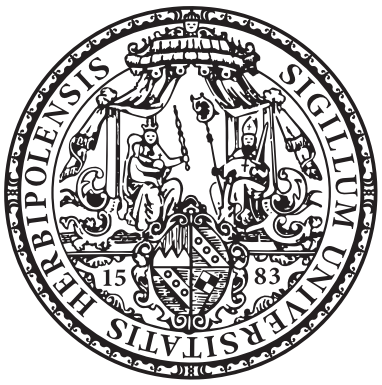


Genome wide expression profiling of *Echinococcus multilocularis*

Genomweite Expressionsanalysen von *Echinococcus multilocularis*

Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Infection and Immunity.

submitted by
Michaela Herz
from
Peißenberg

Würzburg, 2019

Submitted on: .....

Office stamp

Members of the Promotionskomitee:

Chairperson: Prof. Dr. Christian Janzen

Primary Supervisor: Prof. Dr. Klaus Brehm

Supervisor (Second): Prof. Dr. Jörg Schultz

Supervisor (Third): Prof. Dr. Nicolai Siegel

Date of Public Defense: .....

Date of Receipt of Certificates: .....

Addavit

I hereby confirm that my thesis entitled "Genome wide expression profiling of *Echinococcus multilocularis*" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis. Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, ...

Date Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation "Genomweite Expressionsanalysen von *Echinococcus multilocularis*" eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben. Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg, ...

Datum Unterschrift

Acknowledgments - Danksagung

An dieser Stelle bedanke ich mich ganz herzlich bei allen, die zum Gelingen dieser Arbeit beigetragen haben:

Prof. Klaus Brehm für die Möglichkeit, in seiner Arbeitsgruppe zu promovieren, die gute Betreuung und Unterstützung, sowie die vielen interessanten Gespräche.

Prof. Jörg Schultz und Prof. Nicolai Siegel für die Übernahme der Gutachten, die anregenden Diskussionen und die Teilnahme an der Prüfungskomission.

Dr. Matt Berriman für die Möglichkeit in seiner Arbeitsgruppe die Grundlagen der Transkriptionsdatenanalyse zu erlernen und die gute Zusammenarbeit.

Prof. Matthias Frosch für die freundliche Aufnahme am Institut für Hygiene und Mikrobiologie.

Der Echinokokkenarbeitsgruppe: Monika Bergmann, Dirk Radloff, Markus Spiliotis, Raphaël Duvoisin und Ruth Hermann für die angenehme Arbeitsatmosphäre, die kollegiale Unterstützung und die vielen interessanten und motivierenden Gespräche.

Der Arbeitsgruppe von Matt Berriman, insbesondere Avril Coghan, für die anregenden Diskussionen und die gute Zusammenarbeit.

Den Azubis für die tatkräftige Unterstützung und Hilfe im Laboralltag.

Allen Mitarbeitern des Instituts für ihre unermüdliche Hilfsbereitschaft.

Meiner Familie und meinen Freunden für ihre moralische Unterstützung.

Contents

1 Summary

Alveolar echinococcosis, which is caused by the metacestode stage of the small fox tapeworm *Echinococcus multilocularis*, is a severe zoonotic disease with limited treatment options. For a better understanding of cestode biology the genome of *E. multilocularis*, together with other cestode genomes, was sequenced previously. While a few studies were undertaken to explore the *E. multilocularis* transcriptome, a comprehensive exploration of global transcription profiles throughout life cycle stages is lacking. This work represents the so far most comprehensive analysis of the *E. multilocularis* transcriptome. Using RNA-Seq information from different life cycle stages and experimental conditions in three biological replicates, transcriptional differences were qualitatively and quantitatively explored. The analyzed datasets are based on samples of metacestodes cultivated under aerobic and anaerobic conditions as well as metacestodes obtained directly from infected jirds. Other samples are stem cell cultures at three different time points of development as well as non-activated and activated protoscoleces, the larval stage that can develop into adult worms. In addition, two datasets of metacestodes under experimental conditions suitable for the detection of genes that are expressed in stem cells, the so-called germinative cells, and one dataset from a siRNA experiment were analyzed. Analysis of these datasets led to expression profiles for all annotated genes, including genes that are expressed in the tegument of metacestodes and play a role in host-parasite interactions and modulation of the host's immune response. Gene expression profiles provide also further information about genes that might be responsible for the infiltrative growth of the parasite in the liver.

Furthermore, germinative cell-specific genes were identified. Germinative cells are the only proliferating cells in *E. multilocularis* and therefore of utmost importance for the development and growth of the parasite. Using a combination of germinative cell depletion and enrichment methods, genes with specific expression in germinative cells were identified. As expected, many of these genes are involved in translation, cell cycle regulation or DNA replication and repair. Also identified were transcription factors, many of which are involved in cell fate commitment. As an example, the gene encoding the telomerase reverse transcriptase (TERT) was studied further. Expression of *E. multilocularis tert* in germinative cells was confirmed experimentally. Cell culture experiments indicate that TERT is required for proliferation and development of the parasite, which makes TERT a potentially interesting drug target for chemotherapy of alveolar echinococcosis.

Germinative cell specific genes in *E. multilocularis* also include genes of densoviral origin. More than 20 individual densovirus loci with information for non-structural and structural densovirus proteins were identified in the *E. multilocularis* genome. Densoviral elements were also detected in many other cestode genomes. Genomic integration of these elements suggests that densovirus-based vectors might be suitable tools for genetic manipulation of tapeworms. Interestingly, only three of more than 20 densovirus loci in the *E. multilocularis* genome are expressed. Since the canonical piRNA pathway is lacking in cestodes, this raises the question about potential silencing mechanisms. Exploration of RNA-Seq information indicated natural antisense transcripts as a potential gene regulation mechanism in *E. multilocularis*. Preliminary experiments further suggest DNA-methylation, which was previously shown to occur in platyhelminthes, as an interesting avenue to explore in future.

The transcriptome datasets also contain information about genes that are expressed in differentiated cells, for example the serotonin transporter gene that is expressed in nerve cells. Cell culture experiments indicate that serotonin and serotonin transport play an important role in *E. multilocularis* proliferation, development and survival.

Overall, this work provides a comprehensive transcription data atlas throughout the *E. multilocularis* life cycle. Identification of germinative cell-specific genes and genes important for host-parasite interactions will greatly facilitate future research. A global overview of gene expression profiles will also aide in the detection of suitable drug targets and the development of new chemotherapeutics against alveolar echinococcosis.

2 Zusammenfassung

Alveoläre Echinokokkose wird durch das Metazestodenstadium des kleinen Fuchsbandwurms *Echinococcus multilocularis* verursacht und medizinisch als eine schwere Zoonose mit begrenzten Behandlungsmöglichkeiten betrachtet. Um ein besseres Verständnis für die Biologie der Zestoden zu erlangen, wurde das Genom von *E. multilocularis*, zusammen mit denen anderer Zestoden, bereits sequenziert. Bisher wurden nur wenige Studien zum Transkriptom von *E. multilocularis* durchgeführt und eine umfassende Analyse der Transkriptionsprofile über verschiedene Stadien des Lebenszykluses hinweg fehlt bislang. Diese Arbeit stellt die bisher umfassendste Untersuchung des Transkriptoms von *E. multilocularis* dar. Unterschiede in der Genexpression in verschiedenen Stadien des Lebenszykluses und unter experimentellen Bedingungen wurden qualitativ und quantitativ untersucht. Dazu wurden Daten aus RNA-Sequenzierungen in drei biologischen Replikaten verwendet. Die untersuchten Datensätze beruhen auf Proben von Metazestoden, die unter aeroben und anaeroben Bedingungen kultiviert, sowie von Metazestoden, die direkt aus Gerbilen isoliert wurden. Weitere Proben umfassen Stammzellkulturen zu drei verschiedenen Entwicklungszeitpunkten sowie nicht-aktivierte und aktivierte Protoskolizes, das Larvenstadium das sich zu Adulten entwickeln kann. Zusätzlich wurden zwei Datensätze von Metazestoden unter experimentellen Bedingungen, die zur Identifizierung stammzellspezifischer (keimzellspezifischer) Gene geignet sind, sowie ein Datensatz von einem siRNA-Experiment untersucht. Die Analyse dieser Datensätze führte zu Genexpressionsprofilen für alle annotierten Gene, unter anderem für Gene, die im Tegument des Metazestoden exprimiert werden und eine Rolle spielen bei Wirt-Parasit-Interaktionen und der Modulierung der Immunantwort des Wirts. Genexpressionsprofile liefern zudem Informationen über Gene, die für das infiltrative Wachstum des Parasiten in der Leber verantwortlich sein könnten.

Des Weiteren wurden keimzellspezifische Gene identifiziert. Keimzellen sind die einzigen proliferierenden Zellen in *E. multilocularis* und daher von essentieller Bedeutung für die Entwicklung und das Wachstum des Parasiten. Durch eine Kombination von Keimzelldepletierungs- und Keimzellanreicherungsverfahren wurden Gene mit keimzellspezifischer Expression identifiziert. Wie erwartet, sind viele dieser Gene in der Translation, der Zellzyklusregulation oder DNA-Replikation und -Reparatur involviert. Darüber hinaus wurden keimzellspezifisch exprimierte Transkriptionsfaktoren detektiert, von denen viele in der Festlegung des Zellschicksals eine Rolle spielen. Als Beispiel eines keimzellspezifischen Genes wurde das Gen, das für die

reverse Transkriptase (TERT) kodiert, genauer untersucht. Die Expression von *E. multilocularis tert* in Keimzellen wurde experimentell bestätigt. Zellkulturexperimente weisen darauf hin, dass TERT für die Proliferation und die Entwicklung essentiell ist. TERT ist daher ein potentiell interessantes Wirkstofftarget für die chemotherapeutische Behandlung der alveolären Echinokokkose.

Zu den keimzellspezifischen Genen in *E. multilocularis* gehören auch Gene densoviralen Ursprungs. Es wurden mehr als 20 Densovirusloci mit Informationen für nicht-strukturelle und strukturelle Densovirusproteine im *E. multilocularis*-Genom identifiziert. Densovirale Elemente wurden auch in vielen anderen Zestodengenomen detektiert. Die genomische Integration dieser Element deutet darauf hin, dass densovirus-basierte Vektoren zur genetischen Manipulation von Zestoden geeignet sein könnten. Interessanterweise sind nur drei von mehr als 20 Densovirusloci im *E. multilocularis*-Genom exprimiert. Da es in Zestoden keinen kanonischen pi-RNA-Signalweg gibt, stellt sich die Frage nach möglichen Genabschaltungsmechanismen. Die Analyse der RNA-Sequenzierdaten ergab Hinweise auf natürliche Antisense-Transkripte als einen möglichen Genregulationsmechanismus in *E. multilocularis*. Vorläufige Experimente und bisherige Studien deuten weiterhin darauf hin, dass DNA-Mehylierung ein Mechanismus der Genregulation und -abschaltung in Zestoden sein könnte.

Die Transkriptionsdaten enthalten auch Informationen zu Genen, die in differenzierten Zellen exprimiert werden, wie zum Beispiel das Serotonintransportergen, das in Nervenzellen exprimiert wird. Zellkulturversuche weisen daruf hin, dass Serotonin und Serotonintransport eine wichtige Rolle bei der Proliferation, der Entwicklung und dem Überleben von *E. multilocularis* spielen.

Insgesamt bietet diese Arbeit einen umfassenden Transkriptionsdatenatlas über die Stadien des Lebenszykluses von *E. multilocularis*. Die Identifizierung von keimzellspezifischen Genen und Genen, die für die Interaktion zwischen Wirt und Parasit wichtig sind, wird die zukünftige Forschung erheblich erleichtern. Ein globaler Überblick über die Genexpressionsprofile wird zudem hilfreich sein bei der Entdeckung geeigneter Wirkstofftargets und bei der Entwicklung neuer Chemotherapeutika gegen die alveoläre Echinokokkose.

3 Introduction

3.1 The small fox tapeworm *Echinococcus multilocularis*

3.1.1 Phylogeny and epidemiology of *E. multilocularis*

Echinococcus multilocularis is the causative agent of alveolar echinococcosis, a severe zoonotic disease. Two other forms of echinococcosis are known to occur in humans: cystic echinococcosis caused by *Echinococcus granulosus* and the less frequent polycystic echinococcosis caused by *Echinococcus vogeli* and *Echinococcus oligarthrus* (Eckert and Deplazes, 2004). *E. multilocularis* is mainly distributed in the northern hemisphere (see Figure 1) (Torgerson et al., 2010), *E. granulosus* is found worldwide (Eckert and Deplazes, 2004) and *E. vogeli* as well as *E. oligarthrus* are restricted to Central and South America (Eckert and Deplazes, 2004).

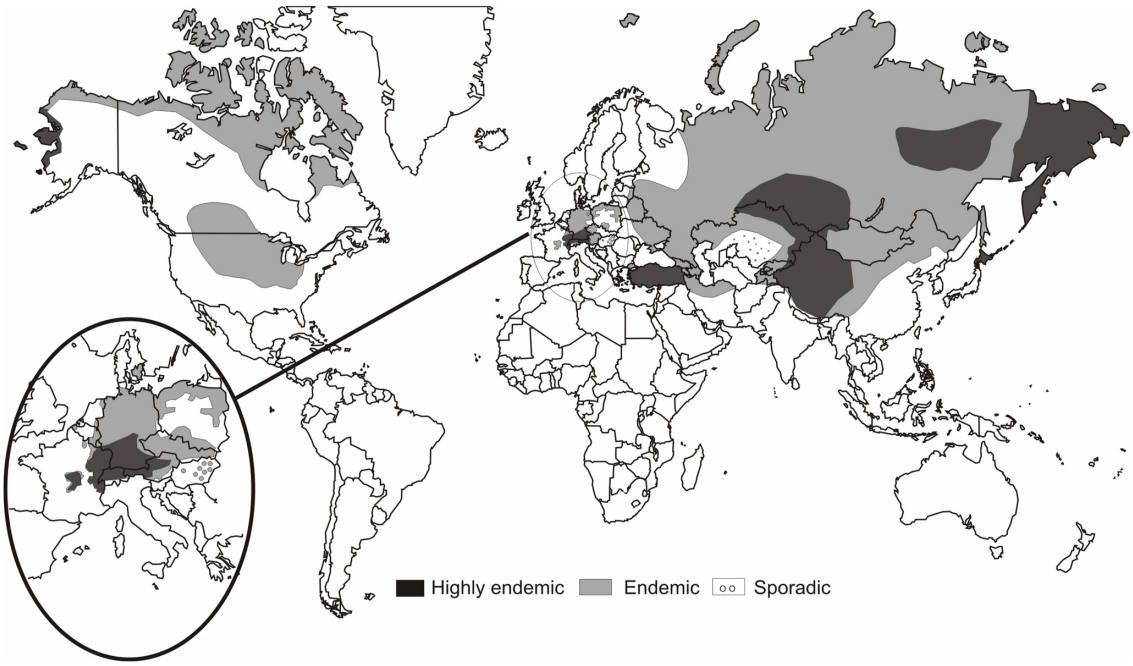


Figure 1: **Global distribution of AE** by (Torgerson et al., 2010), retrieved from https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0000722 Used under Creative Commons Attribution License. **AE**: alveolar echinococcosis.

Historically only four *Echinococcus* species were recognized as valid: *E. granulosus* (Batsch 1786), *E. multilocularis* (Leuckart 1863), *E. oligarthrus* (Diesing 1863) and *E. vogeli* (Rausch and Bernstein, 1972). With the discovery of *Echinococcus shiquicus* (Xiao et al., 2005) and the subdivision of *E. granulosus sensu lato* in *E. granulosus sensu stricto*, *E. felidis*, *E. equinus*, *E. ortleppi*, *E. canadensis*, *E. intermedius* and *E. borealis*, there are now at least ten described species (Romig et al.,

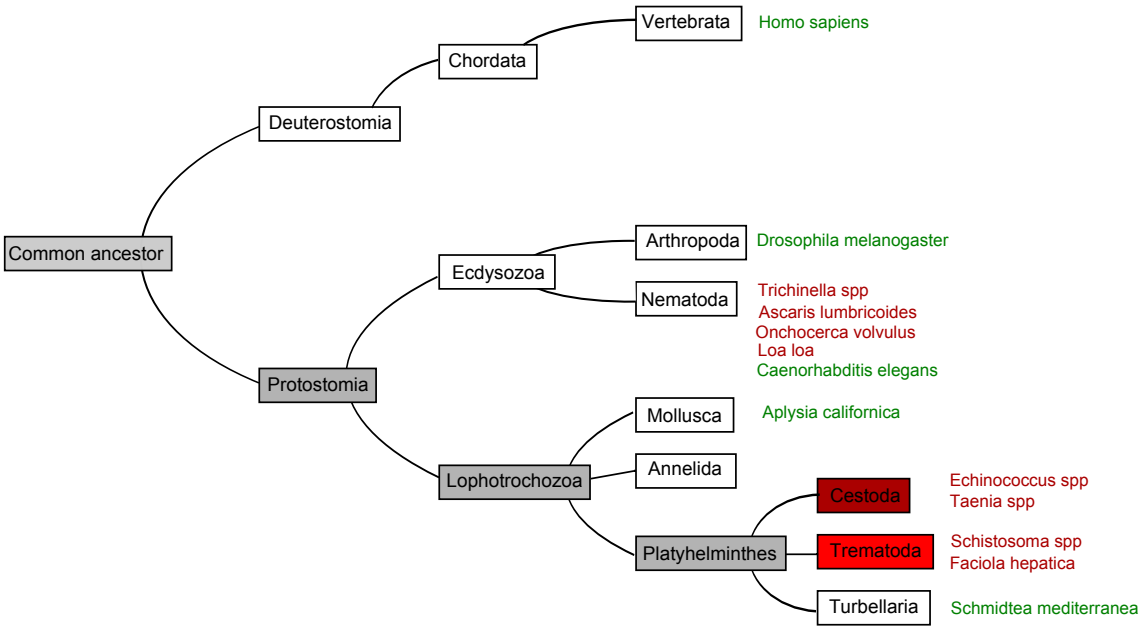


Figure 2: **Phylogenetic position of the genus *Echinococcus*** (schematic taken from Herz, 2015): Parasitic species are written in red, non-parasitic species in green.

2015; Lymbery et al., 2015*b*,*a*; Thompson and Jenkins, 2014). In the following text, *E. granulosus* will always refer to *E. granulosus sensu lato* unless specifically stated otherwise.

The genus *Echinococcus* is a member of the family Taenidae which belongs to the order Cyclophyllidea. Cyclophyllidea are placed in the subclass Eucestda of the class Cestoda which in turn is a member of the phylum Platyhelminthes (see Figure 2). Platyhelminthes consist of the parasitic classes Cestoda and Trematoda and the free living Turbellaria (Ehlers, 1986). Cestoda and Trematoda are monophyletic, while Turbellaria from a paraphylectic taxa (Littlewood et al., 1999). The important model organism for regeneration *Schmidtea mediterranea* (Sanchez Alvarado, 2004) is a member of the Turbellaria. Important members of the Trematoda are the pathogens of schistosomiasis *Schistosoma mansoni* and *Schistosoma haematobium* as well as the common liver fluke *Faciola hepatica*. Platyhelminthes belong to the Lophotrochozoa and are Protostomia (Olson et al., 2012). Protostomia also include the molting Ecdysozoa (Giribet, 2008). Members of the Ecdysozoa are the fruit fly *Drosophila melanogaster* (Arthropoda) and parasitic and non-parasitic worms (Nematoda) such as *Ascaris lumbricoides*, *Onchocerca volvulus*, *Loa loa* and *Caenorhabditis elegans*.

3.1.2 Life cycle and biology of *E. multilocularis*

E. multilocularis is a small tapeworm with a length of up to 4,5 mm. The adult worm resides in the small intestine of definitive hosts, usually foxes. Canids and felids can also be infected as definitive hosts. Typical intermediate hosts are small mammals, especially rodents. Humans are considered aberrant intermediate host as infection of humans disrupts the life cycle (Eckert and Deplazes, 2004). Infection of intermediate hosts occurs by oral uptake of infectious eggs, which harbor the oncosphere. In the small intestine the oncosphere hatches from the egg and penetrates the epithelium. With the blood flow the oncosphere is then transported to the inner organs, particularly the liver, where it transforms into the metacestode (Brehm et al., 2006). The metacestode vesicle consists of an outer acellular layer, the laminated layer that is composed of carbohydrates and proteins (Brehm, 2010), and an inner cellular layer, the germinal layer. Metacestode vesicles develop daughter cysts by budding and infiltrate the liver tissue (Dixon, 1997). Later in infection brood capsules form through invagination of the germinal layer and subsequently protoscoleces develop within the brood capsules (Koziol et al., 2016) (see Figure 3).

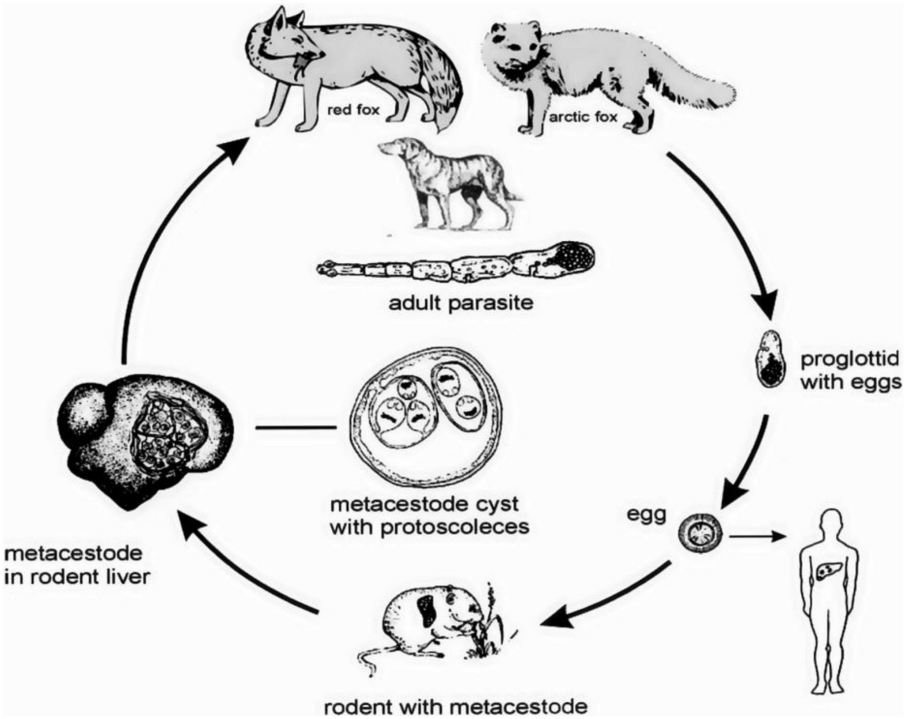


Figure 3: **The life cycle of *Echinococcus multilocularis*** by (Torgerson et al., 2010), retrieved from https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0000722 Used under Creative Commons Attribution License.

When the definitive host takes the prey, the surrounding tissue is digested and the protoscoleces are released and activated. The activated protoscoleces move inside the small intestine and settle between the villi in the crypts of Liberkühn by attaching to the epithelium with their hooks and suckers (Thompson and Eckert, 1983; Smyth, 1968). Protoscoleces then develop into adult worms by growth and segmentation. After self insemination, adult worms develop embryonated eggs containing the 6-hooked oncosphere (Thompson and Eckert, 1982). The last proglottid containing the eggs is shed with the feces and the eggs are distributed in the environment where they are taken up again by intermediate hosts.

The life cycle of the closely related *E. granulosus* is very similar (see Figure 4). One of the major differences between *E. multilocularis* and *E. granulosus* are their host preferences. Definitive hosts of *E. granulosus* are generally dogs or other carnivores and intermediate hosts are unusually ungulates such as sheep, goats or cows Eckert and Deplazes (2004).

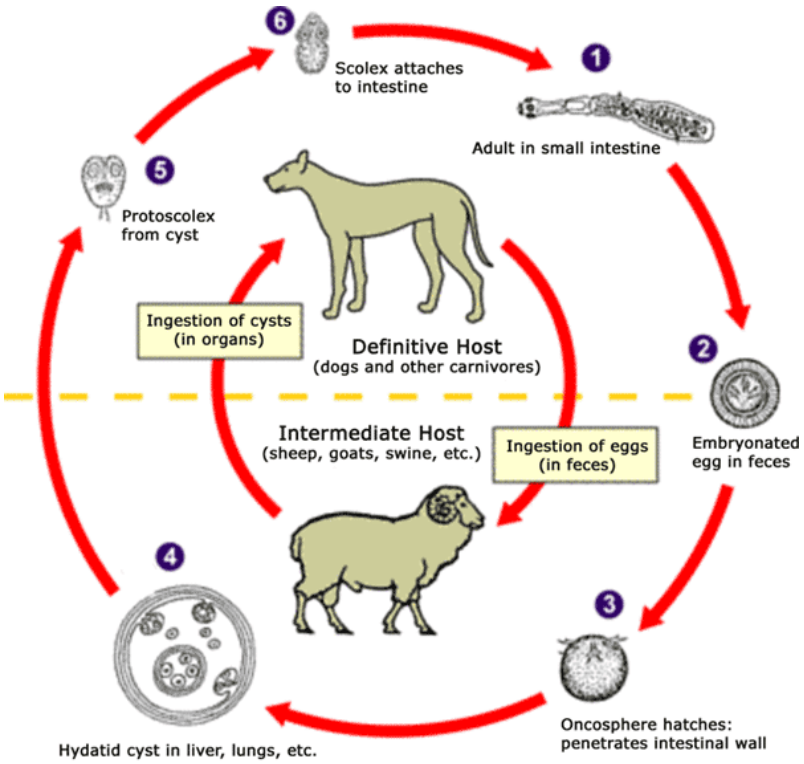


Figure 4: **Life cycle of Echinococcus** by the Centers for Disease Control and Prevention, retrieved from https://commons.wikimedia.org/wiki/File:Echinococcus_-Life_Cycle.png The image is in the public domain. Shown is the *E. granulosus* life cycle.

Another difference is their organ tropism. *E. multilocularis* in humans almost exclusively infects the liver (99 %), whereas *E. granulosus* is also found in the lungs

(17 %) (Eckert and Deplazes, 2004). In contrast to the *E. granulosus* metacestode that forms one massive cyst with a thick laminated layer, the *E. multilocularis* metacestode consists of a multi-vesicular parasite tissue with a thin laminated layer and has the capacity to metastasize into other organs such as brain, heart and kidney (Brehm, 2010).

3.2 Alevolar echinococcosis

Alveolar echinococcosis is the most dangerous parasitosis in Europe. The disease is caused by the metacestode larval stage of *E. multilocularis*. Untreated alveolar echinococcosis causes death in over 90 % of patients within 10-15 years after diagnosis (Craig, 2003). Since the introduction of the reporting obligation in 2001, between 6 and 49 new cases have been reported annually in Germany (Robert Koch-Institut, 06.03.2019). Incidence of cystic echinococcosis is higher than incidence of alveolar echinococcosis (see Figure 5) though it is thought that most cases of cystic echinococcosis were not acquired in Germany. Both alveolar and cystic echinococcosis are rare diseases in Europe.

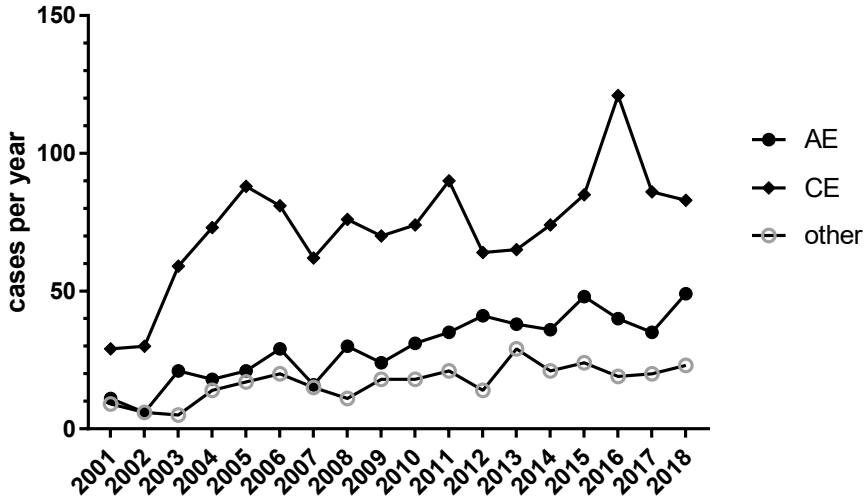


Figure 5: **Echinococcosis cases per year reported in Germany according to the Robert Koch-Institut (06.03.2019)**: (**AE**): alveolar echinococcosis, (**CE**): cystic echinococcosis, (**other**): undetermined, unspecified or other types of echinococcosis.

Highest endemicity for alveolar and cystic echinococcosis has been reported for the Sichuan Province in China with a prevalence of 3.1% for alveolar echinococcosis and 3.2% for cystic echinococcosis (Li et al., 2010). The global burden is estimated at 666.434 disability-adjusted life years (DALYs) for alveolar echinococcosis

and 285.500 DALYs for cystic echinococcosis Wen et al. (2019). Assessment of risk factors for infection with *E. multilocularis* and *E. granulosus* is difficult due to generally low incidence and long asymptomatic incubation periods. Likely risk factors are contact with dogs or foxes and raising livestock (Eckert and Deplazes, 2004; Wen et al., 2019).

Infection of humans occurs through oral uptake of infected eggs containing the oncosphere. After hatching in the small intestine, the oncosphere penetrates the epithelium and is transported to the liver with the blood flow. In the liver the oncosphere undergoes metamorphosis and transforms into the metacestode (Brehm et al., 2006). The incubation period is estimated at 5 to 15 years (Eckert and Deplazes, 2004).

After the initial asymptomatic phase, the infection may be cured spontaneously or progress. In the progressive phase, the *E. multilocularis* metacestode can grow infiltratively like a malignant tumor (Dixon, 1997) and even metastasize (Brehm, 2010). First symptoms include abdominal pain, nausea, vomiting, fever and anemia. With increasing destruction of the liver tissue, symptoms of hepatic dysfunction, such as icterus, cholestasis, portal hypertension and thrombosis of the portal vein (Kern, 2010), occur.

Diagnosis of alveolar echinococcosis is based on clinical findings, imaging techniques, histopathology, nucleic acid detection and serology (Kern, 2010). Imaging techniques are often the first used for diagnosis and combined with laboratory diagnosis, particularly serology, for confirmation (Craig, 2003). Widely used imaging techniques are ultrasonography (Rogan et al., 2006), then computer tomography (CT) and magnetic resonance tomography (MRT) (Moro and Schantz, 2009). Laboratory diagnostics include detection of antigens and parasite specific antibodies with ELISA and Western blot. The spread of the disease can be classified using the PNM-system which has been developed based on the TNM-system for malignant tumors and evaluates the extension of the lesion in the liver (P), the infestation of lymph nodes or neighboring organs (N) and the presence of metastases (M) Kern et al. (2006).

Treatment of alveolar echinococcosis is difficult as the disease can only be cured by complete removal of the parasite tissue (Moro and Schantz, 2009). Radical surgical resection of the lesion is the first line therapy. Nonetheless, curative therapy is only possible in 20-30 % of cases (Robert-Koch-Institut, 2006). Most patients need lifelong drug therapy with benzimidazoles (mebendazole, albendazole) which, however, can only inhibit the growth of the parasite at applied doses and is associated with severe side effects (Reuter et al., 2004). Benzimidazoles inhibit microtubuli formation in the parasite through binding of β-tubulin (Brehm, Jensen and Frosch, 2000;

Jura et al., 1998). Nevertheless, it is likely that the parasite stem cells (so-called germinative cells) which are crucial for proliferation, express a different isoform of β-tubulin and are thus less susceptible to benzimidazoles (Brehm and Koziol, 2014).

Cystic echinococcosis, which is caused by *E. granulosus*, usually manifests as a solitary cyst with a size of 1 to 15 cm (Eckert and Deplazes, 2004). In humans cysts are found in many anatomic sites, most frequently in the liver and lungs. Rupture of cysts can cause anaphylactic reactions (Vuitton, 2004) and secondary cystic echinococcosis through release of protoscoleces (Eckert and Deplazes, 2004). As with alveolar echinococcosis curative treatment is achieved by complete removal of the cyst, which can be accomplished in a high proportion of patients (Wen et al., 2019; Eckert and Deplazes, 2004). Alternative treatment options are the percutaneous "PAIR" (puncture, aspiration, injection, and reaspiration) technique and chemotherapy(Brunetti et al., 2010; Wen et al., 2019). For biologically inactive cysts a "watch-and-wait" approach can be considered (Brunetti et al., 2010; Wen et al., 2019).

3.3 Advances in *E. multilocularis* research

3.3.1 *In vitro* cultivation system of *E. multilocularis*

First *in vitro* cultivation systems for *E. multilocularis* were based on a tissue block system (Hemphill and Gottstein, 1995) or co-cultivation of homogenized parasite material with rat hepatoma cells (Jura et al., 1996). Both systems require the presence of serum and host cells for parasite growth and development. Absence of the so-called feeder cells leads to fast degeneration of the parasite tissue. It is presumed that secreted factors from feeder cells are necessary for parasite growth and development (Spiliotis et al., 2004; Brehm, 2010).

To enable the study of drug effects on parasite tissue without the presence of confounding host cells, an axenic cultivation system was developed for the cultivation of metacestodes (Spiliotis et al., 2004) (see Figure 6 A). In axenic cultivation, reducing conditions (nitrogen atmosphere and reducing agents) are necessary for survival of metacestodes. In presence of serum and under reducing conditions metacestodes survive for several weeks. Growth and development of metacestodes, however, additionally requires the use of feeder-cell-pre-conditioned medium (Spiliotis et al., 2004). Primary cells can be isolated from axenic metacestodes and cultivated in axenic culture without reducing agents (Spiliotis and Brehm, 2009). 2-day-old pri-

mary cell cultures are highly enriched in stem cells, the so-called germinative cells (Koziol et al., 2014). Primary cells aggregate within a couple of days (see Figure 6 B), subsequently form central cavities, so-called "red dots", and after 1-3 weeks develop metacestode vesicles (see Figure 6 C).

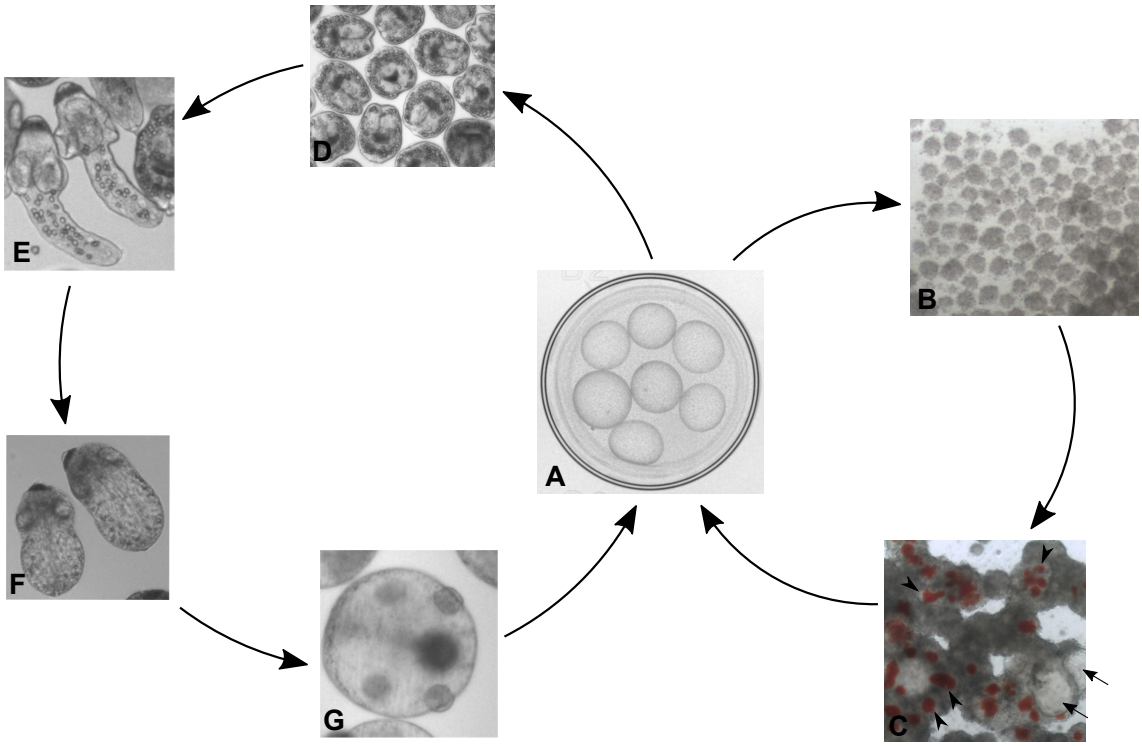


Figure 6: ***In vitro* cultivation of *Echinococcus* larval stages:** **A**: metacestodes; **B**: early primary cell cultures; **C**: late primary cell cultures with "red dots" (arrowheads) and developing metacestode vesicles (arrows); **D**: non-activated protoscoleces; **F**: activated protoscoleces; **F:** re-differentiating protoscoleces; **G:** completely vesicularized protoscolex.

Metacestodes of isolates that still have the capacity to develop protoscoleces (Tsai et al., 2013) can develop brood capsules and protoscoleces *in vitro* (Spiliotis et al., 2008). To obtain sufficient amounts of protoscoleces, however, protoscoleces are isolated directly from parasite material of infected jirds (Brehm, Kronthaler, Jura and Frosch, 2000; Spiliotis and Brehm, 2009) or in case of *E. granulosus* isolated from hydatid fluid of metacestode cysts obtained from an abattoir (Fernandez et al., 2002) (see Figure 6 D). Protoscoleces are activated by mimicking the gastrointestinal passage through treatment with pepsin and low pH, followed by incubation in sodium taurocholate (Fernandez et al., 2002) (see Figure 6 E). Protoscoleces can undergo a redifferentiation process (see Figure 6 F and G) at the end of which they become completely vesicularized per-microcysts and microcysts (Heath and Osborn, 1976; Rodriguez-Caabeiro and Casado, 1988; Walker et al., 2004).

Cultivation of *E. granulosus* is basically limited to protoscoleces. *E. granulosus* metacestodes increase in size but do not proliferate, which makes a constant supply of metacestode material difficult (Hemphill et al., 2010). In contrast, turn proliferation of *E. multilocularis* metacestodes *in vitro* and the primary cell culture system *E. multilocularis* into a suitable model for the study of drug effects and developmental processes.

3.3.2 Molecular and biochemical approaches in *E. multilocularis* research

The first approach to analyze gene expression patterns of different larval stages in *E. multilocularis* was based on spliced leader trans-spliced transcripts (Brehm et al., 2003). In spliced leader trans-splicing a donor RNA, the so-called spliced leader, is connected to the 5' end exon of a RNA transcript (Hastings, 2005; Lasda and Blumenthal, 2011). In *E. multilocularis* around 14 % of genes are trans-spliced (Ference Kiss, personal communication).

The whole genome sequencing project for *E. multilocularis* was started in 2008 by the Wellcome Trust Sanger Institute in cooperation with the group of Klaus Brehm (Würzburg, Germany). In 2013 the *E. multilocularis* genome was published, together with the genomes of three other tapeworms: *E. granulosus*, *Taenia solium*, *Hymenolepis microstoma* (Tsai et al., 2013). An analysis of the *E. granulosus* transcriptome based on approximately 10,000 ESTs was published in 2012 (Parkinson et al.) and a study on the gene expression profile in *E. multilocularis* oncospheres and early metacestodes in 2016 (Huang et al.). Transcriptome data of various *E. multilocularis* and *E. granulosus* larval stages and conditions (Wellcome Trust Sanger Institute in cooperation with the work groups of Klaus Brehm and Cecilia Fernandez) was analyzed in this work. The high quality genome of *E. multilocularis* together with the transcriptome information provide important resources for future research and the development of reverse genetic toolkits. While robust methods for functional genomic analysis of *E. multilocularis* are still lacking, the development of RNA-interference methods for primary cells (Spiliotis et al., 2010) and protoscoleces (Mizukami et al., 2010) offers first approaches for functional investigation of parasite genes.

3.3.3 Evolutionarily conserved signaling pathways in *E. multilocularis*

Due to their early rise in metazoan evolution long before before the Cambrian explosion, the majority of cell-cell communication systems are conserved between

deuterostomes and protostomes (Brehm, 2010). Not only are signaling systems of non-vertebrates and mammals related, they are often also functionally exchangeable. Early studies for example showed the stimulation of the *D. melanogaster* insulin receptor by mammalian insulin (Fernandez et al., 1995). Likewise, the *E. multilocularis* insulin receptor also binds human insulin and IGF-I in addition to the *E. multilocularis* insulin-like peptides and activates the PI3K/Akt signaling pathway (Konrad et al., 2003; Hemer et al., 2014). Further conserved cell-cell communication systems in *E. multilocularis* include EGF/FGF, TGF-β/BMP, nuclear hormone receptor, hedgehog and wingless (wnt) signaling pathways (Brehm and Koziol, 2017). The study of conserved signaling pathways can reveal mechanisms of host-parasite interactions. Host-parasite cross-communication could explain the influence of host factors on parasite growth and development as well as modulation of the host immune response by the parasite.

3.3.4 The *E. multilocularis* stem cell system

Undifferentiated stem cells in cestodes, designated germinative cells, are similar to the so-called neoblasts of free-living flatworms (Koziol et al., 2014). Germinative cells are the only proliferating cells in *E. multilocularis* and the origin of all differentiated cell types (Koziol et al., 2014). While homogeneous in morphology, on a molecular level they appear to consist of sub-populations expressing different molecular markers, such as transcripts encoding homologues of the post-transcriptional regulators *nanos* and *argonaute* (Koziol et al., 2014). Germinative cells can be specifically depleted in *E. multilocularis* metacestodes by 7 day treatment with hydroxyurea (HU), an inhibitor of ribonucleotide reductase, or by 21 day treatment with the polo-like kinase inhibitor Bi2536. Both treatments result in a 90 % reduction of EdU incorporating cells without affecting differentiated cells or the structural integrity of metacestode vesicles (Koziol et al., 2014; Schubert et al., 2014). Enrichment of germinative cells can be achieved through primary cell isolation (Spiliotis et al., 2008; Spiliotis and Brehm, 2009). 2-day-old primary cell cultures consist of 62 % to 83 % germinative cells in comparison to 32 % to 55 % in metacestodes (Koziol et al., 2014). Primary cells are able to develop complete metacestode vesicles *in vitro* (Spiliotis et al., 2008), showing the remarkable regeneration capabilities of germinative cells. As the only proliferating cell type, germinative cells should be responsible for parasite growth, metastasis and recurrence after discontinuation of chemotherapy (Brehm and Koziol, 2014).

3.4 Transcriptomic analysis of stem cells

To identify genes that are specifically expressed in stem cells, the transcriptome of stem cells needs to be compared to the transcriptome of differentiated cells. While stem cells can be separated from differentiated cells by FACS sorting (Iriondo et al., 2015), for non-model organisms specific cell surface markers are often unavailable (Hayashi and Agata, 2012). Subtractive methods are then often used to characterize the gene expression of proliferating cells by comparison of the transcriptome of stem cell depleted organisms with the transcriptome of intact organisms. In planarians and schistosomes depletion of proliferating cells can be achieved by irradiation or RNA interference (Collins et al., 2013; Rossi et al., 2007; Eisenhoffer et al., 2008; Solana et al., 2012; Wagner et al., 2012).

For *E. multilocularis* two methods of germinative cell depletion by drug treatment were developed (Koziol et al., 2014; Schubert et al., 2014) (see section 3.3.4). Recent methods for transcriptomic analysis of stem cell also include a subtractive FACS method for isolation of planarian neoblasts (Hayashi and Agata, 2018) and single cell sequencing of isolated stem cells (Wen and Tang, 2016). While isolation of germinative cells from *E. multilocularis* by FACS sorting was unsuccessful so far (Markus Spiliotis, personal communication), enrichment of germinative cells by primary cell isolation (Spiliotis and Brehm, 2009; Koziol et al., 2014) offers another approach to characterize the transcriptome of germinative cells in *E. multilocularis*.

3.5 Wnt signaling pathway

Wnt ligands are secreted glycoproteins involved in cell-cell signaling and regulation of various developmental processes, such as cell proliferation, cell differentiation and axial patterning in all metazoan organisms (Mikels and Nusse, 2006). They also play a major role in maintenance of stem cell pluripotency (Reya and Clevers, 2005). Depending on the context, Wnt ligands can bind to different receptors, such as receptors of the Frizzled family and to the tyrosin kinases Ror2 and RYK (van Amerongen and Nusse, 2009). In the canonical (β-catenin dependent) pathway, wnt ligands bind to Frizzled/LRP5-6 receptors, causing the recruitment and activation of Dishevelled. This leads to disassembly of the β-catenin destruction complex that includes Axin, APC and GSK-3. β-catenin is no longer phosphorylated and degraded and is trans-located into the nucleus where it regulates transcription of target genes (Almuedo-Castillo et al., 2012). Non-canonical Wnt signaling pathways include the planar cell polarity pathway and the calcium pathway (van Amerongen and Nusse, 2009).

The role of the Wnt signaling pathway is well studied in free-living planarian flatworms (Almuedo-Castillo et al., 2012). In planarians the canonical/β-catenin dependent Wnt signaling pathway specifies posterior identity and regulates head/tail specification during regeneration (Iglesias et al., 2008; Petersen and Reddien, 2008; Gurley et al., 2008). According to *in silico* analysis, all major Wnt signaling components are present in all analyzed flatworms, including in *E. multilocularis* (Riddiford and Olson, 2011). Recently, gene expression of Wnt components and markers of anterior and posterior specification has been studied in *E. multilocularis* and *H. microstoma* (Koziol et al., 2016), revealing that the scolex is the true anterior of tapeworms. According to gene expression analysis, the germinal layer of the *E. multilocularis* metacestode is completely posteriorized tissue and localized expression of Wnt inhibitors precedes the development of brood capsules Koziol et al. (2016).

3.6 DNA methylation

The genome is epigenetically regulated at various levels ranging from structural compartmentalization, to nucleosome positioning, histone and DNA modifications (Du et al., 2015). DNA methylation is the most studied epigenetic modification in eukaryotes and regulates gene expression, X chromosome inactivation, allelic exclusion, embryogenesis and repetitive element silencing (Meier and Recillas-Targa, 2017; Massah et al., 2015; Canovas et al., 2017; Maksakova et al., 2008; Miura et al., 2001).

DNA-methyltransferases (DNMTs) transfer methyl groups to the carbon-5 position of cytosine residues, creating 5-methylcytosines (5-mC) (Du et al., 2015). In vertebrates 5mC-methylation mostly occurs on CpG dinucleotides in a symmetric manner (Meier and Recillas-Targa, 2017). However, also cytosines in CpA, CpT, and CpC dinucleotides have been found to be methylated (Patil et al., 2014). In the classical model DNMT3a and DNMT3b, which do not discriminate between unmethylated and hemimethylated DNA (Jurkowska et al., 2011), are responsible for *de novo* methylation, while DNMT1, which has a preference for hemimethylated DNA (Jeltsch, 2006), plays a role in methylation maintenance. However, the results of various experiments are not compatible with a strict separation of DNMT3 and DNMT1 into *de novo* and maintenance enzymes (Jeltsch and Jurkowska, 2014). A dynamic stochastic model based on local activity of DNMTs, DNA demethylases and DNA replication rate has been proposed to describe DNA methylation(Jeltsch and Jurkowska, 2014). DNMT2 is another member of the DNA methyltransferases

and the most widely conserved in animals, fungi, plants and protists (Schaefer and Lyko, 2010). DNMT2 possesses conserved catalytic motifs predicting DNA-methyltransferase activity and is able to methylate cytosine residues in *Drosophila* without apparent sequence specificity (Schaefer and Lyko, 2010). Besides its believed weak DNA-methyltransferase activity, DNMT2 is known to methylate tRNAs (Schaefer and Lyko, 2010).

Methylated DNA is recognized and bound by proteins of the methyl-CpG-binding domain (MBD) family. The MBD family consists of 11 proteins with five core proteins that bind 5mC (Du et al., 2015). Additionally to the MBD domain, MBD proteins contain various other domains reflecting their sometimes contradictory roles in transcriptional regulation and coordination of cross-talk between DNA methylation, histone modifications and chromatin organization (Du et al., 2015).

DNA demethylation can occur passively through "dilution" by DNA replication or actively through oxidation or deamination of the methylated base and subsequent base excision repair (Bochtler et al., 2017). Ten-eleven translocation enzymes convert 5mC into 5-hydroxymethylcytosine (5hmC) and 5hmC into 5-formylcytosine and 5-carboxylcytosine, which in turn are recognized and replaced with unmodified cytosines in thymine-DNA-glycosylase-mediated base excision repair (Meier and Recillas-Targa, 2017).

While it was long believed that platyhelminthes were lacking DNA methylation, first evidence for functional DNA methylation was established in 2011 for *S. mansoni* (Geyer et al.). Since then 5mC DNA methylation has been described for species throughout the phylum, including *E. multilocularis* (Geyer et al., 2013). A computational approach has identified thousands of CpG islands in the genomes of *E. canadensis*, *E. multilocularis* and *E. granulosus* (Maldonado et al., 2017). Conserved DNMT2 and MBD2/3 candidates have been reported in all analyzed platyhelminth species Geyer et al. (2013). The *S. mansoni* DNMT2 has been shown to be responsible for 5mC DNA methylation in *S. mansoni* (Geyer et al., 2011). Furthermore, treatment with 5-azacytidine, a DNMT inhibitor, inhibits cytosine methylation and disrupts egg production and egg maturation in *S. mansoni* (Geyer et al., 2011). Besides DNA methylation, 5-azacytidine also disrupts transcription, translation, repetitive element maintenance and corresponding downstream processes, including stem cell activities, likely affecting schistosome oviposition in a pleiotropic manner (Geyer, Munshi, Vickers, Squance, Wilkinson, Berrar, Chaparro, Swain and Hoffmann, 2018). Depletion of the *S. mansoni* MBD2/3 with RNA interference also

reduces egg production and affects proliferating neoblasts (Geyer, Munshi, Whiteland, Fernandez-Fuentes, Phillips and Hoffmann, 2018). It is unclear, however, if the effects of *S. mansoni* MBD2/3 depletion are based on DNA methylation or other mechanisms. The schistosome MBD2/3 is localized in the nucleus and capable of 5mC binding, indicating a DNA methylation dependent function (Geyer, Munshi, Whiteland, Fernandez-Fuentes, Phillips and Hoffmann, 2018). In contrast, the *S. mediterranea* MBD2/3 protein likely controls stem cell pluripotency in a DNA methylation independent manner (Jabbar et al., 2011). A possible explanation for this discrepancy are amino acid substitutions critical for 5mC binding in the *S. mediterranea* MBD2/3 and proposed diverse roles of MBD2/3 functions in platyhelminthes (Geyer, Munshi, Whiteland, Fernandez-Fuentes, Phillips and Hoffmann, 2018).

3.7 Parvoviridae

Parvoviridae belong to the smallest animal viruses and also their genome is one of the smallest (Martynova et al., 2016). The family of Parvoviridae can be divided into two subfamilies: the Parvovirinae that infect vertebrates and the Densovirinae that infect invertebrates (Cotmore et al., 2014). The genome of densoviruses is between 4 and 6 kb long and encodes two sets of functionally different proteins: the nonstructural proteins (NS) and the viral particles (VP) that form the capsid (Cotmore et al., 2014). Most densoviruses have two non-structural proteins NS1 and NS2 with the ORF encoding NS1 completely containing the ORF encoding NS2 (Martynova et al., 2016). Ambisense densoviruses additionally have an ORF that codes for NS3 (Martynova et al., 2016). NS1 proteins have endonuclease, helicase and ATPase activities, enhance promoter activity and are required for replication of the virus genome (Han et al., 2013; Yang et al., 2008). Originating from the inverted terminal repeats the virus genome is replicated through a "rolling hairpin" mechanism (Cotmore and Tattersall, 2013). The replication generally takes place in mitotically active host cells (Tijssen et al., 2016). Parvoviruses and densoviruses have invaded the germline of various animals (Liu et al., 2011). The fact that parvoviruses and densoviruses integrate into the host genome makes them interesting tools for genetic manipulation. Indeed, they have been used as autonomous vectors for the genetic manipulation of vertebrate and invertebrate species (Afanasiev et al., 1999; Bossin et al., 2003; Dupont et al., 1994). To retain infectivity, densovirus clones require an intact ORF for NS1 and at least one structurally integer inverted terminal repeat (ITR) (Afanasiev et al., 1999).

3.8 Mobile genetic elements in cestodes

In many metazoans the piRNA pathway mediates silencing of mobile genetic elements (MGEs) in the germline (Ozata et al., 2019). While the proteins of the piRNA pathway are generally highly conserved, no true piwi and vasa orthologues have been found in the genomes of cestodes (Tsai et al., 2013; Skinner et al., 2014; Fontenla et al., 2017) suggesting that cestodes lack a canonical piRNA pathway. This raises the question which alternative mechanisms control MGEs in cestodes instead. Several MGEs have been described in cestode genomes, such as inactive copies of GYPSY class of LTR retrotransposons and Merlin DNA transposons (Tsai et al., 2013; Bae, 2016). Recently, terminal repeat retrotransposons in miniature (TRIM) have been reported to be massively expressed in germinative cells of *E. multilocularis* (Koziol et al., 2015). It is therefore expected that cestodes employ some kind of protective mechanism (Skinner et al., 2014). A first step toward the detection of such a mechanism would be the identification of active and silenced MGEs. To address this, densovirus sites in flatworm genomes were analyzed.

