluke *Fasciola hepatica.*

First Digital GGL Annual Conference 2020, 29-30 September 2020, Giessen, Germany.
Talk: **Puckelwaldt, O.**, Hoek, A., Kopke, J., Samakovlis, C., Goesmann, A., Grevelding, CG., Häberlein, S. (2020). Single-cell and spatial transcriptomics to create a cell atlas of the liver fluke *Fasciola hepatica* for unraveling developmental processes.

11. Declarations

Selbstständigkeitserklärung

Ich (Oliver Puckelwaldt) erkläre, dass die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den in der Dissertation angegebenen Hilfsmitteln angefertigt worden ist. Alle Stellen, die veröffentlichten Schriften wörtlich oder sinngemäß entnommen sind, sowie alle Angaben, die auf mündlichen Auskünften beruhen, habe ich als solche kenntlich gemacht. Ich bin mit einer eventuellen Überprüfung meiner Dissertation durch eine Plagiatssoftware einverstanden. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis beachtet, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind.

_____ _____

Datum **Unterschrift**