3.9 Telomeres and telomerase reverse transcriptase

Telomeres are repetitive DNA sequences and their associated proteins that cap and protect the end of linear eukaryotic chromosomes. The telomere sequence varies in different organisms: TTAGGG in vertebrates, TTGGGG in *Tetrahymena*, TTAGGG(T/C) in the apicomplexan parasite *Plasmodium*; TTTAGGG in *Arabidopsis thaliana*; TTAGGC in the parasitic roundworm *Ascaris lumbricoides* and TGTGGGTGTGGTG in *Saccharomyces cerevisiae* (Witzany, 2008). In platyhelminthes, including the trematode *S. mansoni* (Hirai and LoVerde, 1996), the planarian *Polycelis tenuis* (Joffee et al., 1996) and the cestodes *Nippotaenia mogurndae* (Bombarova et al., 2009), *E. multilocularis*, *E. granulosus*, *H. microstoma* and *T. solium* (Tsai et al., 2013), the telomere repeat sequence is TTAGGG, as in vertebrates.
Critically short telomere sequences are associated with cellular senescence and early mortality in humans (Parks et al., 2009). Various studies have shown associations of smoking, obesity, age and psychosocial stress, possible mediated through cortisone (Vasunilashorn and Cohen, 2014), with shorter telomeres (Blackburn, 2001). While the exact mechanism is unclear, it is assumed that oxidative stress and inflammation contribute to telomere shortening (Li et al., 2014). The length of the telomere repeats is species-specific. Human telomeres are 2-10 kbp long (Lansdorp, 2005), whereas for example *Tetrahymena thermophila* telomeres are only 300 bp long (Sandin and Rhodes, 2014). In proliferating cells, telomere maintenance is required

for continued viability. Incomplete genome replication by DNA-dependent DNA polymerases leads to shortening of the terminal repeats (end replication problem) (Collins and Mitchell, 2002). Telomere length is maintained by the ribonucleoprotein telomerase or independently of telomerase by alternative lengthening of telomeres (Gomes et al., 2010). The telomerase consists of two different subunits: a protein subunit, the so-called telomerase reverse transcriptase (TERT) and a RNA subunit, the telomerase RNA (TERC) (Brazvan et al., 2018). TERC subunits from different organisms are very divergent in length and structure. Regardless, the catalytically pseudoknot-template core domain and in vertebrates also the trans-activation domain are structurally conserved (Sandin and Rhodes, 2014). The TERT subunit contains four domains: the N-terminal extension domain that binds single-stranded telomeric DNA, the telomerase RNA binding domain that interacts with the trans-activation domain of TERC, the reverse transcriptase domain and the C-terminal domain (Sandin and Rhodes, 2014). In humans the telomerase is expressed in embryonic tissues and germline cells but not in normal somatic cells (Gomes et al., 2010). While adult human stem cells, such as hematopoietic stem cells, activate telomerase activity when dividing, the activity is not enough to prevent telomere shortening and senescence (Gomes et al., 2010). It is presumed that replicative aging functions as an anti-tumor mechanism in humans, as most tumor have up-regulated telomerase activity. This has lead to the development of telomerase inhibitors as drugs for cancer treatment (Agrawal et al., 2012). Other organisms, for example lobsters and marine demosponges, express telomerase in somatic and germinal tissues (Gomes et al., 2010). In sexual worms of *S. mediterranea tert* is expressed in ovaries and testes, while mature asexual worms show expression in the transparenchyma around the pharynx, in the adult somatic stem cells (Tan et al., 2012).

In planarians, the role of telomerase in asexually and sexually reproducing worms has been studied. Sexual worms of *S. mediterranea* show age-correlated decline in telomere length and seem to lengthen their telomeres only through the sexual reproduction process, whereas asexual worms have longer telomeres than sexual worms and maintain their telomere length (Tan et al., 2012). In both asexual and sexual worms telomerase activity is up-regulated during regeneration. However, regeneration decreases telomere length in sexual animals while it increases telomere length in asexual worms (Tan et al., 2012). Similarly, *Dugesia ryukyuensis* asexual worms maintain their telomere length, whereas sexual worms exhibit telomere shortening (Tasaka et al., 2013). Experimentally sexualized worms (originally asexual) exhibit the same ability to maintain telomere length as innate sexual planarians suggesting that telomere maintenance capabilities are determined by the mode of reproduction

at birth (Tasaka et al., 2013).

3.10 Serotonin

Serotonin is a widespread neurotransmitter and paracrine hormone that modulates a variety of biological processes (Berger et al., 2009; Berumen et al., 2012). As a neurotransmitter and neuromodulator serotonin influences *inter alia* sleep, appetite, pain sensation and mood regulation. Lack of serotonin or a serotonin precursor can cause depression which can be treated with so-called selective serotonin re-uptake inhibitors (Terry et al., 2008; Baganz and Blakely, 2013). Furthermore, serotonin is involved in regulation of blood pressure and remodeling processes after myocardial infarction (Berger et al., 2009), platelet aggregation McCloskey et al. (2008); Yuan et al. (2006), the immune system (Baganz and Blakely, 2013) and many other processes. As a morphogenic factor serotonin can cause cell differentiation, proliferation or apoptosis depending on the context (Seuwen and Pouyssegur, 1990; Fanburg and Lee, 1997; Whitaker-Azmitia et al., 1996; Renaud et al., 1983; Sarrouilhe et al., 2015).

Serotonin is synthesized from the amino acid tryptophan (Liu et al., 2008). First, the enzyme tryptophan hydroxylase catalyzes the hydroxylation of tryptophan to hydroxytryptophan in a rate limiting step (Cote et al., 2003), then hydroxytryptophan is decarboxylated to serotonin (5-Hydroxytryptophan) by the amino acid decarboxylase. Degradation of serotonin in vertebrates is mostly catalyzed by the monoamine oxidase (Culpepper, 2012). Serotonin binds to a wide range of receptors, most of them G-protein coupled receptors (GPCRs). Seven main classes of receptors with different structures and binding mechanisms have been described (Hoyer et al., 1994; Berumen et al., 2012). Binding of serotonin to the 5-HT_3 receptor, a Na^+/K^+ ion channel, leads to neuronal excitation and neurotransmitter release (Berumen et al., 2012). All other classes of serotonin receptors are GPCRs of the *Rhodopsin* family and activate different G proteins (Fredriksson et al., 2003; Berumen et al., 2012) thereby fulfilling a variety of functions. Serotonin uptake into cells is mediated by the serotonin transporter, a Na^+/Cl^--dependent neurotransmitter transporter. The serotonin transporter is a member of the solute carrier 6 family (Andersen et al., 2009) and consists of 12 transmembrane domains. 10 of the transmembrane domains form the protein core with an outer and an inner ring where the latter binds the substrate (Kristensen et al., 2011). In humans the serotonin transporter is responsible for the uptake of serotonin from the blood plasma into platelets and the removal of serotonin from the synaptic cleft, thereby terminating

the transmitter action (Lesurtel et al., 2008). In the treatment of depression, the serotonin receptor represents a drug target for tricyclic antidepressants or serotonin selective re-uptake inhibitors (SSRI) (Horschitz et al., 2001; Neubauer et al., 2006).

In flatworms serotonin is an important neurotransmitter. The serotonergic nervous system has been characterized for several cestodes and trematodes (Gustafsson et al., 1985; Fairweather et al., 1987; Gustafsson, 1987; Fairweather et al., 1988; McKay et al., 1991; Gustafsson et al., 1995; Terenina et al., 2009), including *E. multilocularis* (Koziol et al., 2013) and *E. granulosus* (Brownlee et al., 1994; Fairweather et al., 1994; Camicia et al., 2013). In *S. mansoni* serotonin causes contraction of muscle fibers (Day et al., 1994) and increases motility of sporocysts (Boyle et al., 2000). Serotonin also increases motility of *E. granulosus* protoscoleces (Camicia et al., 2013) and *Mesocestoides corti* tetrathyridia (Camicia et al., 2018). Besides motility, serotonin influences metabolic processes, such as glucose uptake, glycolysis and excretion of lactic acid in *S. mansoni* (Estey and Mansour, 1987; Boyle and Yoshino, 2005; de Saram et al., 2013). The effect of serotonin on glucose uptake and lactate excretion appears to depend on serotonin uptake as the serotonin transporter inhibitor fluoxetine abolishes it (Harder, Abbink, Andrews and Thomas, 1987; Harder, Andrews and Thomas, 1987). Furthermore, serotonin regulates developmental processes, such as traumatic regeneration of the planarian *Polycelis tenuis* (Franquinet, 1979) and miracidial transformation in *S. mansoni* (Kawamoto et al., 1989; Taft et al., 2010).

In previous studies, I have shown that endogenously supplied serotonin induces redifferentiation of *E. granulosus* protoscoleces towards the metacestode stage (Herz, 2015; Camicia et al., 2013) and stimulates vesicle formation of *E. multilocularis* primary cell cultures (Herz, 2015). In contrast, the serotonin transporter inhibitor citalopram reduces survival of *E. granulosus* protoscoleces (Herz, 2015; Camicia et al., 2013), prevents vesicle formation of *E. multilocularis* primary cell cultures and affects structural integrity of *E. multilocularis* metacestodes (Herz, 2015). In this work, the role of serotonin in developmental processes is further studied.

3.11 Objectives

Germinative cells are the only proliferating cells in *E. multilocularis* and give rise to all other cell types (Koziol et al., 2014). As the only proliferating cells, germinative cells are also responsible for the remarkable transitions of the parasite during its life cycle. Genomic information (Tsai et al., 2013) together with detailed genome wide expression profiling can lead to a better understanding of the parasite and its developmental processes.

The general **objective** of this work is the characterization of the *E. multilocularis* transcriptome.

The specific **questions** are

- Which genes are expressed in different *E. multilocularis* life cycle stages?
- Does *in vitro* cultivation of metacestodes change gene expression?
- Which genes are specifically expressed in germinative cells?
- Do transcriptome data correlate with proteome data?
- Are there natural antisense transcripts in *E. multilocularis*?
- Does DNA methylation play a role in *E. multilocularis* development?
- Are densoviral genes specifically expressed in germinative cells?
- Are densoviruses still active in *E. multilocularis*?
- Could densoviruses be used for genetic manipulation?
- Is the telomerase reverse transcriptase gene expressed in germinative cells?
- Is the telomerase reverse transcriptase required for proliferation?
- Where are the serotonin transporter and tryptophan hydroxylase genes expressed?
- Do the serotonin transporter and the tryptophan hydroxylase play a role in developmental processes?
- Does serotonin stimulate proliferation in metacestodes?

4 Materials and Methods

4.1 Equipment

- Autoclave: Selectomat S2000 (Münchner Medizin Mechanik (MMM),München)
- ChemiDoc^{TM} MP System (Bio-Rad, München)
- Centrifuges: Bench-top Centrifuge Mikro 200 (Hettich, Tuttlingen), Refrigerated Centrifuge 3K30 (Sigma, München)
- Confocal microscope:Leica TCS SP5, Leica Microsystems
- Gelelectrophoresis camber (Bio-Rad, München)
- Heating block: DB-3 (Techne, Cambridge, UK), Heizblock (Liebisch, Bielefeld)
- Heating stirrer: Typ RCT (Jahnke & Kunkel, Staufen i. Br.)
- Incubator: Heraeus (Thermo Electron, Langenselbold)
- Laminar flow hood: BSB 6A (Gelaire Flow Laboratories, Meckenheim), HERA safe Heraeus (Thermo Electron, Langenselbold)
- Microscope: Leica IRB (Leica Microsystems, Wetzlar)
- NanoDrop 1000 (PeqLab Biotechnologie, Erlangen)
- Neubauer counting chamber: Neubauer Precicolor, depth 0.1 mm, 0.0025 mm^2
- Pipettes: 0,5-10 µl, 10-100 µl, 1-1000 µl (Eppendorf, Hamburg)
- Scales: 10-1000 g (Sartorius, Göttingen), R160 (accuracy weighing scale) (Sartorius, Göttingen)
- Shaking incubator: TH30 (Edmund BÃijhler, Hartensein, Würzburg), G24 (New Brunswick Scientic, Edison, N.J., USA)
- Spectrophotometer U-2000 (Hitachi, NY, USA)
- StepOnePlus Real-Time PCR-Systems (Thermo Scientific, Schwerte, Germany)
- TECAN ELISA reader (Tecan Group, Crailsheim)
- Thermocycler: Trio-Thermoblock^{TM} heated lid, Trio-Thermoblock^{TM} oil, T-Gradient (Biometra, Göttingen)
- Vacuum pump (ILMVAC)
- Voltage generator: Power Pack P24 and P25 (Biometra, Göttingen)
- Vortex mixer: L46 (Gesellschaft für Laborbedarf, Würzburg)

4.2 Consumables

- 12-well plates (Sarstedt, Nuembrecht)
- 24-well plates (Sarstedt, Nuembrecht)
- 96-well plates (Sarstedt, Nuembrecht)

- Cell culture flasks: 25 cm^2, 75 cm^2, 175 cm^2 (NUNC, Wiesbaden; Sarstedt, Nümbrecht)
- Nitrocellulose membrane (GE Healthcare, München)
- Safe-lock tubes 0,5, 1,5 and 2 ml (Eppendorf, Hamburg)
- Semi-micro cuvettes (Sarstedt, Nümbrecht)
- Sterile filter (Nalgene, New York, USA)
- Sterile tubes, 15 and 50 ml (Greiner, Nürtingen)
- Syringes and canula, sterile (Braun Melsungen AG, Melsungen)

4.3 Chemicals and reagents, enzymes, media, kits

Chemicals and reagents

- 4-Chloro-DL-phenylalanine (Sigma-Aldrich, München)
- Agarose (ROTH, Karlsruhe)
- Albumin fraction V (pH 7) Blotting grade (BSA) (AppliChem, Darmstadt)
- Anti-digoxigenin, AP conjucated (Roche, Basel)
- Anti-digoxigenin, POD conjucated (Roche, Basel)
- Aqua demin. (VE-water)
- Ampuwa (Fresenius, Bad Homburg)
- Ampicillin (Sigma-Aldrich, München)
- 5-Aza-2'-deoxycytidine (Sigma-Aldrich, München)
- Bathocuproine disulfonic acid (Sigma-Aldrich, München)
- β-Mercaptoethanol (Sigma-Aldrich, München)
- Bi2536 (Selleckchem,München)
- BIBR 1532 (Selleckchem,München)
- Blocking reagent (Roche, Basel)
- Bovine serum albumin (Sigma-Aldrich, München)
- CHAPS (Sigma-Aldrich, München)
- Chloroform (Merck, Darmstadt)
- Citalopram hydrobromide (Sigma-Aldrich, München)
- Denhardt´s solution (Sigma-Aldrich, München)
- Diethyl pyrocarbonate (DEPC) (Applichem, Darmstadt)
- Dimethyl formamide (Sigma-Aldrich, München)
- DIG RNA labeling mix 10x (Roche, Basel)
- Dimethyl sulfoxide (DMSO) (Sigma-Aldrich, München)
- dNTP lyophilised (ROTH, Heidelberg)
- 5-Ethynyl-2'-deoxyuridine (EdU, Life Technologies, Darmstadt)
- Fetal Calf Serum (FCS) (Invitrogen, Darmstadt)

- Fluoprep (Biomérieux, Nurtingen)
- Formamide (Sigma-Aldrich, München)
- Gel Loading Dye Blue 6x (New England Biolabs, Schwalbach)
- Heparin (Sigma-Aldrich, München)
- HOT FIREPol®EvaGreen® qPCR Mix (ROX)(Solis Biodyne, Düsseldorf)
- Hydroxyurea (Sigma-Aldrich, München)
- L-Cystein (Sigma-Aldrich, München)
- Midori Green Direct (NIPPON Genetics Europoe, Düren)
- Midori Green Advance (NIPPON Genetics Europoe, Düren)
- NHS-Fluorescein (Sigma-Aldrich, München)
- nuclease-free water (Qiagen, Hilden)
- Oligonucleotides (Sigma-Aldrich, München)
- Paraformaldehyde (Sigma-Aldrich, München)
- Paroxetine maleate salt (Sigma-Aldrich, München)
- Penicillin/Streptomycin (Invitrogen, Darmstadt)
- Resazurin (Sigma-Aldrich, München)
- Serotonin hydrochloride (Sigma-Aldrich, München)
- RNase-Exitus Plus (Applichem, Darmstadt)
- Sheep serum (Sigma-Aldrich, München)
- SmartLadder (Eurogentec, Köln)
- Sodium carbonate (ROTH, Karlsruhe)
- Sodium taurocholate (Sigma-Aldrich, München)
- Tavanic® (Tava, active component levofloxacin, 5mg/ml) (Aventis)
- Torula RNA (Sigma-Aldrich, München)
- Triethanolamine (Sigma-Aldrich, München)
- Triethylamine (Sigma-Aldrich, München)
- Triton® X-100 (Sigma-Aldrich, München)
- Trizol® Reagent (Invitrogen, Darmstadt)
- Tween® 20 (Sigma-Aldrich, München)
- Tyramine (Sigma-Aldrich, München)

Enzymes

- Pepsin (Sigma-Aldrich, München)
- Proteinase K (Fermentas, St. Leon-Rot)
- RQ1 RNase-Free DNase (Promega, Mannheim)
- SP6 polymerase (New England Biolabs, Frankfurt/Main)
- T7 polymerase (New England Biolabs, Frankfurt/Main)

- Taq-Polymerase (New England Biolabs, Schwalbach)
- Trypsin/EDTA solution (0,05%/0,02% (w/v) in PBS w/o Ca^{2+}, Mg^{2+}) (Biochrom, Berlin)

Media

- c-DMEM-A (prepared according to Spiliotis and Brehm (2009); see section 10.10.1)
- Dulbecco's Modified Eagle Medium, high glucose (4.5 g/l) GlutaMAX (Life Technologies 10566-016)

Kits

- Direct-zol™ RNA MiniPrep (Zymo Research,Freiburg)
- cDNA Synthesis: Omniscript® RT Kit (QIAGEN, Hilden), SuperScript®III Reverse Transcriptase (Invitrogen, Darmstadt), SuperScript®IV Reverse Transcriptase (Invitrogen, Darmstadt)
- DNA Purification Kits: Easy Pure DNA Purification Kit (Biozym, Hessisch Oldendorf), NucleoSpin® Extract II (Macherey-Nagel, DÃijren)
- Plasmid Isolation Kit: NucleoSpin® Plasmid (Macherey-Nagel)
- PCR Cloning Kits: PCR Cloning Kits (QIAGEN, Hilden), CloneJET™ PCR Cloning Kit (Fermentas, St. Leon-Rot)
- PCR amplification: Phusion© High-Fidelity PCR Master Mix with HF buffer (New England Biolabs, Frankfurt/Main), KOD Hot Start DNA Polymerase (Merck, Darmstadt)
- EdU Detection Kit: Click-iT® EdU Alexa Fluor® 555 Imaging Kit (Life Technologies, Darmstadt)

4.4 Oligonucleotids

***E. multilocularis sert* specific primers**

SERT_A_dw	5'-GAA TGC TGT AGA TGT GGT TAT GG-3'
SERT_1_dw	5'-GAT GCC GTT GTG GTG GAG AC-3'
SERT_2_dw	5'-GCA ATC AAC TGT GGG ACC AG-3'
SERT_2_up	5'-CTG GTC CCA CAG TTG ATT GC-3'
SERT_Z_up	5'-GAT TGG TGC AAT GGG GAG-3'

***E. multilocularis tph* specific primers**

TPH_1_dw	5'-GAC GCT GGT GAT GTC GTA ATT C-3'
TPH_2_dw	5'-GAG TTG GGT ATC GCC TCT CTG-3'
TPH_2_up	5'-CAG AGA GGC GAT ACC CAA CTC-3'
TPH_3_up	5'-GGA TCC GAG GGC TTG ACG-3'
TPH_Z_up	5'- GAG GTT AAA TGA TGC GGT GC-3'

***E. multilocularis tert* specific primers**

TRT_1_dw	5'-GAC TTC TGA GGC CAA CGA GAG-3'
TRT_2_up	5'-CAC AAC AGA CCG AGA GAC TCC-3'

***E. multilocularis dnmt2* specific primers**

DNMT-dw	5'-GCG CTC TTT CCA TAC CTG ATA G-3'
DNMT-up	5'-CTA ACC AAG AAA AGA TTA CAC TGT CG-3'

***Emu*DNV-NS1 specific primers**

NS1a-fw	5'-GGC GTT CCA CTA CAA G-3'
NS1-rev	5'-GCC AAC AAT TCA TAA ATG G-3'

Primers for sequencing of *Emu*DNV-NS1 genome sites

g_EmuJ_000013900	5'-GGA AAC CTC CTC CGA CA-3'
DV_EmuJ_000013900	5'-GAT AGT CTG CCA TTA GGC G-3'
g_EmuJ_002195700	5'-GAT AGT TTG TTC CAC CAT TGA-3'
DV_EmuJ_002195700	5'-GCT TAT TCA TTC TGC GGT TTT-3'
g_EmuJ_000388600	5'-GGT GCT TTT TCA TAT TCT CGT-3'
DV_EmuJ_000388600	5'-GAT TTC ATT GGC TGA AAA CAT-3'
g2_EmuJ_000329200	5'-GGC TCA ACA ACC GAC GTA AT-3'
DV2_EmuJ_000329200	5'-GGC TCG AGG AAG GTA GTT GTC-3'

***E. multilocularis elp* specific primers**

Em10 15	5'-AAT AAG GTC AGG GTG ACT AC-3'
Em10 16	5'-TTG CTG GTA ATC AGT CGA TC-3'

cDNA synthesis primer

cd3rt	5'-ATC TCT TGA AAG GAT CCT GCA GGA CT_{23}VX-3'
RandomOct	5'-NNN NNN NN-3'

pDrive specific primers

SP6	5'-CCA TTT AGG TGA CAC TAT AGA ATA C-3'
T7	5'-GCT CTA ATA CGA CTC ACT ATA GG-3'

CloneJET1.2 primers

pJET1.2 forward sequencing primer	5'-CGA CTC ACT ATA GGG AGA GCG GC-3'
pJET1.2 reverse sequencing primer	5'-AAG AAC ATC GAT TTT CCA TGG CAG-3'

Primers for WMISH probe synthesis from CloneJET1.2

T7 Plus 2	5'-AGA AGA GTA ATA CGA CTC ACT ATA GG-3'
5-SP6+pJET1.2.Rev-3	5'-ATA ATT TAG GTG ACA CTA TAG AAC ATC GAT ' TTT CCA TGG CAG-3'

Primer for RACE PCR of *E. multilocularis tph*

IG 4-5-SPR2	5'-CTT ATG ATG TGC CAG ATT ATG-3'

qPCR primers

primer name	concentration in reaction	annealing temperature	sequence
F_000388600 (*Emu*DNV-NS1)	200 nM	60°C	5'-CAA CCA GCA GGA TCT CAA GCA-3'
R_000388600 (*Emu*DNV-NS1)	200 nM	60°C	5'-CAT CTA CCC TCT ATG GCG GCT-3'
TERT-qPCR4-dw	300 nM	60°C	5'-GAA GGA CTG TTT GTT CCG A-3'
TERT-qPCR4-up	300 nM	60°C	5'-GAG GTG ACC GAA ATA CAA G-3'
qPCR-DNMT-up	200 nM	62°C	5'-ACC AGT GTT GAC GAT CTG-3'
qPCR-DNMT-dw	200 nM	62°C	5'-GAT TCT CAG TTG ATG CTA TG-3'
Elp F	300 nM	60°C	5'-TGA TGA AAG TGA AGC CAA GGA ACT TGA G-3'
Elp R	300 nM	60°C	5'-TTC GTC TGG AGC GTC TCA TTC TTA GAG-5'

4.5 Maintenance and cultivation of *E. multilocularis* parasite material

Parasite material was maintained through serial peritoneal passages in Mongolian jirds (*Meriones unguiculatus*) (Spiliotis and Brehm, 2009). All animal experiments were carried out in accordance with German and European regulations on the protection of animals (*Tierschutzgesetz*) with ethical approval from the local ethics committee of the Government of Lower Franconia (permit no. 55.2-2531.01/61/13). Metacestode vesicles were obtained by co-cultivation of parasite material with rat Reuber hepatoma cells (RH^- cells) (Spiliotis et al., 2004).

4.5.1 Isolation and activation of protoscoleces

Parasite material was isolated as described previously (Spiliotis and Brehm, 2009). Protoscoleces were obtained as described in (Brehm, Kronthaler, Jura and Frosch, 2000) with slight modifications: PBS was added to washed parasite material to a volume of 25 ml. To free protoscoleces, material was shaken vigorously for 10 min. Suspension was filtered over a 150 µm gauze filter. Then the flow-through was filtered over a 30 µm gauze filter. Retained protoscoleces were collected from the surface of the filter by pipetting and resuspending in PBS and were transferred to a petri dish. To remove calcium bodies, petri dish was rotated slowly. The protoscoleces concentrated in the middle of the dish and were transferred to a 50 ml Falcon tube. In order to activate protoscoleces, the gastrointestinal passage was mimicked. Protoscoleces were incubated in 30 ml DMEM with 0,05 % pepsin (pH 2, w/o FBS, sterile filtrated) for 30 min at 37°C and 125 rpm. After washing with PBS three times, protoscoleces were incubated in 30 ml DMEM with 0,2 % sodium taurocholate (pH 7,4; w/o FBS, sterile filtrated) for 3 h at 37°C and 125 rpm. Protoscoleces were washed again with PBS.

4.5.2 Isolation of primary cells from metacestode vesicles

Primary cells were isolated from metacestode vesicles after at least three month of co-cultivation. Co-cultivated metacestode vesicles were transferred to axenic cultivation for three days before primary cell isolation (modified from Spiliotis and Brehm (2009); protocol see section 10.10). The amount of primary cells was calculated using a units system. OD_{600} = 0,02 of 1 ml diluted cell suspension was considered to be 1 Unit.

4.5.3 Inhibitors used in *in vitro* experiments

Substances and stock solutions were prepared as following:

Substance	Solvent	Stock solution
4-Chloro-DL-phenylalanine	H_2O	4 mM
5-Aza-2'-deoxycytidine	DMSO	50 mM
Bi2536	DMSO	10 mM
BIBR1532	DMSO	100 mM
Hydroxyurea	DMEM	2 M
Paroxetine	ethanol	50 mM

All solutions were sterile filtrated using a 0,02 µm filter prior to use in cell culture.

4.5.4 Drug experiments on primary cells

For drug experiments, primary cells were cultivated in 24-well-plates (100 Units per well) with 1 ml c-DMEM-A medium (see section 10.10.1) without reducing agents in a nitrogen atmosphere (Spiliotis and Brehm, 2009). 500 µl medium was changed three times a week and the drugs substituted accordingly. Formation of aggregates and vesicles was observed with the optical microscope Leica IRB using bright-field and phase-contrast settings. The number of vesicles was counted at several time points as an indicator for proliferation. All experiments were performed with three technical replicates using the solvent of the drugs for controls. Three independent experiments were performed for the BIBR1532 inhibitor studies and statistical differences of the number of developed vesicles were analyzed with GraphPadPrism7 for Windows (GraphPad Software, La Jolla California, USA) using an two-way-ANOVA (repeated measurements for the time component) with a Dunnett's multiple comparisons test (multiplicity adjusted P values) for each time point comparing all concentrations with the control.

4.5.5 Resazurin assay with primary cells

Primary cells were seeded in a 96-well-plate with 10 Units/well. 100 µl c-DMEM-A (see section 10.10.1) supplemented with the respective drug was added to each well. As positive control for cell death 1 % triton was used. A negative control with the drug solvent and a medium control without the cells was included. The plate was incubated at 37°C in a nitrogen atmosphere for two days. The resazurin stock solution

(2 mg/ml in H_2O, stored at 4°C) was diluted 1:100 in PBS and 100 µl of the dilution was added to each well. Then the plate was incubated for 3 h at 37°C in a nitrogen atmosphere. Fluorescence was measured at 540 nm (reference 595 nm) using TECAN ELISA reader. Average medium florescence was subtracted from fluorescence values. With the exception of the resazurin assay for 5-aza-2'-deoxycytidine (only one experiment there), three independent experiments with each three technical replicates were performed. Statistical differences were analyzed with GraphPadPrism7 using an one-way-ANOVA with a a Dunnett's multiple comparisons test to compare all samples with the negative control.

4.5.6 Drug experiments on metacestode vesicles

Metacestode vesicles co-cultivated for at least three month were transferred to axenic cultivation for three days prior to use in experiments (see section 10.10.2). 4-7 vesicles with a size of 3-4 mm for each sample were cultivated in a 12-well-plate with 2 ml c-DMEM-A with reducing agents (see section 10.10.1) in a nitrogen atmosphere (Spiliotis and Brehm, 2009). Medium and the respective drugs were completely exchanged three times a week. Structural integrity of metacestode vesicles was observed with the optical microscope Leica IRB using bright-field settings. Completely collapsed vesicles were considered to have lost their structural integrity. Experiments were performed with three biological replicates using the solvent of the drugs for controls. For the BIBR1532 inhibitor studies three different isolates with each three biological replicates were used. Statistical differences of the percentages of collapsed vesicles were analyzed with GraphPadPrism7 using an two-way-ANOVA (repeated measurements for the time component) with a Dunnett's multiple comparisons test (multiplicity adjusted P values) for each time point comparing all concentrations with the control.

4.5.7 Depletion of germinative cells with hydroxyurea or Bi2536

Hydroxyurea and the polo-like kinase inhibitor Bi2536 are known to deplete germinative cells effectively and specifically (Koziol et al., 2014; Schubert et al., 2014). Therefore hydroxyurea and BI2536 were used for germinative cell depletion of metacestodes. Approximately 30 previously axenized metacestode vesicles with a diameter of 2-3 mm were cultivated in T-25 25 cm^2 cell culture flasks (vertically positioned) with 7 ml c-DMEM-A with reducing agents (see section 10.10.1) in a nitrogen atmosphere (Spiliotis and Brehm, 2009). Treatment was either performed with 40 mM hydroxyurea (Sigma Aldrich, München, Germany) for 7 days or with 150 nM Bi2536 (Axon Medchem, Groningen, The Netherlands) for 21 days. Due to instability of

hydroxyurea at 37°C, it was added every day. In both treatments, the medium was completely exchanged every second day. After treatment, metacestode vesicles were washed extensively with PBS before use in WMISH or RNA-isolation.

4.5.8 EdU detection of treated metacestode vesicles

After a defined time of drug exposure, vesicles were incubated in DMEM (10 % FCS) with 50 µM 5-ethynyl-2'deoxyuridine (EdU) for 5 hours. Afterwards vesicles were fixed in 4 % PFA as described for WMISH. Fluorescence detection of EdU was performed with the Click-iT® EdU Alexa Fluor® 555 Imaging Kit according to manufacturer's instructions. Samples were analyzed by confocal microscopy (Leica TCS SP5, Leica Microsystems). For quantification of EdU positive cells, random fields were photographed (2 pictures per vesicle for the serotonin experiment, one picture per vesicle for the BIBR1532 experiment, both with 4 vesicles per replicate and three biological replicates per concentration). Fiji (Schindelin et al., 2012, 2015) was used to analyze images. Total number of cells was determined by a custom script identifying the cell nuclei in the Dapi staining (see section 10.9). EdU-positive cells (from the previously identified cells) were counted manually. Statistical differences of percentages of EdU-positive cells were analyzed with GraphPadPrism7 using an ordinary one-way-ANOVA with a Dunnett's multiple comparison test comparing all concentrations with the control.

4.6 Working with nucleic acids

4.6.1 Precautions for working with RNA

Working bench and pipettes were cleaned with RNase Exitus Plus, then rinsed with 70 % ethanol. Plastic material and chemicals were used exclusively for working with RNA. Solutions were treated with diethyl pyrocarbonate (DEPC) when possible or prepared with DEPC-treated water. For enzymatic reactions nuclease-free double distilled water (Qiagen) was used.

4.6.2 RNA isolation

Axenized vesicles or primary cell were washed with PBS. Vesicles were opened with a tip. Vesicles or primary cells were transferred to a 1,5 ml tube and centrifuged at 500 g for 1 min. The PBS was removed and primary cells were resuspended in 500 µl Trizol® Reagent, vesicles in 1 ml, vortexed briefly and incubated at room temperature for 5 min. RNA extraction was performed using the Direct-zol™ RNA

MiniPrep according manufacturer's instructions (including DNAse treatment). Integrity of RNA was evaluated by agarose gel electrophoresis (no degradation of ribosomal RNA).

4.6.3 DNA isolation

Axenized metacestode vesicles were washed with PBS and opened with a tip. After washing, parasite material was centrifuged for 10 min at 5000 g. The supernatant was removed and the pellet re-suspended in lysis buffer (100 mM NaCl; 10 mM Tris-HCL (pH 8,0); 50 mM EDTA (pH 8,0); 0,5 % SDS, 20 µ µ/ml RNase A; 0,1 mg/ml Proteinase K; 1,2 ml/100 mg pellet). Overnight incubation was carried out at 50°C, followed by a standard phenol-chloroform extraction and ethanol precipitation.

4.6.4 Determination of RNA and DNA concentrations

RNA and DNA concentrations were determined spectrophotometrically at a wavelength of 260 nm using Nano Drop 1000. Purity of nucleic acids was analyzed on the basis of the ratios of absorbance: 260 nm/280 nm for protein impurity (1,8 - 2,0 indicating pure nucleic acids), 260 nm/ 230 nm for salt impurity (above 2,0 indicating pure nucleic acids).

4.6.5 Reverse transcription

Reverse transcription was performed using Omniscript® RT Kit, SuperScript®III Reverse Transcriptase or SuperScript®IV Reverse Transcriptase according manufacturers' instructions with Oligo-dT primer cd3rt (5'-ATC TCT TGA AAG GAT CCT GCA GGA CT_{23}VX-3') or a combination of cd3rt and a random octamer primer. When required for later quantitative real-time PCR, a RT^- - control (same reaction mixture without reverse transcriptase) was included. PCR on the reference gene *elp* (EmuJ_000485800) was performed to determine if reverse transcription had been successful and sample free of genomic DNA (primers Em10 15 5'-AAT AAG GTC AGG GTG ACT AC-3' and Em10 16 5'-TTG CTG GTA ATC AGT CGA TC-3'). A PCR product of 400 bp was expected from cDNA, while a 500 bp product would indicate the presence of genomic DNA.

4.6.6 Quantitative real-time PCR (qPCR)

qPCR was performed with StepOnePlus Real-Time PCR-Systems. The reaction mixture contained 200 nM or 300 nM of each forward and reverse primer, 1x HOT FIREPol®EvaGreen® qPCR Mix (ROX) and distilled water up to 9 µl. 1µl of 1:5

diluted cDNA (or RT$^-$ -control) was used per reaction. The following program was used: 15 min at 95°C, 40 cycles of: 15 s at 95°C, 20 s at 58°C or 60°C, 20 s at 72°C. Fluorescence was measured at 72°C. Primer sequences, concentrations and annealing temperatures are shown in section 4.4. Amplification product specificity was assessed by melting curve analysis and gel electrophoresis. RT$^-$- controls were included for all samples that did not span an intron. *E. multilocularis elp* (EmuJ_000485800) was used as reference gene. Experiments were performed in three technical and three biological replicates. Efficiency (E) of amplification was calculated using linREG (Ramakers et al., 2003; Ruijter et al., 2009). Comparison of relative gene expression (R) was based on the analysis of crossing points (CP) by the relative RT-PCR method using the formula from Pfaffl (2001):

$$R = \frac{(E_{\text{target}})^{\Delta CP_{\text{target}}(\text{control-sample})}}{(E_{\text{reference}})^{\Delta CP_{\text{reference}}(\text{control-sample})}}$$

Statistical differences were analyzed with a permutation test using fgStatistics (Rienzo, 2012), method RT-PCR Comparison with technical replicates sampled at random and 5000 resampling cycles.

4.6.7 PCR

The reaction mixture for PCR with Taq polymerase contained: 5 µl tenfold Taq buffer, 0,5 µl of each primer (100 µM stock) (see section 4.4), 0,5 µl dNTPs (10 mM), 1-3 µl template, 0,5 µl Taq polymerase, H_2O ad 50 µl. The following program was run: denaturation at 95°C for 5 min; 35 cycles with 30 s of denaturation at 95°C, 30 s at annealing temperature of the primers and elongation at 68°C for 1-3 min depending on the length of the product; final elongation at 68°C for 10 min, hold at 4°C. Templates used were either cDNA or cDNA libraries prepared previously (Hubert et al., 2004). For colony screens PCR was performed with 20 µl total volume, 2 µl colony water as template and 30 instead of 35 cycles. PCRs with KOD or Phusion were performed according to manufacturers' instructions.

4.6.8 PCR for *Emu*DNV-NS1

Primers were designed based on the sequence of the gene copies EmuJ_000034800, EmuJ_000388600, EmuJ_002195700 and EmuJ_000329200 (see section 4.4). PCR was performed with Taq-Polymerase using cDNA of 2-day old primary cells with gDNA and RT-neg as negative controls (same protocol as 4.6.7). *Emu*DNV-NS1 genome sites were sequenced using the primer combinations listed in section 4.4 with gDNA as template (same protocol as in section 4.6.7).

4.6.9 Gel electrophoresis of PCR products or RNA

1 % or 2 % agarose gels and Tris-acetate-EDTA (TAE) buffer were used for gel electrophoresis. In case of double or multiple bands on the gel, gel extraction was performed before purification. For non-denaturing RNA electrophoresis, the chamber was treated with 3 % H_2O_2 for 20 min, then rinsed with RNase-free water before use. Agarose gels and TAE buffer were freshly prepared with RNase-free water. Loading Dye Blue was used for loading and Smart Ladder as ladder. DNA/RNA was visualized with Midori Green Direct or Midori Green Advance using ChemiDoc^{TM} MP System.

4.6.10 Purification and cloning of PCR products

PCR products were purified using NucleoSpin® Extract II or Easy Pure DNA Purification Kit and cloned into pJet1.2 using CloneJET^{TM} PCR Cloning Kit according to manufacturers' protocol. PCR amplification products of *Emu*DNV-NS1 genome sites were ligated into pDrive Cloning Vector using the QIAGEN® PCR Cloning Kit according to manufacturers' instructions.

4.6.11 Chemically competent *Escherichia coli* TOP10

E. coli TOP10 were incubated in 1 ml SOB medium at 37°C and 200 rpm over night. This over night culture was transferred to 250 ml fresh SOB medium and left to grow at 37°C and 170 rpm. When an OD_{600} of 0,5 was reached, culture was centrifuged at 5000 rpm at 4°C for 10 min. Supernatant was discarded and pellet resuspended in 80 ml ice cold DDMB80 buffer (10 mM KOAc, ph 7,0; 80 mM $CaCl_2.2H_2O$; 20 mM $MnCl_2.4H_2O$; 10 mM $MgCl_2.6H_2O$; 10 % glycerin; pH adjusted to 6.4; sterile filtrated; stored at 4°C). After incubation for 20 min on ice, suspension was centrifuged at 5000 rpm at 4°C for 10 min. Supernatant was discarded and pellet resuspended in 8 ml ice cold CCMB80 buffer. After incubation on ice for 20 min, cells were aliquoted into 0,5 ml tubes with each 50 µl cell suspension, frozen on dry ice and stored at −80°C.

4.6.12 Transformation of chemically competent *E. coli* with pJET1.2 Cloning Vector

Competent cells were thawn on ice. 2,5 µl ligation mixture was added to 50 µl *E. coli* and incubated on ice for 15 min. Cells were heat-shocked at 42°C for 45 sec and incubated for 2 min on ice. 200 µl SOC-medium was added to cells and the mixture was incubated at 37°C in a shaker for 45 min. Transformed *E. coli* were plated on

LB-ampicillin plates and incubated at 37°C overnight. For *E. coli* transformed with pDrive, 40 µl x-Galactose (40 mg/ml) was applied to plates before plating.

4.6.13 Colony PCR and overnight culture of transformed *E. coli*

Clones were picked from plates, only white clones in case of pDrive transformed *E. coli*, and solved in 30 µl H_2O. Colony water was used as template for colony PCR with pJET1.2 Forward Sequencing Primer (5'-CGA CTC ACT ATA GGG AGA GCG GC-3') and pJET1.2 Reverse Sequencing Primer (5'-AAG AAC ATC GAT TTT CCA TGG CAG-3') or the pDrive primers SP6 (5'-CCA TTT AGG TGA CAC TAT AGA AT-3') and T7 (5'-GCT CTA ATA CGA CTC ACT ATA GG-3') to check the length of the insert (see section 4.6.7). Clones that gave rise to a PCR product of the expected size were used for overnight cultures. 3-4 µl colony water were added to 6 ml LB-medium with ampicillin (1 µl/ml) and incubated in a shaker at 37°C over night. Plasmid isolation was performed using NucleoSpin® Plasmid according manufacturer's instructions. DNA amount was measured at the NanoDrop.

4.6.14 Sequencing of plasmid inserts

Sequencing was carried out at GATC (Konstanz, Germany). The sequencing mixture (final volume 10 µl) contained 400-500 ng plasmid with 2,5 µM pJET1.2 Forward or Reverse Sequencing Primer (see section 4.4).

4.7 Whole-mount *in situ* hybridization (WMISH)

WMISH of *E. multilocularis* protoscoleces and metacestodes was basically performed as described previously (Koziol et al., 2014)(detailed protocol: Koziol (2014)). Digoxigenin-labeled probes were synthesized by PCR-amplification and *in vitro* transcription. cDNA fragments of the target genes were amplified using the primers listed in section 10.8 and cloned into pJET1.2 using CloneJET™ PCR Cloning Kit. Promoter sequences were added by PCR with the primers listed in section 4.4. Sense and anti-sense digoxigenin-labeled probes were synthesized by *in vitro* transcription with SP6 and T7 polymerases according to manufacturer's instructions (overview of probes in 10.8). Probes were quantified as described in (Koziol, 2014). For co-localization studies, EdU-incorporation (described in section 4.5.8) was performed before WMISH and EdU-detection (described in section 4.5.8) after WMISH. Samples were analyzed by confocal microscopy (Leica TCS SP5, Leica Microsystems). In case of protoscoleces, z-stack images were generated. For quantification of cells in

metacestode vesicles, two random fields were photographed per vesicle. Fiji (Schindelin et al., 2012, 2015) was used for image processing. Z-stacks of protoscoleces were converted using z-projections (maximum intensity) and analyzed individually as z-stacks and z-projections. *sert* and *tph* positive cells were counted and analyzed according to position. On metacestode images, cells were identified on the Dapi channel and marked using a custom macro (see section 10.9). Positive cells (EdU and/or target gene) were counted manually based on marked cells. Percentages of EdU-positive, target gene positive and co-stained cells were calculated.

4.8 Bioinformatic analysis of *E. multilocularis sert*, *tph*, *tert* and *dnmt2*

4.8.1 Primer Design and Sequencing

Genes encoding the *E. multilocularis* serotonin transporter (SERT), tryptophan hydroxylase (TPH) (Camicia et al., 2013; Herz, 2015) and DNA-methyltransferase (DNMT2) (Geyer et al., 2013) were already identified previously. To identify genes encoding the telomerase reverse transcriptase (TERT) in *E. multilocularis*, annotated TERT protein sequences of other organisms (see section 10.7) were used for blastp searches against predicted *E. multilocularis* protein sequences. Blast hits were confirmed by blastp searches against the non-redundant protein sequences (nr) and the SwissProt/UniProt databases at NCBI. Primer design was based on available genome sequences and gene prediction data (EmuJ_000391300.1 for *sert*, EmuJ_000069500.1 for *tph*, EmuJ_001039300.1 and EmuJ_001038500.1 for *tert* and EmuJ_001185500.1 for *dnmt2*). The primers used for PCRs are shown in 4.4. After sequencing, the partially overlapping fragments were assembled in BioEdit 7.2.5 (Hall, 1999). The complete sequences were deposited at the EMBL Nucleotide Sequence Database under the accession numbers LT934126.1 (*sert*), LT934127.1 (*tph*), LR594027.1 (*tert*) and LR585068.1 (*dnmt2*).

4.8.2 Sequence analysis

In order to determine exon-intron-boundaries, the CDS sequences were analyzed with blastn searches against the *E. multilocularis* genome at WormBaseParaSite Version 9.0 (Tsai et al., 2013; Howe et al., 2016, 2017). CDS sequences were translated into amino acid sequences using BioEdit 7.2.5 (Hall, 1999) and domain analysis was performed with SMART 8.0 (Letunic et al., 2015; Letunic and Bork, 2018). To analyze conservation, blastp searches with the complete protein sequences or the sequences of the active domains were performed against the non-redundant protein

sequences of selected organisms in the National Center for Biotechnology Center (NCBI) database (cut-off $1e^{-5}$). Multiple sequence alignments were generated in BioEdit 7.2.5 (Hall, 1999) using MUSCLE (Edgar, 2004*a*,*b*).

4.9 Computational analysis of RNA-Seq data

4.9.1 RNA extraction and sequencing (performed by others)

Sample collection, RNA extraction and sequencing were performed by other members of the work groups Brehm at the Institute for Hygiene and Microbiology in Würzburg, Rosenzvit at the "Instituto de Microbiología y Parasitología Médica" in Buenos Aires and Berriman, "Parasite Genomics", at the Wellcome Trust Sanger Institute in Hinxton. In brief, samples of *E. multilocularis* from different larval stages and experiments (group Brehm), and samples of *E. granulosus* protoscoleces and metacestodes (group Rosenzvit) were collected and frozen in ten volumes of Tri-Reagent (5 PRIME). RNA isolation and processing for sequencing was performed at the Wellcome Trust Sanger Institute. RNA was isolated with TRIzol (Invitrogen, UK), treated with TURBO DNA-free DNAse (Ambion) to remove contaminating DNA, precipitated in ethanol and re-suspended in nuclease free water. Agilent RNA 6000 Nano-Bioanalyzer was used for assessment of RNA quality. RNA-Seq libraries (400-500 bp fragments) were produced using polyadenylated mRNA. Libraries were sequenced with Illumina Genome Analyser IIx or HiSeq following Illumina RNA-seq protocols, producing paired-end reads with 100 bp length. Image deconvolution and calculation of quality values were performed using Illumina GA pipeline v1.6.

4.9.2 Genomes and gene annotation

The current genome versions and gene annotations of *E. multilocularis* and *E. granulosus* (Tsai et al., 2013) were downloaded from WormBaseParaSite (WBPS7) (Howe et al., 2016) (see section 10.1). GO-terms were obtained at WormBaseParaSite BioMart.

4.9.3 RNA-Seq datasets

All datasets used for transcriptome data analysis are listed in 10.2. Abbreviations for *E. granulosus* datasets always start with "EG". Additionally to the unpublished datasets, the RNA-Seq datasets of *E. multilocularis* and *E. granulosus* from Tsai et al. (2013) were included as they contained information on further life-cycle stages and were not mapped to the current genome versions or analyzed with the current gene models. With the exception of EG_MCvivo (1 biological replicate), all new

samples consist of three biological replicates, while the published data used one biological replicate per sample.

4.9.4 Mapping and calculation of expression levels

Sequencing reads were mapped to their respective reference genomes using Hisat2 v2.0.5 (Kim et al., 2015) with a maximum intron length of 40000. Sequencing replicates of the same biological replicate were merged at this step. The output SAM files were converted to BAM files using samtools 1.2 (Li et al., 2009; Li, 2011). Reads per transcript were counted with HTSeqCount v0.7.1 (Anders et al., 2015) using a minimum quality score of 30 to filter out low quality or multiple mapped reads. The number of reads for each transcript was used to calculate TPM values (Transcripts Per kilobase Million).

4.9.5 Estimation of expression levels with Kallisto

To estimate expression for multiple copy genes or genes with identical sequence regions, Kallisto v0.43.1 (Bray et al., 2016) was used. Reads were "pseudo-mapped" to the predicted transcripts and expression was estimated in TPM values. Sequencing replicates of the same biological replicate were merged at the "pseudo-mapping" step.

4.9.6 Differential expression

For differential expression analysis only datasets with three biological replicates were used. Pairwise comparisons are listed in section 10.3. Differential expression was calculated using DESeq2 v1.16.1 (Love et al., 2014) on the statistical computation platform R v3.4.3 (R Development Core Team, 2008) based on read outs from HTSeqCount (Anders et al., 2015). After the integrated, independent filtering using genefilter v1.58.1 (Gentleman et al., 2017), the Benjamini-Hochberg procedure was performed to adjust for multiple testing (false discovery rate 0.05). For quality control, the fitting of the dispersion curve was evaluated by plotting the dispersions using the DESeq2 plotDispEsts function. Outlier detection was assessed by plotting of the Cook's distances.

For a more realistic representation of the actual biological fold changes of the expression levels, the lfcShrink function was used to calculate maximum a posteriori (MAP) log2fold changes (LFCs) additionally to the already calculated unshrunken maximum likelihood estimate (MLE) LFCs. Both MLE and MAP LFCs were visualized by plotting.

For visualization and interpretation, datasets were normalized using the rlog (regularized logarithm) function. To evaluate the data transformation, the variance of the normalized data was plotted using the meanSdPlot function from the vsn package (Huber et al., 2002). The transformed data was then used to used to generate heatmaps for the 20 highest expressed genes using the pheatmap package (Kolde, 2017). To observe similarities and differences within biological replicates and the degree of separation between different life-stages or treatments, PCA plots and sample-to-sample distances were calculated with the DESeq2 plotPCA and dist functions. Additionally to the pairwise comparisons, PCA plots and/or sample-to-sample distances were generated for the following groups with indicated samples as reference level:

Samples	Reference
PC1, PC2, PC3, MCnoBC	PC1
MCvivo, MCvitro, MCnoBC	MCvivo
MCvivo, MCnoBC, naPS, aPS	MCvivo
all Em samples	MCvivo

Session info is listed in 10.4.

4.9.7 GO-enrichment

Significantly differentially expressed genes (p_{adjust}<0,05) were analyzed concerning GO-enrichment using topGO_2.28.0 (Alexa and Rahnenfuhrer, 2016) on biological processed (BP) under the Fisher statistic and the Kolmogorov-Smirnov test with the algorithm "weight01" and a node size of 5. The gene universe was defined as all genes with a non-zero base mean.

In order to learn more about genes which are potentially specifically expressed in germinative cells, a list of genes was generated using the following criteria: (i) Significantly (p_{adjust}<0,05) lower expression in samples depleted of germinative cells with hydroxyurea (HU versus MCanaerob) and (ii) significantly (p_{adjust}<0,05) lower expression in samples treated with Bi2536 (Bi2536 versus mcDMSO) and (iii) higher expression in early primary cell cultures (sample PC1), which is enriched in germinative cells, compared to expression in later primary cell cultures (sample PC2) and to expression in metacestodes (sample MCnoBC). GO-enrichment analyses was performed with the Fisher statistic as described above.

To determine if gene expression was affected differently by treatment with either hydroxyurea or Bi2536, GO-enrichment analysis was performed with the Fisher

test on genes only affected by one of the treatments. Genes with significantly (p_{adjust}<0,05) lower expression in Bi2536 treated samples and constant or higher expression in samples treated with hydroxyurea (compared to their respective controls) were considered to be Bi2536 specifically downregulated and vice versa. Session info is listed in 10.5.

4.9.8 Reverse spliced reads

To get an idea of natural antisense transcripts, spliced reads were analyzed. As an unstranded protocol was used for sequencing, it was not possible to determine the orientation of reads by analyzing only their sequence. However, for spliced reads with canonical splice sites it was possible to ascertain their orientation. Spliced reads were extracted from BAM files with bamtools 2.3.0 (Barnett et al., 2011) filtering for each the tags XS:+ and XS:-. Hisat2 uses the tags to mark reads with canonical splice sites (GT..AG) in genome orientation (XS:+) and in reverse orientation (XS:-). The resulting BAM files were re-sorted by name using samtools 1.2 (Li et al., 2009; Li, 2011). Reads per transcript were counted with HTSeqCount v0.7.1 (Anders et al., 2015) with a minimum quality score of 30. Read counts for transcripts in genome orientation were based on spliced reads reverse to genome orientation and vice versa. Percentages of reverse spliced reads per transcript based on the total read counts for each transcript (as determined before, see section 4.9.4) were calculated.

4.10 Proteomic analysis (performed by others)

For proteome analysis, 2-day-old primary cell cultures from three isolates (H95, GH09 and Ingrid) were prepared by members of the work group Brehm. Sample processing and measurement with LC-MS/MS (Liquid Chromatography - Tandem Mass Spectrometry) was performed at the Leiden University Medical Center in the work group of Ron Hokke and by George Janssen. In brief, samples were labeled with TMT, prefractionated by high pH Reversed Phase C18 chromatography into 8 fractions and each fraction was measured on a LUMOS mass spectrometer. Sample quality and quantity was reported to be sufficient. Based on at least two unique tryptic peptides, 3400 proteins were identified using Mascot against predicted protein sequences of *E. multilocularis* downloaded from WormBaseParaSite Version 9.0 (Tsai et al., 2013; Howe et al., 2016, 2017) and their abundances estimated.

4.11 Comparison of *E. multilocularis* primary cell transcriptome and proteome

The proteome data of 2-day-old primary cells was compared to the transcriptome information of 2-day-old primary cells (sample PC1) to identify genes with high transcription levels without corresponding detected proteins and abundant proteins without transcripts (cut-off 10 TPMs).

4.12 Bioinformatic analysis of densovirus sequences in platyhelminth genomes

4.12.1 Identification of genes coding for homologues of densovirus non-capsid protein 1 in the *E. multilocularis* genome

To identify genes that are specifically expressed in germinative cells, transcriptome data analysis was performed (see section 4.9). The predicted gene EmuJ_000388600 annotated as "non-capsid protein 1" was detected by this method and analyzed further. The predicted protein sequence of EmuJ_000388600 and the downstream ORF EmuJ_000388500 served as query for blastp searches (E-value $< 1e^{-10}$, identities > 20 %, coverage > 50 %) against the SWISSPROT database at GenomeNET and for domain analysis with pfam (E-value < 1e-10) (Finn et al., 2016). Using MUSCLE v3.8.31 (Edgar, 2004*a*,*b*) a multiple sequence alignment was generated for EmuJ_000388600 and the first two blast hits.

Further putative densovirus non-capsid protein 1 sequences in the *E. multilocularis* genome were identified by blast searches: The protein sequence for EmuJ_000388600 was used as query for blastp searches (E-value $< 1e^{-10}$, identities > 80 %, coverage > 30 %) against the predicted protein sequences of *E. multilocularis* (downloaded from WormBaseParaSite WBPS 14) (Howe et al., 2016, 2017). Retrieved sequences were confirmed by local blastp searches against the non-redundant sequences (nr) database from NCBI (E-value $< 1e^{-10}$, identities > 90 %, coverage > 90 %). Confirmed sequences served then as queries for tblastn searches (E-value $< 1e^{-10}$, identities > 70 %, coverage > 20 %) against the *E. multilocularis* genome. Retrieved nucleotide sequences were confirmed by blastx searches (E-value $< 1e^{-10}$, identities > 80 %, coverage > 80 %) against the non-redundant sequences (nr) database from NCBI.

4.12.2 Analysis of densovirus sites in the *E. multilocularis* genome

The longest detected nucleotide sequences for the designated *Emu*DNV-NS1 (*E. multilocularis* densovirus non-capsid protein 1 gene) were presumed to be complete gene copies and were used for local blastn searches against the *E. multilocularis* genome to determine start and end positions of all *Emu*DNV-NS1 gene copies. After individual curation of start and end positions, frameshift mutations were detected by BioEdit six-frame translation (Hall, 1999). Often, another ORF was found downstream of *Emu*DNV-NS1. The ORF was presumed to be coding for a capsid protein (VP) and hence named *Emu*DNV-VP (*E. multilocularis* densovirus capsid protein gene). To identify further *Emu*DNV-VP gene copies, the *E. multilocularis* genome was searched with blastn (E-value $< 1e^{-10}$, identities > 90 %, coverage > 10 %) using the longest *Emu*DNV-VP gene version as query. Start and end positions of *Emu*DNV-VP gene copies were curated manually. BioEdit six-frame translation (Hall, 1999) was used to find frameshift mutations and to translate the *Emu*DNV-VP nucleotide sequences into amino acid sequences. The EmuDNV-NS1 and EmuDNV-VP protein sequences were used for domain analyses with pfam (E-value $< 1e^{-10}$) (Finn et al., 2016) and for blastp searches (E-value $< 1e^{-10}$, identities $>$ 20 %, coverage > 90 %) against the non-redundant protein sequences (nr) database (organism viruses) and the SwissProt/UniProt database at NCBI. Previously, putative promoter elements, such as TATA-boxes and activator elements, have been described for the *Penaeus stylirostris* densovirus (Rai et al., 2011). Similar promoter structures were detected by individual inspection of the upstream regions of *Emu*DNV-NS1 and *Emu*DNV-VP. The shown alignment of the promoter region was performed with MUSCLE (Edgar, 2004*a*,*b*). To find ITRs, *Emu*DNV-NS1 nucleotide sequences together with 5000 bp flanking regions on both sides were as input for the computer program einverted (Richard Durbin, modification by Peter Rice; http://www.bioinformatics.nl/cgi-bin/emboss/einverted) (maximum extent of repeats 2000 bp, > 80 % matches, loop < 100 bp). The longest found ITR sequence (370 bp long) was then used as query for local blastn searches (E-value $< 1e^{-5}$, identities > 80 %, coverage > 10 %) against the *E. multilocularis* genome to detect incomplete ITR sequences nearby *Emu*DNV-NS1 genes. Analysis of neighboring genes and the genomic location of densovirus sites was performed with genome browser Ensemble at WormBaseParaSite (WBPS10) (Howe et al., 2016, 2017).

4.12.3 Identification of genes coding for homologues of densovirus non-capsid protein 1 in cestode genomes

To detect putative densovirus non-capsid protein 1 gene sequences in other tapeworm genomes, all available tapeworm genomes as well as the genome of the trematode *S. mansoni* were downloaded from WormBaseParaSite WBPS 14:

Species	Assembly version	Citation
Dibothriocephalus latus	D_latum_Geneva_0011_upd	(International-Helminth-Genomes-Consortium, 2019)
Echinococcus canadensis	ECANG7	(Maldonado et al., 2017)
E. granulosus (1)	EGRAN001	(Tsai et al., 2013)
E. granulosus (2)	ASM52419v1	(Zheng et al., 2013)
E. multilocularis	EMULTI002	(Tsai et al., 2013)
Hydatigera taeniaeformis	H_taeniaeformis_Canary_Islands_0011_upd	(International-Helminth-Genomes-Consortium, 2019)
Hymenolepis diminuta	H_diminuta_Denmark_0011_upd	(International-Helminth-Genomes-Consortium, 2019)
Hymenolepis microstoma	HMN_v3	(Tsai et al., 2013)
Hymenolepis nana	H_nana_Japan_0011_upd	(International-Helminth-Genomes-Consortium, 2019)
Mesocestoides corti	M_corti_Specht_Voge_0011_upd	(International-Helminth-Genomes-Consortium, 2019)
Schistocephalus solidus	S_solidus_NST_G2_0011_upd	(International-Helminth-Genomes-Consortium, 2019)
Spirometra erinaceieuropaei	S_erinaceieuropaei	(Bennett et al., 2014)
Taenia asiatica (1)	T_asiatica_South_Korea_0011_upd	(International-Helminth-Genomes-Consortium, 2019)
Taenia asiatica (2)	Taenia_asiatica_TASYD01_v1	(Wang et al., 2016)
Taenia multiceps	ASM192302v3	(Li et al., 2018)
Taenia saginata	ASM169307v2	(Wang et al., 2016)
Taenia solium	Tsolium_Mexico_v1	(Tsai et al., 2013)
S. mansoni	Smansoni_v7	(Berriman et al., 2009; Protasio et al., 2012)

The downloaded genomes were searched by local tblastn searches (E-value $< 1e^{-5}$, identities > 30 %, coverage > 30 %) with the putative non-capsid protein 1 EmuJ_000388600 as query. Subsequently, blast hits were confirmed by local blastx searches

(E-value $< 1e^{-5}$, identities > 35 %, coverage > 90 %) against the non-redundant sequences (nr) database from NCBI. Confirmed sequences were then used as queries for blastn searches (E-value $< 1e^{-10}$, identities > 70 %, coverage > 30 %) against the downloaded genomes. Hits overlapping more than 30 % of the sequence on the genomic location were merged before reciprocal blastx searches (E-value $< 1e^{-5}$, identities > 35 %, coverage > 90 %) against the non-redundant sequences (nr) database from NCBI.

4.12.4 Phylogenetic analysis of densovirus NS1 genes in cestode genomes

Confirmed nucleotide sequences (coverage > 50 % ofpf version EmuJ_000388600) were aligned with MUSCLE (Edgar, 2004*a*,*b*). Phylogenetic analysis was performed using MEGA-X (Kumar et al., 2018) with the Neighbor-Joining method (Saitou and Nei, 1987) using pairwise deletion for gaps and 1000 bootstrap replications (Felsenstein, 1985). Branches with bootstrap reproduction rates less than 50 % were collapsed.

5 Results

5.1 Genome wide expression analysis

To understand developmental changes throughout the *E. multilocularis* and *E. granulosus* life cycle, genome wide expression analysis was performed using the accessible life stages of the two tapeworms. The transcriptome of 8 developmental stages and tissues of *E. multilocularis* were analyzed: Three stages of primary cell cultures (early (PC1), middle (PC2) and late stage (PC3)), metacestodes without brood capsules (MCnoBC), non-activated (naPS) and activated (aPS) protoscoleces as well as metacestodes cultivated under aerob conditions with feeder cells (MCvitro) and metacestode material isolated directly from infected jirds (MCvivo). *E. granulosus* material was retrieved from an abattoir in form of complete cysts. Non-activated protoscoleces (EG_naPS) and cyst wall material (EG_MCvivo) were used as samples. Additionally, *E. multilocularis* samples from experiments were analyzed to learn more about germinative cell specific genes and about the influence of β-catenin on gene expression.

Sample collection, RNA extraction and sequencing were performed by members of the work groups Brehm at the Institute for Hygiene and Microbiology in Würzburg, Rosenzvit at the "Instituto de Microbiología y Parasitología Médica" in Buenos Aires and Berriman, "Parasite Genomics", at the Wellcome Trust Sanger Institute in Hinxton. The subsequent bioinformatic analyses were performed and evaluated by me.

The already published RNA-Seq datasets (see section 10.2.2)(Tsai et al., 2013) were also updated as they included samples of additional life cycle stages for *E. multilocularis*: metacestodes with brood capsules (MC_LateBC), pregravid (EmPreAWDog) and gravid (EmAdultGravide) adults.

In the following, *E. granulosus* samples will always be declared as such (abbreviation "EG" in sample names) while *E. multilocularis* samples will simply be referred to by their sample names. Unless specified otherwise, expression levels (in TPMs) will always refer to information obtained with the Hisat2/HTSeqCount method.

5.1.1 Quality control and summarized results

5.1.1.1 Mapping and pseudo-mapping results in high alignment rates

Mapping of the reads to their reference genomes with Hisat2 (Kim et al., 2015) resulted in overall alignment rates of over 93 % for most datasets. Of the *E. multilocularis* datasets only those of the sample MCvivo, that was isolated directly from jirds, and the adult samples (EmPreAWDog and EmAdultGravide) had lower overall alignment rates, which was attributed to host contamination of the samples. The overall alignment rates for *E. granulosus* datasets ranged from 61 to 78 % (see table 1). Percentages of pseudo-aligned reads from Kallisto (Bray et al., 2016) were generally lower. This was expected as the pseudo-alignment process was based exclusively on the sequences of the gene models. For most datasets pseudo-alignment rates ranged from 51 to 80 %. Only the samples MCvivo, EmPreAWDog and EmAdultGravide showed lower pseudo-alignment rates (see table 1).

Sample	Alignment rate with Hisat2	Pseudo-aligned reads with Kallisto
PC1.1	95,79 %	75,60 %
PC1.2	96,58 %	76,09 %
PC1.3	97,19 %	79,11 %
PC2.1	95,19 %	75,17 %
PC2.2	96,56 %	76,67 %
PC2.3	97,40 %	75,68 %
PC3.1	93,72 %	77,16 %
PC3.2	96,19 %	76,83 %
PC3.3	97,19 %	80,28 %
MCnoBC.1	94,67 %	76,12 %
MCnoBC.2	95,56 %	73,62 %
MCnoBC.3	96,98 %	74,21 %
naPS.1	94,92 %	67,85 %
naPS.2	95,61 %	71,58 %
naPS.3	97,07 %	70,85 %
aPS.1	96,49 %	74,90 %
aPS.2	96,56 %	73,97 %
aPS.3	98,05 %	70,40 %
MCvitro.1	96,02 %	74,58 %
MCvitro.2	96,45 %	73,25 %
MCvitro.3	97,30 %	73,15 %

MCvivo.1	24,49 %	16,84 %
MCvivo.2	41,37 %	25,18 %
MCvivo.3	36,69 %	24,27 %
HU.1	95,74 %	71,79 %
HU.2	96,23 %	67,99 %
HU.3	96,31 %	73,77 %
MCanaerob.1	95,79 %	70,44 %
MCanaerob.2	96,36 %	66,80 %
MCanaerob.3	96,11 %	73,44 %
Bi2536.1	97,94 %	76,93 %
Bi2536.2	98,11 %	76,59 %
Bi2536.3	97,78 %	77,59 %
mcDMSO.1	98,03 %	75,93 %
mcDMSO.2	97,72 %	75,39 %
mcDMSO.3	98,00 %	75,27 %
beta.1	99,05 %	76,29 %
beta.2	98,99 %	76,65 %
beta.3	99,00 %	75,57 %
Neg.1	98,65 %	76,47 %
Neg.2	98,54 %	75,66 %
Neg.3	98,92 %	76,22 %
EG_MCvivo	70,23 %	66,73 %
EG_naPS.1	78,83 %	71,55 %
EG_naPS.2	74,72 %	67,49 %
EG_naPS.3	74,35 %	67,20 %
PC_2d	96,74 %	72,18 %
PC_11d	96,41 %	67,21 %
MC_noBC	96,15 %	66,29 %
MC_LateBC	96,76 %	51,22 %
PS_noact	96,77 %	60,11 %
PS_act	96,96 %	65,72 %
EmPreAWDog	86,76 %	35,92 %
EmAdultGravide	88,31 %	32,03 %
EG_PS_noact	61,48 %	58,00 %

Table 1: **Mapping statistics:** Shown are the overall alignment rates of mapping with Hisat2 (Kim et al., 2015) and the percentages of pseudo-aligned reads with Kallisto (Bray et al., 2016) for each biological replicate. Number behind the sample name indicates the biological replicate. `New samples:` **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds; **HU**: metacestodes treated with hydroxyurea; **MCanaerob**: control for HU treatment; **Bi2536**: metacestodes treated with Bi2536; **mcDMSO**: control for Bi2536 treatment; **beta**: primary cells treated with β-catenin siRNA; **Neg**: control for β-catenin siRNA treatment; **EG_MCvivo**: cyst wall of *E. granulosus*; **EG_naPS**: non-activated protoscoleces of *E. granulosus*. `Old samples:` **PC_2d**: primary cells, 2-day-old; **PC_11d**: primary cells, 11 days old; **MC_noBC**: metacestode vesicles without brood capsules; **MC_LateBC**: metacestodes with brood capsules; **PS_noact**: non-activated protoscoleces; **PS_act**: activated protoscoleces; **EmPreAWDog**: pregravid adult; **EmAdult Gravide**: gravid adults; **EG_PS_noact**: non-activated protoscoleces of *E. granulosus*.

5.1.1.2 All datasets used for comparisons are of satisfactory quality

For statistical comparisons, only datasets with three biological replicates were used. Therefore, the old datasets and the *E. granulosus* datasets were excluded. The quality of the data was evaluated at several steps in the analysis. The number of genes with nonzero total read counts for each comparison was between 9958 (MCnoBC versus MCvivo) and 10276 (aPS versus naPS) (see table 2). Fitting of the dispersion curve looked similar for all pairwise comparisons. Exemplary, the dispersion plot of the comparison aPS versus naPS is shown (see Figure 7).

Cook's distances for outlier detection showed no sample with high numbers of outliers in any of the comparisons. The highest number observed was 30 outliers (less than 0.3 %) in the comparison naPS versus MCnoBC (see table 2). Also, there was no biological replicate with consistently higher Cook's distances than the others in any of the comparisons (see Figure 8 as an example).

Independent filtering was applied to exclude genes with low counts from adjustment for multiple testing to increase statistical power (see Figure 9 as an example). In the analysis for differential gene expression, between 196 (naPS versus MCnoBC) and 1192 (PC3 versus PC2) genes were considered to have low counts (see table 2).

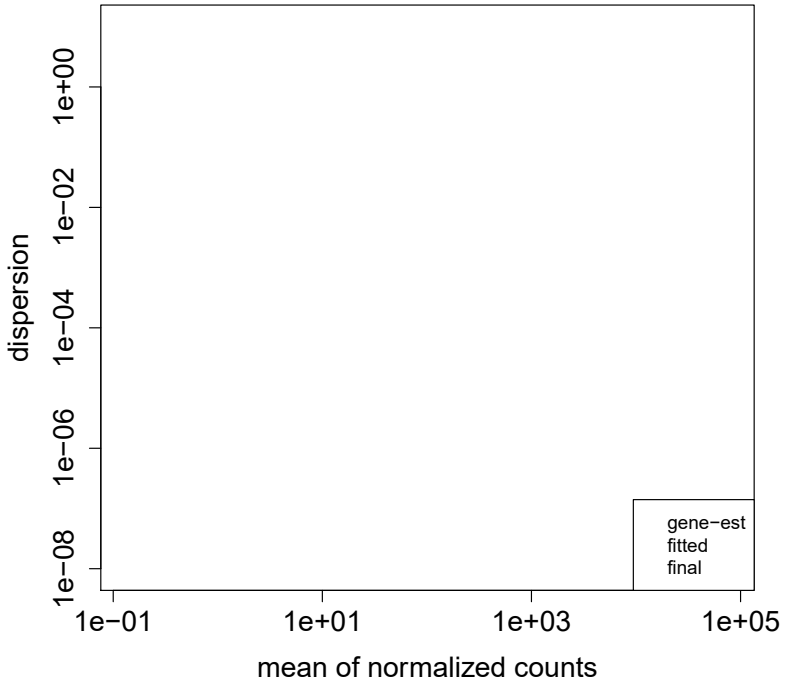


Figure 7: **Dispersion plot for the comparison aPS versus naPS:** Gene-wise dispersion estimates are shown in black, the fitting curve in red and the final estimates in blue. Gene-wise estimates flagged as outliers and therefore not shrunken towards the fitting curve are circled in blue. **aPS**: activated protoscoleces; **naPS**: non-activated protoscoleces.

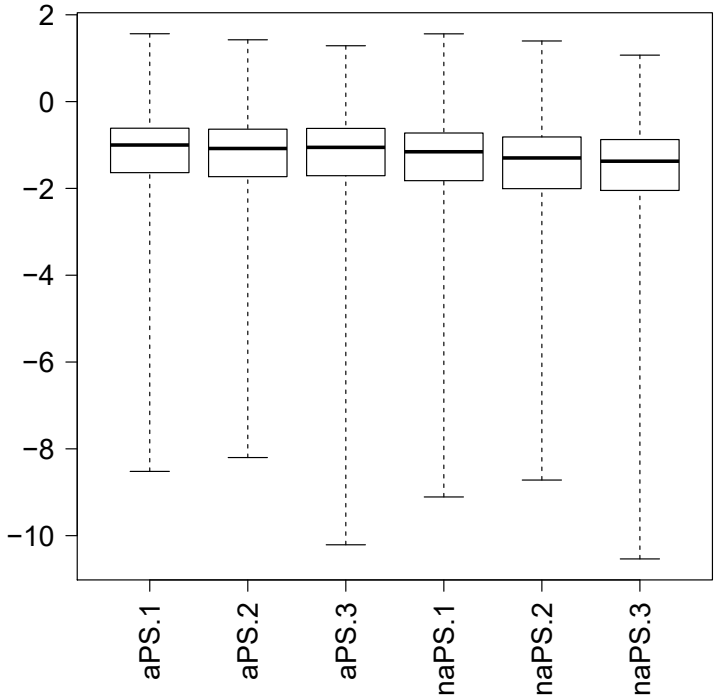


Figure 8: **Cook's distances:** Shown is a boxplot of the Cook's distances for each biological replicate in the comparison naPS versus MCnoBC. **naPS**: non-activated protoscoleces; **MCnoBC**: metacestode without brood capsules.

Comparison	nonzero read counts	LFC > 0	LFC < 0	outliers	low counts
PC1 versus MCnoBC	10102	1627	1832	15	781
PC2 versus PC1	10218	1551	1479	13	791
PC3 versus PC2	10264	189	199	16	1192
MCnoBC versus PC3	10189	1282	1475	26	592
naPS versus MCnoBC	10229	2773	2583	30	196
naPS versus MCvivo	10214	2963	2591	20	593
aPS versus naPS	10276	795	430	25	796
MCvitro versus MCvivo	9994	1756	1450	5	1159
MCnoBC versus MCvivo	9958	1744	1570	11	962
MCnoBC versus MCvitro	10002	419	817	8	1160
HU versus MCanaerob	9995	1736	1853	3	773
Bi2536 versus mcDMSO	10017	2768	2662	0	387
beta versus Neg	10059	1314	1513	2	972

Table 2: **Summary of DESeq2 results**: Second mentioned in comparisons is the reference level. Shown are the number of genes with nonzero total read counts, the number of genes significantly ($p_{adjust} < 0{,}05$) higher expressed, the number of genes significantly ($p_{adjust} < 0{,}05$) lower expressed (equals higher expression in reference sample), the number of outlier genes in the comparisons and the number of genes that were excluded from the adjustment of p-values due to low counts (as determined by independent filtering). **LFC**: log2fold change; **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds; **HU**: metacestodes treated with hydroxyurea; **MCanaerob**: control for HU treatment; **Bi2536**: metacestodes treated with Bi2536; **mcDMSO**: control for Bi2536 treatment; **beta**: primary cells treated with β-catenin siRNA; **Neg**: control for β-catenin siRNA treatment.

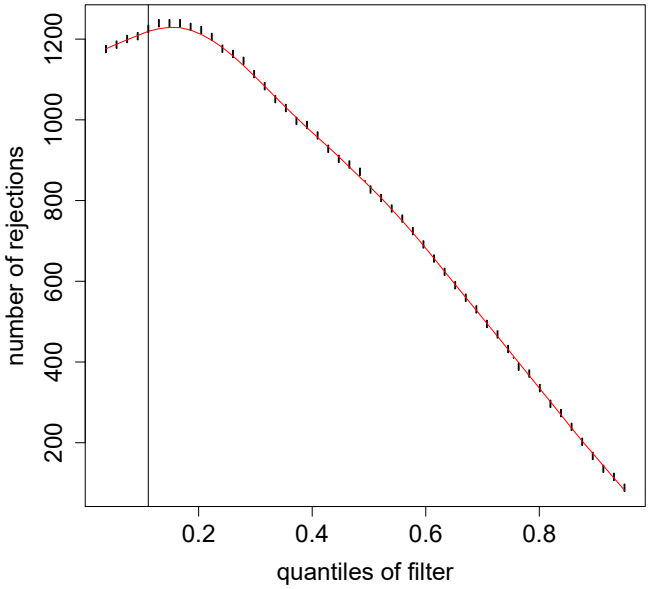
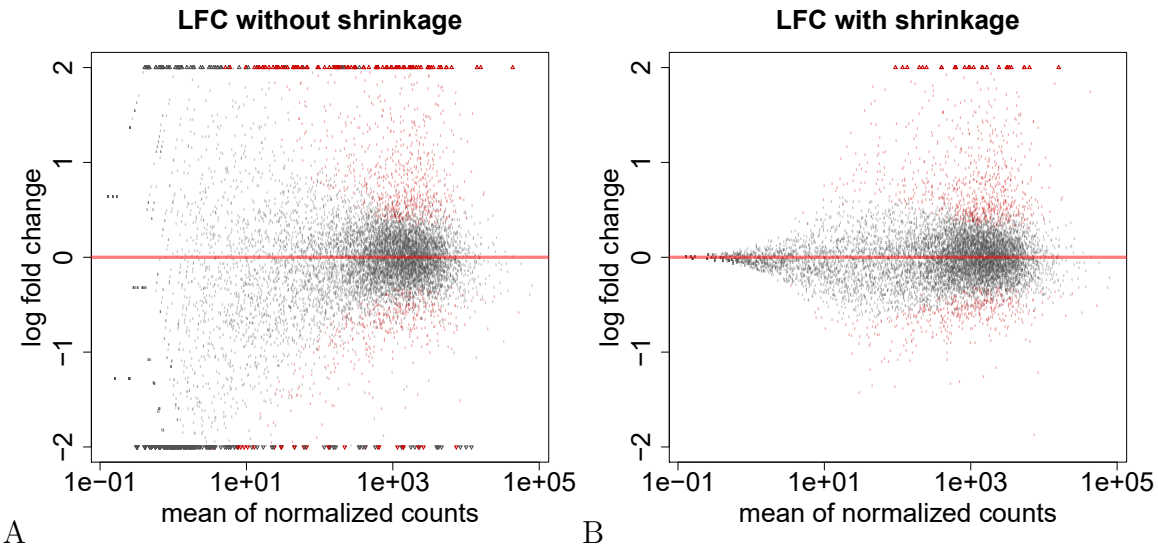


Figure 9: **Independent filtering:** Shown are the numbers of rejections ($p_{adjust} < 0{,}05$) over the quantiles of a filter statistic (based on mean of normalized counts) for the comparison aPS versus naPS. The chosen filtering threshold is represented by the vertical line. **aPS**: activated protoscoleces; **naPS**: non-activated protoscoleces.

Numbers of significantly ($p_{adjust} < 0{,}05$) differentially expressed genes varied between 388 (PC3 versus PC2) and 5554 (naPS versus MCvivo) with similar numbers of genes being up- and downregulated within each comparison (see table 2).

As maximum likelihood estimate (MLE) LFCs highly depend on mean read counts, shrunken maximum a posteriori (MAP) LFCs were calculated for all comparisons to obtain more robust values less likely to overestimate actual biological changes. As an example, the MLE and MAP LFCs of the comparison aPS versus naPS are shown, plotted against the mean of normalized read counts (see Figure 10). As in the example, in all comparisons LFCs of genes with lower mean of normalized read counts were more affected by shrinkage than higher expressed genes.

Data normalization with rlog for all comparisons removed the dependence of the variance on the mean, thereby reducing standard deviations in the lower count range (see Figure 11 as an example). Heatmaps generated from normalized data using the highest expressed genes visualized genes with differential expression (see Figure 12 6^{th} row) and genes with more constant expression. Sometimes, the influence of the sample isolate on gene expression could be observed, as seen for the first gene in figure 12 which is higher expressed in the isolate Ingrid than in the isolates GH09 or MS1010.



A B

Figure 10: **Shrinkage of LFCs:** Shown are LFC values without (**A**) and with (**B**) shrinkage from the comparison aPS versus naPS plotted against the mean of the normalized read counts. Significantly differentially ($p_{adjust} < 0{,}05$) expressed genes are shown in red. **LFC:** log2fold change; **aPS**: activated protoscoleces; **naPS**: non-activated protoscoleces.

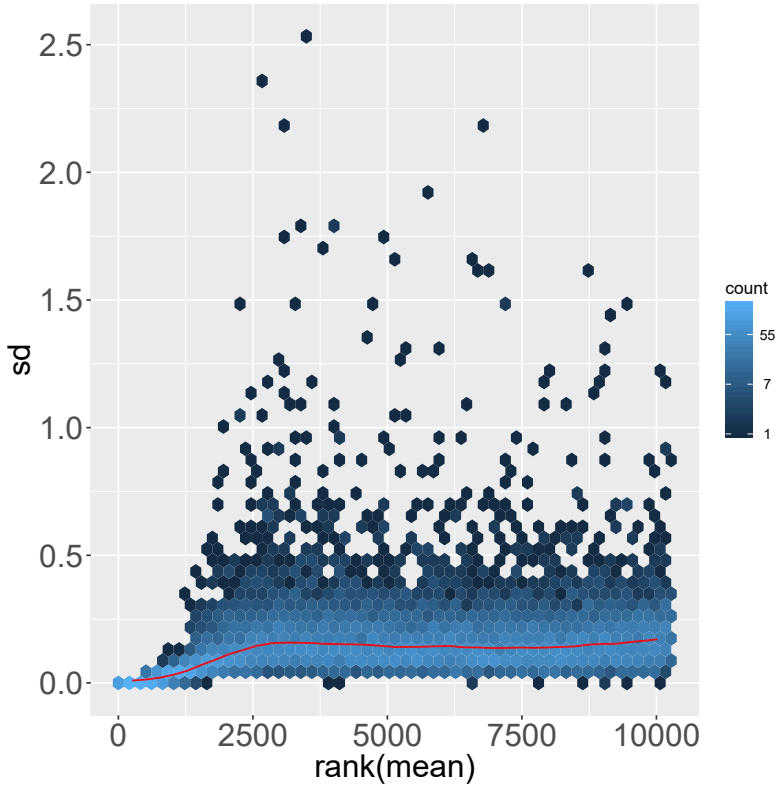


Figure 11: **Data normalization with regularized logarithm:** The standard deviation across samples of the transformed data from the comparison aPS versus naPS is plotted against the mean. **sd**: standard deviation; **aPS**: activated protoscoleces; **naPS**: non-activated protoscoleces.

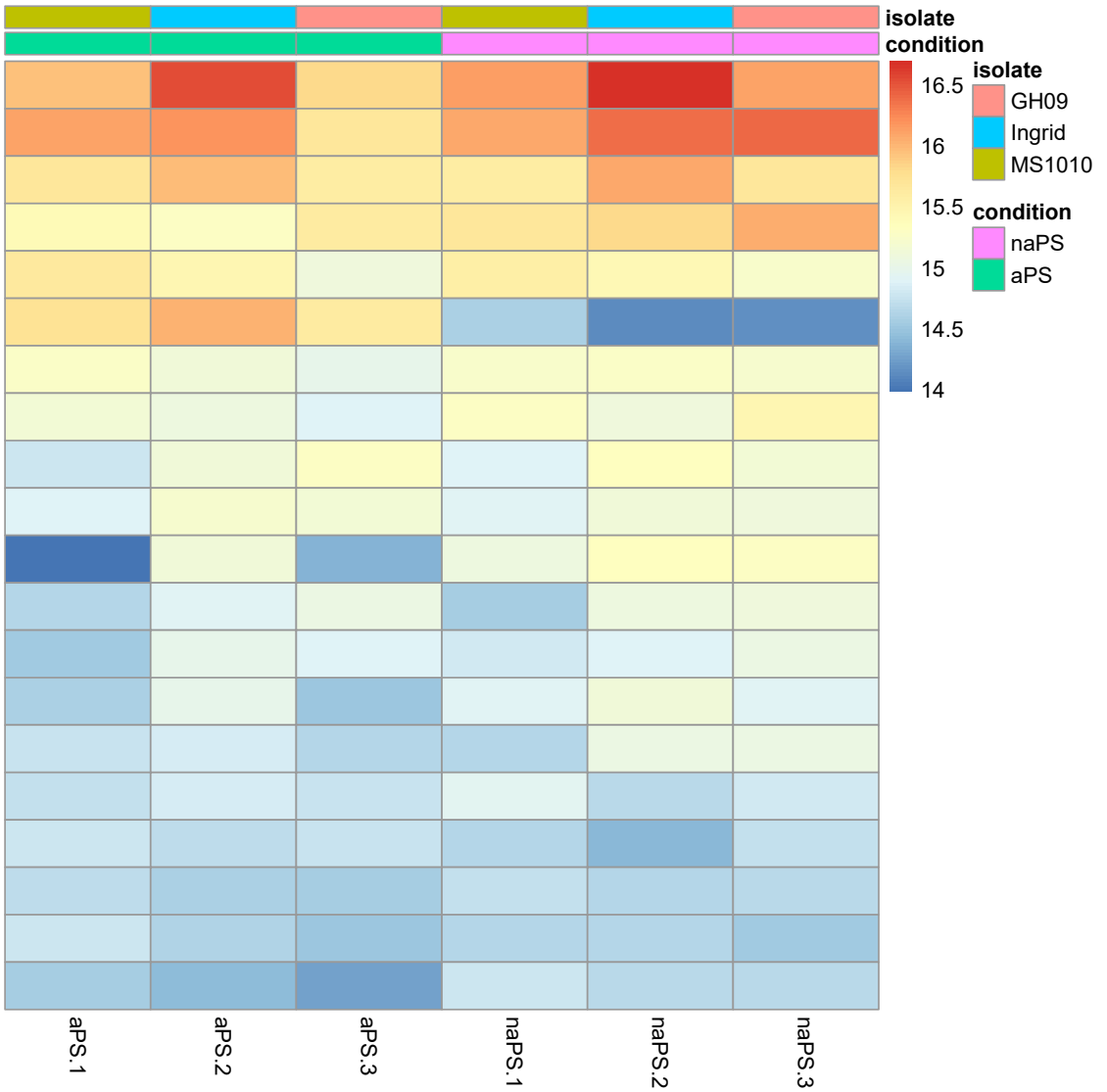


Figure 12: **Heatmap of the highest expressed genes in the comparison aPS versus naPS:** Shown is a heatmap of the 20 genes with the highest base mean in the comparison aPS versus naPS. Each column represents a biological replicate. Abbreviations of the replicates are written below the columns. The conditions and isolates of the replicates are color coded in the first two rows. **aPS**: activated protoscoleces; **naPS**: non-activated protoscoleces.

5.1.1.3 GO-enrichment analyses

GO-enrichment analyses were performed for all pairwise comparisons, each for the significantly ($p_{adjust} < 0{,}05$) higher and lower expressed genes with both the Fisher and the Kolmogorov-Smirnov test. The numbers of annotated genes were between 4388 (MCnoBC versus MCvitro) and 4439 (PC3 versus PC2) (see table 3). The numbers of annotated genes varied slightly as the gene universe for each comparison was restricted to genes with a nonzero base mean in the compared datasets. The numbers of significant (and annotated) genes and the resulting significant terms are listed in table 3.

		higher expressed			lower expressed		
	annotated genes	sig genes	sig terms Fisher	sig terms KS	sig genes	sig terms Fisher	sig terms KS
PC1 versus MCnoBC	4410	826	800	917	882	757	917
PC2 versus PC1	4431	621	722	917	812	842	917
PC3 versus PC2	4439	95	288	917	83	397	917
MCnoBC versus PC3	4426	673	752	917	648	688	917
naPS versus MCnoBC	4428	1137	826	917	1407	858	917
naPS versus MCvivo	4426	1234	750	917	1452	889	917
aPS versus naPS	4431	401	579	918	213	580	918
MCvitro versus MCvivo	4396	864	778	917	763	812	917
MCnoBC versus MCvivo	4390	819	753	917	775	789	917
MCnoBC versus MCvitro	4388	171	443	915	409	705	915
HU versus MCanaerob	4393	847	757	915	920	818	915
Bi2536 versus mcDMSO	4398	1351	880	915	1290	863	915
beta versus Neg	4405	663	758	914	751	782	914

Table 3: **GO-enrichment analyses for pairwise comparisons**: Shown are the numbers of annotated genes, significant (**sig**) genes and the resulting significant terms for the Fisher and the Kolmogorov-Smirnov (**KS**) tests both for the higher and lower expressed genes of each comparison.

Additionally, GO-enrichment analysis was performed for genes that are likely specifically expressed in germinative cells (see table 4). To learn about differences between HU and Bi2536 germinative cell depletion, 150 HU and 322 Bi2536 specifically and significantly down-regulated genes (of 4424 annotated genes) were evaluated with GO-enrichment analyses (see table 4).

	annotated genes	sig genes	sig terms
germinative cell specific genes	4452	395	702
HU specific genes	4424	150	429
Bi2536 specific genes	4424	322	663

Table 4: **Further GO-enrichment analysis**: Shown are the numbers of annotated genes, significant (**sig**) genes and the resulting significant terms with the Fisher test. **germinative cell specific**: genes significantly downregulated in HU (compared to MCanaerob) and Bi2536 (compared to mcDMSO) and up-regulated in PC1 (compared to PC2 and to MCnoBC); **HU specific**: genes significantly downregulated in HU (compared to MCanaerob) and constant or up-regulated in Bi2536 (compared to mcDMSO); **Bi2536 specific**: genes significantly downregulated in Bi2536 (compared to mcDMSO) and constant or up-regulated in HU (compared to MCanaerob).

5.1.2 Over 40 % of genes are expressed in all examined life-cycle stages

Of 10669 predicted *E. multilocularis* genes, 4608 were expressed ($\geq$ 10 TPM) in each analyzed stage of the life-cycle (including the old datasets) and of 10275 predicted *E. granulosus* genes 6674 were expressed ($\geq$ 10 TPM) in both protoscoleces and cyst wall according to HTSeqCount (Anders et al., 2015) results. The highest expressed genes in *E. multilocularis* and *E. granulosus* included genes coding for cytoskeleton elements, Polyubiquitin, 14-3-3 as well as genes encoding proteins involved in metabolism (e.g. glycolysis), protein generation and folding (elongation factors, heat shock proteins) (see table 5). Another highly expressed gene was coding for ferritin which is considered a potential candidate for diagnostics and therapy of echinococcosis. *E. granulosus* ferritin is present in the hydatid fluid of cysts and shows potential as an immunodiagnostic tool (Ersfeld and Craig, 1995; Aziz et al., 2011) and as an antigen for vaccines (Wang et al., 2015). *E. multilocularis* ferritin was also shown to potentially interact with the antimalarial drug mefloquine and might be a suitable drug target (Küster et al., 2015).

Life cycle stage	Highest expressed genes
PC1	Elongation factor 1-alpha, Actin cytoplasmic type 5, Heat shock protein 90, Ferritin, Phosphoenolpyruvate carboxykinase, Peptidyl-prolyl cis-trans isomerase, Ubiquitin ribosomal protein L40, Glyceraldehyde-3-phosphate dehydrogenase, Glutathione S-transferase, Citrate synthase

PC2	Ferritin, Ubiquitin ribosomal protein L40, Heat shock protein 90, Hypothetical transcript, Polyubiquitin, Heat shock 70 kDa protein 4, Actin cytoplasmic type 5, Expressed conserved protein, Phosphoenolpyruvate carboxykinase, Elongation factor 1-alpha
PC3	Elongation factor 1-alpha, Ubiquitin ribosomal protein L40, Actin cytoplasmic type 5, Heat shock protein 90, Phosphoenolpyruvate carboxykinase, Tubulin alpha chain, Ferritin, Polyubiquitin, Glyceraldehyde-3-phosphate dehydrogenase, Peptidyl-prolyl cis-trans isomerase
MCnoBC	Antigen B, Phosphoenolpyruvate carboxykinase, Expressed protein, Tetraspanin, Ferritin, Expressed protein, Expressed protein, Tegumental protein, Glyceraldehyde-3-phosphate dehydrogenase, Ubiquitin ribosomal protein L40
naPS	EG19 antigen, Metal transporter Nramp1, Senescence associated protein, Phosphoenolpyruvate carboxykinase, Myosin regulatory light chain, Ubiquitin ribosomal protein L40, Heat shock protein 90, Glyceraldehyde-3-phosphate dehydrogenase, Expressed protein, Elongation factor 1-alpha
aPS	TSP1, EG19 antigen, Polyubiquitin, Ubiquitin ribosomal protein L40, Phosphoenolpyruvate carboxykinase, Myosin regulatory light chain, Expressed protein, Elongation factor 1-alpha, Tetraspanin, Tetraspanin
MCvitro	Antigen B, Phosphoenolpyruvate carboxykinase, Ferritin, Oxalate:formate antiporter, Tetraspanin, Expressed protein, Glyceraldehyde-3-phosphate dehydrogenase, Expressed protein, Fructose-bisphosphate aldolase, Antigen B
MCvivo	Antigen B, Ferritin, Phosphoenolpyruvate carboxykinase, Glyceraldehyde-3-phosphate dehydrogenase, Elongation factor 1-alpha, Actin cytoplasmic type 5, Heat shock protein 90, Citrate synthase, Peptidyl-prolyl cis-trans isomerase, Expressed protein

PC_2d	Elongation factor 1-alpha, Actin cytoplasmic type 5, Heat shock protein 90, Peptidyl-prolyl cis-trans isomerase, 40S ribosomal protein S27, Ribosomal protein lp1, Expressed protein, Expressed protein, Sensor histidine kinase, Antigen B
PC_11d	Expressed conserved protein, Heat shock protein 90, Hypothetical transcript, Elongation factor 1-alpha, Ferritin, Peptidyl-prolyl cis-trans isomerase, Histone H3, Tegumental protein, Ubiquitin ribosomal protein L40, Actin cytoplasmic type 5
MC_noBC	Antigen B, Phosphoenolpyruvate carboxykinase, Expressed conserved protein, Expressed protein, Expressed protein, Tetraspanin, Glyceraldehyde-3-phosphate dehydrogenase, Antigen B, Tegumental protein, Expressed protein
MC_LateBC	Tegumental protein, Expressed protein, Ferritin, Phosphoenolpyruvate carboxykinase, Phospholipid transporting ATPase VA, Tetraspanin, Glyceraldehyde-3-phosphate dehydrogenase, Fructose-bisphosphate aldolase, Peptidyl-prolyl cis-trans isomerase, Oxalate:formate antiporter
PS_nonact	EG19 antigen, Tegumental protein, Heat shock protein 90, NADH dehydrogenase subunit 4L, Elongation factor 1-alpha, Phospholipid transporting ATPase VA, Expressed protein, Phosphoenolpyruvate carboxykinase, Glyceraldehyde-3-phosphate dehydrogenase, Myosin regulatory light chain
PS_act	TSP1, Elongation factor 1-alpha, Cupin 2 barrel domain containing protein, Expressed protein, Tegumental protein, Polyubiquitin, EG19 antigen, Tetraspanin, Expressed protein, Expressed protein
EmPreAWDog	Cytochrome c oxidase subunit II, NADH dehydrogenase subunit 4L, Histone acetyltransferase myst4, Ubiquitin ribosomal protein L40, NADH dehydrogenase subunit 6, Polyubiquitin, Expressed protein, Diagnostic antigen gp50, Glutathione S-transferase, Universal minicircle sequence binding protein UMSBP

EmAdultGravide	Ubiquitin ribosomal protein L40, Polyubiquitin, Expressed conserved protein, Expressed conserved protein, Tegumental protein, Expressed protein, Cytochrome c oxidase subunit II, Heat shock 70 kDa protein 4, Expressed protein, Proteinase inhibitor I2 Kunitz metazoa
EG_MCvivo	Actin-3, Elongation factor 1-alpha, Glyceraldehyde 3 phosphate dehydrogenase, Fructose-bisphosphate aldolase, Cysticercus cellulosae specific antigenic, TSP5, Peptidyl-prolyl cis-trans isomerase, Tegumental protein, Phosphoenolpyruvate carboxykinase, Inhibitor of apoptosis protein
EG_naPS	Cysticercus cellulosae specific antigenic, Elongation factor 1-alpha, TSP5, Myosin essential light chain, Fructose-bisphosphate aldolase, Actin-3, Glyceraldehyde 3 phosphate dehydrogenase, Fatty acid-binding protein homolog 1, Large subunit ribosomal protein L27Ae, Expressed protein
EG_PS_nonact	Cysticercus cellulosae specific antigenic, Elongation factor 1-alpha, Ribosomal protein S25, Large subunit ribosomal protein L27Ae, Actin-3, TSP1, Peptidyl-prolyl cis-trans isomerase, DnaJ subfamily A, Phospholipid transporting ATPase VA, Myosin essential light chain

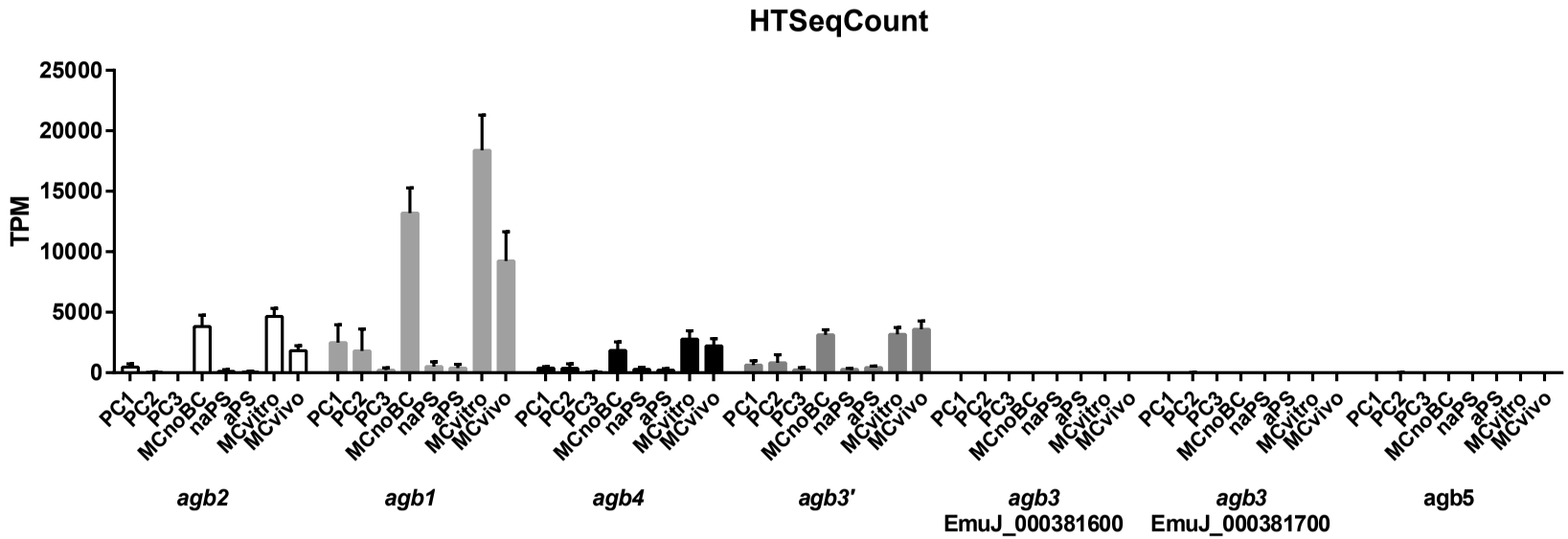
Table 5: **Highest expressed genes for each life cycle stage** according to calculation of expression levels with HTSeqCount (Anders et al., 2015). **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds; **PC_2d**: primary cells, 2-day-old; **PC_11d**: primary cells, 11 days old; **MC_noBC**: metacestode vesicles without brood capsules; **MC_LateBC**: metacestodes with brood capsules; **PS_noact**: non-activated protoscoleces; **PS_act**: activated protoscoleces; **EmPreAWDog**: pregravid adult; **EmAdult Gravide**: gravid adults; **EG_MCvivo**: *E. granulosus* cyst wall; **EG_naPS**: *E. granulosus* non-activated protoscoleces; **EG_PS_noact**: non-activated protoscoleces of *E. granulosus*.

5.1.3 Tapeworm specific antigen B is one of the highest expressed genes in *E. multilocularis* and *E. granulosus*

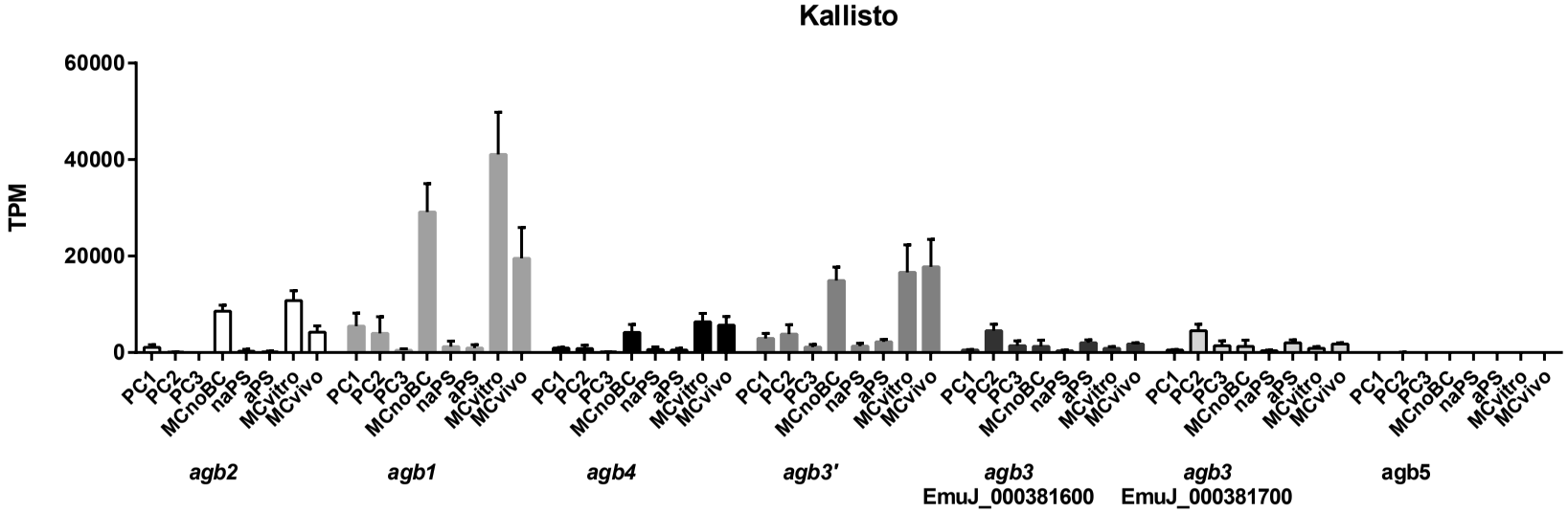
Among the highest expressed genes in metacestodes are two taenid specific gene families: the fatty acid binding proteins (FABPs) and the hydrophobic ligand bind-

ing proteins, to which antigen B (AgB) belongs. Flatworms and flukes are unable to synthesize fatty acids and cholesterol *de novo* and therefore take them up from the host (Frayha, 1971; Berriman et al., 2009) using fatty acid transporters as well as FABPs and probably the lipoprotein AgB. In *E. multilocularis* the antigen B gene family consists of at least seven clustering genes: EmuJ_000381100 (*agb2*), EmuJ_000381200 (*agb1*), EmuJ_000381400 (*agb4*), EmuJ_000381500 (*agb3'*), EmuJ-_000381600 (*agb3*), EmuJ_000381700 (*agb3*) and EmuJ_000381800 (*agb5*) with the two gene sequences for *agb3* being identical and *agb3'* being highly similar to them (Olson et al., 2012). According to expression levels with the HTSeqCount method (new datasets), *E. multilocularis agb1* was the highest expressed antigen B gene and showed especially high expression in the metacestode samples (MCnoBC, MCvitro and MCvivo) (see Figure 13 A). The considerably lower expressed *agb2*, *agb4* and *agb3'* were also highest expressed in the metacestode samples , suggesting a high importance of fatty acid uptake at the metacestode stage. According to the HTSeq-Count method, *agb3* and *agb5* were expressed very lowly or not at all (see Figure 13 A). However, the HTSeqCount method (with chosen parameters) discards multiple mapped reads, which can lead to lower expression levels for genes with (stretches of) identical sequences. Therefore expression levels were additionally estimated with Kallisto (Bray et al., 2016). The Kallisto results showed that both *agb3* copies were expressed, while *agb5* was only very lowly expressed (see Figure 13 B). Note, that it is not possible to conclude from the Kallisto data that both *agb3* genomic locations are actually transcribed. As the gene sequences of EmuJ_000381600 and EmuJ_000381700 are absolutely identical, Kallisto distributes all reads matching to both locations equally between them. However, in the old datasets (HTSeqCount method) both genomic locations showed expression at the adult stages (see Figure 14 A) indicating actual transcription of both genomic locations. What is more, in the Kallisto results expression in the adult stages for both *agb3* copies exceeded expression levels of all other antigen B gene copies in any life-cycle stage (see Figure 14 B), indicting that AgB3 has a specialized role in the adult worms (pregravid and gravid) where it was expressed in vast amounts.

In summary, *E. multilocularis* antigen B genes were extremely highly expressed in metacestodes and adult worms. *agb1*, *agb2*, *agb3'* and *agb4* showed especially high expression in metacestodes while both copies of *agb3* were particularly highly expressed in adults suggesting diversified functions of antigen B in different life cycle stages.

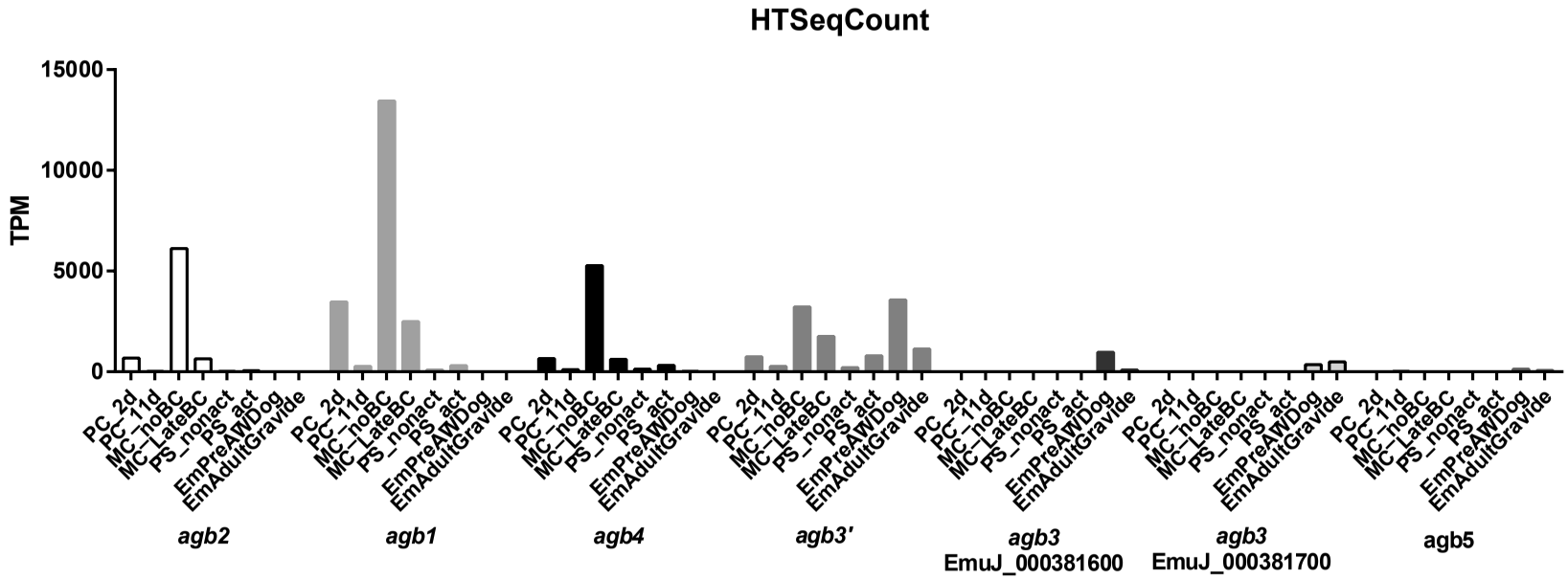


A

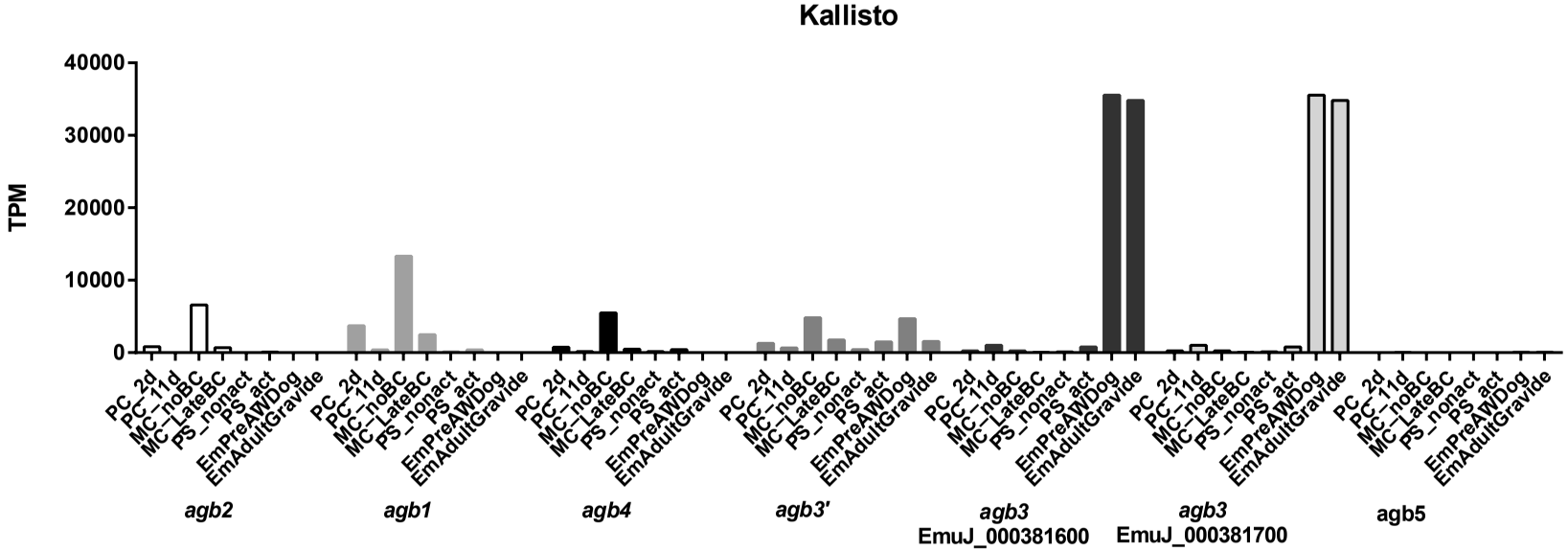


B

Figure 13: **Gene expression of *E. multilocularis agb* as calculated with HTSeqCount (Anders et al., 2015) (A) and Kallisto (Bray et al., 2016)(B):** Expression is shown in Transcripts Per Million (**TPM**). Error bars represent 1 SD. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds; ***agb***: *E. multilocularis* antigen B gene.



A



B

Figure 14: **Gene expression of *E. multilocularis agb* as calculated with HTSeqCount (Anders et al., 2015) (A) and Kallisto (Bray et al., 2016)(B) using the old datasets:** Expression is shown in Transcripts Per Million (**TPM**). **PC_2d**: primary cells, 2-day-old; **PC_11d**: primary cells, 11 days old; **MC_noBC**: metacestode vesicles without brood capsules; **MC_LateBC**: metacestodes with brood capsules; **PS_noact**: non-activated protoscoleces; **PS_act**: activated protoscoleces; **EmPreAWDog**: pregravid adult; **EmAdult Gravide**: gravid adults; ***agb***: *E. multilocularis* antigen B gene.

The situation in *E. granulosus* was highly similar: Seven antigen B gene copies clustered on scaffold 22, in synteny with the *E. multilocularis* antigen B gene cluster: EgrG_000381100 (*agb2*), EgrG_000381200 (*agb1*), EgrG_000381400 (*agb4*), EgrG_-000381500 (*agb3-2*), EgrG_000381600 (*agb3-3*), EgrG_000381700 (*agb3-1*), EgrG_-000381800 (*agb5*) with the three *agb3* gene copies being similar to each other (Olson et al., 2012). Highest expressed was *agb1*, followed by *agb4*. Both were higher expressed in metacestodes than in protoscoleces. *agb2* showed some expression in metacestodes and *agb3-3* in metacestodes and protoscoleces. The other gene copies were expressed at low levels (*agb3-2*, *agb5*) or not at all. Higher expression in metacestodes (compared to other life cycle stages) for both *E. granulosus* and *E. multilocularis agb1*, *agb2* and *agb4* suggests a specialized function of these antigen B isoforms in metacestodes.

Apart from the clusters, two additional antigen B related gene sequences were found in each *E. multilocularis* (EmuJ_000550500 and EmuJ_000525400) and *E. granulosus* (EgrG_000550500 and EgrG_000525400). While EmuJ_000550500 showed some expression in adults and EmuJ_000525400 was lowly expressed in activated protoscoleces, none of them were highly expressed.

5.1.4 From primary cells to metacestodes

5.1.4.1 Primary cell isolation from metacestodes increases expression of developmental genes

Primary cells were isolated from axenically cultivated metacestodes (MCnoBC) and grown for two days. Early stage primary cells (PC1) are highly enriched in germinative cells and therefore able to regenerate metacestode cysts. As the process of the primary cell isolation destroys the tegument, comparison of the transcriptome of metacestodes and primary cells can reveal genes that are expressed in the tegument. 1627 genes were significantly ($p_{adjust} < 0{,}05$) up-regulated in PC1 (reference MCnoBC) and 1832 downregulated. According to GO-enrichment analysis, up-regulated genes were especially involved in DNA replication and repair, DNA biosynthetic process, microtuble-based movement, telomere maintenance, DNA recombination, histone modifications, transcription and translation regulation, cell cycle process, intracellular signal transduction and amino acid transmembrane transport. Down-regulated genes (corresponds to higher expression in MCnoBC) included genes for various transport processes such as protein transport, vesicle-mediated transport, proton transport and carbohydrate derivative transport, which might be expressed in the tegument. Corresponding to the loss of tegument, the genes en-

coding the known markers for the metacestode tegument, alkaline phosphatase 1 and 2 (EmuJ_000393300, EmuJ_000393400) as well as mucin1 (EmuJ_000742900) (Koziol et al., 2014), were strongly downregulated in PC1 compared to MCnoBC. Other downregulated genes included dynein light chain genes and a Kunitz protease inhibitor gene (EmuJ_000419100) which is the orthologue of *E. granulosus* EgrG_000419100, encoding a chymotrypsin inhibitor (Flo et al., 2017), or were involved in cell communication, endocytosis, membrane fusion and cytoskeleton organization.

5.1.4.2 Development of metacestodes from primary cells is accompanied by up-regulation of genes involved in multicellular organism development

After approximately one week of cultivation primary cell cultures showed fused aggregates with central cavities (sample PC2). In comparison to early primary cells (sample PC1), 1551 genes were significantly ($p_{adjust} < 0{,}05$) up-regulated in PC2 and 1479 were significantly ($p_{adjust} < 0{,}05$) downregulated. Genes downregulated in PC2 (compared to PC1) were involved in translation, DNA repair, DNA replication, histone modification and metabolic processes. Up-regulated genes were involved in neuropeptide signaling pathway, G-protein coupled receptor signaling pathway and cell communication. Among the highest up-regulated genes were also several encoding transcription factors that regulate multicellular organism development such as the T box transcription factor tbx2 (EmuJ_001171900) which in humans is involved in transcriptional regulation of genes required for mesoderm differentiation (Wang et al., 2012; Martin et al., 2012) and the Forkhead box protein J1 A (EmuJ_000660100) which is the key transcription factor for motile ciliogenesis in *Danio rerio* (Yu et al., 2008).

Continued cultivation of primary cells led to formation of small metacestode vesicles (sample PC3). While PC2 and PC3 represent different degrees of metacestode development, the rates of development varied between biological replicates and within each culture. Therefore, samples for each PC2 and PC3 were not harvested at specific time points but rather at comparable degrees of development. This somewhat arbitrary approach lead to high heterogeneity between biological replicates (see Figure 16) and to overlapping clusters of PC2 and PC3 in the principal component analysis (see Figure 15). A direct comparison between PC2 and PC3 resulted therefore in a low number of differentially expressed genes and was not considered to be of (high) biological relevance. Rather, together with PC1 the samples could be used to observe gene expression tendencies during primary cell development from PC1 to PC2 to PC3.

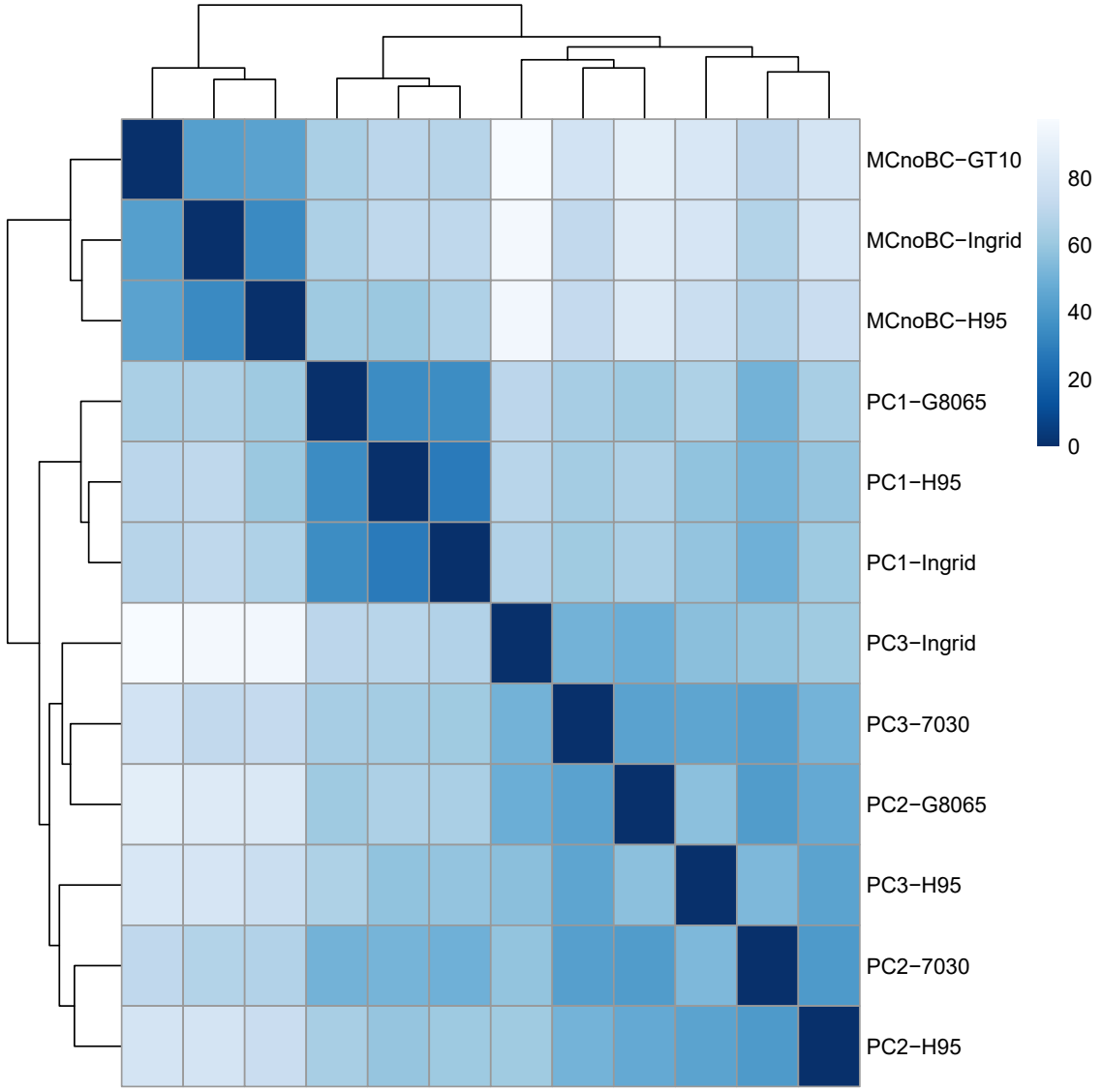


Figure 15: **Heatmap of sample-to-sample distances of primary cell and metacestode samples:** Samples are described as following: sample name - isolate. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules.

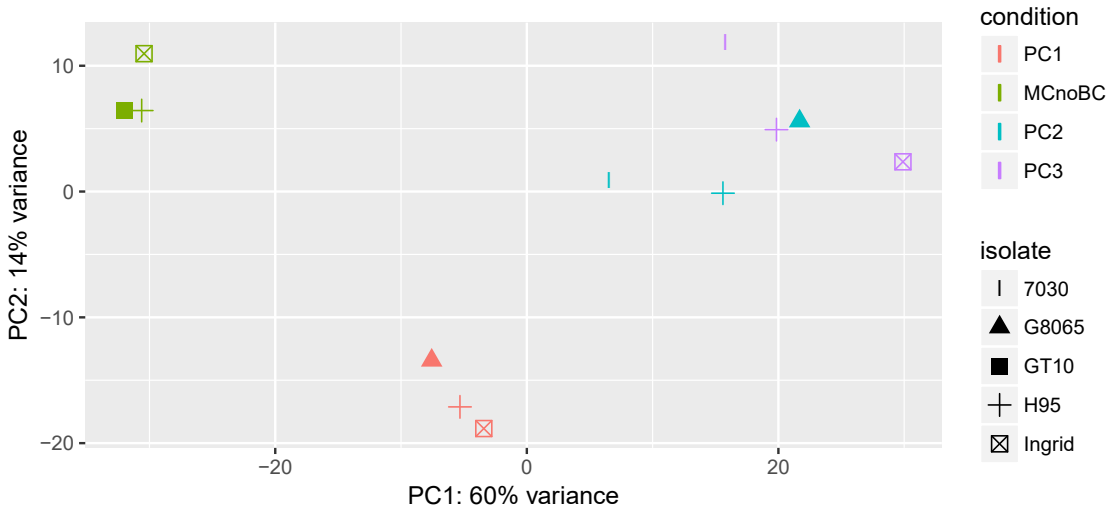


Figure 16: **Principal component analysis of primary cell and metacestode samples:** Conditions or life cycle stages are color coded and biological isolates are indicated by the shape. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules.

When comparing the metacestode culture (MCnoBC) to the furthest developed primary cell culture (sample PC3), 1282 significantly ($p_{adjust} < 0{,}05$) higher expressed genes and 1475 significantly ($p_{adjust} < 0{,}05$) lower expressed genes were detected. Genes higher expressed in MCnoBC were involved in cell communication, membrane lipid metabolic process, intracellular and transmembrane transport. The highest up-regulated genes in MCnoBC belonged to the antigen B family. Other highly up-regulated genes encoded the tegument markers alkaline phosphatase 1 and 2. Genes that were significantly ($p_{adjust} < 0{,}05$) higher expressed in PC3 (compared to MCnoBC) included genes involved in neuropeptide signaling pathway, multicellular organism development, axonemal dynein complex assembly and Wnt signaling pathway.

5.1.4.3 *wnt* and *sfrp* genes are higher expressed in late primary cells than in metacestodes

In the principal component analysis the metacestode samples MCnoBC were more closely related to early primary cell samples (PC1) than to further developed primary cell samples (PC2 and PC3) (see Figure 16). One of the differences between late primary cells and metacestodes was the higher expression of genes of the Wnt signaling pathway in late primary cells. Therefore, gene expression profiles of *wnt* and *sfrp* genes as well as markers of anterior and posterior specification were analyzed. According to Koziol et al. (2016), *E. multilocularis sfrp*, *sfl*, *six3/6*, *wnt2* and

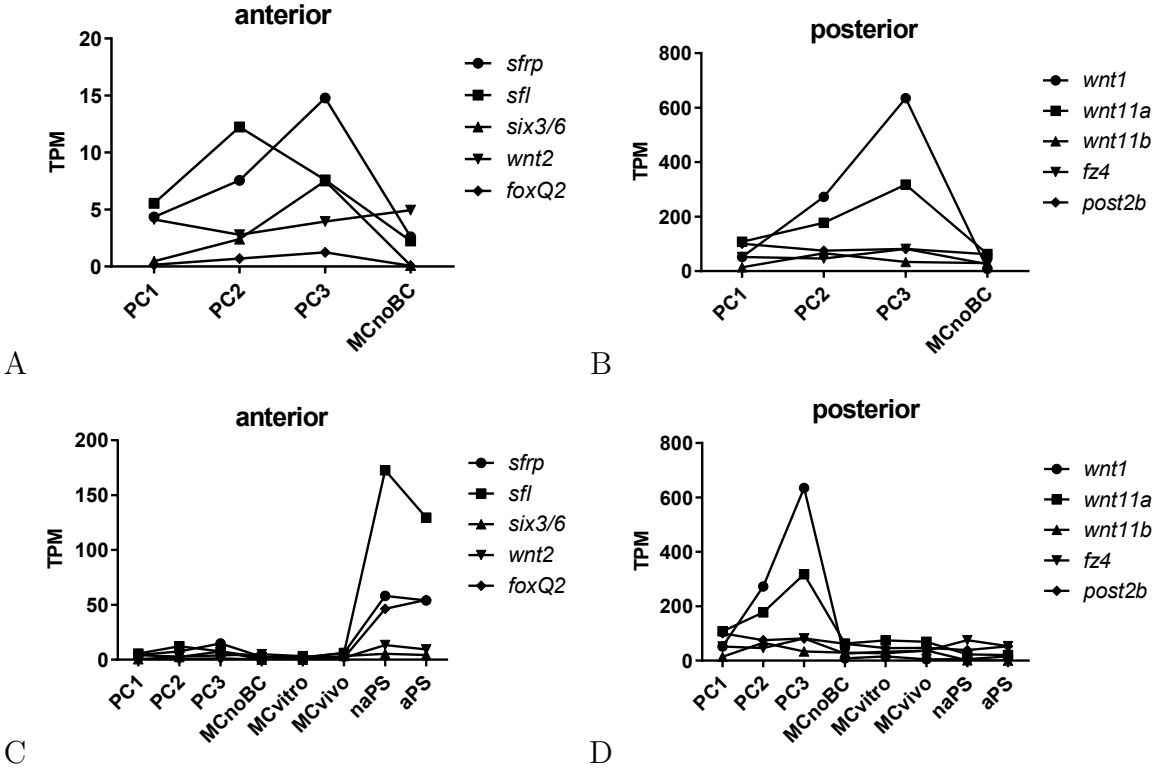


Figure 17: **Gene expression profiles of genes expressed in anterior (A, C) and posterior (B, D) regions in *E. multilocularis*:** Mean expression is shown in Transcripts Per Million (**TPM**). **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces.

foxQ2 are expressed in anterior regions of brood capsules and protoscoleces while *wnt1*, *wnt11a*, *wnt11b*, *fz4* and *post2b* are expressed in posterior regions of brood capsules and protoscoleces and/or in the germinal layer of the metacestode. The "posterior" *wnt* genes *wnt1* and *wnt11a* as well as the Wnt inhibitor genes *sfrp* and *sfl* were up-regulated in the later primary cell stages (PC2 and PC3) compared to early primary cells (PC1) or metacestodes (MCnoBC) (see Figure 17 A and B). The expression of markers for anterior and posterior specification was relatively low and constant in early and late primary cells with only a slight increase of *six3/6* in PC3. In comparison to expression in protoscoleces, expression of Wnt inhibitor genes in primary cells was minor, while expression of *wnt1* and *wnt11a* in primary cells by far exceeded expression in any other larval stage (see Figure 17 C and D).

5.1.4.4 Knock-down of β-catenin leads to anteriorization of primary cells

In order to learn more about the effect of β-catenin in primary cells, knock-down of β-catenin was performed using RNAi (by Raphaël Duvoisin, repeated by Ruth Herrmann). Phenotypically, this resulted in the so-called "red dot phenotype" with primary cells developing massive amounts of red colored central cavities and relatively few vesicles (Raphaël Duvoisin and Ruth Herrmann, personal communication). On a gene expression level, primary cells with knock-down of β-catenin showed significantly ($p_{adjust} < 0{,}05$) lower expression of genes involved in cell adhesion, signal transduction, cytoskeleton organization and transport than controls. Significantly ($p_{adjust} < 0{,}05$) higher expressed genes were involved in DNA replication, system development, cytokinesis, multicellular organism development, Wnt signaling pathway and histone modifications. Among the significantly higher expressed genes were also the Wnt inhibitor *sfrp* ($p_{adjust}=4{,}8e^{-16}$) and the "posterior" *wnt11a* ($p_{adjust}=0{,}01$) (see Figure 18). *sfl*, *six3/6*, *wnt2* and *foxQ2*, which are expressed in anterior regions of brood capsules and protoscoleces (Koziol et al., 2016), were also higher expressed though not significantly. Significantly lower expressed were the "posterior" *wnt11b* ($p_{adjust}=4{,}6e^{-13}$) and the marker for posterior specification *post2b* ($p_{adjust}=4{,}0e^{-39}$) (see Figure 18). Taken together, this suggests an overall anteriorization of β-catenin siRNA treated primary cells.

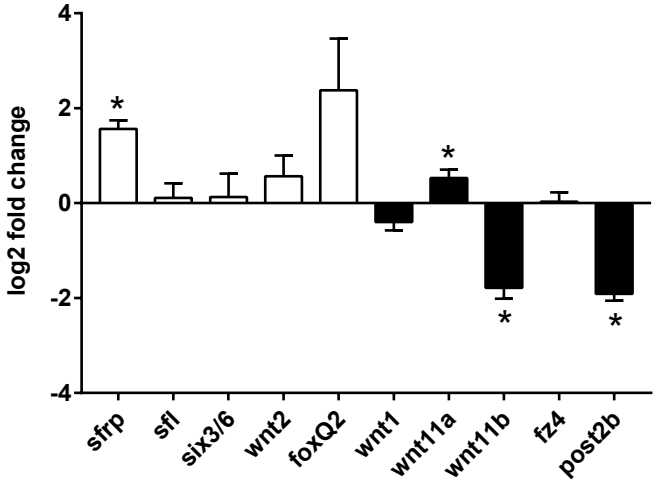


Figure 18: **Gene expression changes after treatment with β-catenin siRNA:** Shown are log2 fold changes (maximum likelihood) in expression (compared to control) of genes expressed in anterior (white) and posterior (black) regions. Genes with significantly different expression ($p_{adjust} < 0{,}05$) are marked with stars. Error bars represent 1 SEM.

5.1.5 Both metacestode *in vitro* culture conditions are similarly close to the *in vivo* sample

Cultivation of metacestodes *in vitro* is possible under aerob conditions together with RH$^-$ feeder cells (sample MCvitro) or without feeder cells under anaerobe conditions and with reducing agents (sample MCnoBC) (Spiliotis and Brehm, 2009). In order to find out which condition better represents *in vivo* conditions (MCvivo), gene expression analyses were performed. The principal component analysis showed three separate clusters, one for each sample, with each MCvitro and MCnoBC being at similar distances to MCvivo and slightly closer to each other (see Figure 19). The heatmap of the sample-to-sample distances depicted even clearer that the two *in vitro* samples were closer related to each other than to the *in vivo* sample. Interestingly, *in vitro* samples of the isolate H95 showed smaller distances to the *in vivo* samples, which were all of the isolate H95, than *in vitro* samples of other isolates (see Figure 20), indicating that also the isolate plays an important role in gene expression.

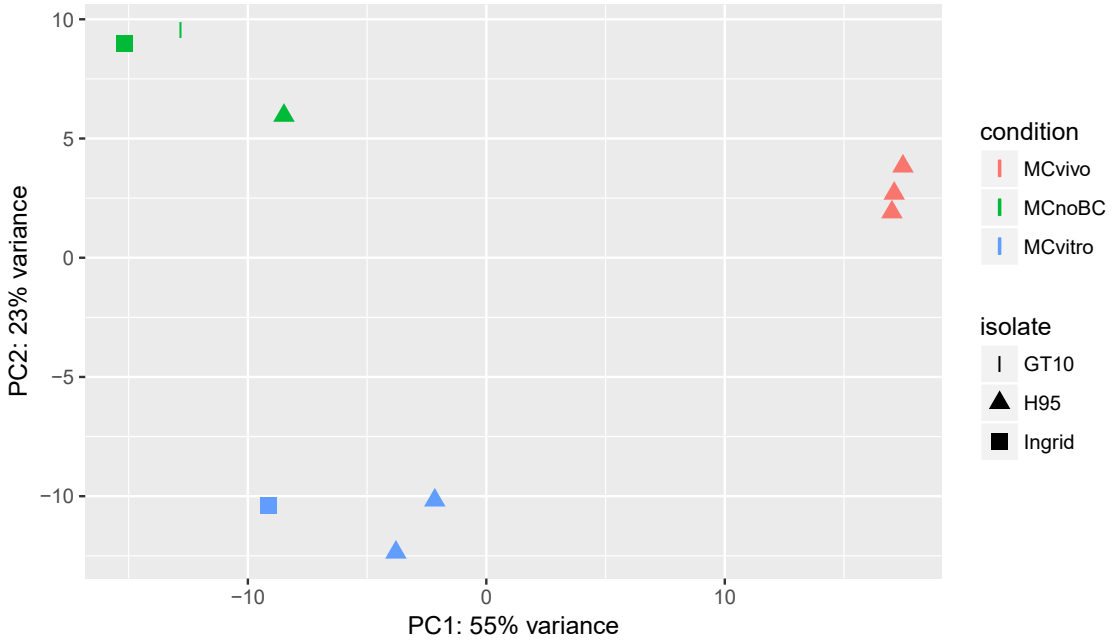


Figure 19: **Principal component analysis of metacestode samples:** Conditions are color coded and isolates are indicated by the shape. **MCnoBC**: metacestode without brood capsules; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds.

Genes significantly ($p_{adjust} < 0{,}05$) higher expressed *in vivo* (MCvivo) than *in vitro* (MCnoBC and MCvitro) were involved in DNA replication, mismatch repair, methylation, telomere maintenance and cytokinesis. Higher expression of genes for DNA

synthesis and proliferation in MCvivo suggests that *in vitro* conditions are not able to support proliferation and development to the same extent as *in vivo* conditions.

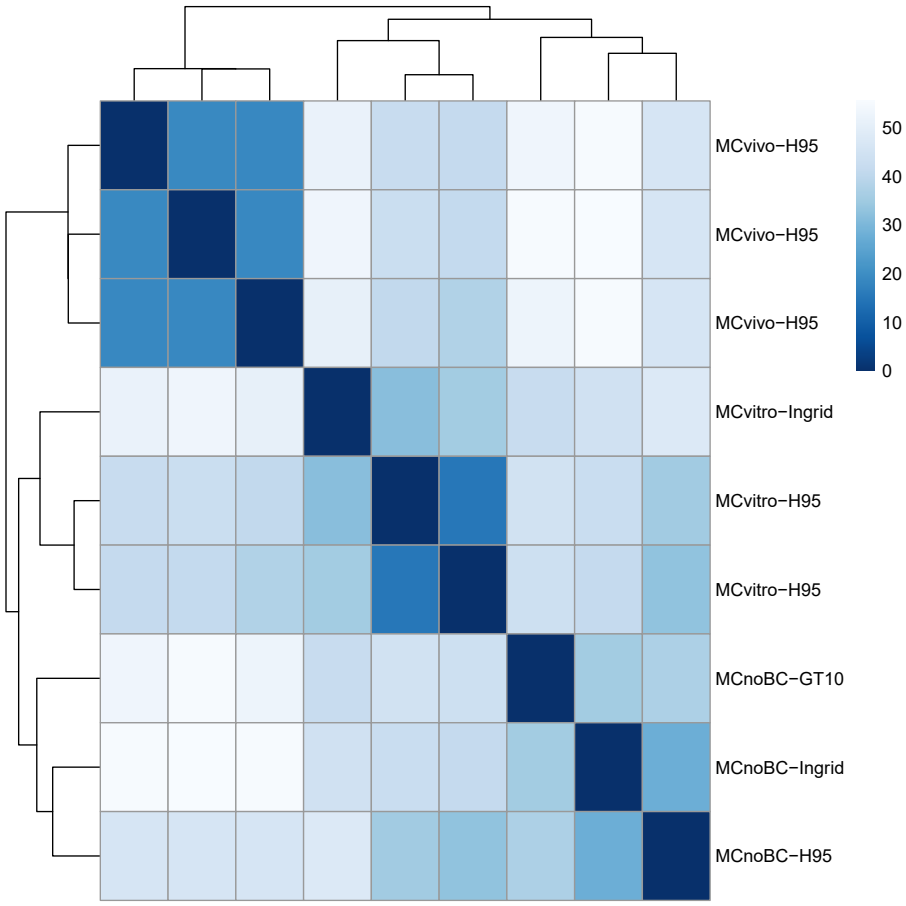


Figure 20: **Heatmap of sample-to-sample distances of metacestode samples:** Samples are described as following: sample name - isolate. **MCnoBC**: metacestode without brood capsules; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds.

Among the highest up-regulated genes in MCvivo (compared to MCnoBC and MCvitro) was the oncosphere specific antigen *em95* (EmuJ_000368620, designated "EG95" in WormBaseParaSite). The other three designated "EG95" genes were not expressed at relevant levels. *em95* (EmuJ_000368620) showed increasing expression levels during primary cell development (from PC1 to PC2 to PC3) indicating an important role in this developmental process. In metacestodes *em95* was expressed in the sample MCvivo but not in MCvitro and MCanaerob (see Figure 21). Exclusive expression in MCvivo might be explained by continuous budding and formation of new vesicles or infiltration of host tissue *in vivo* which does not occur *in vitro*.

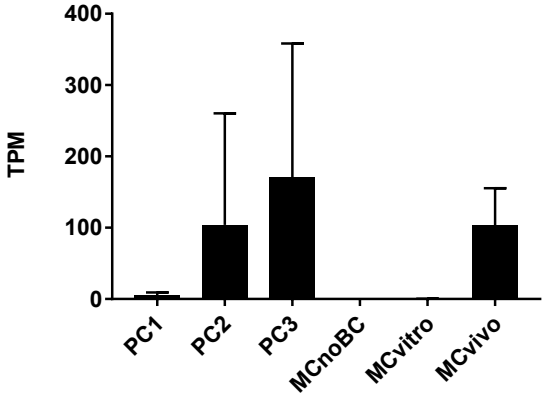


Figure 21: **Gene expression of *em95*:** Expression is shown in Transcripts Per Million (**TPM**). Error bars represent 1 SD. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds.

5.1.6 Wnt inhibitor genes are higher expressed in non-activated protoscoleces than in metacestodes

During later stages of infection protoscoleces develop in brood capsules within metacestode vesicles Koziol et al. (2016). To compare gene expression of metacestodes and non-activated protoscoleces (naPS), metacestode samples needed to be free of contaminating protoscoleces. The naPS sample was therefore compared to the *in vitro* cultivated metacestodes without brood capsules (and thus also without protoscoleces) (sample MCnoBC). As protoscoleces were isolated from *in vivo* material, the sample naPS was also compared to *in vivo* metacestodes (MCvivo). Biological replicates of MCvivo were all of the isolate H95, which has mostly lost the ability to develop protoscoleces (Tsai et al., 2013).

In both comparisons over 5000 genes were differentially expressed. One of the genes with considerably higher expression in both metacestodes samples compared to the naPS sample was the metacestode tegument marker alkaline phosphatase 1 (EmuJ_000393300) (Koziol et al., 2014) which showed almost no expression in non-activated protoscoleces. Alkaline phosphatase 2 (EmuJ_000393400) and mucin 1 (EmuJ_000742900), which are further markers for the metacestode tegument (Koziol et al., 2014), were also considerably higher expressed in metacestodes compared to non-activated protoscoleces. In contrast, expression of alkaline phosphatase 3 (EmuJ_000752700), which has been shown to be only expressed in protoscoleces Koziol et al. (2014), was significantly ($p_{adjust} < 0{,}05$) higher expressed in non-activated protoscoleces than in both metacestode samples (almost no expression there), indi-

cating that the metacestode samples were free of contamination with protoscoleces.

Genes significantly ($p_{adjust} < 0{,}05$) higher expressed in non-activated protoscoleces than in metacestodes were involved in ion transport and signaling processes, such as the Wnt signaling pathway and the neuropeptide signaling pathway. The Wnt inhibitors *sfl* and *sfrp*, which are involved in anterior development Koziol et al. (2016), showed substantially higher expression in non-activated protoscoleces than in metacestodes (see Figure 17). According to the old transcriptome datasets, expression of *sfrp* and *sfl* was already somewhat increased in metacestode vesicles with brood capsules compared to metacestodes without brood capsules before reaching highest expression in the non-activated protoscolex sample (see Figure 22).

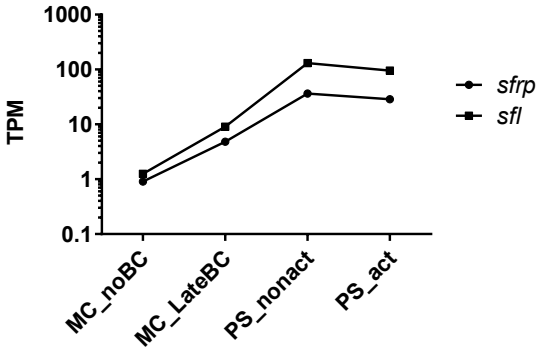


Figure 22: **Gene expression of the Wnt inhibitors *sfrp* and *sfl*:** Expression according to old RNA-Seq datasets is shown in Transcripts Per Million (**TPM**). **MC_noBC**: metacestode vesicles without brood capsules; **MC_LateBC**: metacestodes with brood capsules; **PS_noact**: non-activated protoscoleces; **PS_act**: activated protoscoleces.

Kunitz inhibitors were also significantly ($p_{adjust} < 0{,}05$) higher expressed in non-activated protoscoleces than in metacestodes: EmuJ_000419100 and EmuJ_000534800 which were orthologues to the *E. granulosus* chymotrypsin inhibitor encoding genes EgrG_000419100 and EgrG_000534900 (Flo et al., 2017), respectively, as well as EmuJ_000549400, EmuJ_000548800 and EmuJ_000534700, which were orthologues to the *E. granulosus* trypsin inhibitor encoding genes EgrG_000549400, EgrG_000548800 and EgrG_000534700 Flo et al. (2017), respectively.

Genes significantly higher expressed in metacestodes than in non-activated protoscoleces were involved in metabolic processes like the tricarboxylic acid cycle and the carbohydrate metabolic process. Significantly higher expressed in metacestodes were also genes encoding antigen B (see also section 5.1.3) and FABPs.

5.1.7 Activation of protoscoleces is accompanied by higher expression of transporter encoding genes

Infection of definitive hosts occurs through digestion of parasite metacestode material containing non-activated protoscoleces. During the digestion process protoscoleces are freed from their surrounding brood capsules and activated. This process can be mimicked in the laboratory by treatment with pepsin and bile salts (Fernandez et al., 2002). Comparison of activated with non-activated protoscoleces resulted in a relatively small number of differentially expressed genes. On a gene expression level, non-activated and activated protoscoleces were closely related, while protoscoleces and metacestodes showed greater distances in the principal component analysis (see Figure 23). Interestingly, activation of protoscoleces lead to significant ($p_{adjust} < 0{,}05$) downregulation of only 419 genes but up-regulation of 817 genes. This asymmetry is likely due to the "dormant" stage of non-activated protoscoleces.

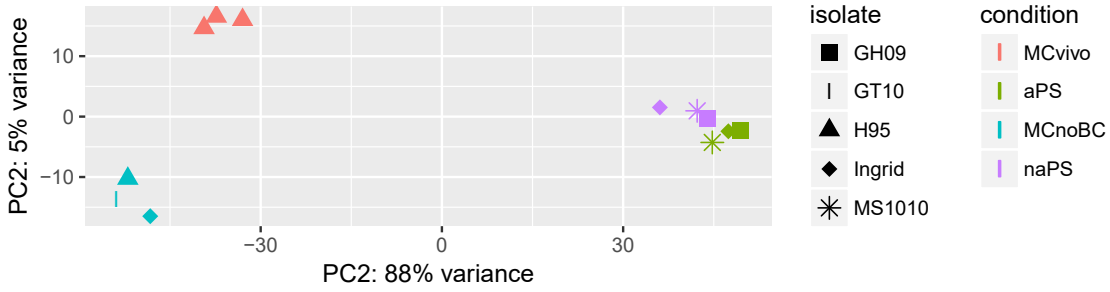


Figure 23: **Principal component analysis of metacestode and protoscolex samples:** Life-cycle stages and conditions are color coded and isolates are indicated by the shape. **MCnoBC**: metacestode without brood capsules; **MCvivo**: metacestodes extracted from infected jirds; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces.

Significantly ($p_{adjust} < 0{,}05$) up-regulated genes in activated protoscoleces compared to non-activated protoscoleces were especially involved in transport and metabolic processes, particularly lipid metabolic processes. Strongly up-regulated were genes coding for amino acid, glucose and nucleoside transporters.Other strongly up-regulated genes coded for homologues of the *Taenia* antigen GP50 (Hancock et al., 2004; Gomez-Puerta et al., 2019). Slightly downregulated genes in activated protoscoleces compared to metacestodes included genes involved in mismatch repair, cell cycle regulation and DNA replication.

5.1.8 Adults express different gene sets than larval stages

Inside the small intestine of definitive hosts, protoscoleces differentiate into a strobilar direction and develop into sexually mature adults (Smyth, 1968; Thompson et al., 1990). The principal component analysis showed the two available adult samples (1 biological replicate each) rather removed from the larval stage samples, which formed two clusters (see Figure 24). One of the larval stage clusters contained the protoscolex samples while the other included the metacestode and primary cell samples. The gravid adult sample was farthest from the larval stage samples whereas the pregravid adult sample was located in between, indicating that gene expression of adults, especially gravid adults, vastly differs from gene expression of larval stages. Genes highly expressed in adults included genes of the antigen B (see section 5.1.3), FABP and Kunitz inhibitor families.

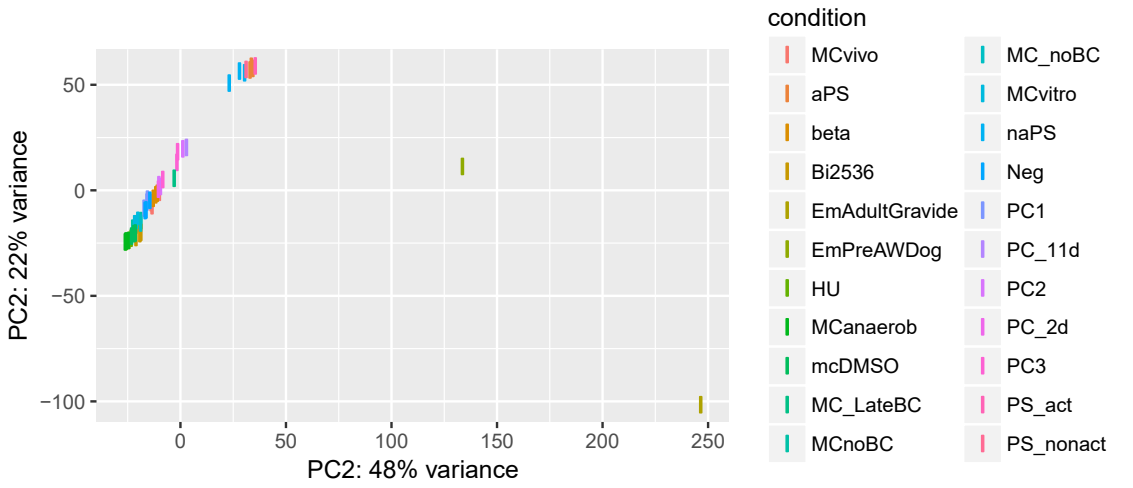


Figure 24: **Principal component analysis of all analyzed *E. multilocularis* samples:** Conditions or life cycle stages are color coded. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces; **MCvitro**: metacestodes in aerobic culture; **MCvivo**: metacestodes extracted from infected jirds; **PC_2d**: primary cells, 2-day-old; **PC_11d**: primary cells, 11 days old; **MC_noBC**: metacestode vesicles without brood capsules; **MC_LateBC**: metacestodes with brood capsules; **PS_noact**: non-activated protoscoleces; **PS_act**: activated protoscoleces; **EmPreAWDog**: pregravid adult; **EmAdult Gravide**: gravid adults; **HU**: metacestodes treated with hydroxyurea; **MCanaerob**: control for HU treatment; **Bi2536**: metacestodes treated with Bi2536; **mcDMSO**: control for Bi2536 treatment; **beta**: primary cells treated with β-catenin siRNA; **Neg**: control for β-catenin siRNA treatment.

5.1.9 Identification of a stem cell-specific gene set in *E. multilocularis*

To identify genes that might be specifically expressed in germinative cells, two different approaches were used: depletion of germinative cells and enrichment of germinative cells. Germinative cells can be specifically depleted by treatment with hydroxyurea (HU) or the polo-like kinase inhibitor Bi2536 (Koziol et al., 2014; Schubert et al., 2014). Comparison of gene expression in metacestodes treated with HU or Bi2536 compared to their respective controls can therefore reveal genes that are potentially germinative cell specifically expressed. However, it cannot be completely excluded that other cells are also affected by treatment with HU or Bi2536, especially the direct progeny of germinative cells. Treatment with HU was carried out for 7 days, treatment with Bi2536 for three weeks (Koziol et al., 2014; Schubert et al., 2014). It hence follows that also the direct progeny of germinative cells might be affected by these treatments. Therefore the second approach was enrichment of germinative cells. 2-day-old primary cell cultures are known to be highly enriched in germinative cells (62 - 83 % of all cells) compared to metacestodes (32 - 55 %) (Koziol et al., 2014). Higher expression in 2-day-old primary cell cultures (PC1) than in later primary cells cultures (PC2) and metacestodes (MCnoBC) was therefore included as criterium for germinative cell specifically expressed genes. Genes were considered to be potentially specifically expressed in germinative cells if they fulfilled all three of the following criteria: their gene expression was significantly ($p_{adjust} < 0{,}05$) downregulated after treatment with (i) HU and (ii) Bi2536 compared to their respective controls and they were higher expressed in (iii) PC1 than in PC2 and MCnoBC.

1853 genes were significantly ($p_{adjust} < 0{,}05$) downregulated in metacestode samples treated with HU compared to control samples while treatment of metacestodes with Bi2536 resulted in 2662 significantly ($p_{adjust} < 0{,}05$) downregulated genes. 1212 genes were significantly ($p_{adjust} < 0{,}05$) downregulated in both treatments. In the germinative cell enriched sample PC1 4851 genes were higher expressed than in PC2 and 5220 higher than in MCnoBC with 2969 genes being higher expressed in PC1 in both comparisons. In total 717 genes fulfilled all criteria and were considered candidates for germinative cell specific expression(see Figure 25).

GO-enrichment analysis of putative germinative cell specific genes showed that most of the GO terms with the highest significance were related to DNA replication and repair or cell cycle regulation, as would be expected (see table 6). Interestingly, the most significant GO-term was translation indicating that not only DNA but also protein synthesis plays an important role in germinative cells.

GO ID	GO Term	Annotated genes	Significant genes	Expected significant genes	p-value
GO:0006412	translation	242	61	21,47	1,10E-19
GO:0006260	DNA replication	71	23	6,3	1,70E-05
GO:0032508	DNA duplex unwinding	32	11	2,84	5,40E-05
GO:0006281	DNA repair	75	23	6,65	5,40E-05
GO:0006270	DNA replication initiation	8	5	0,71	0,00024
GO:0007093	mitotic cell cycle checkpoint	5	4	0,44	0,00028
GO:0030071	regulation of mitotic metaphase/anaphase...	7	4	0,62	0,00172
GO:0006298	mismatch repair	7	4	0,62	0,00172
GO:0000723	telomere maintenance	7	4	0,62	0,00172
GO:0000910	cytokinesis	5	3	0,44	0,00605
GO:0007268	chemical synaptic transmission	11	4	0,98	0,01127
GO:0006302	double-strand break repair	6	3	0,53	0,01131
GO:0007156	homophilic cell adhesion via plasma memb...	40	8	3,55	0,02191
GO:0035235	ionotropic glutamate receptor signaling ...	13	4	1,15	0,02286
GO:0006414	translational elongation	33	7	2,93	0,0233
GO:0034220	ion transmembrane transport	200	20	17,74	0,02363
GO:0000075	cell cycle checkpoint	9	6	0,8	0,04105
GO:0007049	cell cycle	86	20	7,63	0,04346

Table 6: **Gene Ontology (GO) analysis of genes that are potentially germinative cell specific**

To evaluate specificity of germinative cell depletion by treatment with HU or Bi2536, gene expression of known markers was analyzed. The tegument markers alkaline phosphatase 1 (EmuJ_000393300), alkaline phosphatase 2 (EmuJ_000393400) and mucin 1 (EmuJ_000742900) (Koziol et al., 2014) were not downregulated after either treatment. In contrast, *nos1* (EmuJ_000861500.1), *nos2* (EmuJ_000606200.1)

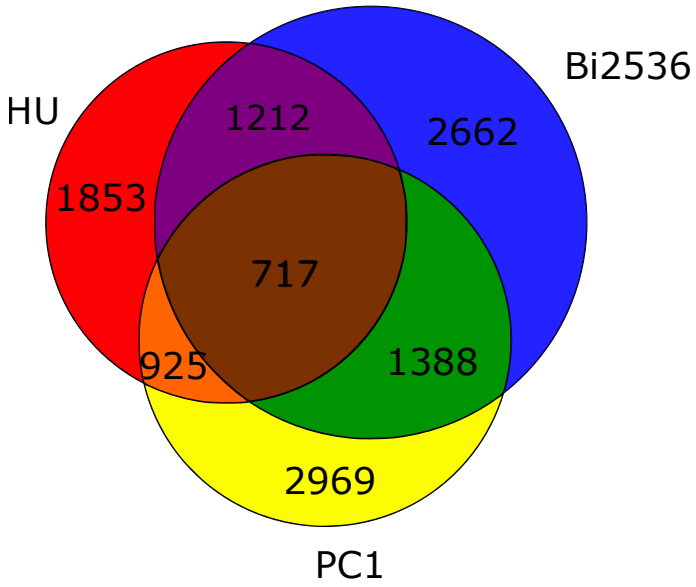


Figure 25: **Venn diagram of potentially germinative cell specific genes:** Shown are the number of genes significantly ($p_{adjust} < 0{,}05$) lower expressed in HU treated samples compared to their controls (red), significantly lower expressed in Bi2536 treated samples compared to their controls (blue) and higher expressed in PC1 compared to PC2 and MCnoBC (yellow) and the numbers of overlaps. **HU**: metacestodes treated with hydroxyurea; **Bi2536**: metacestodes treated with Bi2536; **PC1**: primary cells, stage 1.

and the polo-like kinase *plk1* (EmuJ_000471700.1), which are expressed in subpopulations of germinative cells (Koziol et al., 2014; Schubert et al., 2014), were strongly and significantly ($p_{adjust} < 1e^{-16}$) downregulated (see table 7). In this context, lower expression of *nos2* in PC1 than in MCnoBC was surprising but might have occurred due to its expression in only a very small proportion of germinative cells (Koziol et al., 2014).

Furthermore, of the putative germinative cell specific genes, the non capsid protein 1 genes were among the most downregulated genes after HU and Bi2536 treatment (for detailed analysis see section 5.4.1). Other strongly downregulated genes encoded transcription factors: the DNA binding protein inhibitor ID 4 (EmuJ_000457300.1) and MYB (EmuJ_000909600.1) that determine cell proliferation and differentiation (Roy and Zhuang, 2018; Oh and Reddy, 1999; Prouse and Campbell, 2012); the GATA binding factor 2 (EmuJ_000215900.1) that regulates maintenance and expansion of hematopoietic stem cells (Fujiwara, 2017); the Basic helix loop helix dimerization region bHLH (EmuJ_000451500.1) that have widespread roles in development and differentiation (Robinson and Lopes, 2000); the Neurogenic differentiation factor 1 (EmuJ_000344700.1) that controls differentiation of neurons and

metabolism (Shalabi et al., 2013) and the Bhlh factor math6 (EmuJ_000098000.1) that is required for early embryonic development and involved in tissue-specific differentiation processes (Lynn et al., 2008).

	PC1 vs MCnoBC	PC1 vs PC2	HU	Bi2536	Gene ID
alp1	**-5,2**	0,5	**1,4**	**0,9**	EmuJ_000393300.1
alp2	**-3,9**	**-0,6**	**0,5**	**1,0**	EmuJ_000393400.1
muc1	**-4,5**	**-1,0**	0,4	**0,9**	EmuJ_000742900.1
nos1	**0,8**	**0,5**	**-3,2**	**-3,0**	EmuJ_000861500.1
nos2	**-0,8**	0,6	**-2,8**	**-2,8**	EmuJ_000606200.1
plk1	**1,2**	**0,6**	**-2,7**	**-3,0**	EmuJ_000471700.1

Table 7: **Gene expression of known markers**: Shown are log2fold changes of the comparisons PC1 compared to MCnoBC, PC1 to PC2, HU to control and Bi2536 to control. Significant ($p_{adjust} < 0{,}05$) changes are printed in fat. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **MCnoBC**: metacestode without brood capsules; **HU**: metacestodes treated with hydroxyurea; **Bi2536**: metacestodes treated with Bi2536.

Besides transcription factors, several genes encoding kinases involved in proliferation and cell-cycle regulation were found among the most downregulated genes after HU or Bi2536 treatment of the putative germinative cell specific genes. These kinases included cyclin-dependent kinases (such as EmuJ_000786500.1 and CDKD1.1), the aurora kinase A (EmuJ_001059700.1) and the polo-like kinase 1 (EmuJ_000471700.1). Also strongly downregulated was a mitogen-activated protein kinase *emSSY* (EmuJ_000139200.1) which is a paralog of the *E. multilocularis mpk2* p38 MAPK (Gelmedin et al., 2008). In contrast to the universally conserved Thr-Gly-Tyr motif in the kinase activation loop of p38 MAPK (Han et al., 1998), *emSSY* has an unique Ser-Ser-Tyr motif (Gelmedin, 2008). Atypical biochemical properties together with germinative cell specific gene expression make *emSSY* an interesting target for future research.

To find out if HU and Bi2536 treatment influenced gene expression differently, genes only affected by one of the treatments were analyzed. Genes significantly ($p_{adjust}<0{,}05$) downregulated after HU treatment compared to controls and with constant or higher expression in Bi2536 treated samples compared to controls were considered to be HU specifically downregulated. According to GO-enrichment analysis HU specifically downregulated genes were involved in carbohydrate derivative and organophosphate catabolic processes, acetyl-CoA and carbohydrate metabolic

processes, proteolysis and multicellular organism processes. Among the HU specifically downregulated genes were also *wnt4* (EmuJ_000211300.1) and *wnt11a* (EmuJ_000907500.1). Many cells expressing *wnt11a* are muscle cells while there is no significant expression in germinative cells (Koziol et al., 2016). Consequently, also muscle cells might be affected by HU treatment.

On the other hand, genes significantly (p_{adjust}<0,05) lower expressed in Bi2536 treated metacestodes compared to controls, and constant or higher expression in HU treated samples compared to controls, were considered to be Bi2536 specifically downregulated. Many of these genes were involved in methylation, metabolic processes, mRNA splicing, transcription and transcription regulation, DNA repair and cell division. Down-regulation of genes for DNA repair and cell division was expected for germinative cell depleted samples. Therefore, this suggests that treatment with Bi2536 stronger depletes germinative cells than treatment with HU.

5.2 Indication of natural antisense transcripts in *E. multilocularis*

To get a first impression of natural antisense transcripts in *E. multilocularis*, reverse spliced reads of the available RNA-Seq data were analyzed. As an unstranded protocol was used for RNA-Seq, it was only possible to ascertain the orientation of reads with canonical splice sites. Spliced reads with canonical splice sites reverse to the orientation of the corresponding gene prediction were designated reverse spliced reads. Many genes with numerous reverse spliced reads were declared as "expressed protein" without known homologues. Of the genes with annotated function, the following had the most reverse spliced reads: Ubiquitin ribosomal protein L40, Tapeworm specific antigen B, Polyubiquitin, ATP synthase F0 subunit 6, Annexin, Serine:threonine protein kinase pak, Metal transporter Nramp1, Protein AHNAK2 and L-lactate dehydrogenase. Further genes with reverse spliced reads included gene copies encoding the densovirus NS1 and VP proteins, the tryptophan hydroxylase and transcripts antisense to ribosomal RNA protein. These results indicate that natural antisense transcripts are present in *E. multilocularis* and might play a role in regulation of gene expression.

5.3 5-Aza-2'-deoxycytidine inhibits development of metacestode vesicles in primary cell cultures

Another known mechanism of gene regulation and silencing is DNA methylation. The essential DNA methylation machinery elements (here DNMT2 and MBD2/3) are present in platyhelminthes, including *E. multilocularis* (Geyer et al., 2013). The gene encoding the DNMT2 in *E. multilocularis* was cloned and sequenced (LR585068.1). The *E. multilocularis dnmt2* was comprised of 9 exons and 8 introns and encoded a protein of 361 amino acids. Analysis of the protein sequence with SMART 8.0 showed a DNA-methylase domain. The *E. multilocularis* DNMT2 showed significant homologies to *S. mansoni* (61 % similarities) and human (49% similarities) DNA-methyltransferase 2. A detailed analysis of conserved motifs and target recognition domains of flatworm DNMT2s and MBD2/3s, including the predicted sequences of *E. multilocularis*, can be found in Geyer et al. (2013).

The analyzed transcriptome data showed that the *E. multilocularis dnmt2* (EmuJ_001185500.1) and *mbd2/3* (EmuJ_001033700.1) genes are expressed in all larval stages with their lowest expression in the metacestode stage (see Figure 26 A and B), indicating that DNA methylation plays a role in all larval stages. Interestingly, in the transcriptome data *dnmt2* showed significantly (p=0,0008) lower gene expression in HU treated metacestodes (isolate Ingrid) compared to controls while treatment of metacestodes (isolate Ingrid) with Bi2536 caused only a very slight, not significant, decrease in gene expression (also see Figure 26 C). *mbd2/3* expression was not significantly affected by either treatment (also see Figure 26 D), suggesting that *mbd2/3* is not specifically expressed in germinative cells. To find out if *dnmt2* might be specifically expressed in germinative cells, quantitative RT-PCR based on cDNA of HU treated metacestode vesicles and corresponding controls was performed. Quantitative RT-PCR showed almost equal relative quantities for HU treated and untreated control samples without significant difference (see Figure 27), indicating that *dnmt2* is not specifically expressed in stem cells. To analyze this further, a preliminary WMISH experiment for *dnmt2* was performed. Unfortunately, many vesicles (including all sense controls) were not suitable for analysis due to extremely high background. In the remaining three vesicles, expression of *dnmt2* was detected in 38 % of all cells (see Figure 28). 17 % of *dnmt2*-positive cells were also stained with EdU. Of the 6 % EdU-positive cells, 82 % were also positive for *dnmt2*, suggesting *dnmt2* is expressed in most of the proliferating cells but also in other cells.

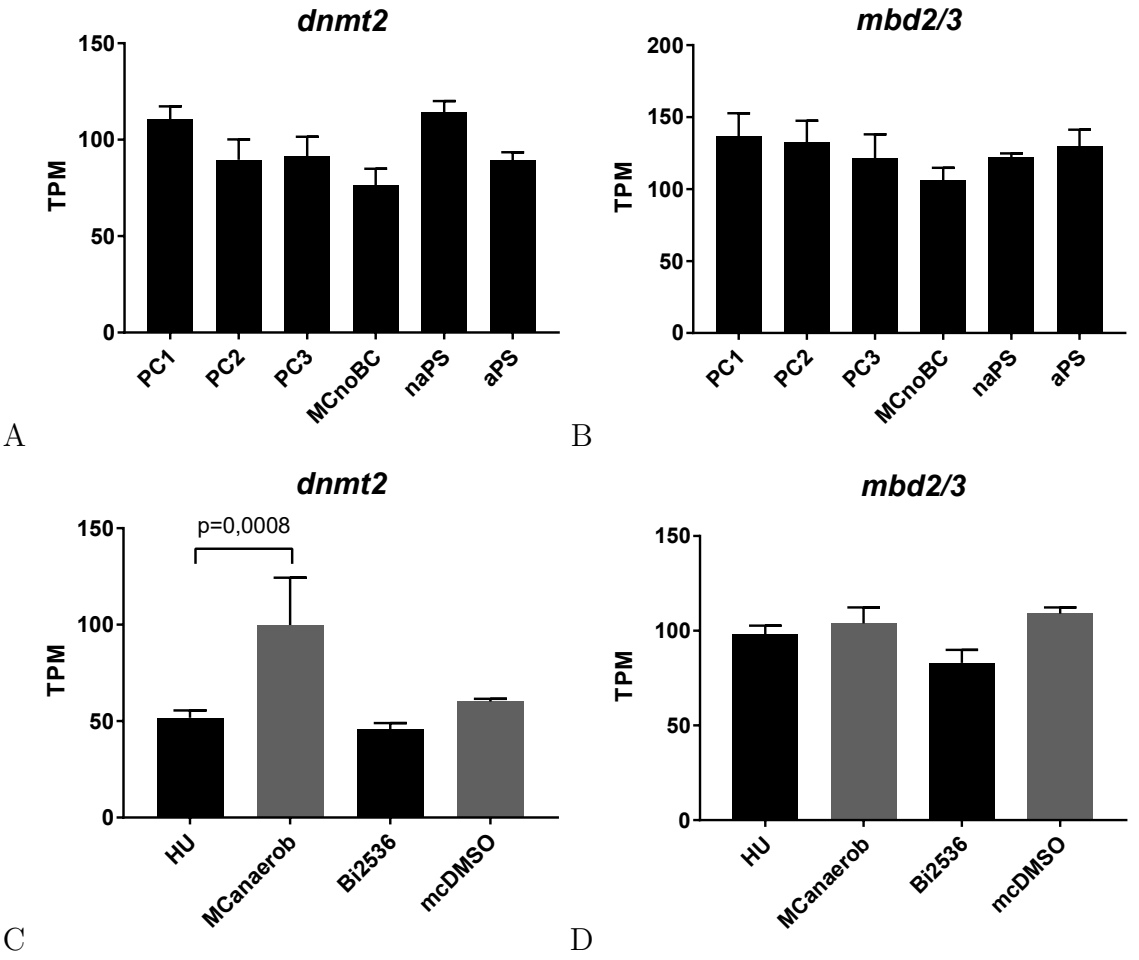


Figure 26: **Gene expression of *E. multilocularis dnmt2*** (**A,C**) **and *mbd2/3*** (**B,D**): Expression is shown in Transcripts Per Million (**TPM**). Error bars represent 1 SD. Depicted p-value is the adjusted p-value of the DESeq2 analysis. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces; **HU**: metacestodes treated with hydroxyurea; **MCanaerob**: control for HU treatment; **Bi2536**: metacestodes treated with Bi2536; **mcDMSO**: control for Bi2536 treatment.

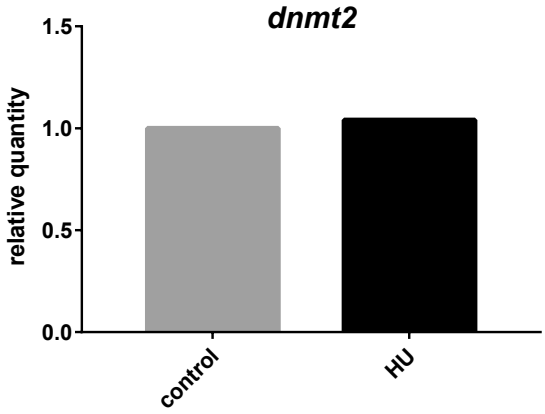


Figure 27: **Gene expression of *E. multilocularis dnmt2* after treatment with hydroxyurea:** qPCR was performed for *E. multilocularis dnmt2* with cDNA from metacestode vesicles depleted of germinative cells by treatment with hydroxyurea (HU). Experiment was performed with three technical and three biological replicates (isolates GH09, H95, J2012). Shown is the mean relative gene expression obtained by fgStatistics.

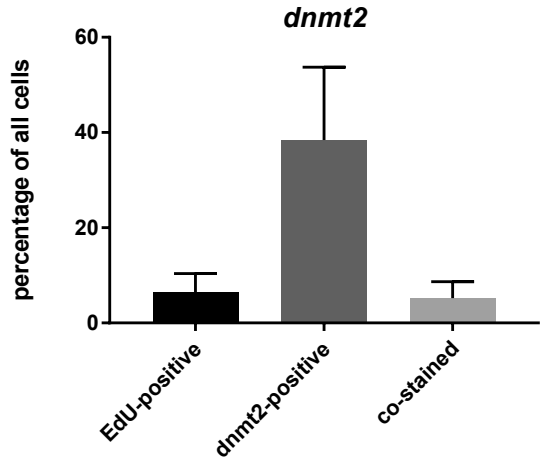


Figure 28: **Gene expression of *E. multilocularis dnmt2* in metacestode vesicles:** Shown are preliminary results of one WMISH experiment based on three metacestode vesicles (2 images each). Mean percentages of EdU-positive cells, *dnmt2*-positive cells and co-stained cells were calculated. Error bar is 1 SD.

To functionally study the role of DNA methylation in *E. multilocularis*, the DNMT inhibitor 5-aza-2'-deoxycytidine (100 µM, 10 µM, 1 µM and 0,1 µM) was applied to primary cell cultures. In a preliminary experiment, short time cultivation with 5-aza-2'-deoxycytidine for two days had no effect on primary cell viability (see Figure 29). A preliminary experiment with prolonged incubation, however, showed differences regarding vesicle formation. Untreated primary cells started to develop metacestode

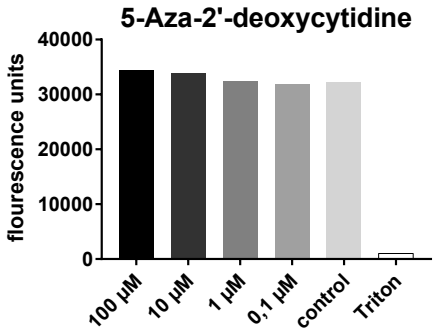


Figure 29: **Effect of 5-aza-2'-deoxycytidine on primary cell viability:**: Viability was measured using a resazurin assay after two day treatment with indicated concentrations of 5-aza-2'-deoxycytidine. Triton was used as cytotoxic control. Mean fluorescence units are shown. Experiment was performed with three technical replicates.

vesicles from day 10 day on (see Figure 30). From day 14 on cultures treated with 5-aza-2'-deoxycytidine, even at the lowest concentration of 0,1 µM, contained less vesicles than control cultures (see Figure 30). These results suggest that 5-aza-2'-deoxycytidine inhibits development of metacestode vesicles in primary cell cultures.

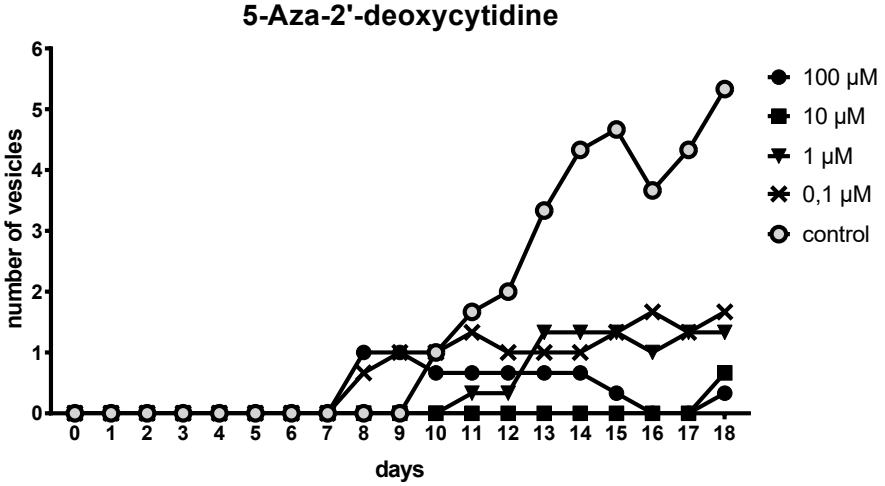


Figure 30: **Effect of 5-aza-2'-deoxycytidine on metacestode formation in primary cell cultures:**: Time course of vesicle development from primary cells in the presence of 5-aza-2'-deoxycytidine. Shown are the numbers of developed metacestode vesicles. Experiment was performed with three technical replicates.

5.4 Densovirus sites in flatworm genomes

5.4.1 Densovirus sequences are present in the *E. multilocularis* genomes

One of the genes that was detected by the transcriptome data analysis to be specifically expressed in germinative cells was EmuJ_000388600, which is annotated as "non-capsid protein 1". The protein sequence of EmuJ_000388600 was 392 amino acids long and showed highest similarities (24 % identical and 34 % similarities) to insect densovirus NS1 proteins and was therefore designated EmuDNV-NS1. The downstream ORF EmuJ_000388500 coded for a protein with weak (below cut-off) homologies to a viral capsid protein of the *Pea enation* mosaic virus and was designated EmuDNV-VP. Domain analysis revealed a PPV_E1_C domain (Papillomavirus helicase E1 C-terminus) in the EmuJ_000388600 EmuDNV-NS1 protein whereas no domains were detected in the protein EmuJ_000388500 EmuDNV-VP. Taken together, these results indicate the identification of a densovirus locus.

With blast searches a total of 26 *Emu*DNV-NS1 gene version were identified in the *E. multilocularis* genome. For detailed analysis of densoviral sequences, all putative *Emu*DNV-NS1 gene sequences were curated individually and translated into amino acid sequences. The results of the blastp searches with these sequences suggested that all sequences encoded full-length or truncated versions of one protein, the designated EmuDNV-NS1. The longest EmuDNV-NS1 version was 431 amino acids long and was encoded by loci on the contigs 0155 (EmuJ_000368400), 0221 (EmuJ_000048100), 0266 (EmuJ_000369300 and EmuJ_000368900) and 0868 (EmuJ_000007400) (see Figure 31). Some other gene versions, including EmuJ_000388600, appeared to be full length gene copies but contained frameshift mutations. Frameshift mutations were mostly detected in gene regions encoding the N-terminus of NS1 (6 gene versions). Only one gene sequence (EmuJ_000329200) contained a frameshift mutation in a gene region encoding the C-terminal region of NS1. The other gene versions were truncated at the 5' end (11 sequences), the 3' end (2 sequences) or at both sides (1 sequences) (see Figure 31).

Blastp searches of EmuDNV-NS1 sequences against the nr database (organism: viruses) revealed high homologies (26 %/43 %) to the NS1 protein of the infectious hypodermal and hematopoietic necrosis virus (IHHNV) isolated from *Penaeus stylirostris* (blue shrimps) (Shike et al., 2000). In accordance with these results, blast searches of EmuDNV-NS1 versions against the SWISSPROT database at NCBI showed the highest homologies to the NS1 protein of the *Aedes* densonucleosis virus (23 % identical, 42 % similar residues) and the *Aedes albopictus* denso-

Figure 31: **Schematic overview of densoviral genes in *E. multilocularis*:** Complete gene versions are represented by long black arrows. Crossed out black arrows indicate frameshift mutations. Gene versions with truncated 3' ends are depicted as short black arrows, gene versions with truncated 5' ends as white arrows. Corresponding gene IDs from WormBaseParaSite are shown above the arrows, when available. **P1:** potential promoter for *Emu*DNV-NS1 gene; **P2:** potential promoter for *Emu*DNVA-VP gene; **X:** no promoter; **ITR:** ITR sequences; **(ITR):** Remnants of ITR sequences; **His:** near histone cluster.

virus (24 %/43 %). Furthermore, domain analysis with pfam showed a PPV_E1_C domain and an overlapping Parvo_NS1 domain at the C-terminus of all complete EmuDNV-NS1 versions. The Parvo_NS1 domain constitutes the helicase domain of the parvoviral NS1 protein which is essential for viral DNA replication (Iseki et al., 2005), suggesting again a parvoviral/densoviral origin of EmuDNV-NS1. Alignment of the EmuDNV-NS1 sequence with the NS1 sequences of *Aedes* densonucleosis virus and *Aedes albopictus* densovirus revealed high homologies within the PPV_E1_C domain while the rest of the EmuDNV-NS1 sequence was less conserved and contained long stretches of gaps in the alignment (see Figure 32), indicating conservation in functional regions and a shortened and condensed EmuDNV-NS1 protein compared to other densovirus NS1 proteins.

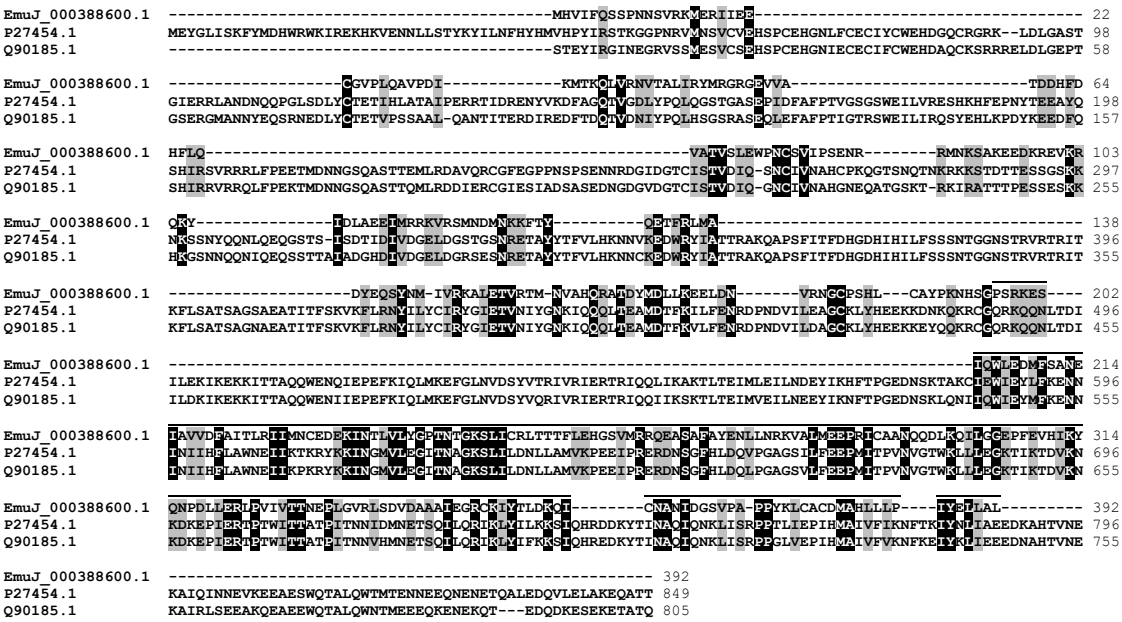


Figure 32: **Comparison of densoviral NS1 sequences**: Alignment of amino acid sequences was performed with MUSCLE (Edgar, 2004*a*,*b*). Identical amino acids are depicted white on black background, biochemically similar amino acids black on grey background. The line indicates the location of the predicted PPV_E1_C domain in EmuDNV-NS1 EmuJ_000388600.1. P27454.1: *Aedes* densonucleosis virus NS1; Q90185.1: *Aedes albopictus* parvovirus NS1.

Six-frame translations of regions near *Emu*DNV-NS1 genes led to the detection of downstream ORFs which corresponded to the *Emu*DNV-VP EmuJ_000388500. The ORFs were located 67 nt downstream of many *Emu*DNV-NS1 gene versions and encoded a 321 amino acid protein. To find further *Emu*DNV-VP gene versions, blast searches were performed which led to the detection of 26 *Emu*DNV-VP versions. 13 of them appeared to be complete (see Figure 31). Frameshift mutations were

detected in two gene versions (EmuJ_000388500 and EmuJ_000368800), which were otherwise very similar to those encoding the 321 aa protein. Interestingly, 9 of 11 truncated *Emu*DNV-VP gene versions were missing their 3' end. Together with missing 5' ends of *Emu*DNV-NS1 gene versions, this resulted in a pattern that left the middle part of the virus genome intact while often both ends were truncated in the genome assembly (see Figure 31). Protein structure analysis with complete EmuDNV-VP versions did not lead to significant results, suggesting that the VP is less conserved than the NS1 or that less information is available.

To identify ITR sequences, flanking regions of *Emu*DNV-NS1 were analyzed with the program "einverted". Several repeat sequences of different length were detected. The longest identified ITR sequence was 370 nt long (165 nt stem with 89 % matches and a 37 nt loop) and located 37 nt downstream of the *Emu*DNV-VP gene version EmuJ_000329300. The best conserved repeat structures were detected upstream of *Emu*DNV-NS1 EmuJ_000048300 and downstream of *Emu*DNV-VP EmuJ_000048400 on contig 0221 with 100 % matches within and between the ITR sequences (lengths 228 and 229 nt, stem 95 and 96 nt, loop 37 nt, respectively). According to blast results all identified ITR sequences were based on one sequence with slight variations. In addition to ITRs with maintained repeat structure, blast searches detected remnants of ITR sequences near several densoviral genes (see Figure 31).

To learn more about densovirus sites, neighboring genes and scaffold/contig locations of densoviral sequences were analyzed. Densovirus sequences were localized in isolated or repeat rich regions of the genome (see Figure 33 as examples), often in a head-to-tail concatemer configuration. Densovirus loci were sometimes located

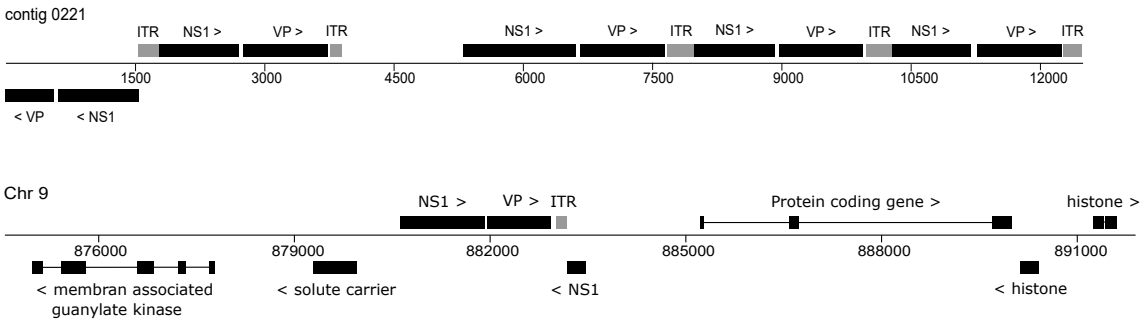


Figure 33: **Densovirus sites:** Schematic overview of densovirus loci on contig 221 and chromosome 9. Numbers are genome positions in bp. Genes are depicted as black boxes (exons) and lines (introns), ITRs as grey boxes. Arrow heads show gene orientation. NS1: *Emu*DNV-NS1; VP: *Emu*DNV-VP; ITR: inverted terminal repeats.

within or near histone clusters (see Figure 33). To evaluate the correctness of the genome assembly at densovirus sites, PCR analysis was performed for selected sites (gene versions EmuJ_000013900, EmuJ_002195700, EmuJ_000388600 and EmuJ_-000329200). Using one primer annealing to one *Emu*DNV-NS1 gene version and the other to a predicted neighboring tapeworm gene, densovirus loci were amplified and sequenced (see Figure 34). All sequences had highest similarities (> 99,5 % identities) to the sequence of their respective genomic location thus corroborating the correctness of the genome assembly at these sites.

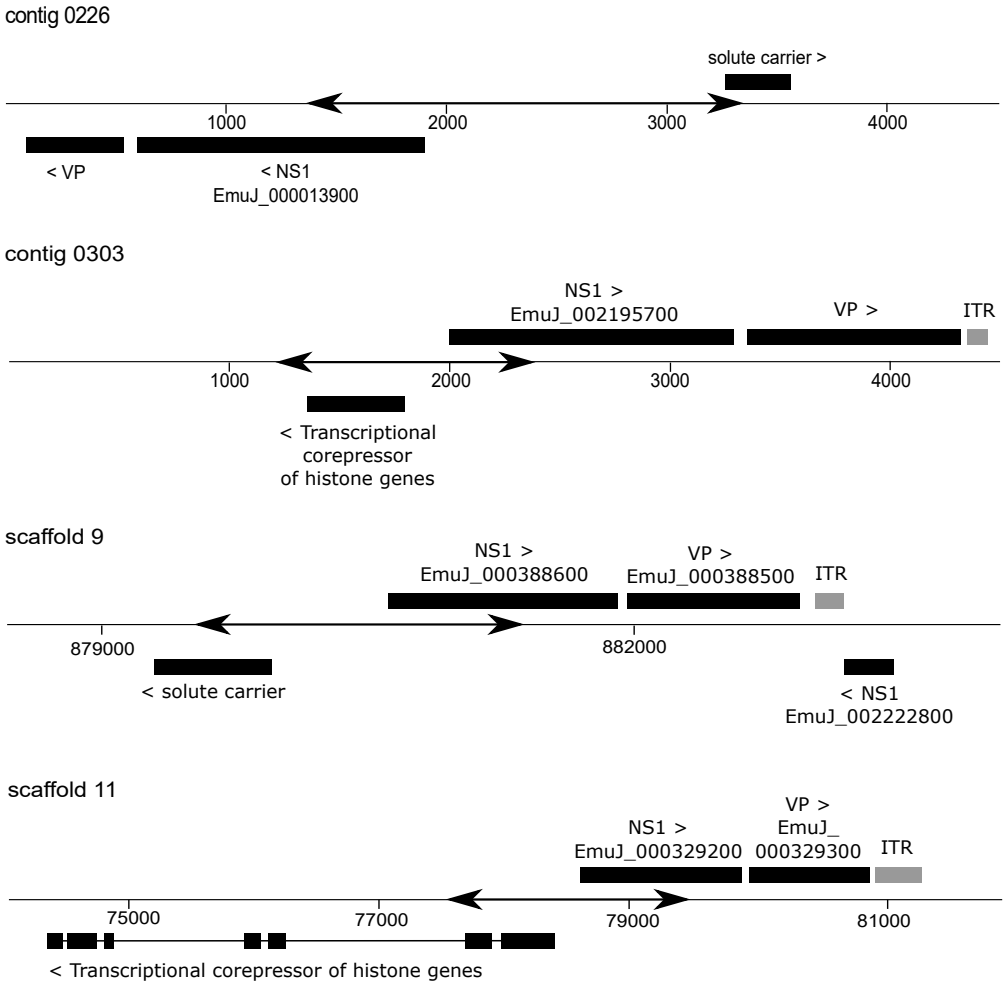


Figure 34: **Densovirus sites in the *E. multilocularis* genome:** Schematic representation of the complete contigs 0226 and 0303 as well as parts of the scaffold 09 and 11. Numbers are genome positions in bp. Genes are depicted as black boxes (exons) and lines (introns), ITRs as grey boxes. Arrow heads show gene orientation. The sequenced regions are marked by double headed arrows. NS1: *Emu*DNV-NS1; VP: *Emu*DNV-VP; ITR: inverted terminal repeats.

5.4.2 Densoviral genes are expressed in *E. multilocularis*

When searching for similar promoter elements as described for the *Penaeus stylirostris* densovirus (Rai et al., 2011), TATA-boxes were detected upstream of all *Emu*DNV-NS1 and *Emu*DNV-VP gene versions with complete 5' ends. The TATA-boxes of the 14 *Emu*DNV-NS1 gene versions with complete 5' ends were located 53 or 54 nt upstream of their putative start codons. Additionally, putative initiation of transcription sequences (CATTCA) and DPE-like boxes (downstream promoter element) were detected upstream of the putative start codons of *Emu*DNV-NS1 gene versions (see Figure 35 as examples). The TATA-boxes of the 24 *Emu*DNV-VP gene versions with complete 5' ends were located 34 or 35 nt upstream of their putative start codons, which corresponds to 28 or 29 nt upstream of their putative initiation of transcription sequences (CACATT) (see Figure 35 as examples).

Interestingly, only three *Emu*DNV-NS1 and two *Emu*DNV-VP gene copies were actually expressed (> 10 TPMs) according to transcriptome data. All expressed gene versions showed their highest expression in early primary cells (see Figure 36). Highest overall expression levels were observed for EmuJ_002222800. The expressed versions of *Emu*DNV-NS1 contained frameshift mutations (EmuJ_000034800 and EmuJ_000388600) or were truncated at the 3' end (EmuJ_002222800). Two of the expressed *Emu*DNV-NS1 genes had corresponding *Emu*DNV-VP gene copies that were both expressed. One of the expressed *Emu*DNV-VP gene versions was complete (EmuJ_000034900), while the other contained a frameshift mutation (EmuJ_000388500). No differences regarding possible promoter elements, completeness of gene copies, genomic sites or ITR sequences were found between expressed and not expressed gene versions.

To confirm gene expression, *Emu*DNV-NS1 was amplified with primers annealing without mismatches to the gene versions EmuJ_000034800, EmuJ_000388600, EmuJ_002195700 and EmuJ_000329200 using cDNA of 2-day old primary cells as template. After ligation and transformation, eight clones were chosen for sequencing. Six of the resulting sequences were identical, the other two sequences differed in only one nucleotide and were therefore considered variations of the same sequence. The sequence was deposited at the EMBL Nucleotide Sequence Database under the accession number LR029140. The 1103 nt long sequence showed 99,8 % identities (2 mismatches) to the *Emu*DNV-NS1 gene version EmuJ_000388600 and at least 16 mismatches to other *Emu*DNV-NS1 versions. It is therefore likely that the sequence originated from the gene version EmuJ_000388600, verifying gene expression of *Emu*DNV-NS1 in *E. multilocularis*.

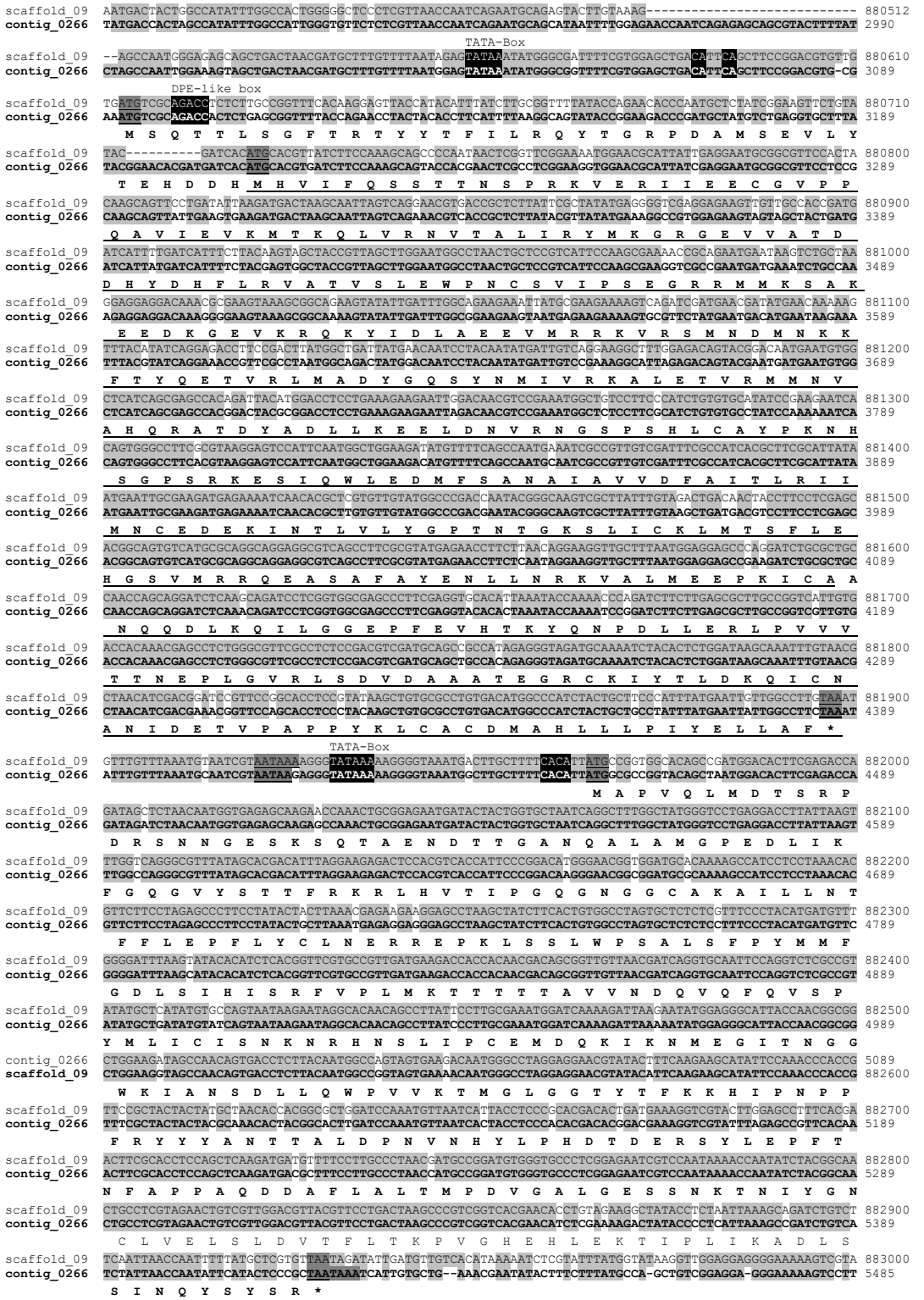


Figure 35: **Densovirus sequences in the *E. multilocularis* genome**: Shown are densoviral sequences of two genomic loci aligned with MUSCLE (Edgar, 2004*a*,*b*). Genomic positions are indicated. Identical nucleotides are depicted on light gray background. Putative start codons, stop codons and polyadenylation signals are underlined on dark grey background. Putative initiation of transcription (CA elements), TATA-boxes and DPE-like boxes are depicted white on black background. Amino acid sequences encoded by possible ORFs on contig 266 are shown below the alignment. The sequence arising from an alternative start codon is underlined.

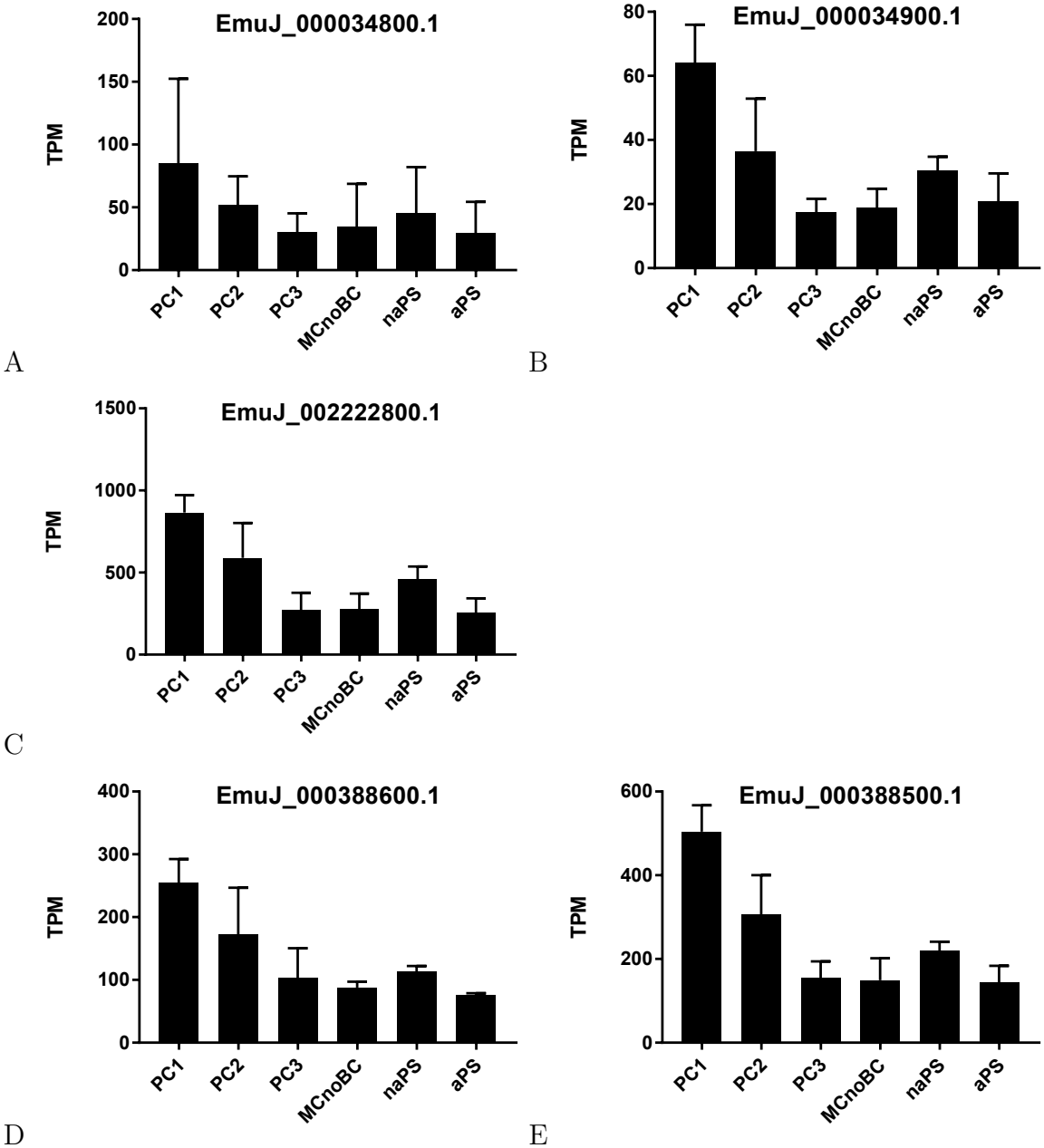


Figure 36: **Expression of *Emu*DNV-NS1 (A,C,D) and corresponding *Emu*DNV-VP (B,E) gene versions**: Expression is shown in Transcripts Per kilobase of exon per Million transcripts mapped (**TPM**). Error bars represent 1 SD. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces.

5.4.3 Densoviral genes are specifically expressed in *E. multilocularis* germinative cells

Transcriptome data of expressed *Emu*DNV-NS1 and *Emu*DNV-VP gene versions showed higher expression in 2-day-old primary cells, which are enriched in germinative cells (Koziol et al., 2014), than in metacestodes (see Figure 36). Additionally, compared to untreated controls gene expression levels of the expressed *Emu*DNV-NS1 and *Emu*DNV-VP gene versions were significantly and considerably reduced by treatment with HU or Bi2536, which depletes germinative cells (Koziol et al., 2014; Schubert et al., 2014) (see Figure 38). Such a transcription profile is typical for genes that are specifically expressed in germinative cells, indicating that expression of *Emu*DNV-NS1 and *Emu*DNV-VP is germinative cell specific. In quantitative RT-PCR experiments, gene expression levels of *Emu*DNV-NS1 (primers annealing to the gene versions EmuJ_000034800, EmuJ_000388600, EmuJ_000329200) were significantly lower in metacestodes depleted of germinative cell with HU than in untreated controls (see Figure 37), confirming the information of the RNA-Seq analysis. Taken together, these results strongly suggest that *E. multilocularis* *Emu*DNV-NS1 and *Emu*DNV-VP genes are specifically expressed in germinative cells.

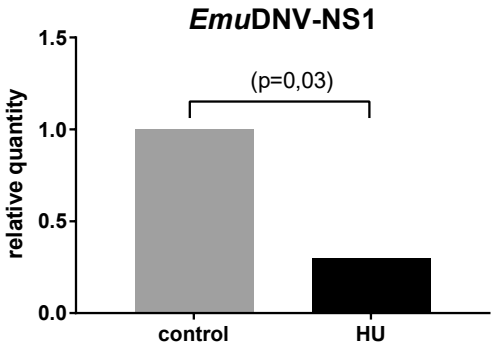


Figure 37: **Relative gene expression of *Emu*DNV-NS1 in HU treated metacestodes:** Quantitative RT-PCR for *Emu*DNV-NS1 was carried out using cDNA from metacestode vesicles treated with hydroxyurea (HU) for germinative cell depletion and untreated controls. Experiment was performed with three technical and three biological replicates. Shown are the mean relative gene expression and the p-value obtained by fgStatistics.

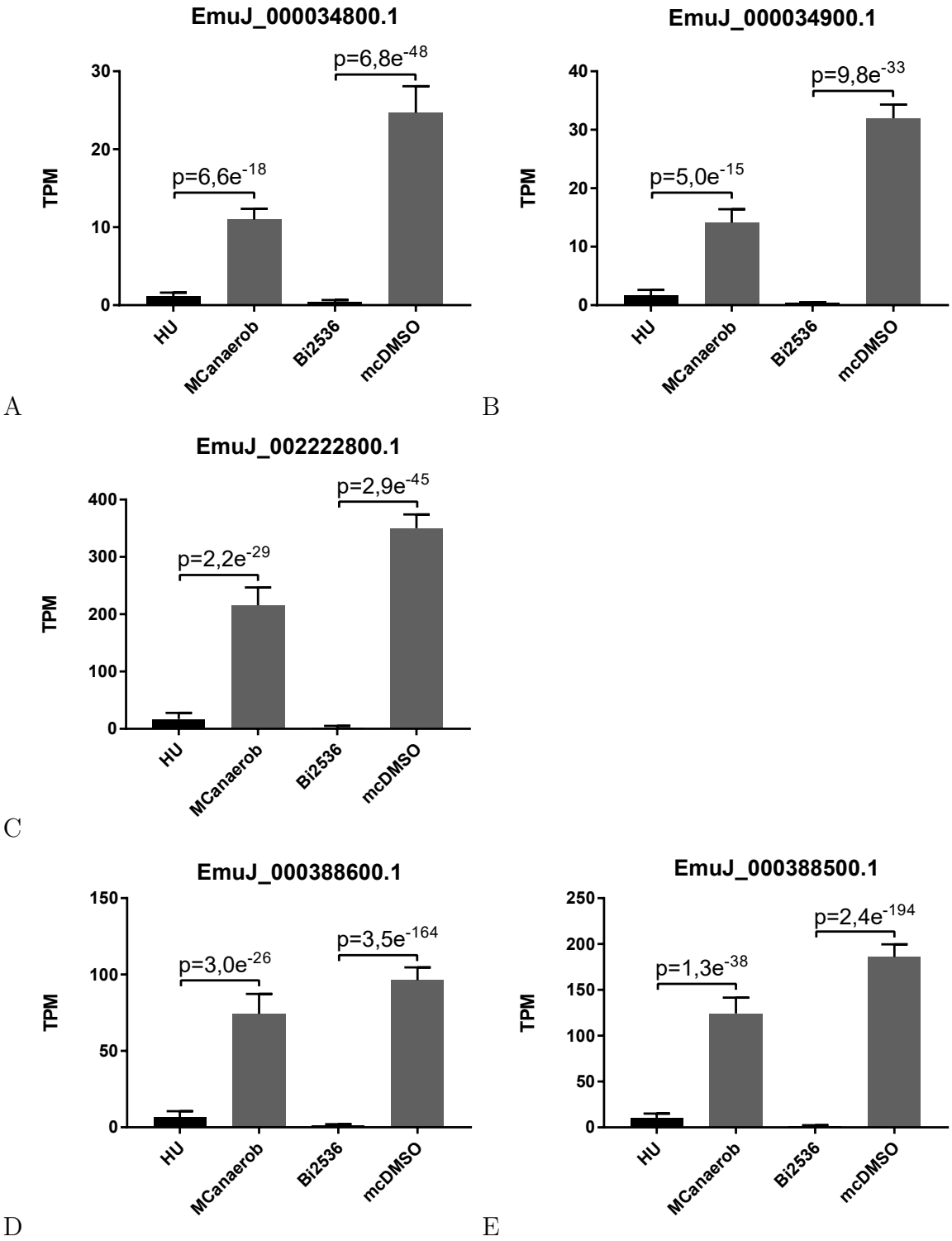


Figure 38: **Expression of *Emu*DNV-NS1 (A,C,D) and corresponding *Emu*DNV-VP (B,E) gene versions**: Expression is shown in Transcripts Per kilobase of exon per Million transcripts mapped (**TPM**). Error bars represent 1 SD. Adjusted p-values obtained by analysis with DESeq2 are depicted. **HU**: metacestodes treated with hydroxyurea; **MCanaerob**: control for HU treatment; **Bi2536**: metacestodes treated with Bi2536; **mcDMSO**: control for Bi2536 treatment.

5.4.4 Densovirus NS1 gene sequences are present in many cestode genomes

To identify densovirus NS1 sequences in other cestodes, the genomes of *D. latus*, *E. canadensis*, *E. granulosus*, *E. multilocularis*, *H. taeniaeformis*, *H. diminuta*, *H. microstoma*, *H. nana*, *M. corti*, *S. solidus*, *S. erinaceieuropaei*, *T. asiatica*, *T. multiceps*, *T. saginata* and *T. solium* together with the genome of the trematode *S. mansoni* were analyzed by blast searches. A total of 211 putative NS1 gene sequences were detected. The highest number of sequences was found in the genome of *H. diminuta* (37). Putative NS1 sequences were also identified in the genomes of *E. canadensis* (24), *E. multilocularis* (23), *T. asiatica* (Taenia_asiatica_TASYD01_v1) (23), *T. multiceps* (21), *H. microstoma* (19), *H. nana* (17), *T. asiatica* (T_asiatica_-South_Korea_0011_upd) (12), *T. saginata* (12), *E. granulosus* (6 each in EGRAN001 and ASM52419v1), *H. taeniaeformis* (4), *T. solium* (4) and *S. mansoni* (3). In contrast, no sequences were detected in the genomes of *D. latus*, *M. corti*, *S. solidus* and *S. erinaceieuropaei*. To learn more about the phylogenetic relationship and the origin of these sequences, phylogenetic analysis was performed. For the neighbor-joining tree only sequences with a coverage > 50 % of EmuDNV-NS1 EmuJ_000388600 were used (83 in total, listed in section 10.6). The phylogenetic tree shows clustering of sequences within and between species (see Figure 39). One large cluster is formed by sequences from *E. multilocularis* and *E. canadensis* with connected, less closely related sequences from *E. granulosus* (see Figure 39). Within this cluster, sequences cluster by species, indicating expansion of NS1 sequences also after species separation. Another large cluster consists of sequences from *T. asiatica*, *T. saginata*, *T. multiceps*, *H. diminuta* and *H. microstoma*. This cluster has two main branches, one containing the sequences from *Hymenolepis spp.* and the other the sequences from *Taenia spp.*. The branch containing the sequences from *Taenia spp.* consists of two separate branches for sequences from *T. multiceps* on one side and sequences from *T. asiatica* and *T. saginata* on the other side (see Figure 39). Clustering of sequences from different species and genera suggests that densoviral sequences might have originally integrated into the genome of a common cestode ancestor.

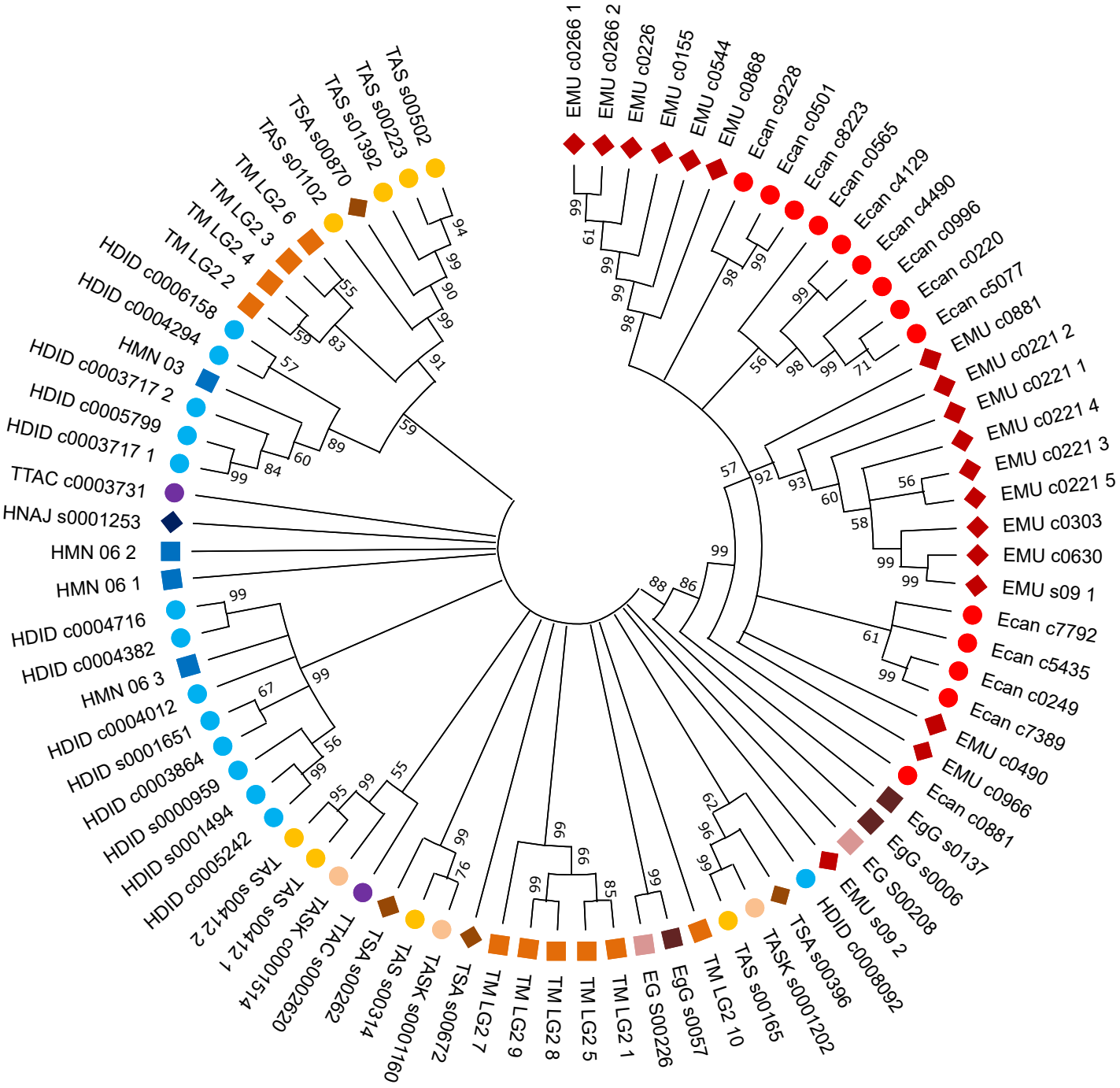


Figure 39: **Phylogenetic analysis of densovirus NS1 genes:** Nucleotide sequences with a coverage > 50 % of the gene version EmuJ_00038860 were aligned with MUSCLE (Edgar, 2004*a*,*b*). The bootstrap consensus tree was inferred using the Neighbor-Joining method (Saitou and Nei, 1987) with 1000 replications (Felsenstein, 1985) using MEGA-X (Kumar et al., 2018). Numbers are bootstrap values. Branches reproduced in less than 50 % bootstrap replications were collapsed. Ecan: *E. canadensis*, EgG: *E. granulosus* (PRJEB121), EG: *E. granulosus* (PRJNA182977), EMU: *E. multilocularis*, TTAC: *H. taeniaeformis*, HDID: *H. diminuta*, HMN: *H. microstoma*, HNAJ: *H. nana*, TASK: *T. asiatica* (PRJEB532), TAS: *T. asiatica* (PRJNA2998719, TM: *T. multiceps*, TSA: *T. saginata*.

5.5 Telomerase reverse transcriptase in *E. multilocularis*

5.5.1 The *E. multilocularis tert* genes are expressed in germinative cells

According to the transcriptome data analysis for genes specifically expressed in germinative cells, two *E. multilocularis* genes predicted to encode the telomerase reverse transcriptase (TERT) fulfilled all defined criteria for germinative cell specific expression. Expression profiles revealed highest expression of *E. multilocularis tert* genes in early primary cells (PC1) with decreased expression in later primary cells and metacestodes as well as low expression in protoscoleces (see Figure 40 A and B). Metacestodes depleted of germinative cells by treatment with HU or Bi2536 showed lower expression levels than controls (see Figure 40 C and D). As the sequences of EmuJ_001038500 and EmuJ_001039300 were in large parts identical, the actual transcription levels were better represented by Kallisto data, which is shown in figure 40. The information obtained with the Hisat2/HTSeqCount method was used for statistical analysis in DESeq2, which showed significant reductions in gene expression after depletion of germinative cells with HU and Bi2536.

Reciprocal blast searches with known TERT protein sequences from other organisms (see section 10.7) did detect only the two predicted *E. multilocularis* TERT proteins mentioned above. No additional TERT sequences were identified. The *E. multilocularis tert* was cloned from cDNA and sequenced (LR594027.1). Blast searches matched the sequence to two genomic locations on chromosome 2, corresponding to the locations of the two gene predictions. Sequence identity to the genomic locus 10935396 to 10939540 bp was 100 %, while the sequence alignment to the locus 10872004 to 10867860 bp contained one mismatch in exon 3. The *E. multilocularis tert* sequence consisted of 13 exons and encoded a 574 aa protein. SMART analysis of the translated protein sequence revealed a Telomerase_RBD (Telomerase ribonucleoprotein complex - RNA binding domain; position 41 to 163 aa) and a RVT_1 (reverse transcriptase domain; position 188 to 402 aa), both with a score below the threshold. Blastp searches against non-platyhelminth sequences did not give significant results, indicating low conservation.

To analyze if *E. multilocularis tert* is indeed expressed in germinative cells, as was suggested by the RNA-Seq transcriptome data, quantitative RT-PCR was performed with primers binding to both genes. Metacestode samples were depleted of germinative cells by treatment with HU or Bi2536. Both treatments lead to a significant reduction in *tert* gene expression compared to untreated controls (see Figure 41). Samples depleted of germinative cells had mean relative *tert* gene expressions of

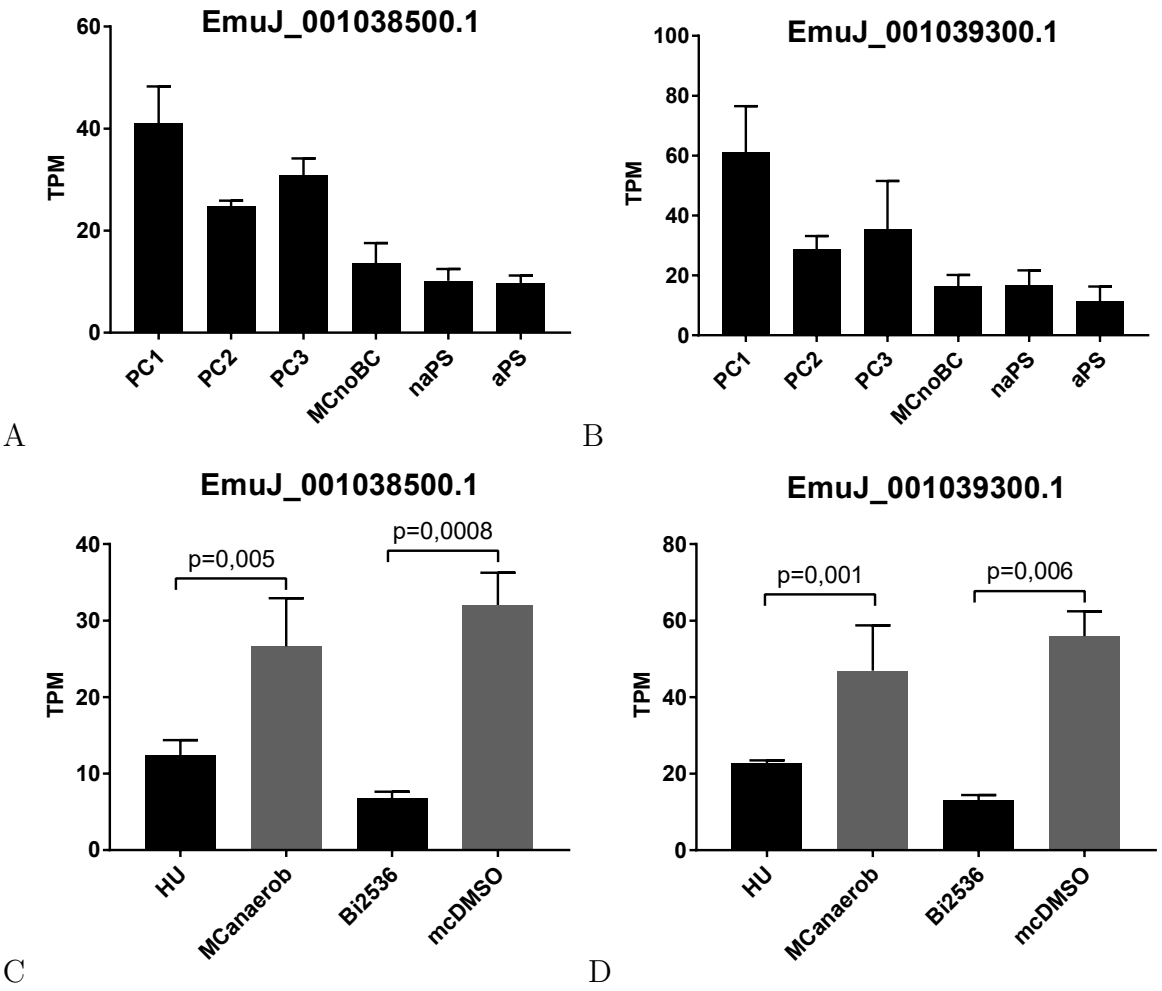


Figure 40: **Expression of *E. multilocularis tert* EmuJ_001038500.1 (A,C) and EmuJ_001039300.1 (B,D)**: Expression levels estimated with Kallisto are shown in Transcripts Per Million (**TPM**). Shown p-values were obtained by statistical analysis of the Hisat2/HTSeqCount data in DESeq2. Error bars represent 1 SD. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces; **HU**: metacestodes treated with hydroxyurea; **MCanaerob**: control for HU treatment; **Bi2536**: metacestodes treated with Bi2536; **mcDMSO**: control for Bi2536 treatment.

0,57 (HU treatment) and 0,56 (Bi2536 treatment) compared to controls. However, the relative *tert* expressions strongly varied between the examined isolates (0,34 to 0,89 for HU treatment and 0,40 to 0,87 for Bi2536 treatment compared to controls), leading to a high variation between replicates (see Figure 41). Regardless of variation between different isolates, generally reduced *tert* expressions in germinative cell depleted samples suggests that *tert* is expressed in germinative cells.

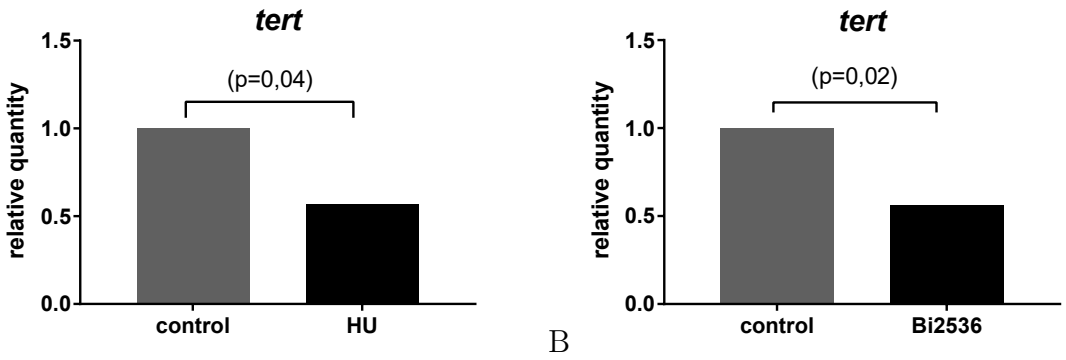


Figure 41: **Relative gene expression of *E. multilocularis tert* in HU (A) and Bi2536 (B) treated metacestodes:** Quantitative RT-PCR was carried out using cDNA from metacestode vesicles treated with hydroxyurea (HU) or the polo-like kinase inhibitor Bi2536 for germinative cell depletion and untreated controls. Experiments were performed with three technical and three biological replicates. Shown are the mean relative gene expressions and the p-values obtained by fgStatistics.

The expression patterns of *tert* in *E. multilocularis* metacestodes and protoscoleces were studied through WMISH experiments. WMISH was combined with EdU incorporation and detection to stain proliferating cells. In protoscoleces of the isolate MS1010 (n=10), many cells expressing *tert* were located in the posterior region. Several *tert*-positive cells were also situated in the rostellar area and the suckers (see Figure 42 B). The EdU stained proliferating cells showed a similar pattern with some positive cells in rostellar and sucker regions and many positive cells in the posterior region (see Figure 42 C). Co-localizations of *tert*- and EdU-signals were observed in several cells in the protoscoleces (see Figure 42 D). The quantification of co-localization in protoscoleces, however, was not possible due to confluent signals. Co-localization of *tert*- and EdU-signals were also seen in metacestode vesicles (see Figure 43). For quantification 8 vesicles from two independent experiments (isolates MS1010 and J2012) were used. Of 4686 analyzed cells 8,0 % were EdU-positive, 18,0 % *tert*-positive and 6,5 % co-stained (see Figure 44). The high proportion of co-stained cells in EdU-positive cells suggests that *E. multilocularis tert* is expressed in most proliferating germinative cells.

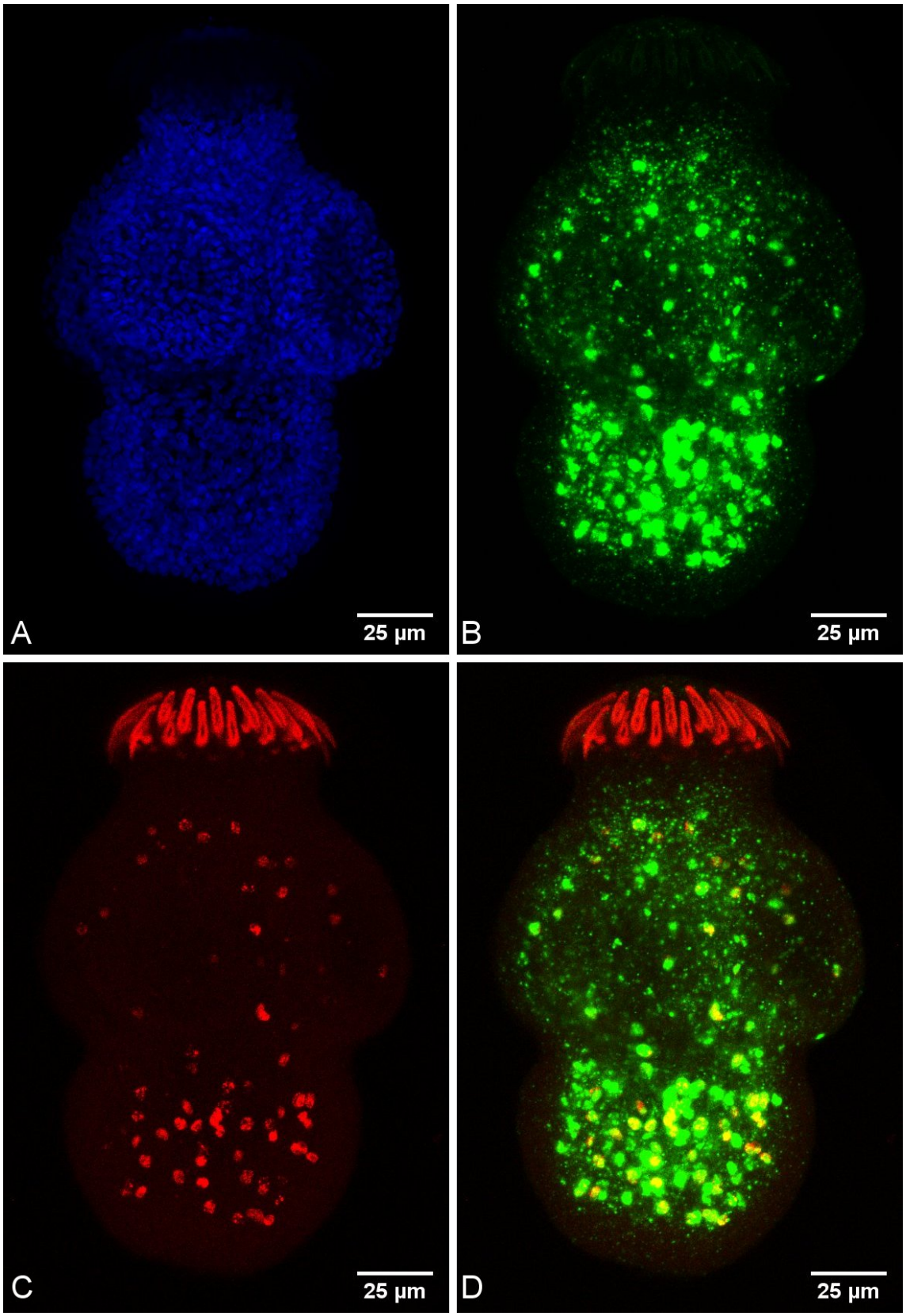


Figure 42: **WMISH of protoscolex for *E. multilocularis tert*:** Micrographs shown are Z-projections (maximum intensity) of 16 focal planes. The region of the protoscolex was cut out, rotated, converted to RGB and made into a montage. **A**: Dapi staining; **B**: WMISH with *tert* probe; **C**: EdU staining; **D**: composite of B and C.

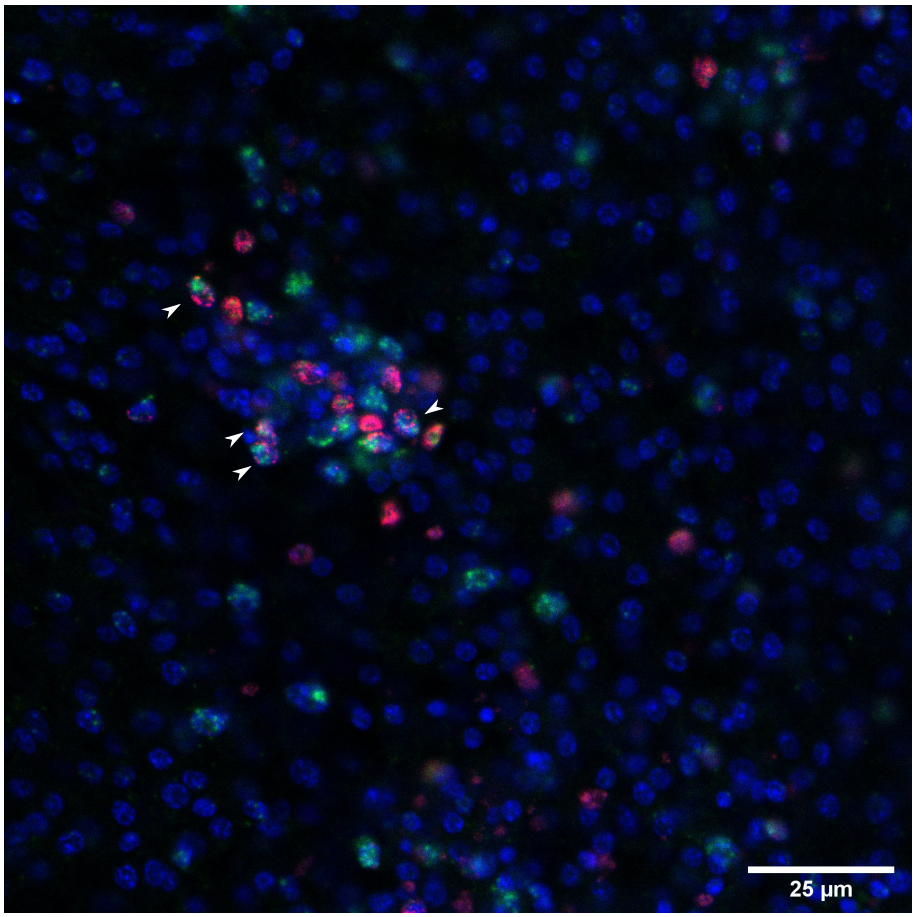


Figure 43: **WMISH of metacestode vesicle for *E. multilocularis tert*:** Micrograph shown is a composite image of the channels for the Dapi staining (blue), EdU detection (red) and the *tert* probe of the WMISH (green). Image was flattened and converted to RGB. Arrowheads mark co-stained (EdU- and *tert*-positive) cells.

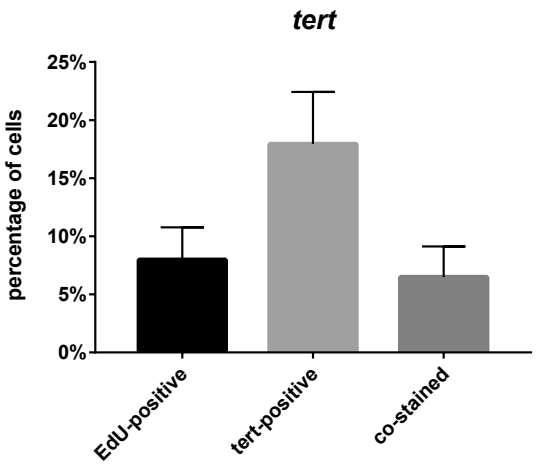


Figure 44: **Gene expression of *E. multilocularis tert* in metacestode vesicles:** Shown are the results of two independent WMISH experiments with each four metacestode vesicles (2 images each). Mean percentages of EdU-positive cells, *tert*-positive cells and co-stained cells were calculated. Error bar is 1 SD.

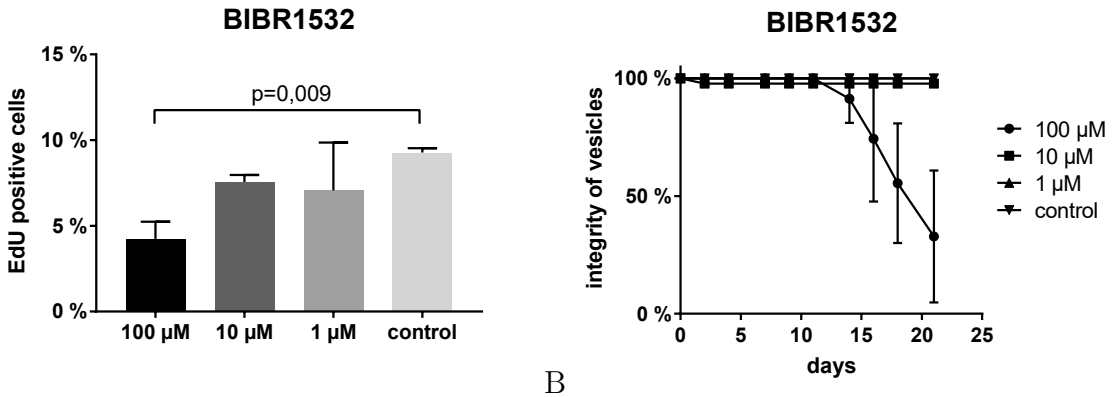
5.5.2 BIBR1532 inhibits *E. multilocularis* development

To functionally study the role of the telomerase reverse transcriptase in *E. multilocularis*, the telomerase inhibitor BIBR1532 was applied to primary cell cultures and metacestode vesicles. Short term treatment with BIBR1532 (100 µM, 10 µM and 1 µM) for two days had no effect on primary cell viability (see Figure 45 A). Prolonged incubation, however, inhibited the development of metacestode vesicles in primary cell cultures in a dose dependent manner. After 21 days, primary cells incubated with 10 µM BIBR1532 had developed significantly less metacestode vesicles than controls and cultures treated with 100 µM BIBR1532 had formed no metacestode vesicles at all (see Figure 45 B). Treatment with 1 µM BIBR1532 resulted in a non-significant reduction of developed vesicles.

A B

Figure 45: **Effect of BIBR1532 on primary cells. A: Viability.** Viability was measured using a resazurin assay after two day treatment with indicated concentrations of BIBR1532 (3 independent experiments, three technical replicates each). Triton was used as cytotoxic control. Mean fluorescence units are shown. Error bars represent 1 SD. **B: Metacestode formation.** Shown are the mean numbers of developed metacestode vesicles in primary cell cultures after 21 days of treatment with indicted concentrations BIBR1532 (3 independent experiments, three technical replicates each). Error bars are 1 SD.

Metacestode vesicles were treated with 100 µM, 10 µM and 1 µM BIBR1532 for 7 days. Proliferation of germinative cells was measured by EdU incorporation. Metacestodes treated with 100 µM BIBR1532 contained significantly less EdU-positive cells than controls, 4,2 % compared to 9,3 % in controls. Treatment with the lower concentrations 10 and 1 µM BIBR1532 also reduced the number of EdU-positive cells (7,6 and 7,1 %, respectively), tough the effect was not statistically significant (see Figure 46 A). Long term incubation with BIBR1532 caused structural damage of metacestodes. After two days of treatment first damages, such as holes in the vesicles and disruption of the barrier (uptake of color from the medium)



A B

Figure 46: **Effect of BIBR1532 on metacestodes**: **A: Proliferation.** Metacestode vesicles were treated with depicted concentrations of BIBR1532 for 7 days (3 biological replicates). Shown are the percentages of EdU-positive cells as an indicator for proliferation. Error bars are 1 SD. **B: Structural integrity.** Time course experiment in the presence of BIBR1532 at indicated concentrations. Shown are the mean percentages of intact, not collapsed, vesicles from three different isolates with each three biological replicates. Error bars are 1 SD.

were observed in samples treated with 100 and 10 µM BIBR1532. From day 16 on metacestodes treated with 100 µM BIBR1532 contained significantly (p=0,0003 and below) more collapsed vesicles than untreated controls.

Taken together, the results of the cell culture experiments suggest that BIBR1532 inhibits primary cell development and metacestode proliferation and in the long term even affects structural integrity of metacestodes.

5.6 Serotonin in *E. multilocularis*

5.6.1 *E. multilocularis* serotonin transporter and tryptophan hydroxylase

Almost the complete canonical serotonergic pathway is encoded in the *E. multilocularis* genome: the tryptophan hydroxylase (rate limiting enzyme in serotonin biosynthesis), the aromatic-L-amino-acid decarboxylase (also involved in serotonin biosynthesis), the G-protein coupled serotonin receptors, the serotonin transporter and the vesicular monoamine transporter (Camicia et al., 2013). The only exceptions are the monoamine oxidase, the enzyme responsible for degradation of serotonin in most organisms, and the serotonin receptor 3, a serotonin gated ion channel. Neither were found in the genomes of *E. multilocularis* and *E. granulosus* (Camicia et al., 2013). In this work the focus is on the serotonin transporter (SERT) and the tryptophan hydroxylase (TPH). The genes coding for the serotonin transporter (*E.*

multilocularis sert: LT934126.1) and the tryptophan hydroxylase (*E. multilocularis tph*: LT934127.1) were cloned and sequenced. *E. multilocularis sert* was localized on chromosome 9, comprised of 13 exons and 12 introns. The corresponding protein sequence of *E. multilocularis* SERT was 640 amino acids long. SMART analysis of *E. multilocularis* SERT revealed a sodium neurotransmitter symporter family (SNF) domain (PF00209) with 12 transmembrane domains. This domain structure is typical for neutransporters with sodium symporter activity. Within the SNF domain, *E. multilocularis* SERT showed high homologies to human SERT (69 % similarities). Conservation within flatworms was even higher: 82 % similarities to the SNF domain of *S. mansoni* SERT and 99 % similarities to the SNF domain of *E. granulosus* SERT (see Figure 47). Mutagenesis studies on mammalian SERTs have identified binding sites for citalopram and paroxetine. Of 8 amino acid residues that have been reported to be involved in the binding of citalopram (Barker et al., 1998, 1999; Henry et al., 2006; Mortensen et al., 2001; Andersen et al., 2009), 5 were conserved in *E. multilocularis* SERT. Compared to human SERT, Y95 was replaced with phenylalanine (F54), I172 with threonine (T131) and M180 with isoleucine (I139). These substitutions were also present in *E. granulosus* SERT and *S. mansoni* SERT (see Figure 47) and have already been described previously for the predicted sequences of *E. multilocularis* SERT, *E. granulosus* SERT and *S. mansoni* SERT (Herz, 2015). For paroxetine, 5 binding sites have been reported (Larsen et al., 2004; Henry et al., 2006; Mortensen et al., 2001). Three of them are conserved in *E. multilocularis* SERT, *E. granulosus* SERT and *S. mansoni* SERT (see Figure 47). The substitutions I172 to T131 and M180 to I139 also affect paroxetine binding.

Sequencing of *E. multilocularis tph* resulted in a partial sequence. Even with 5'RACEs (rapid amplification of cDNA ends) it was not possible to obtain the complete coding sequence up to the start codon. The partial sequence of *E. multilocularis tph* (missing 5' end) was localized on chromosome 3 and consisted of 12 exons and 11 introns. The translated protein sequence *E. multilocularis* TPH was 445 amino acids long. Domain analysis with SMART (Letunic et al., 2015; Letunic and Bork, 2018) showed that *E. multilocularis* TPH contained a complete Biopterin_H domain (Biopterin-dependent aromatic amino acid hydroxylase)(PF00351) which is characteristic for all aromatic amino acid hydroxylases. Within the Biopterin_H domain, *E. multilocularis* TPH showed high homologies to the the human TPH 2 (81 % similarities) and to the *S. mansoni* TPH (82 % similarities).

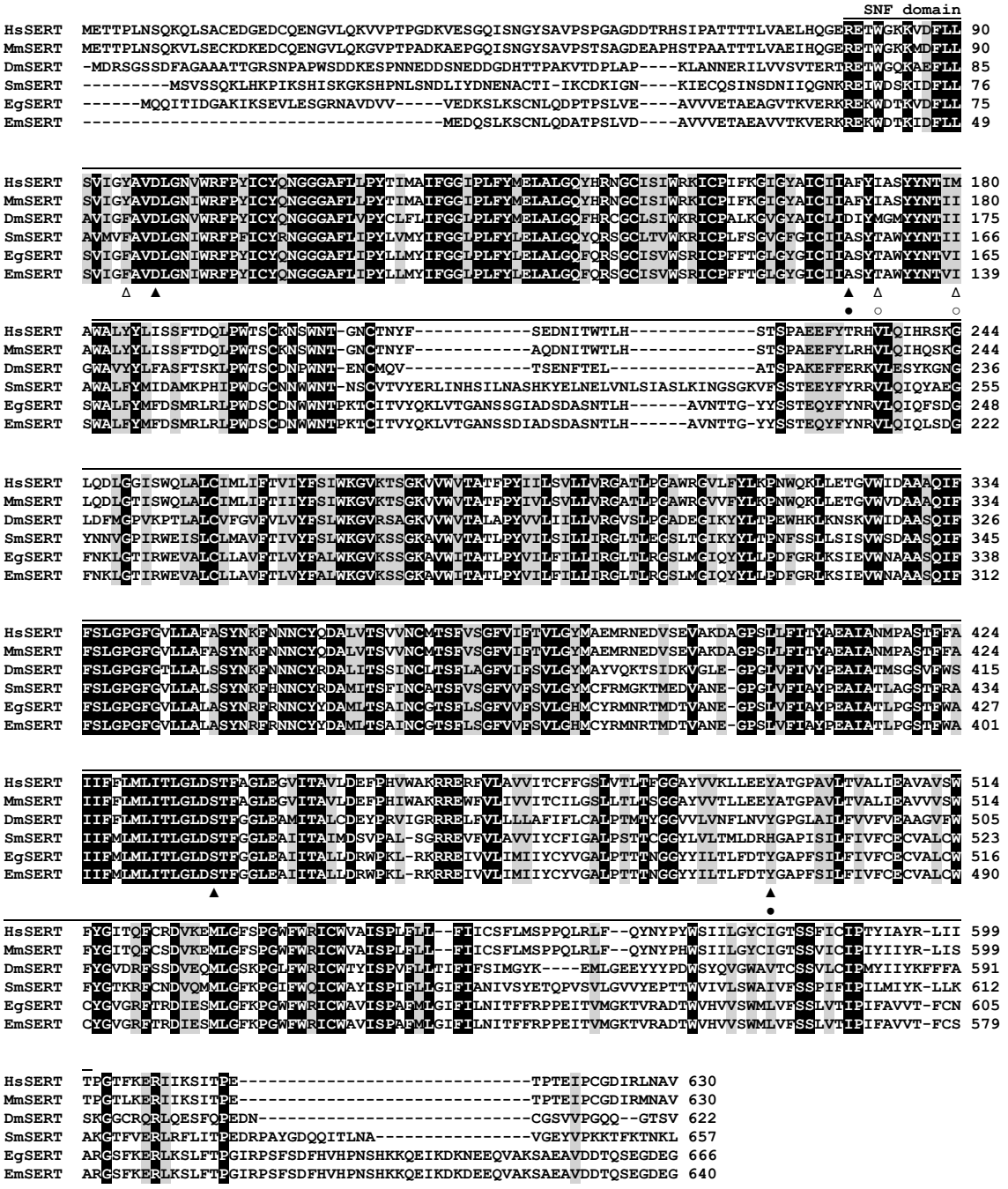
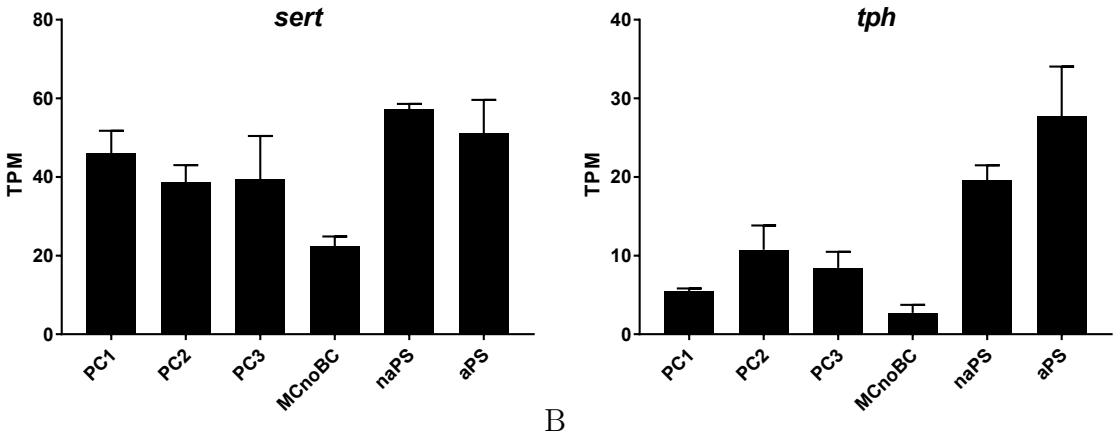


Figure 47: **Comparison of SERTs:** MUSCLE alignment of SERT protein sequences from *Homo sapiens* (HsSERT, NP_001036.1), *Mus musculus* (MmSERT, AAB67172.1), *Drosophila melanogaster* (DmSERT, NP_523846.2), *S. mansoni* (SmSERT, EF061308), *E. granulosus* (EgSERT, EUB59773.1) and *E. multilocularis* (EmSERT, LT934126.1). Highly conserved amino acids are printed white on black background, biochemically similar amino acids are printed black on grey background. Black triangles (citalopram) and circles (paroxetine) identify conserved binding sites for SSRIs in *E. multilocularis* SERT. White triangles and circles indicate substitutions. The SNF domain is underlined.

5.6.2 *E. multilocularis sert* and *E. multilocularis tph* might be expressed in the nervous system of the protoscolex

Analysis of the new transcriptome data showed that *E. multilocularis sert* was generally higher expressed than *E. multilocularis tph*. Both genes showed their highest expression in activated and non-activated protoscoleces (see Figure 48). *E. multilocularis sert* was also quite highly expressed in primary cells and showed low expression in metacestodes. *E. multilocularis tph* was lowly expressed in primary cells and almost not at all in metacestodes. In this context it is worth mentioning, that a serotonergic nervous system has been described for protoscoleces of both *E. multilocularis* and *E. granulosus* (Brownlee et al., 1994; Camicia et al., 2013; Fairweather et al., 1994; Koziol et al., 2013) but no serotonergic nerve cells were found in metacestodes (Camicia et al., 2013; Koziol et al., 2013).



A B

Figure 48: **Gene expression of *E. multilocularis sert*** (**A**) **and *E. multilocularis tph*** (**B**): Expression is shown in Transcripts Per Million (**TPM**). Error bars represent 1 SD. **PC1**: primary cells, stage 1; **PC2**: primary cells, stage 2; **PC3**: primary cells, stage 3; **MCnoBC**: metacestode without brood capsules; **naPS**: non-activated protoscoleces; **aPS**: activated protoscoleces.

In order to learn more about the expression patterns of *E. multilocularis sert* and *E. multilocularis tph*, whole mount *in situ* hybridizations (WMISH) were performed (n=18, isolates DPZ and MS1010). The median number of cells per protoscolex was 14,5 for the expression of *E. multilocularis sert* and 15 for the expression of *E. multilocularis tph*. Positions of *E. multilocularis sert*- and *tph*-expressing cells in protoscoleces coincided with the location of the serotonergic nervous system as described in Koziol et al. (2013). Cells expressing *E. multilocularis sert* were localized in the rostellar ring (median 3, range 0-4), the lateral ganglia (median 3, range 1-4), the area of the anterior ring commissure (median 1, range 0-2), the posterior lateral ganglia (median 3, range 2-5), the lateral nerve cords (median 4, rage 2-6)

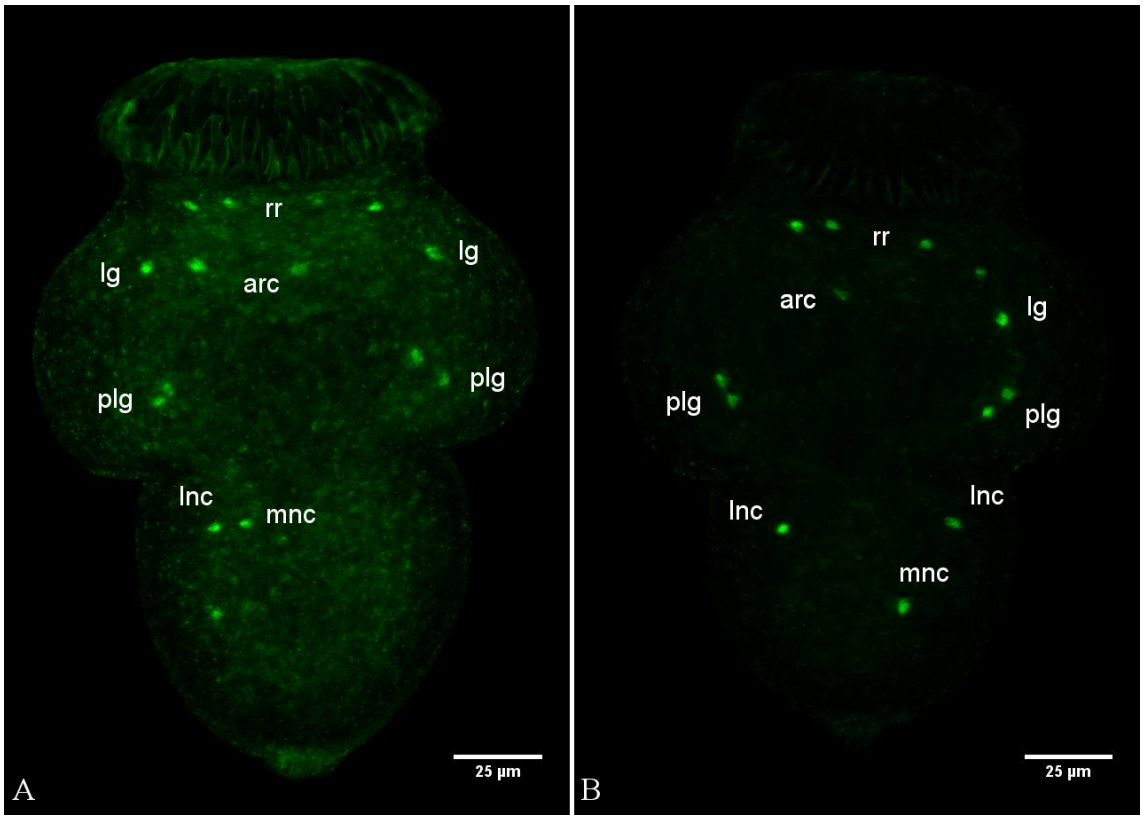


Figure 49: **Whole mount *in situ* hybridizations of *E. multilocularis sert* (A) and *E. multilocularis tph* (B) in protoscoleces:** Shown are Z-projections (maximum intensity) of several focal planes. **rr**: rostellar ring; **lg**: lateral ganglion; **arc**: anterior ring commissure; **plg**: posterior lateral ganglion; **lnc**: lateral nerve cords; **mnc**: medial nerve cords.

and seldom in the area of the posterior ring commissure (median 0, range 0-1) and the medial nerve cords (median 0, range 0-1) (see Figure 49 A). Similarly, cells expressing *E. multilocularis tph* were localized in the rostellar ring (median 3, range 0-8), the lateral ganglia (median 2, range 0-5), the area of the anterior ring commissure (median 1, range 0-2), the posterior lateral ganglia (median 4, range 3-4), the lateral nerve cords (median 4, rage 2-6) and rarely in the area of the posterior ring commissure (median 0, range 0-1) and the medial nerve cords (median 0, range 0-3) (see Figure 49 B). In metacestodes WMISH for *E. multilocularis sert* and *tph* did not succeed, which might be due to low expression levels for both genes at the metacestode stage.

In summary, transcriptome data analysis together with the results of WMSIH experiments indicate that *E. multilocularis sert* and *tph* might be expressed in the (serotonergic) nervous system of the protoscolex.

5.6.3 Serotonin induced proliferation in *E. multilocularis* metacestode vesicles

To study the role of exogenous serotonin on metacestode development, metacestode vesicles were treated with 100 µM, 10 µM and 1 µM serotonin for 7 days. Proliferation was measured by incorporation of EdU into newly synthesized DNA of germinative cells in the S-phase. Incubation with lower concentrations of serotonin did either not increase the percentage of EdU-positive cells (1 µM serotonin) or the effect was not statistically significant (10 µM serotonin). Treatment with 100 µM serotonin, however, significantly increased the number of EdU-positive cells compared to control (see Figure 50), suggesting that serotonin stimulates proliferation in *E. multilocularis* metacestodes.

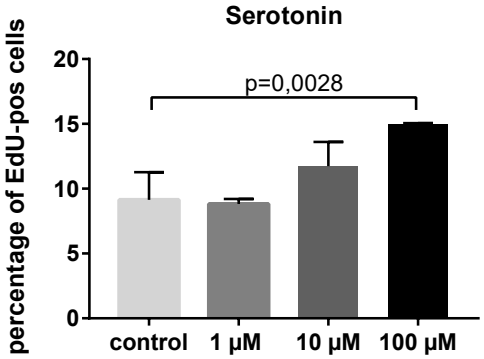


Figure 50: **Proliferation of serotonin treated metacestode vesicles:** Metacestode vesicles were treated with indicated concentrations of serotonin for 7 days (three biological replicates). EdU-incorporation was measured as an indicator for proliferation. Error bars represent 1 SD.

5.6.4 Paroxetine affects structural integrity of metacestodes and primary cell viability

To further study the influence of serotonin on *E. multilocularis*, different inhibitors were applied to the parasite's larval stages. Metacestodes were incubated with 1 µM, 10 µM and 100 µM paroxetine for 14 days. Metacestodes incubated with 10 µM and 100 µM paroxetine lost their structural integrity and collapsed. Treatment with 10 µM paroxetine resulted in significantly more collapsed vesicles compared to control cultures from day 10 on (p=0,0318). Treatment with 100 µM paroxetine already caused significant structural damage after three days (p=0,0005) compared to controls (see Figure 51). In contrast, control cultures and metacestodes treated with 1 µM paroxetine showed no structural damage.

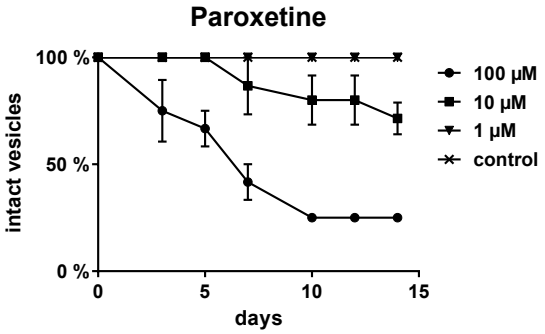
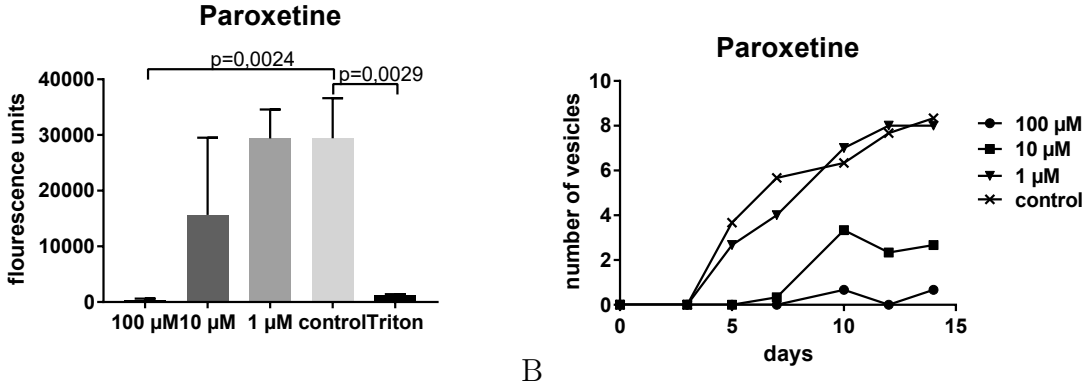


Figure 51: **Effect of paroxetine on structural integrity of metacestode vesicles:** Shown are the percentages of intact vesicles treated with indicated concentrations of paroxetine over the time of 14 days (three biological replicates). Error bars are 1 SD.

In order to determine the effects of paroxetine on cell viability, a resazurin assay was performed on primary cells. In the resazurin assay, cells incubated with 100 µM paroxetine for two days showed strongly reduced cell viability compared to untreated controls (p=0,0024). The fluorescence levels of cells incubated with 100 µM paroxetine were at a similar level as cells treated with 1 % triton, the cytotoxic control (see Figure 52 A). Incubation with 10 µM paroxetine affected cell viability to a lesser extent and differences compared to controls were not statistically significant. Treatment with 1 µM paroxetine showed no effect on cell viability (see Figure 52 A).



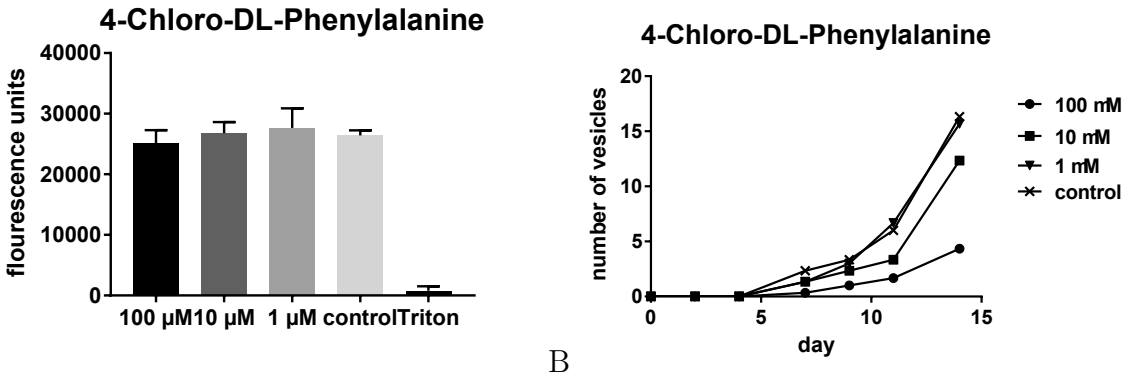
A B

Figure 52: **Effect of paroxetine on primary cells: A.** Viability of primary cells as determined by resazurin assay after two day treatment with shown concentrations of paroxetine and triton as cytotoxic control (three independent experiments with three technical replicates). Fluorescence values are shown. Error bars are 1 SD. **B.** Metacestode development from primary cells in the presence of indicated concentrations paroxetine (three technical replicates). Numbers of developed vesicles are shown.

In a preliminary experiment with prolonged incubation, primary cell cultures treated with 10 µM paroxetine and 100 µM paroxetine developed less vesicles than control cultures, while primary cells treated with 1 µM paroxetine showed no differences compared to control cultures. Taken together, these results indicate that paroxetine greatly affects the parasite's integrity and viability.

5.6.5 4-chloro-DL-phenylalanine might inhibit metacestode development from primary cells

4-chloro-DL-phenylalanine, a known inhibitor of the tryptophan hydroxylase, was used to investigate the role of the endogenous serotonin biosynthesis in *E. multilocularis*. Incubation with 1 µM, 10 µM and 100 µM 4-chloro-DL-phenylalanine for three weeks did not affect structural integrity of metacestodes vesicles. Nor did treatment with 4-chloro-DL-phenylalanine for two days show any effects on primary cell viability compared to untreated controls (see Figure 53). A preliminary experiment with prolonged incubation, however, showed a reduced number of developed metacestode vesicles in primary cell cultures treated with 10 µM and 100 µM 4-chloro-DL-phenylalanine compared to untreated cells. Though 4-chloro-DL-phenylalanine does not affect cell viability of primary cells or structural integrity of metacestode vesicles, it appears to inhibit vesicle development from primary cell cultures.



A B

Figure 53: **Effect of 4-chloro-DL-phenylalanine on primary cells: A.** Viability of primary cells as determined by resazurin assay after two days of incubation with shown concentrations of 4-chloro-DL-phenylalanine and triton as cytotoxic control (three independent experiments with three technical replicates). Shown are fluorescence units. Error bars are 1 SD. **B.** Development of metacestode vesicles in the presence of indicated concentrations 4-chloro-DL-phenylalanine (three technical replicates). Numbers of developed vesicles over the course of 14 days are shown.

6 Discussion

6.1 Transcriptome data analysis

To better understand developmental changes throughout the life cycle of *E. multilocularis* and *E. granulosus*, transcriptional profiles for various life cycle stages and experimental conditions were generated through RNA-Seq and transcriptome analyses. In addition to comparison of different life cycle stages, samples from experimental conditions, such as different cultivation methods for *E. multilocularis* metacestodes and primary cell samples with knockdown of *β-catenin*, were analyzed. This study also analyzes the transcriptome of germinative cells on a global scale.

6.1.1 Diversified functions of antigen B

The polymeric lipoprotein antigen B is known as an extremely abundant component in the hydatid fluid of metacestode cysts (Oriol and Oriol, 1975). The transcriptome data analyses showes that a set of antigen B gene copies are expressed at extremely high levels in the metacestode stages of *E. multilocularis in vivo* and *in vitro* as well as in metacestodes of *E. granulosus*, while a completely different set of gene copies is massively expressed in pregravid and gravid adults of *E. multilocularis*. This suggests specialized functions of different antigen B isoforms in different life cycle stages. While the actual function of antigen B is continuously researched, reports of immunogenic and immunmodulatory properties of antigen B indicate a role in evasion of the immune system (Chemale et al., 2001; Siracusano et al., 2008) which could be one of the functions of antigen B in the metacestode stage where the parasite is in direct contact with the mammalian host. Another proposed function of antigen B is the involvement in lipid storage and transport. Biochemical investigations demonstrate binding of hydrophobic ligands to antigen B (Chemale et al., 2005) and the capacity of antigen B to transfer fatty acids to membranes (Silva-Alvarez et al., 2015). As tapeworms are unable to synthesize fatty acids and cholesterol *de novo* (Frayha, 1971; Berriman et al., 2009), retrieval of lipids from the host is of great importance and might therefore account for the extremely high expression levels of antigen B in metacestodes and adults.

6.1.2 From primary cells to metacestodes

Primary cells isolated from metacestodes have the capacity to regenerate and form new metacestodes vesicles (Spiliotis et al., 2008). Compared to metacestodes, early primary cell cultures show high expression of genes required for DNA replication and

cell division. This is in agreement with enrichment of undifferentiated germinative cells in early primary cell cultures (Koziol et al., 2014). Consistent with the loss of tegument during primary cell preparation, the expression of tegument markers is significantly reduced in early primary cells. Also down-regulated in primary cells compared to metacestodes are genes encoding various transporters and secreted proteins. It is likely that they are expressed in the tegument and involved in host-parasite interactions, uptake of nutrients and modulation of the host´s immune response.

During the development of metacestodes from primary cells the expression of genes required for DNA replication decreases while expression of genes involved in multicellular organism development increases, suggesting that after a phase of initial proliferation, cell differentiation becomes more prominent. Higher expression of genes involved in neuropeptide signaling pathways in later primary cell stages might indicate a role of nerve cells and/or neurotransmitters during primary cell development as has already been proposed for the role of serotonin on metacestode vesicle formation (Herz, 2015).

The principal component analysis of primary cell and metacestode samples suggests that late primary cells are not an intermediate state of early primary cells (PC1) and mature metacestodes (MCnoBC) on a gene expression level but form a separate entity. Compared to metacestodes and early primary cells, later primary cell stages show considerably higher expression of the Wnt inhibitor genes *sfrp* and *sfl* as well as the "posterior" *wnt* genes *wnt1* and *wnt11a*. It has been shown by Koziol et al. (2016) that *sfrp* and *sfl* are expressed in anterior regions of brood capsules and protoscoleces but not in the laminated layer of metacestodes, while the "posterior" *wnt* genes *wnt1*, *wnt11a* are expressed in the laminated layer of metacestodes as well as in posterior regions of brood capsules and protoscoleces. Interestingly, expression of *wnt1* and *wnt11a* is substantially higher in late primary cells than in any other analyzed life-cycle stage, indicating an important role in later primary cells and possibly their development. Why they are expressed at these extremely high levels and accompanied by an increase in wnt inhibitor expression remains unclear. It can be speculated that the complete destruction of tissue organization during primary cell isolation causes a similar first response as a wound inflicted upon planaria: expression of *wnt1* and *sfrp* (Almuedo-Castillo et al., 2012).

Knock-down of β-catenin in primary cells results in higher expression of wnt inhibitors and markers for anterior specification and lower expression of "posterior"

wnt genes and markers for posterior specification, suggesting an overall shift towards anterior. Silencing of β-catenin-1 in *S. mediterranea* leads to a complete loss of posterior and causes the development of a completely anteriorized animal (Iglesias et al., 2008). It is possible that the "red dot" phenotype of β-catenin siRNA treated primary cells in *E. multilocularis* represents fully anteriorized primary cells. Taking this a step further, it could be hypothesized that "red dots" in normally developing primary cell cultures could also be anterior tissue. However, β-catenin does not only act in the Wnt signaling pathway but is also involved in cell adhesion (Brembeck et al., 2006). It is therefore possible that other roles of β-catenin contribute to the phenotype.

6.1.3 Metacestodes

Metacestodes of *E. multilocularis* can be cultivated *in vitro*. While cultivation of metacestodes with RH^- feeder cells is possible under aerob conditions, prolonged axenic cultivation requires reducing conditions (Spiliotis et al., 2004). Analysis of gene expression shows that both *in vitro* cultivation conditions are similarly "near" to *in vivo* conditions. However, higher expression of genes for DNA replication and proliferation in the MCvivo sample compared to both *in vitro* cultivation conditions indicates that *in vitro* conditions lack important factors beneficial for growth and proliferation. While secreted factors from RH^- cells are required for growth and differentiation of metacestodes *in vitro* (Spiliotis et al., 2004), additional factors not present under current *in vitro* culture conditions likely substantially contribute to growth and proliferation *in vivo*.

The *E. granulosus* oncosphere-specific antigen EG95 is located in the penetration gland of non-activated oncospheres (Jabbar et al., 2011) and thought to play a role in providing or organizing extracellular matrix for the oncosphere-metacestode transition Olson et al. (2012). Higher expression of *E. multilocularis em95*, which is homologous to the *eg95* gene, in metacestodes *in vivo* compared to *in vitro* cultivated metacestodes might therefore be attributed to a role in extracellular matrix organization. Another explanation could be that *em95* is involved in development of new vesicles, which occurs through exogenous budding of metacestodes *in vivo* but not in current *in vitro* cultivation systems (Jura et al., 1996; Spiliotis et al., 2004). Increased expression during primary cell development would be consistent with a role in vesicle formation.

6.1.4 Protoscoleces and adults

Protoscoleces develop in brood capsules within metacestode vesicles (Koziol et al., 2016). Fully developed, not yet activated protoscoleces are considered to be mostly a resting stage. In contrast, metacestodes fulfill a variety of functions including uptake of nutrients and energy generation as well as protection against the immune system (Frayha, 1971; Wakelin, 1997; Xiao et al., 1995). As expected, a large number of genes is differentially expressed between protoscoleces and metacestodes. Markers for the metacestode tegument are not or only lowly expressed in protoscoleces. In contrast, the alkaline phosphatase 3 gene, a marker for protoscoleces (Koziol et al., 2014), is exclusively expressed in protoscoleces and not in any metacestode sample suggesting that metacestodes are free of protoscoleces. As anticipated, metacestodes show higher expression of genes involved in metabolic processes than protoscoleces, including genes encoding antigen B and FABPs (also see section 6.1.1). During the formation of brood capsules and protoscoleces in metacestodes, Wnt inhibitors are expressed locally in the anterior regions of the developing brood capsule buds and of the eventually emerging protoscoleces but not in the germinal layer of metacestodes (Koziol et al., 2016). Expression of Wnt inhibitors exclusively in anterior regions can explain their increasing expression levels from metacestodes without brood capsules, where only posteriorized tissue is present, to metacestodes with brood capsules to non-activated protoscoleces. Increased expression of genes involved in neuropeptide signaling pathways in protoscoleces compared to metacestodes can be attributed to their expression in the remarkably complex nervous system of the protoscolex (Koziol et al., 2013). Higher expression of putative Kunitz chymotrypsin and trypsin inhibitors in protoscoleces than in metacestodes might indicate a role in protection against digestion enzymes.

In contrast to the transition from metacestodes to protoscoleces, the activation of protoscoleces results in a small number of differentially expressed genes. Up-regulation of genes involved in transport and metabolic processes in activated protoscoleces compared to non-activated protoscoleces likely allows uptake of nutrients from the host and their processing after activation of protoscoleces through gastrointestinal passage. The concurrent slight decrease in expression of genes involved in mismatch repair and DNA replication is surprising. Proliferation of fully developed, non-activated protoscoleces has been described as very low with a dramatic increase in DNA replication after protoscolex activation (Koziol et al., 2014). One possible explanation would be that the total, absolute gene expression strongly increases with activation of protoscoleces and as a result expression of genes involved in DNA

replication relatively decreases slightly. Another possibility is that the RNA is already generated before the activation and translated afterwards leading to increased proliferation after protoscolex activation.

Inside the small intestine activated protoscoleces develop into adults (Smyth, 1968; Thompson et al., 1990). On a gene expression level the two available adult samples, especially the gravid adult sample, greatly differ from the samples of the larval stages. This suggests the conclusion that adults and larval stages express different sets of genes as is the case with the antigen B gene copies (see section 6.1.1). To verify this hypothesis further biological replicates of adult samples would be required.

6.1.5 Gene expression of germinative cells

The only proliferating cells in *E. multilocularis* are the pluripotent germinative cells (Koziol et al., 2014). In order to identify genes that are likely specifically expressed in germinative cells, transcriptomic analysis of germinative cell depleted and germinative cell enriched samples were performed. Depletion of proliferating cells by irradiation or RNAi has been used in planarians and schistosomes to characterize gene expression of proliferating cells (Collins et al., 2013; Rossi et al., 2007; Eisenhoffer et al., 2008; Solana et al., 2012; Wagner et al., 2012). In *E. multilocularis* germinative cells can be specifically depleted by treatment with HU or Bi2536 (Koziol et al., 2014; Schubert et al., 2014). However, treatment with HU for depletion of germinative cells takes 7 days (Koziol et al., 2014), treatment with Bi2536 even 21 days (Schubert et al., 2014). As differentiation of germinative cells to somatic cells can occur within seven days (Koziol et al., 2014), it is highly likely that treatment with HU or Bi2536 at least partially depletes the differentiating, direct progeny of germinative cells. To obtain candidate genes that are specifically expressed in germinative cells and not in the progeny, the criterium of higher expression in germinative cell enriched samples was introduced. Compared to metacestodes (32 - 55% germinative cells), 2-day-old primary cells are highly enriched in germinative cells (62 - 83 %) (Koziol et al., 2014) while later stages of primary cells are likely to contain more progeny and differentiated cells (see section 6.1.2). Higher expression in 2-day-old primary cells compared to both later primary cells (PC2) and metacestodes (MCnoBC) should therefore exclude genes expressed in progeny of germinative cells. Consequently, genes higher expressed in 2-day-old primary cells (compared to both PC2 and MCnoBC) and additionally lower expressed in samples depleted of germinative cells (both HU and Bi2536 compared to their respective controls) are good candidates for specific expression in germinative cells. The expression profiles of known marker genes for somatic and germinative cells corroborate this assumption.

To further reassert that candidate genes are indeed specifically expressed in germinative cells, WMISH experiments could be performed for selected genes. In this work, the telomerase reverse transcriptase subunit gene *tert*, a candidate for specific expression in germinative cells according to transcriptomic analysis, was studied in detail. WMISH of *tert* indicated expression in almost all currently replicating cells and likely in all germinative cells (see section 44) confirming the assessment of the transcriptome analysis.

As a future approach, germinative cells could be isolated via FACS sorting in order to obtain purified RNA of germinative cells for sequencing without contamination of other cell types. For *Mesocestoides corti* FACS isolation of cells in G0/G1 and G2/M phase has been established (Dominguez et al., 2014). However, isolation of proliferating cells by DNA amount was unsuccessful in our laboratory (pers. communication Markus Spiliotis). Current efforts focus on finding a suitable cell surface marker for isolation of germinative cells. Another possibility is single cell sequencing, which has the advantage that sorting is not absolutely necessary because cells can be identified by their transcriptomic signature. Additionally, single cell sequencing allows the investigation of subpopulations of germinative cells (Wen and Tang, 2016; Kumar et al., 2017) that have been proposed for *E. multilocularis* (Koziol et al., 2014).

Treatment of *E. multilocularis* with HU or Bi2536 probably causes specific changes in gene expression besides depletion of germinative cells. Comparison of gene expression after HU and Bi2536 treatment indicates that treatment with HU might also affect other cells types, particularly muscle cells expressing *wnt11a*. In contrast, treatment with Bi2536 appears to have a stronger effect on depletion of germinative cells. However, the higher number of downregulated genes after Bi2536 treatment compared to HU could indicate more unspecific, not germinative cell depletion related, effects. It is therefore not possible to give a general recommendation which method is more suitable for germinative cell depletion.

6.2 Transcriptomics and proteomics

Comparison of transcriptome and proteome information of 2-day-old primary cells reveals proteins whose abundances do not correlate with their corresponding transcript levels. High protein abundances without corresponding transcripts, as is the case for certain histones, could be caused by differences in half lives (Haider and Pal, 2013). In contrast, high transcription levels without corresponding proteins for

densoviral genes in *E. multilocularis* might relate to a post-transcriptional silencing mechanism.

6.3 Natural antisense transcripts

Natural antisense transcripts are recognized as important regulators of gene expression at almost all levels including transcription, mRNA processing and translation (Murray and Mellor, 2016; Villegas and Zaphiropoulos, 2015) with significant roles in DNA repair, cancer, neurological disorders, diabetes and cardiovascular disease (Barman et al., 2019). Analysis of RNA-Seq data regarding reverse spliced reads indicates that natural antisense transcripts are present in *E. multilocularis*. As only limited information can be gathered from RNA-Seq data obtained with an unstranded sequencing protocol, further experiments are required to obtain more detailed information on natural antisense transcripts, e.g. stranded RNA sequencing or direct RNA sequencing with Nanopore. Note, that natural antisense transcripts could also provide a mechanism for silencing of MGEs in *E. multilocularis*, such as densoviral sequences.

6.4 DNA methylation

The required elements for DNA methylation (DNMT and MBD) are encoded in the *E. multilocularis* genome (Geyer et al., 2013) and expressed in all analyzed larval stages. Wide expression of *dnmt2* in germinative and differentiated cells as well as in different larval stages might indicate an important role of *dnmt2* in the parasite. In a preliminary experiment, primary cell cultures treated with the DNMT inhibitor 5-aza-2'-deoxycytidine showed reduced vesicle formation which proposes a role of DNA methylation in developmental processes. In *S. mansoni* the related DNMT inhibitor 5-azacytidine inhibits oviposition in a pleiotropic manner (Geyer et al., 2011; Geyer, Munshi, Vickers, Squance, Wilkinson, Berrar, Chaparro, Swain and Hoffmann, 2018). Furthermore, preliminary experiments of Laura Campos (Campos, 2018) show that global DNA methylation levels depend on the larval stage (or possibly the isolate) of *E. multilocularis* suggesting that DNA methylation might be involved in transition of larval stages. To further investigate this, bisulfite sequencing could be used to determine methylation patterns and to correlate these with gene expression profiles. Direct comparison of gene expression in parasite material with demethylated DNA (e.g. through treatment with siRNA for *dnmt2* or a DNMT inhibitor) and in untreated controls could reveals genes that are regulated by DNA methylation.

6.5 Densovirus integrations into cestode genomes

A remarkable characteristic of tapeworms is their apparent lack of a canonical piRNA-pathway (Tsai et al., 2013; Skinner et al., 2014; Fontenla et al., 2017), which in many animals silences MGEs in the germline (Siomi et al., 2011). So far, GYPSY class of Long Terminal Repeats retrotransposons (Bae, 2016), *Merlin* DNA transposons (Tsai et al., 2013) and TRIMs, which are massively expressed in *E. multilocularis* germinative cells (Koziol et al., 2015), have been described in cestode genomes. Since uncontrolled MGEs cause genomic instability (Kines et al., 2014; Levine et al., 2016), cestodes are expected to have alternative mechanisms to silence MGEs (Skinner et al., 2014). To investigate the molecular nature of these mechanisms, information about MGEs, especially silenced ones, is required first.

In this work densovirus-related elements in *E. multilocularis* and other cestode genomes were analyzed. In *E. multilocularis* identified sequences show clear homologies to parvo- and densovirus elements in other organisms: ORFs coding for proteins with homologies to non-structural and structural densovirus proteins, flanked by ITRs. Densovirus sequences in *E. multilocularis* are often located in isolated areas of the genome or near histone clusters. PCR-analysis for selected densovirus sites confirms the genome assembly at these sites, indicating that densoviral elements are indeed present in the genome. Identification of densovirus-related sequences in 13 of 17 examined tapeworm genomes suggests widespread presence of densoviral sequences in cestodes. This is in accordance with endogenization of densoviruses and parvoviruses in many animal genomes, including flatworm genomes (Liu et al., 2011). Highly variable numbers of detected densoviral NS1 sequences in the investigated species might relate to various factors and not necessarily represent the real numbers of densovirus sequences in the respective genomes. One important factor could be the quality and method of the genome assembly. Repeats are often collapsed in the genome assembly (Biscotti et al., 2015), resulting in an under-representation of repetitive sequences in the genome assembly. In consequence, less densovirus sequences would be detected, leading to an underestimation of the real number of densoviral sequences in the genome. This could explain why only 6 NS1 sequences were found in the *E. granulosus* genomes compared to 23 in the high quality reference genome of *E. multilocularis* (Tsai et al., 2013). Other factors influencing the number of NS1 sequences in the genomes might be the time-point and number of original densovirus integration events as well as the subsequent spread of densoviral sequences. The phylogenetic analysis shows clustering of sequences from different species suggesting that first infections and genome integrations might have occurred

before the species separation. Furthermore, clustering of sequences within species indicates that densoviral elements have expanded since and might still be active.

Requirements for active replication of densoviruses are expression of functioning NS1 proteins and structurally integer ITR sequences. The NS1 is involved in the excision of the virus genome and the hairpin structure of the ITRs serve as primers (Afanasiev and Carlson, 2000). ITR sequences have been detected at several densovirus sites in the *E. multilocularis* genome. Regarding NS1, the transcribed *Emu*DNV-NS1 gene versions contain frameshift mutations or are truncated. Whether functional NS1 proteins are available and whether densovirus elements are active in *E. multilocularis* is therefore questionable. Regardless, *Aedes aegypti* densovirus-based vectors have already been successfully used for transduction into living *Aedes aegypti* mosquitoes (Afanasiev et al., 1999) and *Junonia coenia* densovirus-based vectors have been shown to stably integrate into the genome of insect cells with persistent transgene expression (Bossin et al., 2003). Presence of densovirus sequences in the *E. multilocularis* and other cestode genomes indicates that densovirus-based vectors for genetic manipulation might be an promising avenue to explore in future.

While the majority of densovirus genes in *E. multilocularis* appear to be transcriptionally silent, transcriptome data shows transcription of three densovirus loci. Cloning and sequencing of *Emu*DNV-NS1 further confirms transcription of *Emu*DNV-NS1 EmuJ_000388600. At the same time, no sequences were obtained for three other *Emu*DNV-NS1 gene versions with equal primer annealing properties indicating that they are not or only lowly expressed. Transcriptome information agrees with this conclusion showing low expression levels for one of them and no expression for the other two. Interestingly, transcriptome data and RT-PCR suggest that densoviral genes are specifically expressed in germinative cells, which would explain their maintenance in the parasite's germline. Why, however, three densovirus loci are transcribed when most are transcriptionally silent, remains unclear. No differences between transcribed and silent loci concerning putative promoter elements, sequence integrity or genomic location were discovered. An epigenetic silencing mechanism, as has been proposed for parvovirus B19 (Bonvicini et al., 2012), could be an explanation for the observed differences in transcription. DNA-Methylation has recently been described in cestodes (Geyer et al., 2013) and comparison of silent and transcriptionally active densovirus genes could give fist indications regarding a potential role of DNA-methylation in silencing of densovirus genes and potentially other MGEs in *E. multilocularis* and other cestodes.

6.6 Telomerase reverse transcriptase

According to the RNA-Seq analysis, the *E. multilocularis tert* genes are candidates for specific expression in germinative cells. Decreased expression of *tert* in metacestodes treated with HU and Bi2536 (compared to controls) in quantitative RT-PCR supports this conclusion. High variation of the relative decrease in expression between *E. multilocularis* isolates can be attributed to different expression patterns or to differences in the effectiveness of the germinative cell depletion, which might depend on the age of the culture, the isolate, the thickness of the laminated layer etc. Co-localization of EdU- and *tert*-signals in WMISH of *E. multilocularis* protoscoleces and metacestodes shows expression of *tert* in proliferating cells. The high proportion of co-stained cells in the EdU-positive cell population indicates that *tert* is expressed in most germinative cells. 32 % to 55 % of all cells in metacestodes are germinative cells (Koziol et al., 2014). The 18 % *tert*-positive cells in metacestodes could therefore be germinative cells not currently replicating their DNA. However, it cannot be ruled out that *tert* might also be expressed in differentiated cells. In asexual worms of the planarian *S. mediterranea* telomerase activity is thought to be mainly restricted to proliferating adult stem cells and/or their progeny (Tan et al., 2012). Based on the results of the experiments, it is likely that the same is true for *E. multilocularis*. In conclusion, the results of the quantitative RT-PCR and WMISH support the hypothesis from the RNA-Seq analysis that *tert* is expressed in germinative cells. Whether expression of *tert* is restricted to stem cells or not remains uncertain.

For cell culture experiments the selective telomerase inhibitor BIBR1532, which is a non-competitive inhibitor (Pascolo et al., 2002), was used. Long term cultivation of cell lines derived from several human cancers with 10 µM BIBR1532 led to telomere erosion and slowing of cell growth (Damm et al., 2001). Concentrations above 20 µM BIBR1532 have been shown to also have a direct anti-proliferative and cytotoxic effect on primary leukemia cells, possibly through disruption of the capping function of telomeres (El-Daly et al., 2005). Growth of normal human progenitor cells was only severely affected by concentrations above 160 µM BIBR1532 (El-Daly et al., 2005). Concentrations higher than 100 µM BIBR1532 also have been reported to inhibit RNA polymerases (Damm et al., 2001). In contrast to leukemia cells (El-Daly et al., 2005), short term treatment of *E. multilocularis* primary cells with BIBR1532 does not cause cytotoxic effects, even at the high concentration of 100 µM. Long term cultivation of primary cells with BIBR1532, however, leads to reduced vesicle formation. Treatment with 10 µM BIBR1532 inhibits vesicles formation but does not

completely block it, whereas treatment with 100 µM BIBR1532 completely prevents vesicle development in primary cell cultures. Whether telomere erosion or telomere dysfunction play a role in the inhibitory effects of BIBR1532 on vesicle formation has to be determined by further experiments. In *E. multilocularis* metacestodes stained with EdU, BIBR1532 shows an anti-proliferative effect, which could also be the cause of reduced or absent vesicle formation in primary cell culture treated with BIBR1532. Prolonged incubation of metacestodes with 100 µM BIBR1532 leads to collapse of metacestode vesicles. Overall, BIBR1532 inhibits proliferation and development in *E. multilocularis* and affects structural integrity of metacestode vesicles.

Modulation of telomerase and telomere dynamics is a promising research area in regard to the therapeutic control of cancers, as telomere maintenance is essential for almost all cancers (Sugarman et al., 2019). The results of the *E. multilocularis* cell culture experiments suggest that telomerase inhibitors have potential as drugs for the treatment of echinococcosis, especially as *E. multilocularis tert* is expressed in germinative cells, the cell type responsible for proliferation and probably also for the recurrence of the parasite after discontinuation of chemotherapy (Brehm and Koziol, 2014).

6.7 Serotonin

The results of analyses and experiments of this work together with previous studies indicate that serotonin plays an important role in *E. multilocularis* larval development and survival. The sequenced *E. multilocularis sert* and *tph* code for proteins that show high homologies to their human and *S. mansoni* homologues, especially within their functional domains. However, important binding sites for the well known serotonin transporter inhibitors citalopram and paroxetine are substituted in *E. multilocularis* SERT compared to human SERT. Especially the double substitution at Y95 and I172 is known to strongly reduce inhibitory potency of citalopram (Henry et al., 2006). As paroxetine binding properties are considerably less affected by these substitutions (Henry et al., 2006), paroxetine was used as a SERT inhibitor for cell culture experiments in this work. Studies of the *S. mansoni* SERT, which shows the same substitutions as *E. multilocularis* SERT, showed higher IC_{50} values for paroxetine, and even higher ones for citalopram, compared to human SERT (Fontana et al., 2009). This is the reason for the use of high concentrations paroxetine (up to 100 µM) in cell culture experiments of this work. While unspecific toxic effects due to high concentrations cannot be ruled out, effects were already observed

at 10 µM paroxetine. Together with the above described amino acid substitutions, this strongly suggests that paroxetine effects are specific.

Various studies in vertebrates and invertebrates suggest that activation of the serotonin transporter is necessary for mitogenic effects of serotonin (Fanburg and Lee, 1997; Buznikov et al., 2001; Tutton and Barkla, 1982). Inhibition of the serotonin transporter in mouse mesenchyme cells, in bovine smooth muscle cells and in colonic tumors in rats blocks mitogenic effects of serotonin (Lee et al., 1991, 1994; Buznikov et al., 2001; Tutton and Barkla, 1982). In the trematode *S. mansoni* SSRIs are known to reduce miracidial transformation rates (Taft et al., 2010) and in *E. granulosus* citalopram inhibits vesicularization and re-differentiation of protoscoleces and affects their viability (Herz, 2015; Camicia et al., 2013). In a preliminary experiment paroxetine inhibits metacestode development from primary cell cultures suggesting that serotonin transport is required for *E. multilocularis* larval development. Paroxetine also affects viability of primary cell cultures and structural integrity of mature metacestode vesicles. It was likewise observed that citalopram impairs the structural integrity of metacestodes (Herz, 2015), although the structural damage occurred later than observed with paroxetine and only at the highest concentration applied (100 µM). These results indicate that serotonin uptake is required for the development of the parasite and its survival.

To study the role of endogenous serotonin biosynthesis, the tryptophan hydroxylase, which catalyzes the rate limiting step in serotonin synthesis (Lovenberg et al., 1967), was inhibited with 4-Chloro-DL-phenylalanine. While no effect was observed on the integrity of mature metacestode vesicles or the viability of primary cell, a preliminary experiment shows reduced formation of metacestode vesicles in primary cell cultures treated with 4-Chloro-DL-phenylalanine suggesting that endogenous serotonin synthesis might play a role in developmental processes.

WMISH was used to investigate the locations of *E. multilocularis sert* and *tph* in protoscoleces. Cells expressing *E. multilocularis sert* and *tph* were positioned in locations corresponding to the serotonergic nervous system as described in Koziol et al. (2013). Not only the positions, but also the frequency of cells expressing *sert* and *tph* in each position closely resembled that of serotonin immunoreactive cells as reported in Koziol et al. (2013), suggesting that *E. multilocularis sert* and *tph* are expressed in the nervous system of the protoscolex. Similarly, immunofluorescence analysis and WMISH in the planarian *Dugesia japonica* shows expression of TPH in the nervous system (Nishimura et al., 2007). Transcriptome data of

various *E. multilocularis* larval stages show highest expression levels for *E. multilocularis sert* and *tph* in protoscoleces which might be attributed to expression in the nervous system. Likewise, comparatively high expression in primary cells, especially for *E. multilocularis sert*, could be explained by the presence of nerve cells in primary cell preparations (Koziol et al., 2014). Low expression levels of *E. multilocularis tph* in metacestodes likely are due to a lack of serotonergic nervous system in the metacestode stage (Koziol et al., 2013). Attempts to detect *E. multilocularis sert* and *tph* in metacestodes using WMISH were unsuccessful, which could be attributed to low or ubiquitous expression at this larval stage. Low or no expression of *E. multilocularis tph* in metacestodes could also explain why inhibition of TPH with 4-Chloro-DL-phenylalanine is of no consequence in metacestodes. In contrast, inhibition of SERT with paroxetine affects structural integrity of metacestodes suggesting that *E. multilocularis* SERT is expressed and essential in metacestodes. As no serotonergic nervous system is present in the metacestode Koziol et al. (2013), expression of *E. multilocularis sert* in other cell types, at least at the metacestode stage, appears likely. This would be in accordance with a predominant neuronal function and a secondary role in exogenous serotonin uptake, as has been proposed for *S. mansoni* SERT (Patocka and Ribeiro, 2013).

Serotonin is a known regulator of developmental processes across phyla (Lauder, 1993; Turlejski, 1996; Azmitia, 2001). Previously, serotonin was shown to stimulate metacestode development in *E. multilocularis* primary cell cultures (Herz, 2015). In this work the effect of serotonin on metacestodes was studied. Exogenously supplied serotonin induces proliferation in *E. multilocularis* metacestodes, the larval stage that causes alveolar echinococcosis in humans. Serotonin could therefore be a relevant developmental signal for the formation and the growth of the metacestode in the liver. However, basal serotonin levels in blood (Fanburg and Lee, 1997) and tissues (Mossner and Lesch, 1998) are generally lower than 100 nM and the proliferative effects in cell culture experiments were observed at considerably higher concentrations. While inflammation can lead to higher levels of serotonin (Mossner and Lesch, 1998), developmental effects of serotonin might also be exerted by endogenous serotonin. Reduction of vesicle formation in primary cell cultures treated with the tryptophan hydroxylase inhibitor 4-Chloro-DL-phenylalanine indicates that endogenous serotonin plays a role in the development of metacestode vesicles. Another possibility is that cellular uptake of serotonin is required for developmental and mitogenic effects, as has been proposed previously (Tutton and Barkla, 1982; Fanburg and Lee, 1997; Buznikov et al., 2001). Given the expressions patterns of *E. multilocularis sert* and *tph*, a role of the serotonergic nervous system in the reg-

ulation of developmental processes seems likely. In summary, serotonin stimulates developmental processes and proliferation in *E. multilocularis* larval stages.

7 Conclusion

This work represents the most comprehensive exploration of the *E. multilocularis* transcriptome so far. Information about global gene expression patterns throughout life cycle stages and experimental conditions is an important resource for future research on host-parasite interactions as well as on parasite growth and development. This resource will also facilitate the identification of suitable drug targets and the development of new drugs against alveolar echinococcosis.

8 Bibliography

Afanasiev, B. and Carlson, J. (2000), 'Densovirinae as gene transfer vehicles', *Contrib Microbiol* **4**, 33–58.

Afanasiev, B. N., Ward, T. W., Beaty, B. J. and Carlson, J. O. (1999), 'Transduction of *Aedes aegypti* mosquitoes with vectors derived from *Aedes* densovirus', *Virology* **257**(1), 62–72.

Agrawal, A., Dang, S. and Gabrani, R. (2012), 'Recent patents on anti-telomerase cancer therapy', *Recent Pat Anticancer Drug Discov* **7**(1), 102–17.

Alexa, A. and Rahnenfuhrer, J. (2016), *topGO: Enrichment Analysis for Gene Ontology*. R package version 2.28.0.

Almuedo-Castillo, M., Sureda-Gomez, M. and Adell, T. (2012), 'Wnt signaling in planarians: new answers to old questions', *Int J Dev Biol* **56**(1-3), 53–65.

Anders, S., Pyl, P. T. and Huber, W. (2015), 'HTSeq–a Python framework to work with high-throughput sequencing data', *Bioinformatics* **31**(2), 166–9.

Andersen, J., Taboureau, O., Hansen, K. B., Olsen, L., Egebjerg, J., Stromgaard, K. and Kristensen, A. S. (2009), 'Location of the antidepressant binding site in the serotonin transporter: importance of Ser-438 in recognition of citalopram and tricyclic antidepressants', *J Biol Chem* **284**(15), 10276–84.

Aziz, A., Zhang, W., Li, J., Loukas, A., McManus, D. P. and Mulvenna, J. (2011), 'Proteomic characterisation of *Echinococcus granulosus* hydatid cyst fluid from sheep, cattle and humans', *J Proteomics* **74**(9), 1560–72.

Azmitia, E. C. (2001), 'Modern views on an ancient chemical: serotonin effects on cell proliferation, maturation, and apoptosis', *Brain Res Bull* **56**(5), 413–24.

Bae, Y. A. (2016), 'Evolutionary characterization of Ty3/gypsy-like LTR retrotransposons in the parasitic cestode *Echinococcus granulosus*', *Parasitology* **143**(13), 1691–1702.

Baganz, N. L. and Blakely, R. D. (2013), 'A dialogue between the immune system and brain, spoken in the language of serotonin', *ACS Chem Neurosci* **4**(1), 48–63.

Barker, E. L., Moore, K. R., Rakhshan, F. and Blakely, R. D. (1999), 'Transmembrane domain I contributes to the permeation pathway for serotonin and ions in

the serotonin transporter.', *The Journal of neuroscience : the official journal of the Society for Neuroscience* **19**, 4705–17.

Barker, E. L., Perlman, M. A., Adkins, E. M., Houlihan, W. J., Pristupa, Z. B., Niznik, H. B. and Blakely, R. D. (1998), 'High affinity recognition of serotonin transporter antagonists defined by species-scanning mutagenesis. An aromatic residue in transmembrane domain I dictates species-selective recognition of citalopram and mazindol', *J Biol Chem* **273**(31), 19459–68.

Barman, P., Reddy, D. and Bhaumik, S. R. (2019), 'Mechanisms of Antisense Transcription Initiation with Implications in Gene Expression, Genomic Integrity and Disease Pathogenesis', *Noncoding RNA* **5**(1).

Barnett, D. W., Garrison, E. K., Quinlan, A. R., Stromberg, M. P. and Marth, G. T. (2011), 'Bamtools: a C++ API and toolkit for analyzing and managing BAM files', *Bioinformatics* **27**(12), 1691–2.

Bennett, H. M., Mok, H. P., Gkrania-Klotsas, E., Tsai, I. J., Stanley, E. J., Antoun, N. M., Coghlan, A., Harsha, B., Traini, A., Ribeiro, D. M., Steinbiss, S., Lucas, S. B., Allinson, K. S., Price, S. J., Santarius, T. S., Carmichael, A. J., Chiodini, P. L., Holroyd, N., Dean, A. F. and Berriman, M. (2014), 'The genome of the sparganosis tapeworm *Spirometra erinaceieuropaei* isolated from the biopsy of a migrating brain lesion', *Genome Biol* **15**(11), 510.

Berger, M., Gray, J. A. and Roth, B. L. (2009), 'The expanded biology of serotonin', *Annu Rev Med* **60**, 355–66.

Berriman, M., Haas, B. J., LoVerde, P. T., Wilson, R. A., Dillon, G. P., Cerqueira, G. C., Mashiyama, S. T., Al-Lazikani, B., Andrade, L. F., Ashton, P. D., Aslett, M. A., Bartholomeu, D. C., Blandin, G., Caffrey, C. R., Coghlan, A., Coulson, R., Day, T. A., Delcher, A., DeMarco, R., Djikeng, A., Eyre, T., Gamble, J. A., Ghedin, E., Gu, Y., Hertz-Fowler, C., Hirai, H., Hirai, Y., Houston, R., Ivens, A., Johnston, D. A., Lacerda, D., Macedo, C. D., McVeigh, P., Ning, Z., Oliveira, G., Overington, J. P., Parkhill, J., Pertea, M., Pierce, R. J., Protasio, A. V., Quail, M. A., Rajandream, M. A., Rogers, J., Sajid, M., Salzberg, S. L., Stanke, M., Tivey, A. R., White, O., Williams, D. L., Wortman, J., Wu, W., Zamanian, M., Zerlotini, A., Fraser-Liggett, C. M., Barrell, B. G. and El-Sayed, N. M. (2009), 'The genome of the blood fluke *Schistosoma mansoni*', *Nature* **460**(7253), 352–8.

Berumen, L. C., Rodriguez, A., Miledi, R. and Garcia-Alcocer, G. (2012), 'Serotonin receptors in hippocampus', *ScientificWorldJournal* **2012**, 823493.

Biscotti, M. A., Olmo, E. and Heslop-Harrison, J. S. (2015), 'Repetitive dna in eukaryotic genomes', *Chromosome Res* **23**(3), 415–20.

Blackburn, E. H. (2001), 'Switching and signaling at the telomere.', *Cell* **106**(6), 661–673.

Bochtler, M., Kolano, A. and Xu, G. L. (2017), 'DNA demethylation pathways: Additional players and regulators', *Bioessays* **39**(1), 1–13.

Bombarova, M., Vitkova, M., Spakulova, M. and Koubkova, B. (2009), 'Telomere analysis of platyhelminths and acanthocephalans by FISH and Southern hybridization', *Genome* **52**(11), 897–903.

Bonvicini, F., Manaresi, E., Di Furio, F., De Falco, L. and Gallinella, G. (2012), 'Parvovirus b19 dna cpg dinucleotide methylation and epigenetic regulation of viral expression', *PLoS One* **7**(3), e33316.

Bossin, H., Fournier, P., Royer, C., Barry, P., Cerutti, P., Gimenez, S., Couble, P. and Bergoin, M. (2003), '*Junonia coenia* densovirus-based vectors for stable transgene expression in Sf9 cells: influence of the densovirus sequences on genomic integration', *J Virol* **77**(20), 11060–71.

Boyle, J. P. and Yoshino, T. P. (2005), 'Serotonin-induced muscular activity in *Schistosoma mansoni* larval stages: importance of 5-HT transport and role in daughter sporocyst production', *J Parasitol* **91**(3), 542–50.

Boyle, J. P., Zaide, J. V. and Yoshino, T. P. (2000), '*Schistosoma mansoni*: effects of serotonin and serotonin receptor antagonists on motility and length of primary sporocysts *in vitro*', *Exp Parasitol* **94**(4), 217–26.

Bray, N. L., Pimentel, H., Melsted, P. and Pachter, L. (2016), 'Near-optimal probabilistic RNA-seq quantification', *Nat Biotechnol* **34**(5), 525–7.

Brazvan, B., Ebrahimi-Kalan, A., Velaei, K., Mehdipour, A., Aliyari Serej, Z., Ebrahimi, A., Ghorbani, M., Cheraghi, O. and Nozad Charoudeh, H. (2018), 'Telomerase activity and telomere on stem progeny senescence', *Biomed Pharmacother* **102**, 9–17.

Brehm, K. (2010), 'The role of evolutionarily conserved signalling systems in *Echinococcus multilocularis* development and host-parasite interaction', *Med Microbiol Immunol* **199**(3), 247–59.

Brehm, K., Jensen, K. and Frosch, M. (2000), 'mRNA trans-splicing in the human parasitic cestode *Echinococcus multilocularis*', *J Biol Chem* **275**(49), 38311–8.

Brehm, K. and Koziol, U. (2014), 'On the importance of targeting parasite stem cells in anti-echinococcosis drug development', *Parasite* **21**, 72.

Brehm, K. and Koziol, U. (2017), '*Echinococcus*-Host Interactions at Cellular and Molecular Levels', *Adv Parasitol* **95**, 147–212.

Brehm, K., Kronthaler, K., Jura, H. and Frosch, M. (2000), 'Cloning and characterization of beta-tubulin genes from *Echinococcus multilocularis*', *Mol Biochem Parasitol* **107**(2), 297–302.

Brehm, K., Spiliotis, M., Zavala-Gongora, R., Konrad, C. and Frosch, M. (2006), 'The molecular mechanisms of larval cestode development: first steps into an unknown world', *Parasitol Int* **55 Suppl**, S15–21.

Brehm, K., Wolf, M., Beland, H., Kroner, A. and Frosch, M. (2003), 'Analysis of differential gene expression in *Echinococcus multilocularis* larval stages by means of spliced leader differential display', *Int J Parasitol* **33**(11), 1145–59.

Brembeck, F. H., Rosario, M. and Birchmeier, W. (2006), 'Balancing cell adhesion and wnt signaling, the key role of beta-catenin', *Curr Opin Genet Dev* **16**(1), 51–9.

Brownlee, D. J., Fairweather, I., Johnston, C. F. and Rogan, M. T. (1994), 'Immunocytochemical localization of serotonin (5-HT) in the nervous system of the hydatid organism, *Echinococcus granulosus* (Cestoda, Cyclophyllidea)', *Parasitology* **109 (Pt 2)**, 233–41.

Brunetti, E., Kern, P. and Vuitton, D. A. (2010), 'Expert consensus for the diagnosis and treatment of cystic and alveolar echinococcosis in humans', *Acta Trop* **114**(1), 1–16.

Buznikov, G. A., Lambert, H. W. and Lauder, J. M. (2001), 'Serotonin and serotonin-like substances as regulators of early embryogenesis and morphogenesis', *Cell Tissue Res* **305**(2), 177–86.

Camicia, F., Celentano, A. M., Johns, M. E., Chan, J. D., Maldonado, L., Vaca, H., Di Siervi, N., Kamentezky, L., Gamo, A. M., Ortega-Gutierrez, S., Martin-Fontecha, M., Davio, C., Marchant, J. S. and Rosenzvit, M. C. (2018), 'Unique pharmacological properties of serotoninergic G-protein coupled receptors from cestodes', *PLoS Negl Trop Dis* **12**(2), e0006267.

Camicia, F., Herz, M., Prada, L. C., Kamenetzky, L., Simonetta, S. H., Cucher, M. A., Bianchi, J. I., Fernandez, C., Brehm, K. and Rosenzvit, M. C. (2013), 'The nervous and prenervous roles of serotonin in *Echinococcus spp*', *Int J Parasitol* .

Campos, L. (2018), Molecular characterization of DNA methylation in *Echinococcus multilocularis*, Master's thesis.

Canovas, S., Ross, P. J., Kelsey, G. and Coy, P. (2017), 'DNA Methylation in Embryo Development: Epigenetic Impact of ART (Assisted Reproductive Technologies)', *Bioessays* **39**(11).

Chemale, G., Ferreira, H. B., Barrett, J., Brophy, P. M. and Zaha, A. (2005), '*Echinococcus granulosus* antigen B hydrophobic ligand binding properties', *Biochim Biophys Acta* **1747**(2), 189–94.

Chemale, G., Haag, K. L., Ferreira, H. B. and Zaha, A. (2001), '*Echinococcus granulosus* antigen B is encoded by a gene family', *Mol Biochem Parasitol* **116**(2), 233–7.

Collins, J. J., r., Wang, B., Lambrus, B. G., Tharp, M. E., Iyer, H. and Newmark, P. A. (2013), 'Adult somatic stem cells in the human parasite *Schistosoma mansoni*', *Nature* **494**(7438), 476–9.

Collins, K. and Mitchell, J. R. (2002), 'Telomerase in the human organism.', *Oncogene* **21**(4), 564–579.

Cote, F., Thevenot, E., Fligny, C., Fromes, Y., Darmon, M., Ripoche, M. A., Bayard, E., Hanoun, N., Saurini, F., Lechat, P., Dandolo, L., Hamon, M., Mallet, J. and Vodjdani, G. (2003), 'Disruption of the nonneuronal tph1 gene demonstrates the importance of peripheral serotonin in cardiac function', *Proc Natl Acad Sci U S A* **100**(23), 13525–30.

Cotmore, S. F., Agbandje-McKenna, M., Chiorini, J. A., Mukha, D. V., Pintel, D. J., Qiu, J., Soderlund-Venermo, M., Tattersall, P., Tijssen, P., Gatherer, D. and Davison, A. J. (2014), 'The family parvoviridae', *Arch Virol* **159**(5), 1239–47.

Cotmore, S. F. and Tattersall, P. (2013), 'Parvovirus diversity and DNA damage responses', *Cold Spring Harb Perspect Biol* **5**(2).

Craig, P. (2003), '*Echinococcus multilocularis*', *Curr Opin Infect Dis* **16**(5), 437–44.

Culpepper, L. (2012), 'The use of monoamine oxidase inhibitors in primary care', *J Clin Psychiatry* **73 Suppl 1**, 37–41.

Damm, K., Hemmann, U., Garin-Chesa, P., Hauel, N., Kauffmann, I., Priepke, H., Niestroj, C., Daiber, C., Enenkel, B., Guilliard, B., Lauritsch, I., Muller, E., Pascolo, E., Sauter, G., Pantic, M., Martens, U. M., Wenz, C., Lingner, J., Kraut, N., Rettig, W. J. and Schnapp, A. (2001), 'A highly selective telomerase inhibitor limiting human cancer cell proliferation', *EMBO J* **20**(24), 6958–68.

Day, T. A., Bennett, J. L. and Pax, R. A. (1994), 'Serotonin and its requirement for maintenance of contractility in muscle fibres isolated from *Schistosoma mansoni*', *Parasitology* **108 (Pt 4)**, 425–32.

de Saram, P. S., Ressurreicao, M., Davies, A. J., Rollinson, D., Emery, A. M. and Walker, A. J. (2013), 'Functional mapping of protein kinase A reveals its importance in adult *Schistosoma mansoni* motor activity', *PLoS Negl Trop Dis* **7**(1), e1988.

Dixon, J. B. (1997), 'Echinococcosis', *Comp Immunol Microbiol Infect Dis* **20**(1), 87–94.

Dominguez, M. F., Koziol, U., Porro, V., Costabile, A., Estrade, S., Tort, J., Bollati-Fogolin, M. and Castillo, E. (2014), 'A new approach for the characterization of proliferative cells in cestodes', *Exp Parasitol* **138**, 25–9.

Du, Q., Luu, P. L., Stirzaker, C. and Clark, S. J. (2015), 'Methyl-CpG-binding domain proteins: readers of the epigenome', *Epigenomics* **7**(6), 1051–73.

Dupont, F., Tenenbaum, L., Guo, L. P., Spegelaere, P., Zeicher, M. and Rommelaere, J. (1994), 'Use of an autonomous parvovirus vector for selective transfer of a foreign gene into transformed human cells of different tissue origins and its expression therein', *J Virol* **68**(3), 1397–406.

Eckert, J. and Deplazes, P. (2004), 'Biological, epidemiological, and clinical aspects of echinococcosis, a zoonosis of increasing concern', *Clin Microbiol Rev* **17**(1), 107–35.

Edgar, R. C. (2004*a*), 'MUSCLE: a multiple sequence alignment method with reduced time and space complexity', *BMC Bioinformatics* **5**, 113.

Edgar, R. C. (2004*b*), 'MUSCLE: multiple sequence alignment with high accuracy and high throughput', *Nucleic Acids Res* **32**(5), 1792–7.

Ehlers, U. (1986), 'Comments on a phylogenetic system of the Platyhelminthes', *Hydrobiologia* **132**(1), 1–12.

Eisenhoffer, G. T., Kang, H. and Sanchez Alvarado, A. (2008), 'Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea*', *Cell Stem Cell* **3**(3), 327–39.

El-Daly, H., Kull, M., Zimmermann, S., Pantic, M., Waller, C. F. and Martens, U. M. (2005), 'Selective cytotoxicity and telomere damage in leukemia cells using the telomerase inhibitor BIBR1532', *Blood* **105**(4), 1742–9.

Ersfeld, K. and Craig, P. S. (1995), 'Cloning and immunological characterisation of *Echinococcus granulosus* ferritin', *Parasitol Res* **81**(5), 382–7.

Estey, S. J. and Mansour, T. E. (1987), 'Nature of serotonin-activated adenylate cyclase during development of *Schistosoma mansoni*', *Mol Biochem Parasitol* **26**(1-2), 47–59.

Fairweather, I., Macartney, G. A., Johnston, C. F., Halton, D. W. and Buchnan, K. D. (1988), 'Immunocytochemical demonstration of 5-hydroxytryptamine (serotonin) and vertebrate neuropeptides in the nervous system of excysted cysticercoid larvae of the rat tapeworm, *Hymenolepis diminuta* (Cestoda, Cyclophyllidea)', *Parasitol Res* **74**(4), 371–9.

Fairweather, I., Maule, A. G., Mitchell, S. H., Johnston, C. F. and Halton, D. W. (1987), 'Immunocytochemical demonstration of 5-hydroxytryptamine (serotonin) in the nervous system of the liver fluke, *Fasciola hepatica* (Trematoda, Digenea)', *Parasitol Res* **73**(3), 255–8.

Fairweather, I., McMullan, M. T., Johnston, C. F., Rogan, M. T. and Hanna, R. E. (1994), 'Serotoninergic and peptidergic nerve elements in the protoscolex of *Echinococcus granulosus* (Cestoda, Cyclophyllidea)', *Parasitol Res* **80**(8), 649–56.

Fanburg, B. L. and Lee, S. L. (1997), 'A new role for an old molecule: serotonin as a mitogen', *Am J Physiol* **272**(5 Pt 1), L795–806.

Felsenstein, J. (1985), 'Confidence Limits on Phylogenies: An Approach Using the Bootstrap', *Evolution* **39**(4), 783–791.

Fernandez, C., Gregory, W. F., Loke, P. and Maizels, R. M. (2002), 'Full-length-enriched cDNA libraries from *Echinococcus granulosus* contain separate populations of oligo-capped and trans-spliced transcripts and a high level of predicted signal peptide sequences', *Mol Biochem Parasitol* **122**(2), 171–80.

Fernandez, R., Tabarini, D., Azpiazu, N., Frasch, M. and Schlessinger, J. (1995), 'The *Drosophila* insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential', *EMBO J* **14**(14), 3373–84.

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J. and Bateman, A. (2016), 'The pfam protein families database: towards a more sustainable future', *Nucleic Acids Res* **44**(D1), D279–85.

Flo, M., Margenat, M., Pellizza, L., Grana, M., Duran, R., Baez, A., Salceda, E., Soto, E., Alvarez, B. and Fernandez, C. (2017), 'Functional diversity of secreted cestode Kunitz proteins: Inhibition of serine peptidases and blockade of cation channels', *PLoS Pathog* **13**(2), e1006169.

Fontana, A. C., Sonders, M. S., Pereira-Junior, O. S., Knight, M., Javitch, J. A., Rodrigues, V., Amara, S. G. and Mortensen, O. V. (2009), 'Two allelic isoforms of the serotonin transporter from *Schistosoma mansoni* display electrogenic transport and high selectivity for serotonin', *Eur J Pharmacol* **616**(1-3), 48–57.

Fontenla, S., Rinaldi, G., Smircich, P. and Tort, J. F. (2017), 'Conservation and diversification of small rna pathways within flatworms', *BMC Evol Biol* **17**(1), 215.

Franquinet, R. (1979), '[The role of serotonin and catecholamines in the regeneration of the Planaria *Polycelis tenvis*]', *J Embryol Exp Morphol* **51**, 85–95.

Frayha, G. J. (1971), 'Comparative metabolism of acetate in the taeniid tapeworms *Echinococcus granulosus*, *E. multilocularis* and *Taenia hydatigena*', *Comp Biochem Physiol B* **39**(1), 167–70.

Fredriksson, R., Lagerström, M. C., Lundin, L. G. and Schiöth, H. B. (2003), 'The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints', *Mol Pharmacol* **63**(6), 1256–72.

Fujiwara, T. (2017), 'GATA Transcription Factors: Basic Principles and Related Human Disorders', *Tohoku J Exp Med* **242**(2), 83–91.

Gelmedin, V. (2008), Targeting flatworm signaling cascades for the development of novel antihelminthic drugs, PhD thesis.

Gelmedin, V., Caballero-Gamiz, R. and Brehm, K. (2008), 'Characterization and inhibition of a p38-like mitogen-activated protein kinase (MAPK) from *Echinococcus multilocularis*: antiparasitic activities of p38 MAPK inhibitors', *Biochem Pharmacol* **76**(9), 1068–81.

Gentleman, R., Carey, V., Huber, W. and Hahne, F. (2017), *genefilter: genefilter: methods for filtering genes from high-throughput experiments*. R package version 1.58.1.

Geyer, K. K., Chalmers, I. W., Mackintosh, N., Hirst, J. E., Geoghegan, R., Badets, M., Brophy, P. M., Brehm, K. and Hoffmann, K. F. (2013), 'Cytosine methylation is a conserved epigenetic feature found throughout the phylum Platyhelminthes', *BMC Genomics* **14**, 462.

Geyer, K. K., Munshi, S. E., Vickers, M., Squance, M., Wilkinson, T. J., Berrar, D., Chaparro, C., Swain, M. T. and Hoffmann, K. F. (2018), 'The anti-fecundity effect of 5-azacytidine (5-azac) on *Schistosoma mansoni* is linked to dis-regulated transcription, translation and stem cell activities', *Int J Parasitol Drugs Drug Resist* **8**(2), 213–222.

Geyer, K. K., Munshi, S. E., Whiteland, H. L., Fernandez-Fuentes, N., Phillips, D. W. and Hoffmann, K. F. (2018), 'Methyl-CpG-binding (SmMBD2/3) and chromobox (SmCBX) proteins are required for neoblast proliferation and oviposition in the parasitic blood fluke *Schistosoma mansoni*', *PLoS Pathog* **14**(6), e1007107.

Geyer, K. K., Rodriguez Lopez, C. M., Chalmers, I. W., Munshi, S. E., Truscott, M., Heald, J., Wilkinson, M. J. and Hoffmann, K. F. (2011), 'Cytosine methylation regulates oviposition in the pathogenic blood fluke *Schistosoma mansoni*', *Nat Commun* **2**, 424.

Giribet, G. (2008), 'Assembling the lophotrochozoan (=spiralian) tree of life', *Philos Trans R Soc Lond B Biol Sci* **363**(1496), 1513–22.

Gomes, N. M., Shay, J. W. and Wright, W. E. (2010), 'Telomere biology in Metazoa', *FEBS Lett* **584**(17), 3741–51.

Gomez-Puerta, L., Vargas-Calla, A., Castillo, Y., Lopez-Urbina, M. T., Dorny, P., Garcia, H. H., Gonzalez, A. E. and O'Neal, S. E. (2019), 'Evaluation of cross-reactivity to taenia hydatigena and *Echinococcus granulosus* in the enzyme-linked immunoelectrotransfer blot assay for the diagnosis of porcine cysticercosis', *Parasit Vectors* **12**(1), 57.

Gurley, K. A., Rink, J. C. and Sanchez Alvarado, A. (2008), 'Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis', *Science* **319**(5861), 323–7.

Gustafsson, M. K. (1987), 'Immunocytochemical demonstration of neuropeptides and serotonin in the nervous systems of adult *Schistosoma mansoni*', *Parasitol Res* **74**(2), 168–74.

Gustafsson, M. K., Fagerholm, H. P., Halton, D. W., Hanzelova, V., Maule, A. G., Reuter, M. and Shaw, C. (1995), 'Neuropeptides and serotonin in the cestode, *Proteocephalus exiguus*: an immunocytochemical study', *Int J Parasitol* **25**(6), 673–82.

Gustafsson, M. K., Wikgren, M. C., Karhi, T. J. and Schot, L. P. (1985), 'Immunocytochemical demonstration of neuropeptides and serotonin in the tapeworm *Diphyllobothrium dendriticum*', *Cell Tissue Res* **240**(2), 255–60.

Haider, S. and Pal, R. (2013), 'Integrated analysis of transcriptomic and proteomic data', *Curr Genomics* **14**(2), 91–110.

Hall, T. A. (1999), 'BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT', *Nucleic Acids Symposium Series* **41**, 95–98.

Han, Y., Wang, Q., Qiu, Y., Wu, W., He, H., Zhang, J., Hu, Y. and Zhou, X. (2013), 'Periplaneta fuliginosa densovirus nonstructural protein ns1 contains an endonuclease activity that is regulated by its phosphorylation', *Virology* **437**(1), 1–11.

Han, Z. S., Enslen, H., Hu, X., Meng, X., Wu, I. H., Barrett, T., Davis, R. J. and Ip, Y. T. (1998), 'A conserved p38 mitogen-activated protein kinase pathway regulates *Drosophila* immunity gene expression', *Mol Cell Biol* **18**(6), 3527–39.

Hancock, K., Pattabhi, S., Greene, R. M., Yushak, M. L., Williams, F., Khan, A., Priest, J. W., Levine, M. Z. and Tsang, V. C. (2004), 'Characterization and cloning of GP50, a *Taenia solium* antigen diagnostic for cysticercosis', *Mol Biochem Parasitol* **133**(1), 115–24.

Harder, A., Abbink, J., Andrews, P. and Thomas, H. (1987), 'Praziquantel impairs the ability of exogenous serotonin to stimulate carbohydrate metabolism in intact *Schistosoma mansoni*', *Parasitol Res* **73**(5), 442–5.

Harder, A., Andrews, P. and Thomas, H. (1987), 'Chlorpromazine, other amphiphilic cationic drugs and praziquantel: effects on carbohydrate metabolism of *Schistosoma mansoni*', *Parasitol Res* **73**(3), 245–9.

Hastings, K. E. (2005), 'SL trans-splicing: easy come or easy go?', *Trends Genet* **21**(4), 240–7.

Hayashi, T. and Agata, K. (2012), 'A unique FACS method to isolate stem cells in planarian', *Methods Mol Biol* **879**, 29–37.

Hayashi, T. and Agata, K. (2018), 'A Subtractive FACS Method for Isolation of Planarian Stem Cells and Neural Cells', *Methods Mol Biol* **1774**, 467–478.

Heath, D. D. and Osborn, P. J. (1976), 'Formation of *Echinococcus granulosus* laminated membrane in a defined medium', *Int J Parasitol* **6**(6), 467–71.

Hemer, S., Konrad, C., Spiliotis, M., Koziol, U., Schaack, D., Förster, S., Gelmedin, V., Stadelmann, B., Dandekar, T., Hemphill, A. and Brehm, K. (2014), 'Host insulin stimulates *Echinococcus multilocularis* insulin signalling pathways and larval development', *BMC Biol* **12**, 5.

Hemphill, A. and Gottstein, B. (1995), 'Immunology and morphology studies on the proliferation of *in vitro* cultivated *Echinococcus multilocularis* metacestodes', *Parasitol Res* **81**(7), 605–14.

Hemphill, A., Stadelmann, B., Scholl, S., Müller, J., Spiliotis, M., Muller, N., Gottstein, B. and Siles-Lucas, M. (2010), '*Echinococcus metacestodes* as laboratory models for the screening of drugs against cestodes and trematodes', *Parasitology* **137**(3), 569–87.

Henry, L. K., Field, J. R., Adkins, E. M., Parnas, M. L., Vaughan, R. A., Zou, M. F., Newman, A. H. and Blakely, R. D. (2006), 'Tyr-95 and Ile-172 in transmembrane segments 1 and 3 of human serotonin transporters interact to establish high affinity recognition of antidepressants', *J Biol Chem* **281**(4), 2012–23.

Herz, M. (2015), Molecular characterization of the serotonin and cAMP-signalling pathways in *Echinococcus*, MD thesis.

Hirai, H. and LoVerde, P. T. (1996), 'Identification of the telomeres on schistosoma mansoni chromosomes by FISH', *J Parasitol* **82**(3), 511–2.

Horschitz, S., Hummerich, R. and Schloss, P. (2001), 'Structure, function and regulation of the 5-hydroxytryptamine (serotonin) transporter', *Biochem Soc Trans* **29**(Pt 6), 728–32.

Howe, K. L., Bolt, B. J., Cain, S., Chan, J., Chen, W. J., Davis, P., Done, J., Down, T., Gao, S., Grove, C., Harris, T. W., Kishore, R., Lee, R., Lomax, J., Li, Y., Muller, H. M., Nakamura, C., Nuin, P., Paulini, M., Raciti, D., Schindelman, G., Stanley, E., Tuli, M. A., Van Auken, K., Wang, D., Wang, X., Williams, G., Wright, A., Yook, K., Berriman, M., Kersey, P., Schedl, T., Stein, L. and Sternberg, P. W. (2016), 'WormBase 2016: expanding to enable helminth genomic research', *Nucleic Acids Res* **44**(D1), D774–80.

Howe, K. L., Bolt, B. J., Shafie, M., Kersey, P. and Berriman, M. (2017), 'WormBase ParaSite - a comprehensive resource for helminth genomics', *Mol Biochem Parasitol* **215**, 2–10.

Hoyer, D., Clarke, D. E., Fozard, J. R., Hartig, P. R., Martin, G. R., Mylecharane, E. J., Saxena, P. R. and Humphrey, P. P. (1994), 'International Union of Pharmacology classification of receptors for 5-hydroxytryptamine (Serotonin)', *Pharmacol Rev* **46**(2), 157–203.

Huang, F., Dang, Z., Suzuki, Y., Horiuchi, T., Yagi, K., Kouguchi, H., Irie, T., Kim, K. and Oku, Y. (2016), 'Analysis on Gene Expression Profile in Oncospheres and Early Stage Metacestodes from *Echinococcus multilocularis*', *PLoS Negl Trop Dis* **10**(4), e0004634.

Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A. and Vingron, M. (2002), 'Variance stabilization applied to microarray data calibration and to the quantification of differential expression', *Bioinformatics* **18 Suppl 1**, S96–104.

Hubert, K., Zavala-Gongora, R., Frosch, M. and Brehm, K. (2004), 'Identification and characterization of pdz-1, a n-ermad specific interaction partner of the echinococcus multilocularis erm protein elp', *Mol Biochem Parasitol* **134**(1), 149–54.

Iglesias, M., Gomez-Skarmeta, J. L., Salo, E. and Adell, T. (2008), 'Silencing of Smed-betacatenin1 generates radial-like hypercephalized planarians', *Development* **135**(7), 1215–21.

International-Helminth-Genomes-Consortium (2019), 'Comparative genomics of the major parasitic worms', *Nat Genet* **51**(1), 163–174.

Iriondo, O., Rabano, M. and Vivanco, M. D. (2015), 'FACS Sorting Mammary Stem Cells', *Methods Mol Biol* **1293**, 63–72.

Iseki, H., Shimizukawa, R., Sugiyama, F., Kunita, S., Iwama, A., Onodera, M., Nakauchi, H. and Yagami, K. (2005), '*Parvovirus* nonstructural proteins induce an epigenetic modification through histone acetylation in host genes and revert tumor malignancy to benignancy', *J Virol* **79**(14), 8886–93.

Jabbar, A., Jenkins, D. J., Crawford, S., Walduck, A. K., Gauci, C. G. and Lightowlers, M. W. (2011), 'Oncospheral penetration glands are the source of the EG95 vaccine antigen against cystic hydatid disease', *Parasitology* **138**(1), 89–99.

Jeltsch, A. (2006), 'On the enzymatic properties of Dnmt1: specificity, processivity, mechanism of linear diffusion and allosteric regulation of the enzyme', *Epigenetics* **1**(2), 63–6.

Jeltsch, A. and Jurkowska, R. Z. (2014), 'New concepts in DNA methylation', *Trends Biochem Sci* **39**(7), 310–8.

Joffee, B. I., Solovei, I. V. and Macgregor, H. C. (1996), 'Ends of chromosomes in *Polycelis tenuis* (platyhelminthes) have telomere repeat ttaggg', *Chromosome Res* **4**(4), 323–4.

Jura, H., Bader, A. and Frosch, M. (1998), '*In vitro* activities of benzimidazoles against *Echinococcus multilocularis* metacestodes', *Antimicrob Agents Chemother* **42**(5), 1052–6.

Jura, H., Bader, A., Hartmann, M., Maschek, H. and Frosch, M. (1996), 'Hepatic tissue culture model for study of host-parasite interactions in alveolar echinococcosis', *Infect Immun* **64**(9), 3484–90.

Jurkowska, R. Z., Jurkowski, T. P. and Jeltsch, A. (2011), 'Structure and function of mammalian DNA methyltransferases', *Chembiochem* **12**(2), 206–22.

Kawamoto, F., Shozawa, A., Kumada, N. and Kojima, K. (1989), 'Possible roles of cAMP and Ca2+ in the regulation of miracidial transformation in *Schistosoma mansoni*', *Parasitol Res* **75**(5), 368–74.

Kern, P. (2010), 'Clinical features and treatment of alveolar echinococcosis', *Curr Opin Infect Dis* **23**(5), 505–12.

Kern, P., Wen, H., Sato, N., Vuitton, D. A., Gruener, B., Shao, Y., Delabrousse, E., Kratzer, W. and Bresson-Hadni, S. (2006), 'WHO classification of alveolar echinococcosis: principles and application', *Parasitol Int* **55 Suppl**, S283–7.

Kim, D., Langmead, B. and Salzberg, S. L. (2015), 'HISAT: a fast spliced aligner with low memory requirements', *Nat Methods* **12**(4), 357–60.

Kines, K. J., Sokolowski, M., deHaro, D. L., Christian, C. M. and Belancio, V. P. (2014), 'Potential for genomic instability associated with retrotranspositionally-incompetent l1 loci', *Nucleic Acids Res* **42**(16), 10488–502.

Kolde, R. (2017), 'Pretty heatmaps: Implementation of heatmaps that offers more control over dimensions and appearance.'.

Konrad, C., Kroner, A., Spiliotis, M., Zavala-Gongora, R. and Brehm, K. (2003), 'Identification and molecular characterisation of a gene encoding a member of the insulin receptor family in *Echinococcus multilocularis*', *Int J Parasitol* **33**(3), 301–12.

Koziol, U. (2014), Molecular and developmental characterization of the *Echinococcus multilocularis* stem cell system, PhD thesis, Universität Würzburg.

Koziol, U., Jarero, F., Olson, P. D. and Brehm, K. (2016), 'Comparative analysis of Wnt expression identifies a highly conserved developmental transition in flatworms.', *BMC Biol* **14**(1), 10.

Koziol, U., Krohne, G. and Brehm, K. (2013), 'Anatomy and development of the larval nervous system in *Echinococcus multilocularis*', *Front Zool* **10**(1), 24.

Koziol, U., Radio, S., Smircich, P., Zarowiecki, M., Fernandez, C. and Brehm, K. (2015), 'A Novel Terminal-Repeat Retrotransposon in Miniature (TRIM) Is Massively Expressed in *Echinococcus multilocularis* Stem Cells', *Genome Biol Evol* **7**(8), 2136–53.

Koziol, U., Rauschendorfer, T., Zanon Rodriguez, L., Krohne, G. and Brehm, K. (2014), 'The unique stem cell system of the immortal larva of the human parasite *Echinococcus multilocularis*', *Evodevo* **5**(1), 10.

Kristensen, A. S., Andersen, J., Jorgensen, T. N., Sorensen, L., Eriksen, J., Loland, C. J., Stromgaard, K. and Gether, U. (2011), 'SLC6 neurotransmitter transporters: structure, function, and regulation', *Pharmacol Rev* **63**(3), 585–640.

Kumar, P., Tan, Y. and Cahan, P. (2017), 'Understanding development and stem cells using single cell-based analyses of gene expression', *Development* **144**(1), 17–32.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018), 'MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms', *Mol Biol Evol* **35**(6), 1547–1549.

Küster, T., Stadelmann, B., Rufener, R., Risch, C., M"uller, J. and Hemphill, A. (2015), 'Oral treatments of *Echinococcus multilocularis*-infected mice with the antimalarial drug mefloquine that potentially interacts with parasite ferritin and cystatin', *Int J Antimicrob Agents* **46**(5), 546–51.

Lansdorp, P. M. (2005), 'Major cutbacks at chromosome ends.', *Trends Biochem Sci* **30**(7), 388–395.

Larsen, M. B., Elfving, B. and Wiborg, O. (2004), 'The chicken serotonin transporter discriminates between serotonin-selective reuptake inhibitors. A species-scanning mutagenesis study.', *The Journal of biological chemistry* **279**, 42147–56.

Lasda, E. L. and Blumenthal, T. (2011), 'Trans-splicing', *Wiley Interdiscip Rev RNA* **2**(3), 417–34.

Lauder, J. M. (1993), 'Neurotransmitters as growth regulatory signals: role of receptors and second messengers', *Trends Neurosci* **16**(6), 233–40.

Lee, S. L., Wang, W. W., Lanzillo, J. J. and Fanburg, B. L. (1994), 'Serotonin produces both hyperplasia and hypertrophy of bovine pulmonary artery smooth muscle cells in culture', *Am J Physiol* **266**(1 Pt 1), L46–52.

Lee, S. L., Wang, W. W., Moore, B. J. and Fanburg, B. L. (1991), 'Dual effect of serotonin on growth of bovine pulmonary artery smooth muscle cells in culture', *Circ Res* **68**(5), 1362–8.

Lesurtel, M., Soll, C., Graf, R. and Clavien, P. A. (2008), 'Role of serotonin in the hepato-gastroIntestinal tract: an old molecule for new perspectives', *Cell Mol Life Sci* **65**(6), 940–52.

Letunic, I. and Bork, P. (2018), '20 years of the SMART protein domain annotation resource', *Nucleic Acids Research* **46**(D1), D493–D496.

Letunic, I., Doerks, T. and Bork, P. (2015), 'SMART: recent updates, new developments and status in 2015', *Nucleic Acids Res* **43**(Database issue), D257–60.

Levine, A. J., Ting, D. T. and Greenbaum, B. D. (2016), 'P53 and the defenses against genome instability caused by transposons and repetitive elements', *Bioessays* **38**(6), 508–13.

Li, H. (2011), 'A statistical framework for snp calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data', *Bioinformatics* **27**(21), 2987–93.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009), 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics* **25**(16), 2078–9.

Li, P., Liu, T., Liu, J., Zhang, Q., Lou, F., Kong, F., Cheng, G., Bjorkholm, M., Zheng, C. and Xu, D. (2014), 'Promoter polymorphism in the serotonin transporter (5-HTT) gene is significantly associated with leukocyte telomere length in Han Chinese', *PLoS One* **9**(4), e94442.

Li, T., Chen, X., Zhen, R., Qiu, J., Qiu, D., Xiao, N., Ito, A., Wang, H., Giraudoux, P., Sako, Y., Nakao, M. and Craig, P. S. (2010), 'Widespread co-endemicity of human cystic and alveolar echinococcosis on the eastern Tibetan Plateau, northwest Sichuan/southeast Qinghai, China', *Acta Trop* **113**(3), 248–56.

Li, W., Liu, B., Yang, Y., Ren, Y., Wang, S., Liu, C., Zhang, N., Qu, Z., Yang, W., Zhang, Y., Yan, H., Jiang, F., Li, L., Li, S., Jia, W., Yin, H., Cai, X., Liu, T., McManus, D. P., Fan, W. and Fu, B. (2018), 'The genome of tapeworm *Taenia multiceps* sheds light on understanding parasitic mechanism and control of coenurosis disease', *DNA Res* **25**(5), 499–510.

Littlewood, D. T. J., Rohde, K. and Clough, K. A. (1999), 'The interrelationships of all major groups of Platyhelminthes: phylogenetic evidence from morphology and molecules', *Biological Journal of the Linnean Society* **66**(1), 75–114.

Liu, H., Fu, Y., Xie, J., Cheng, J., Ghabrial, S. A., Li, G., Peng, Y., Yi, X. and Jiang, D. (2011), 'Widespread endogenization of densoviruses and parvoviruses in animal and human genomes', *J Virol* **85**(19), 9863–76.

Liu, Q., Yang, Q., Sun, W., Vogel, P., Heydorn, W., Yu, X. Q., Hu, Z., Yu, W., Jonas, B., Pineda, R., Calderon-Gay, V., Germann, M., O'Neill, E., Brommage, R., Cullinan, E., Platt, K., Wilson, A., Powell, D., Sands, A., Zambrowicz, B. and Shi, Z. C. (2008), 'Discovery and characterization of novel tryptophan hydroxylase inhibitors that selectively inhibit serotonin synthesis in the gastrointestinal tract', *J Pharmacol Exp Ther* **325**(1), 47–55.

Love, M. I., Huber, W. and Anders, S. (2014), 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol* **15**(12), 550.

Lovenberg, W., Jequier, E. and Sjoerdsma, A. (1967), 'Tryptophan hydroxylation: measurement in pineal gland, brainstem, and carcinoid tumor', *Science* **155**(3759), 217–9.

Lymbery, A. J., Jenkins, E. J., Schurer, J. M. and Thompson, R. C. (2015*a*), 'Response to nakao et al. - is *Echinococcus intermedius* a valid species?', *Trends Parasitol* **31**(8), 343–4.

Lymbery, A. J., Jenkins, E. J., Schurer, J. M. and Thompson, R. C. (2015*b*), '*Echinococcus canadensis*, *E. borealis*, and *E. intermedius*. what's in a name?', *Trends Parasitol* **31**(1), 23–9.

Lynn, F. C., Sanchez, L., Gomis, R., German, M. S. and Gasa, R. (2008), 'Identification of the bHLH factor Math6 as a novel component of the embryonic pancreas transcriptional network', *PLoS One* **3**(6), e2430.

Maksakova, I. A., Mager, D. L. and Reiss, D. (2008), 'Keeping active endogenous retroviral-like elements in check: the epigenetic perspective', *Cell Mol Life Sci* **65**(21), 3329–47.

Maldonado, L. L., Assis, J., Araujo, F. M., Salim, A. C., Macchiaroli, N., Cucher, M., Camicia, F., Fox, A., Rosenzvit, M., Oliveira, G. and Kamenetzky, L. (2017), 'The *Echinococcus canadensis* (G7) genome: a key knowledge of parasitic platyhelminth human diseases', *BMC Genomics* **18**(1), 204.

Martin, N., Benhamed, M., Nacerddine, K., Demarque, M. D., van Lohuizen, M., Dejean, A. and Bischof, O. (2012), 'Physical and functional interaction between PML and TBX2 in the establishment of cellular senescence', *Embo j* **31**(1), 95–109.

Martynova, E. U., Schal, C. and Mukha, D. V. (2016), 'Effects of recombination on densovirus phylogeny', *Arch Virol* **161**(1), 63–75.

Massah, S., Beischlag, T. V. and Prefontaine, G. G. (2015), 'Epigenetic events regulating monoallelic gene expression', *Crit Rev Biochem Mol Biol* **50**(4), 337–58.

McCloskey, D. J., Postolache, T. T., Vittone, B. J., Nghiem, K. L., Monsale, J. L., Wesley, R. A. and Rick, M. E. (2008), 'Selective serotonin reuptake inhibitors: measurement of effect on platelet function', *Transl Res* **151**(3), 168–72.

McKay, D. M., Fairweather, I., Johnston, C. F., Shaw, C. and Halton, D. W. (1991), 'Immunocytochemical and radioimmunometrical demonstration of serotonin- and neuropeptide-immunoreactivities in the adult rat tapeworm, *Hymenolepis diminuta* (Cestoda, Cyclophyllidea)', *Parasitology* **103 Pt 2**, 275–89.

Meier, K. and Recillas-Targa, F. (2017), 'New insights on the role of DNA methylation from a global view', *Front Biosci (Landmark Ed)* **22**, 644–668.

Mikels, A. J. and Nusse, R. (2006), 'Wnts as ligands: processing, secretion and reception', *Oncogene* **25**(57), 7461–8.

Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H. and Kakutani, T. (2001), 'Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*', *Nature* **411**(6834), 212–4.

Mizukami, C., Spiliotis, M., Gottstein, B., Yagi, K., Katakura, K. and Oku, Y. (2010), 'Gene silencing in *Echinococcus multilocularis* protoscoleces using RNA interference', *Parasitol Int* **59**(4), 647–52.

Moro, P. and Schantz, P. M. (2009), 'Echinococcosis: a review', *Int J Infect Dis* **13**(2), 125–33.

Mortensen, O. V., Kristensen, A. S. and Wiborg, O. (2001), 'Species-scanning mutagenesis of the serotonin transporter reveals residues essential in selective, high-affinity recognition of antidepressants.', *Journal of neurochemistry* **79**, 237–47.

Mossner, R. and Lesch, K. P. (1998), 'Role of serotonin in the immune system and in neuroimmune interactions', *Brain Behav Immun* **12**(4), 249–71.

Murray, S. C. and Mellor, J. (2016), 'Using both strands: The fundamental nature of antisense transcription', *Bioarchitecture* **6**(1), 12–21.

Neubauer, H. A., Hansen, C. G. and Wiborg, O. (2006), 'Dissection of an allosteric mechanism on the serotonin transporter: a cross-species study', *Mol Pharmacol* **69**(4), 1242–50.

Nishimura, K., Kitamura, Y., Inoue, T., Umesono, Y., Yoshimoto, K., Takeuchi, K., Taniguchi, T. and Agata, K. (2007), 'Identification and distribution of tryptophan hydroxylase (TPH)-positive neurons in the planarian *Dugesia japonica*', *Neurosci Res* **59**(1), 101–6.

Oh, I. H. and Reddy, E. P. (1999), 'The myb gene family in cell growth, differentiation and apoptosis', *Oncogene* **18**(19), 3017–33.

Olson, P. D., Zarowiecki, M., Kiss, F. and Brehm, K. (2012), 'Cestode genomics - progress and prospects for advancing basic and applied aspects of flatworm biology', *Parasite Immunol* **34**(2-3), 130–50.

Oriol, C. and Oriol, R. (1975), 'Physiocochemical properties of a lipoprotein antigen of *Echinococcus granulosus*', *Am J Trop Med Hyg* **24**(1), 96–100.

Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D. and Zamore, P. D. (2019), 'Piwi-interacting rnas: small rnas with big functions', *Nat Rev Genet* **20**(2), 89–108.

Parkinson, J., Wasmuth, J. D., Salinas, G., Bizarro, C. V., Sanford, C., Berriman, M., Ferreira, H. B., Zaha, A., Blaxter, M. L., Maizels, R. M. and Fernandez, C. (2012), 'A transcriptomic analysis of *Echinococcus granulosus* larval stages: implications for parasite biology and host adaptation', *PLoS Negl Trop Dis* **6**(11), e1897.

Parks, C. G., Miller, D. B., McCanlies, E. C., Cawthon, R. M., Andrew, M. E., DeRoo, L. A. and Sandler, D. P. (2009), 'Telomere length, current perceived stress, and urinary stress hormones in women', *Cancer Epidemiol Biomarkers Prev* **18**(2), 551–60.

Pascolo, E., Wenz, C., Lingner, J., Hauel, N., Priepke, H., Kauffmann, I., Garin-Chesa, P., Rettig, W. J., Damm, K. and Schnapp, A. (2002), 'Mechanism of human telomerase inhibition by BIBR1532, a synthetic, non-nucleosidic drug candidate', *J Biol Chem* **277**(18), 15566–72.

Patil, V., Ward, R. L. and Hesson, L. B. (2014), 'The evidence for functional non-CpG methylation in mammalian cells', *Epigenetics* **9**(6), 823–8.

Patocka, N. and Ribeiro, P. (2013), 'The functional role of a serotonin transporter in *Schistosoma mansoni* elucidated through immunolocalization and RNA interference (RNAi)', *Mol Biochem Parasitol* **187**(1), 32–42.

Petersen, C. P. and Reddien, P. W. (2008), 'Smed-betacatenin-1 is required for anteroposterior blastema polarity in planarian regeneration', *Science* **319**(5861), 327–30.

Pfaffl, M. W. (2001), 'A new mathematical model for relative quantification in real-time RT-PCR.', *Nucleic acids research* **29**, e45.

Protasio, A. V., Tsai, I. J., Babbage, A., Nichol, S., Hunt, M., Aslett, M. A., De Silva, N., Velarde, G. S., Anderson, T. J., Clark, R. C., Davidson, C., Dillon, G. P., Holroyd, N. E., LoVerde, P. T., Lloyd, C., McQuillan, J., Oliveira, G., Otto, T. D., Parker-Manuel, S. J., Quail, M. A., Wilson, R. A., Zerlotini, A., Dunne, D. W. and Berriman, M. (2012), 'A systematically improved high quality genome and transcriptome of the human blood fluke *Schistosoma mansoni*', *PLoS Negl Trop Dis* **6**(1), e1455.

Prouse, M. B. and Campbell, M. M. (2012), 'The interaction between MYB proteins and their target DNA binding sites', *Biochim Biophys Acta* **1819**(1), 67–77.

R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
URL: *http://www.R-project.org*

Rai, P., Safeena, M. P., Karunasagar, I. and Karunasagar, I. (2011), 'Complete nucleic acid sequence of *Penaeus stylirostris* densovirus (PstDNV) from India', *Virus Res* **158**(1-2), 37–45.

Ramakers, C., Ruijter, J. M., Deprez, R. H. L. and Moorman, A. F. M. (2003), 'Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data.', *Neuroscience letters* **339**, 62–6.

Rausch, R. L. and Bernstein, J. J. (1972), '*Echinococcus vogeli sp. n.* (cestoda: Taeniidae) from the bush dog, *Speothos venaticus* (lund)', *Z Tropenmed Parasitol* **23**(1), 25–34.

Renaud, F., Parisi, E., Capasso, A. and De Prisco, P. (1983), 'On the role of serotonin and 5-methoxy-tryptamine in the regulation of cell division in sea urchin eggs', *Dev Biol* **98**(1), 37–46.

Reuter, S., Buck, A., Manfras, B., Kratzer, W., Seitz, H. M., Darge, K., Reske, S. N. and Kern, P. (2004), 'Structured treatment interruption in patients with alveolar echinococcosis', *Hepatology* **39**(2), 509–17.

Reya, T. and Clevers, H. (2005), 'Wnt signalling in stem cells and cancer', *Nature* **434**(7035), 843–50.

Riddiford, N. and Olson, P. D. (2011), 'Wnt gene loss in flatworms', *Dev Genes Evol* **221**(4), 187–97.

Rienzo, J. A. D. (2012), 'fgStatistics. Statistical software for the analysis of experiments of functional genomics.'.
URL: *http://sites.google.com/site/fgStatistics/*

Robert Koch-Institut (06.03.2019), 'SurvStat@RKI 2.0'.
URL: *https://survstat.rki.de*

Robert-Koch-Institut (2006), 'Epidemiologie der Fuchsbandwurmerkrankungen in Deutschland - Daten des Echinokokkose Registers', *Epidemiologisches Bulletin* **15**, 283–7.

Robinson, J. T., Thorvaldsdottir, H., Wenger, A. M., Zehir, A. and Mesirov, J. P. (2017), 'Variant Review with the Integrative Genomics Viewer', *Cancer Res* **77**(21), e31–e34.

Robinson, K. A. and Lopes, J. M. (2000), 'SURVEY AND SUMMARY: *Saccharomyces cerevisiae* basic helix-loop-helix proteins regulate diverse biological processes', *Nucleic Acids Res* **28**(7), 1499–505.

Rodriguez-Caabeiro, F. and Casado, N. (1988), 'Evidence of *in vitro* germinal layer development in *Echinococcus granulosus* cysts', *Parasitol Res* **74**(6), 558–62.

Rogan, M. T., Hai, W. Y., Richardson, R., Zeyhle, E. and Craig, P. S. (2006), 'Hydatid cysts: does every picture tell a story?', *Trends Parasitol* **22**(9), 431–8.

Romig, T., Ebi, D. and Wassermann, M. (2015), 'Taxonomy and molecular epidemiology of *Echinococcus granulosus sensu lato*', *Vet Parasitol* **213**(3-4), 76–84.

Rossi, L., Salvetti, A., Marincola, F. M., Lena, A., Deri, P., Mannini, L., Batistoni, R., Wang, E. and Gremigni, V. (2007), 'Deciphering the molecular machinery of stem cells: a look at the neoblast gene expression profile', *Genome Biol* **8**(4), R62.

Roy, S. and Zhuang, Y. (2018), 'Paradoxical role of id proteins in regulating tumorigenic potential of lymphoid cells', *Front Med* **12**(4), 374–386.

Ruijter, J. M., Ramakers, C., Hoogaars, W. M. H., Karlen, Y., Bakker, O., van den Hoff, M. J. B. and Moorman, A. F. M. (2009), 'Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data.', *Nucleic acids research* **37**, e45.

Saitou, N. and Nei, M. (1987), 'The neighbor-joining method: a new method for reconstructing phylogenetic trees', *Mol Biol Evol* **4**(4), 406–25.

Sanchez Alvarado, A. (2004), 'Regeneration and the need for simpler model organisms', *Philos Trans R Soc Lond B Biol Sci* **359**(1445), 759–63.

Sandin, S. and Rhodes, D. (2014), 'Telomerase structure', *Current Opinion in Structural Biology* **25**, 104–110.

Sarrouilhe, D., Clarhaut, J., Defamie, N. and Mesnil, M. (2015), 'Serotonin and cancer: what is the link?', *Curr Mol Med* **15**(1), 62–77.

Schaefer, M. and Lyko, F. (2010), 'Solving the Dnmt2 enigma', *Chromosoma* **119**(1), 35–40.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. (2012), 'Fiji: an open-source platform for biological-image analysis.', *Nat Methods* **9**(7), 676–682.

Schindelin, J., Rueden, C. T., Hiner, M. C. and Eliceiri, K. W. (2015), 'The ImageJ ecosystem: An open platform for biomedical image analysis.', *Mol Reprod Dev* **82**(7-8), 518–529.

Schubert, A., Koziol, U., Cailliau, K., Vanderstraete, M., Dissous, C. and Brehm, K. (2014), 'Targeting *Echinococcus multilocularis* stem cells by inhibition of the polo-like kinase emplk1', *PLoS Negl Trop Dis* **8**(6), e2870.

Seuwen, K. and Pouyssegur, J. (1990), 'Serotonin as a growth factor', *Biochem Pharmacol* **39**(6), 985–90.

Shalabi, A., Fischer, C., Korf, H. W. and von Gall, C. (2013), 'Melatonin-receptor-1-deficiency affects neurogenic differentiation factor immunoreaction in pancreatic islets and enteroendocrine cells of mice', *Cell Tissue Res* **353**(3), 483–91.

Shike, H., Dhar, A. K., Burns, J. C., Shimizu, C., Jousset, F. X., Klimpel, K. R. and Bergoin, M. (2000), 'Infectious hypodermal and hematopoietic necrosis virus of shrimp is related to mosquito brevidensoviruses', *Virology* **277**(1), 167–77.

Silva-Alvarez, V., Franchini, G. R., Porfido, J. L., Kennedy, M. W., Ferreira, A. M. and Corsico, B. (2015), 'Lipid-free antigen B subunits from *Echinococcus granulosus*: oligomerization, ligand binding, and membrane interaction properties', *PLoS Negl Trop Dis* **9**(3), e0003552.

Siomi, M. C., Sato, K., Pezic, D. and Aravin, A. A. (2011), 'Piwi-interacting small rnas: the vanguard of genome defence', *Nat Rev Mol Cell Biol* **12**(4), 246–58.

Siracusano, A., Margutti, P., Delunardo, F., Profumo, E., Rigano, R., Buttari, B., Teggi, A. and Ortona, E. (2008), 'Molecular cross-talk in host-parasite relationships: the intriguing immunomodulatory role of *Echinococcus* antigen B in cystic echinococcosis', *Int J Parasitol* **38**(12), 1371–6.

Skinner, D. E., Rinaldi, G., Koziol, U., Brehm, K. and Brindley, P. J. (2014), 'How might flukes and tapeworms maintain genome integrity without a canonical piRNA pathway?', *Trends Parasitol* **30**(3), 123–9.

Smyth, J. D. (1968), '*In vitro* studies and host-specificity in *Echinococcus*', *Bull World Health Organ* **39**(1), 5–12.

Solana, J., Kao, D., Mihaylova, Y., Jaber-Hijazi, F., Malla, S., Wilson, R. and Aboobaker, A. (2012), 'Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNAseq, RNA interference and irradiation approach', *Genome Biol* **13**(3), R19.

Spiliotis, M. and Brehm, K. (2009), 'Axenic *in vitro* cultivation of *Echinococcus multilocularis* metacestode vesicles and the generation of primary cell cultures.', *Methods Mol Biol* **470**, 245–262.

Spiliotis, M., Lechner, S., Tappe, D., Scheller, C., Krohne, G. and Brehm, K. (2008), 'Transient transfection of *Echinococcus multilocularis* primary cells and complete *in vitro* regeneration of metacestode vesicles', *Int J Parasitol* **38**(8-9), 1025–39.

Spiliotis, M., Mizukami, C., Oku, Y., Kiss, F., Brehm, K. and Gottstein, B. (2010), '*Echinococcus multilocularis* primary cells: improved isolation, small-scale cultivation and RNA interference.', *Molecular and biochemical parasitology* **174**, 83–7.

Spiliotis, M., Tappe, D., Sesterhenn, L. and Brehm, K. (2004), 'Long-term *in vitro* cultivation of *Echinococcus multilocularis* metacestodes under axenic conditions', *Parasitol Res* **92**(5), 430–2.

Sugarman, E. T., Zhang, G. and Shay, J. W. (2019), 'In perspective: An update on telomere targeting in cancer', *Mol Carcinog* .

Taft, A. S., Norante, F. A. and Yoshino, T. P. (2010), 'The identification of inhibitors of *Schistosoma mansoni* miracidial transformation by incorporating a medium-throughput small-molecule screen', *Exp Parasitol* **125**(2), 84–94.

Tan, T. C., Rahman, R., Jaber-Hijazi, F., Felix, D. A., Chen, C., Louis, E. J. and Aboobaker, A. (2012), 'Telomere maintenance and telomerase activity are

differentially regulated in asexual and sexual worms', *Proc Natl Acad Sci U S A* **109**(11), 4209–14.

Tasaka, K., Yokoyama, N., Nodono, H., Hoshi, M. and Matsumoto, M. (2013), 'Innate sexuality determines the mechanisms of telomere maintenance', *Int J Dev Biol* **57**(1), 69–72.

Terenina, N. B., Poddubnaya, L. G., Tolstenkov, O. O. and Gustafsson, M. K. (2009), 'An immunocytochemical, histochemical and ultrastructural study of the nervous system of the tapeworm *Cyathocephalus truncatus* (Cestoda, Spathebothriidea)', *Parasitol Res* **104**(2), 267–75.

Terry, A. V., J., Buccafusco, J. J. and Wilson, C. (2008), 'Cognitive dysfunction in neuropsychiatric disorders: selected serotonin receptor subtypes as therapeutic targets', *Behav Brain Res* **195**(1), 30–8.

Thompson, R. C., Deplazes, P. and Eckert, J. (1990), 'Uniform strobilar development of *Echinococcus multilocularis in vitro* from protoscolex to immature stages', *J Parasitol* **76**(2), 240–7.

Thompson, R. C. and Eckert, J. (1982), 'The production of eggs by *Echinococcus multilocularis* in the laboratory following *in vivo* and *in vitro* development', *Z Parasitenkd* **68**(2), 227–34.

Thompson, R. C. and Eckert, J. (1983), 'Observations on *Echinococcus multilocularis* in the definitive host', *Z Parasitenkd* **69**(3), 335–45.

Thompson, R. C. and Jenkins, D. J. (2014), '*Echinococcus* as a model system: biology and epidemiology', *Int J Parasitol* **44**(12), 865–77.

Tijssen, P., Penzes, J. J., Yu, Q., Pham, H. T. and Bergoin, M. (2016), 'Diversity of small, single-stranded DNA viruses of invertebrates and their chaotic evolutionary past', *J Invertebr Pathol* **140**, 83–96.

Torgerson, P. R., Keller, K., Magnotta, M. and Ragland, N. (2010), 'The global burden of alveolar echinococcosis', *PLoS Negl Trop Dis* **4**(6), e722.

Tsai, I. J., Zarowiecki, M., Holroyd, N., Garciarrubio, A., Sanchez-Flores, A., Brooks, K. L., Tracey, A., Bobes, R. J., Fragoso, G., Sciutto, E., Aslett, M., Beasley, H., Bennett, H. M., Cai, J., Camicia, F., Clark, R., Cucher, M., De Silva, N., Day, T. A., Deplazes, P., Estrada, K., Fernandez, C., Holland, P. W., Hou, J., Hu, S., Huckvale, T., Hung, S. S., Kamenetzky, L., Keane, J. A., Kiss, F., Koziol,

U., Lambert, O., Liu, K., Luo, X., Luo, Y., Macchiaroli, N., Nichol, S., Paps, J., Parkinson, J., Pouchkina-Stantcheva, N., Riddiford, N., Rosenzvit, M., Salinas, G., Wasmuth, J. D., Zamanian, M., Zheng, Y., Cai, X., Soberon, X., Olson, P. D., Laclette, J. P., Brehm, K. and Berriman, M. (2013), 'The genomes of four tapeworm species reveal adaptations to parasitism', *Nature* **496**(7443), 57–63.

Turlejski, K. (1996), 'Evolutionary ancient roles of serotonin: long-lasting regulation of activity and development', *Acta Neurobiol Exp (Wars)* **56**(2), 619–36.

Tutton, P. J. and Barkla, D. H. (1982), 'Influence of inhibitors of serotonin uptake on intestinal epithelium and colorectal carcinomas', *Br J Cancer* **46**(2), 260–5.

van Amerongen, R. and Nusse, R. (2009), 'Towards an integrated view of Wnt signaling in development', *Development* **136**(19), 3205–14.

Vasunilashorn, S. and Cohen, A. A. (2014), 'Stress responsive biochemical anabolic/catabolic ratio and telomere length in older adults', *Biodemography Soc Biol* **60**(2), 174–84.

Villegas, V. E. and Zaphiropoulos, P. G. (2015), 'Neighboring gene regulation by antisense long non-coding RNAs', *Int J Mol Sci* **16**(2), 3251–66.

Vuitton, D. A. (2004), 'Echinococcosis and allergy', *Clin Rev Allergy Immunol* **26**(2), 93–104.

Wagner, D. E., Ho, J. J. and Reddien, P. W. (2012), 'Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis', *Cell Stem Cell* **10**(3), 299–311.

Wakelin, D. (1997), 'Immune response to *Echinococcus* infection: parasite avoidance and host protection', *Parassitologia* **39**(4), 355–8.

Walker, M., Rossignol, J. F., Torgerson, P. and Hemphill, A. (2004), '*In vitro* effects of nitazoxanide on *Echinococcus granulosus* protoscoleces and metacestodes', *J Antimicrob Chemother* **54**(3), 609–16.

Wang, B., Lindley, L. E., Fernandez-Vega, V., Rieger, M. E., Sims, A. H. and Briegel, K. J. (2012), 'The T box transcription factor TBX2 promotes epithelial-mesenchymal transition and invasion of normal and malignant breast epithelial cells', *PLoS One* **7**(7), e41355.

Wang, S., Wang, S., Luo, Y., Xiao, L., Luo, X., Gao, S., Dou, Y., Zhang, H., Guo, A., Meng, Q., Hou, J., Zhang, B., Zhang, S., Yang, M., Meng, X., Mei, H., Li,

H., He, Z., Zhu, X., Tan, X., Zhu, X. Q., Yu, J., Cai, J., Zhu, G., Hu, S. and Cai, X. (2016), 'Comparative genomics reveals adaptive evolution of asian tapeworm in switching to a new intermediate host', *Nat Commun* **7**, 12845.

Wang, Y., Wang, Q., Lv, S. and Zhang, S. (2015), 'Different protein of *Echinococcus granulosus* stimulates dendritic induced immune response', *Parasitology* **142**(7), 879–89.

Wen, H., Vuitton, L., Tuxun, T., Li, J., Vuitton, D. A., Zhang, W. and McManus, D. P. (2019), 'Echinococcosis: Advances in the 21st Century', *Clin Microbiol Rev* **32**(2).

Wen, L. and Tang, F. (2016), 'Single-cell sequencing in stem cell biology', *Genome Biol* **17**, 71.

Whitaker-Azmitia, P. M., Druse, M., Walker, P. and Lauder, J. M. (1996), 'Serotonin as a developmental signal', *Behav Brain Res* **73**(1-2), 19–29.

Witzany, G. (2008), 'The Viral Origins of Telomeres and Telomerases and their Important Role in Eukaryogenesis and Genome Maintenance', **1**(2), 191–206.

Xiao, N., Qiu, J., Nakao, M., Li, T., Yang, W., Chen, X., Schantz, P. M., Craig, P. S. and Ito, A. (2005), '*Echinococcus shiquicus n. sp.*, a taeniid cestode from tibetan fox and plateau pika in china', *Int J Parasitol* **35**(6), 693–701.

Xiao, S., Feng, J. and Yao, M. (1995), 'Effect of antihydatid drugs on carbohydrate metabolism of metacestode of *Echinococcus granulosus*', *Chin Med J (Engl)* **108**(9), 682–8.

Yang, B., Dong, X., Cai, D., Wang, X., Liu, Z., Hu, Z., Wang, H., Cao, X., Zhang, J. and Hu, Y. (2008), 'Characterization of the promoter elements and transcription profile of *Periplaneta fuliginosa densovirus* nonstructural genes', *Virus Res* **133**(2), 149–56.

Yu, X., Ng, C. P., Habacher, H. and Roy, S. (2008), 'Foxj1 transcription factors are master regulators of the motile ciliogenic program', *Nat Genet* **40**(12), 1445–53.

Yuan, Y., Tsoi, K. and Hunt, R. H. (2006), 'Selective serotonin reuptake inhibitors and risk of upper GI bleeding: confusion or confounding?', *Am J Med* **119**(9), 719–27.

Zheng, H., Zhang, W., Zhang, L., Zhang, Z., Li, J., Lu, G., Zhu, Y., Wang, Y., Huang, Y., Liu, J., Kang, H., Chen, J., Wang, L., Chen, A., Yu, S., Gao, Z.,

Jin, L., Gu, W., Wang, Z., Zhao, L., Shi, B., Wen, H., Lin, R., Jones, M. K., Brejova, B., Vinar, T., Zhao, G., McManus, D. P., Chen, Z., Zhou, Y. and Wang, S. (2013), 'The genome of the hydatid tapeworm echinococcus granulosus', *Nat Genet* **45**(10), 1168–75.

9 List of abbreviations

5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
aPS	activated protoscoleces
beta (as RNA-Seq sample)	primary cells treated with β-catenin siRNA
Bi2536 (as RNA-Seq sample)	metacestodes treated with Bi2536
BC	brood capsules
bp	base pair
CDS	coding sequence
D. melanogaster	*Drosophila melanogaster*
DNMT	DNA-methyltransferase
E. coli	*Escherichia coli*
E. granulosus	*Echinococcus granulosus*
EG_MCvivo	Metacestodes, cyst wall, in vivo
EG_naPS	Non-activated protoscoleces, retrieved from an abattoir
EG_PS_noact	Non-activated protoscoleces, retrieved from an abattoir (published RNA-Seq sample)
E. multilocularis	*Echinococcus multilocularis*
EmAdult Gravide	Gravid adults (published RNA-Seq sample)
EmPreAWDog	Pregravid adults (published RNA-Seq sample)
FBS	fetal bovine serum
GPCR	G-Protein coupled receptor
HU	hydroxyurea
HU (as RNA-Seq sample)	metacestodes treated with hydroxyurea
ITR	inverted terminal repeats
LC-MS/MS	Liquid Chromatography - Tandem Mass Spectrometry
MAO	monoamine oxidase
MBD	methyl-CpG-binding domain
MC_noBC	metacestode vesicle without brood chambers (published RNA-Seq sample)
MC_LateBC	metacestode vesicles with brood chambers (published RNA-Seq sample)
MCanaerob	metacestodes, control for HU
MCnoBC	Metacestodes, vesicle without brood chambers, anaerobic
MCvitro	Metacestodes, aerobic co-culture with Rh feeder cells
MCvivo	Metacestodes, extracted from jird

MGE	mobile genetic element
naPS	non-activated protoscoleces
Neg (as RNA-Seq sample)	primary cells, control or beta
PBS	phosphate buffered saline
PC	primary cells
PC_2d	primary cells, 2 days old (published RNA-Seq sample)
PC_11d	primary cells, 11 days old (published RNA-Seq sample)
PC1	Primary cells 48 h (stage 1)
PC2	Primary cells (stage 2)
PC3	Primary cells (stage 3)
PCA	principal component analysis
PCR	polymerase chain reaction
PS	protoscoleces
PS_noact	Non-activated protoscoleces (published RNA-Seq sample)
PS_act	Activated protoscoleces (published RNA-Seq sample)
RACE	Rapid amplification of cDNA-ends
rlog	regularized logarithm
S. mansoni	*Schistosoma mansoni*
S. mediterranea	*Schmidtea mediterranea*
SEM	standard error of the mean
Ser	serotonin
SERT	serotonin transporter
SD	standard deviation
SLC	solute carrier
TERC	telomerase RNA
TERT	telomerase reverse transcriptase
T. saginata	*Taenia saginata*
T. solium	*Taenia solium*
wnt	wingless
WMISH	whole-mount *in situ* hybridization
TRIM	terminal repeat retrotransposons in miniature

10 Supplement

10.1 Genome sequences and gene predictions for transcriptome data analysis

10.1.1 *E. multilocularis*

Genome assembly EMULTI002, GCA_000469725.3:
ftp://ftp.ebi.ac.uk/pub/databases/wormbase/parasite/releases/WBPS7/species/echinococcus_multilocularis/PRJEB122/echinococcus_multilocularis.PRJEB122.WBPS7.genomic.fa.gz

Gene predictions, annotation Version 2015-12-WormBase:
ftp://ftp.ebi.ac.uk/pub/databases/wormbase/parasite/releases/WBPS7/species/echinococcus_multilocularis/PRJEB122/echinococcus_multilocularis.PRJEB122.WBPS7.annotations.gff3.gz

10.1.2 *E. granulosus*

Genome assembly EGRAN001, GCA_000469785.1:
ftp://ftp.ebi.ac.uk/pub/databases/wormbase/parasite/releases/WBPS7/species/echinococcus_granulosus/PRJEB121/echinococcus_granulosus.PRJEB121.WBPS7.genomic.fa.gz

Gene predictions, annotation Version 2014-05-WormBase:
ftp://ftp.ebi.ac.uk/pub/databases/wormbase/parasite/releases/WBPS7/species/echinococcus_granulosus/PRJEB121/echinococcus_granulosus.PRJEB121.WBPS7.annotations.gff3.gz

10.2 RNA-Seq data sets

10.2.1 *E. multilocularis* new datasets

Abbreviation	ENA accession	Seq. lanes	Description	Isolate
PC1	ERS242838	2	Primary cells 48 h (stage 1)	H95
PC1	ERS242871	2	Primary cells 48 h (stage 1)	Ingrid
PC1	ERS242851	2	Primary cells 48 h (stage 1)	G8065
PC2	ERS242839	2	Primary cells 7 days (stage 2)	7030

PC2	ERS242872	2	Primary cells 9 days (stage 2)	H95
PC2	ERS242852	2	Primary cells 11 days (stage 2)	G8065
PC3	ERS242840	2	Primary cells 22 days (stage 3)	7030
PC3	ERS242873	2	Primary cells 16 days (stage 3)	H95
PC3	ERS242853	2	Primary cells 21 days (stage 3)	Ingrid
MCnoBC	ERS242841	2	metacestodes, vesicle without brood chambers, anaerobic	Ingrid
MCnoBC	ERS242874	2	metacestodes, vesicle without brood chambers, anaerobic	H95
MCnoBC	ERS242854	2	metacestodes, vesicle without brood chambers, anaerobic	GT10
naPS	ERS242848	2	Non-activated protoscoleces	MS1010
naPS	ERS242868	2	Non-activated protoscoleces	Ingrid
naPS	ERS242860	2	Non-activated protoscoleces	GH09
aPS	ERS242849	2	Activated protoscoleces	MS1010
aPS	ERS242869	2	Activated protoscoleces	Ingrid
aPS	ERS242861	2	Activated protoscoleces	GH09
MCvitro	ERS242842	2	metacestodes, aerobic co-culture with Rh feeder cells	Ingrid
MCvitro	ERS242863	2	metacestodes, aerobic co-culture with Rh feeder cells	H95
MCvitro	ERS242855	2	metacestodes, aerobic co-culture with Rh feeder cells	H95
MCvivo	ERS242843	2	metacestodes, extracted from jird	H95
MCvivo	ERS242864	2	metacestodes, extracted from jird	H95
MCvivo	ERS242856	2	metacestodes, extracted from jird	H95
HU	ERS242846	2	metacestodes, axenic culture with hydroxyurea 40 mM, exchanged every 24 hs for 1 week	Ingrid
HU	ERS242866	2	metacestodes, axenic culture with hydroxyurea 40 mM, exchanged every 24 hs for 1 week	Ingrid
HU	ERS242858	2	metacestodes, axenic culture with hydroxyurea 40 mM, exchanged every 24 hs for 1 week	Ingrid
MCanaerob	ERS242845	2	metacestodes, axenic culture, control for HU	Ingrid

MCanaerob	ERS242865	2	metacestodes, axenic culture, control for HU	Ingrid
MCanaerob	ERS242857	2	metacestodes, axenic culture, control for HU	Ingrid
Bi2536		1	metacestodes, axenic culture, treated with Bi2536	Ingrid
Bi2536		1	metacestodes, axenic culture, treated with Bi2536	Ingrid
Bi2536		1	metacestodes, axenic culture, treated with Bi2536	Ingrid
mcDMSO		1	metacestodes, axenic culture, treated with DMSO (control for Bi2536)	Ingrid
mcDMSO		1	metacestodes, axenic culture, treated with DMSO (control for Bi2536)	Ingrid
mcDMSO		1	metacestodes, axenic culture, treated with DMSO (control for Bi2536)	Ingrid
beta		1	Primary cells, treated with beta-catenin siRNA, harvested 7 days after electroporation	Ingrid
beta		1	Primary cells, treated with beta-catenin siRNA, harvested 7 days after electroporation	Ingrid
beta		1	Primary cells, treated with beta-catenin siRNA, harvested 7 days after electroporation	Ingrid
Neg		1	Primary cells, negative mock control, harvested 7 days after electroporation	Ingrid
Neg		1	Primary cells, negative mock control, harvested 7 days after electroporation	Ingrid
Neg		1	Primary cells, negative mock control, harvested 7 days after electroporation	Ingrid

10.2.2 *E. multilocularis* published datasets (Tsai et al., 2013)

Abbreviation	ENA accession	Seq. lanes	Description	Isolate
PC_2d	ERS094037	1	Primary cells 2d	G8065
PC_11d	ERS094038	1	Primary cells 11 d	G8065
MC_noBC	ERS094039	1	metacestode vesicle without brood chambers	G8065
MC_LateBC	ERS016464	1	metacestode vesicles with brood chambers	Cd ..
PS_noact	ERS094036	1	Non-activated protoscoleces	MS1010
PS_act	ERS094035	1	Activated protoscoleces	MS1010
EmPreAWDog	ERS018054	1	Pregravid adults	dog infection
EmAdultGravide	ERS018053	1	Gravid adults	fox origin

10.2.3 *E. granulosus* new datasets

Abbreviation	ENA accession	Seq. lanes	Description	Isolate
EG_MCvivo	ERS242844	2	metacestodes, cyst wall, in vivo	G7
EG_naPS	ERS242847	2	Non-activated protoscoleces, retrieved from an abattoir	G7
EG_naPS	ERS242867	2	Non-activated protoscoleces, retrieved from an abattoir	G7
EG_naPS	ERS242859	2	Non-activated protoscoleces, retrieved from an abattoir	G7

10.2.4 *E. granulosus* published datasets (Tsai et al., 2013)

Abbreviation	ENA accession	Seq. lanes	Description	Isolate
EG_PS_noact	ERS094034	1	Non-activated protoscoleces, retrieved from an abattoir	porcine

10.3 Comparisons with DESeq2

Sample	Reference
PC1	MCnoBC
PC2	PC1
PC3	PC2
MCnoBC	PC3
naPS	MCnoBC
naPS	MCvivo
aPS	naPS
MCvitro	MCvivo
MCnoBC	MCvivo
MCnoBC	MCvitro
HU	MCanaerob
Bi2536	mcDMSO
beta	Neg

10.4 Session Info for DESeq2 analysis

- R version 3.4.3 (2017-11-30), Platform: x86_64-pc-linux-gnu (64-bit)
- locale: LC_CTYPE=de_DE.UTF-8, LC_NUMERIC=C, LC_TIME=de_DE.UTF-8, LC_COLLATE=de_DE.UTF-8, LC_MONETARY=de_DE.UTF-8, LC_MESSAGES=de_DE.UTF-8, LC_PAPER=de_DE.UTF-8, LC_NAME=C, LC_ADDRESS=C, LC_TELEPHONE=C, LC_MEASUREMENT=de_DE.UTF-8, LC_IDENTIFICATION=C
- attached base packages: parallel, stats4, stats, graphics, grDevices utils, datasets, methods, base
- other attached packages: ggplot2_2.2.1, RColorBrewer_1.1-2, pheatmap_1.0.8, hexbin_1.27.1, vsn_3.44.0, DESeq2_1.16.1, SummarizedExperiment_1.6.5, DelayedArray_0.2.7, matrixStats_0.52.2, Biobase_2.36.2, GenomicRanges_1.28.6, GenomeInfoDb_1.12.3, IRanges_2.10.5, S4Vectors_0.14.7, BiocGenerics_0.22.1
- loaded via a namespace (and not attached): tidyr_0.7.2, bit64_0.9-7, splines_3.4.3, Formula_1.2-2, assertthat_0.2.0,affy_1.54.0, latticeExtra_0.6-28, blob_1.1.0, GenomeInfoDbData_0.99.0, RSQLite_2.0, backports_1.1.1, lattice_0.20-

35, limma_3.32.10, glue_1.2.0, digest_0.6.13, XVector_0.16.0, checkmate_1.8.5, colorspace_1.3-2, preprocessCore_1.38.1, htmltools_0.3.6, Matrix_1.2-12, plyr_1.8.4, XML_3.98-1.9, pkgconfig_2.0.1, genefilter_1.58.1, zlibbioc_1.22.0, purrr_0.2.4, xtable_1.8-2, scales_0.5.0, affyio_1.46.0, BiocParallel_1.10.1, htmlTable_1.11.0, tibble_1.3.4, annotate_1.54.0, nnet_7.3-12, lazyeval_0.2.1, survival_2.41-3, magrittr_1.5,memoise_1.1.0, foreign_0.8-69, BiocInstaller_1.26.1, tools_3.4.3, data.table_1.10.4-3, stringr_1.2.0, munsell_0.4.3, locfit_1.5-9.1 ,cluster_2.0.6, AnnotationDbi_1.38.2, bindrcpp_0.2, compiler_3.4.3, rlang_0.1.4, grid_3.4.3, RCurl_1.95-4.8, rstudioapi_0.7, htmlwidgets_0.9, labeling_0.3, bitops_1.0-6, base64enc_0.1-3, gtable_0.2.0, DBI_0.7, R6_2.2.2, gridExtra_2.3, knitr_1.17, dplyr_0.7.4, bit_1.1-12,bindr_0.1, Hmisc_4.0-3, stringi_1.1.6, Rcpp_0.12.14, geneplotter_1.54.0, rpart_4.1-11,acepack_1.4.1

10.5 Session Info for topGO analysis

- R version 3.4.4 (2018-03-15), Platform: x86_64-pc-linux-gnu (64-bit)
- locale:LC_CTYPE=de_DE.UTF-8, LC_NUMERIC=C, LC_TIME=de_DE.UTF-8, LC_COLLATE=de_DE.UTF-8, LC_MONETARY=de_DE.UTF-8, LC_MESSAGES=de_DE.UTF-8, LC_PAPER=de_DE.UTF-8, LC_NAME=C, LC_ADDRESS=C, LC_TELEPHONE=C, LC_MEASUREMENT=de_DE.UTF-8, LC_IDENTIFICATION=C
- attached base packages: grid, stats4, parallel, stats, graphics, grDevices, utils, datasets, methods, base
- other attached packages: Rgraphviz_2.20.0, topGO_2.28.0, SparseM_1.77, GO.db_3.4.1, AnnotationDbi_1.38.2, IRanges_2.10.5, S4Vectors_0.14.7, Biobase_2.36.2, graph_1.54.0, BiocGenerics_0.22.1
- loaded via a namespace (and not attached): bit64_0.9-7, splines_3.4.4, Formula_1.2-3, latticeExtra_0.6-28, blob_1.1.1, GenomeInfoDbData_0.99.0, pillar_1.2.2, RSQLite_2.1.1, backports_1.1.2, lattice_0.20-38, digest_0.6.15, GenomicRanges_1.28.6, RColorBrewer_1.1-2, XVector_0.16.0, checkmate_1.8.5, colorspace_1.3-2, htmltools_0.3.6, Matrix_1.2-14, plyr_1.8.4, DESeq2_1.16.1, XML_3.98-1.11, pkgconfig_2.0.1, genefilter_1.58.1, zlibbioc_1.22.0, xtable_1.8-2, scales_0.5.0, BiocParallel_1.10.1, htmlTable_1.11.2, tibble_1.4.2, annotate_1.54.0, ggplot2_2.2.1, SummarizedExperiment_1.6.5nnet_7.3-12, lazyeval_0.2.1, survival_2.43-3, magrittr_1.5, memoise_1.1.0, foreign_0.8-70, tools_3.4.4, data.table_1.11.2, matrixStats_0.53.1, stringr_1.3.1, munsell_0.4.3, locfit_1.5-9.1, cluster_2.0.7-1,

DelayedArray_0.2.7, compiler_3.4.4, GenomeInfoDb_1.12.3, rlang_0.2.0, RCurl_-1.95-4.10, rstudioapi_0.7, htmlwidgets_1.2, bitops_1.0-6, base64enc_0.1-3, gtable_-0.2.0, DBI_1.0.0, gridExtra_2.3, knitr_1.20, bit_1.1-13, Hmisc_4.1-1, stringi_-1.2.2, Rcpp_0.12.17, geneplotter_1.54.0, rpart_4.1-13, acepack_1.4.1

10.6 Densovirus NS1 gene sequences in cestode genomes

Organism	Abbreviation	Contig/scaffold	Start	End
E. canadensis	Ecan_c0220	E.canG7_contigs_0220	2143	968
E. canadensis	Ecan_c0249	E.canG7_contigs_0249	1013	1
E. canadensis	Ecan_c0501	E.canG7_contigs_0501	4195	3020
E. canadensis	Ecan_c0565	E.canG7_contigs_0565	1768	2943
E. canadensis	Ecan_c0881	E.canG7_contigs_0881	776	1
E. canadensis	Ecan_c0996	E.canG7_contigs_0996	2039	936
E. canadensis	Ecan_c4129	E.canG7_contigs_4129	2238	1063
E. canadensis	Ecan_c4490	E.canG7_contigs_4490	2143	968
E. canadensis	Ecan_c5077	E.canG7_contigs_5077	2238	1063
E. canadensis	Ecan_c5435	E.canG7_contigs_5435	1137	275
E. canadensis	Ecan_c7389	E.canG7_contigs_7389	1733	2908
E. canadensis	Ecan_c7792	E.canG7_contigs_7792	866	1
E. canadensis	Ecan_c8223	E.canG7_contigs_8223	3224	4399
E. canadensis	Ecan_c9228	E.canG7_contigs_9228	44	1219
E. granulosus (1)	EgG_s0006	pathogen_EgG_scaffold_0006	681223	680537
E. granulosus (1)	EgG_s0057	pathogen_EgG_scaffold_0057	76086	75042
E. granulosus (1)	EgG_s0137	pathogen_EgG_scaffold_0137	24475	25080
E. granulosus (2)	EG_S00208	EG_S00208	90970	90299
E. granulosus (2)	EG_S00226	EG_S00226	61657	62542
E. multilocularis	EMU_c0155	pathogen_EMU_contig_0155	5121	6296
E. multilocularis	EMU_c0221_1	pathogen_EMU_contig_0221	1537	614
E. multilocularis	EMU_c0221_2	pathogen_EMU_contig_0221	1764	2690
E. multilocularis	EMU_c0221_3	pathogen_EMU_contig_0221	5420	6595
E. multilocularis	EMU_c0221_4	pathogen_EMU_contig_0221	7966	8892
E. multilocularis	EMU_c0221_5	pathogen_EMU_contig_0221	10262	11188
E. multilocularis	EMU_c0226	pathogen_EMU_contig_0226	1779	604
E. multilocularis	EMU_c0266_1	pathogen_EMU_contig_0266	3209	4384
E. multilocularis	EMU_c0266_2	pathogen_EMU_contig_0266	7333	8508
E. multilocularis	EMU_c0303	pathogen_EMU_contig_0303	2106	3281

E. multilocularis	EMU_c0490	pathogen_EMU_contig_0490	334	1089
E. multilocularis	EMU_c0544	pathogen_EMU_contig_0544	10653	9478
E. multilocularis	EMU_c0630	pathogen_EMU_contig_0630	2725	1550
E. multilocularis	EMU_c0868	pathogen_EMU_contig_0868	1007	2182
E. multilocularis	EMU_c0881	pathogen_EMU_contig_0881	992	1918
E. multilocularis	EMU_c0966	pathogen_EMU_contig_0966	2422	1733
E. multilocularis	EMU_s09_1	pathogen_EmW_scaffold_09	880720	881895
E. multilocularis	EMU_s09_2	pathogen_EmW_scaffold_09	940252	941422
H. taeniaeformis	TTAC_c0003731	TTAC_contig0003731	2359	1403
H. taeniaeformis	TTAC_s0002620	TTAC_scaffold0002620	2880	2204
H. diminuta	HDID_c0003717_1	HDID_contig0003717	1163	6
H. diminuta	HDID_c0003717_2	HDID_contig0003717	1733	2812
H. diminuta	HDID_c0003864	HDID_contig0003864	2342	1203
H. diminuta	HDID_c0004012	HDID_contig0004012	997	2001
H. diminuta	HDID_c0004294	HDID_contig0004294	509	1687
H. diminuta	HDID_c0004382	HDID_contig0004382	1803	2474
H. diminuta	HDID_c0004716	HDID_contig0004716	1167	2265
H. diminuta	HDID_c0005242	HDID_contig0005242	1270	152
H. diminuta	HDID_c0005799	HDID_contig0005799	404	1563
H. diminuta	HDID_c0006158	HDID_contig0006158	664	1270
H. diminuta	HDID_c0008092	HDID_contig0008092	835	173
H. diminuta	HDID_s0000959	HDID_scaffold0000959	11545	12687
H. diminuta	HDID_s0001494	HDID_scaffold0001494	987	157
H. diminuta	HDID_s0001651	HDID_scaffold0001651	1797	2636
H. microstoma	HMN_03	HMN_03_pilon	4325899	4326500
H. microstoma	HMN_06_1	HMN_06_pilon	1900302	1899682
H. microstoma	HMN_06_2	HMN_06_pilon	1906750	1906130
H. microstoma	HMN_06_3	HMN_06_pilon	1910429	1909512
H. nana	HNAJ_s0001253	HNAJ_scaffold0001253	21042	21635
T. asiatica (1)	TASK_c0001514	TASK_contig0001514	4004	4625
T. asiatica (1)	TASK_s0001160	TASK_scaffold0001160	5495	4824
T. asiatica (1)	TASK_s0001202	TASK_scaffold0001202	1053	314
T. asiatica (2)	TAS_s00165	Scaffold00165	111645	110906
T. asiatica (2)	TAS_s00223	Scaffold00223	16513	15770
T. asiatica (2)	TAS_s00314	Scaffold00314	64052	64723
T. asiatica (2)	TAS_s00412_1	Scaffold00412	39078	38428
T. asiatica (2)	TAS_s00412_2	Scaffold00412	43199	43820

T. asiatica (2)	TAS_s00502	Scaffold00502	7503	6760
T. asiatica (2)	TAS_s01102	Scaffold01102	6698	5952
T. asiatica (2)	TAS_s01392	Scaffold01392	7352	8095
T. multiceps	TM_LG2_1	LG2	3486003	3485376
T. multiceps	TM_LG2_2	LG2	5376528	5375914
T. multiceps	TM_LG2_3	LG2	7755045	7755659
T. multiceps	TM_LG2_4	LG2	8024133	8024747
T. multiceps	TM_LG2_5	LG2	8293622	8292995
T. multiceps	TM_LG2_6	LG2	8364844	8365458
T. multiceps	TM_LG2_7	LG2	21908050	21908692
T. multiceps	TM_LG2_8	LG2	22633268	22632641
T. multiceps	TM_LG2_9	LG2	22706632	22707259
T. multiceps	TM_LG2_10	LG2	23063897	23063237
T. saginata	TSA_s00262	Scaffold00262	79502	80173
T. saginata	TSA_s00396	Scaffold00396	38045	37298
T. saginata	TSA_s00672	Scaffold00672	24071	23452
T. saginata	TSA_s00870	Scaffold00870	5484	4738

10.7 TERT protein sequences used for blast searches

Accession number	organism
ERG79348.1	*Ascaris suum*
NP_492374.1	*Caenorhabditis elegans*
NP_001093605.1	*Ciona intestinalis*
NP_001077335.1	*Danio rerio*
NP_001026178.1	*Gallus gallus*
NP_937983.2	*Homo sapiens*
NP_033380.1	*Mus musculus*
NP_445875.1	*Rattus norvegicus*
AEK12106.1	*Schmidtea mediterranea*
NP_001079102.1	*Xenopus laevis*

10.8 Digoxigenin labeled probes for WMISH

gene	PCR-primers	length	synthesis	orientation
tph	IG 4-5-SPR2 and TPH_2_up	849 bp	SP6 polymerase	antisense
tph	IG 4-5-SPR2 and TPH_2_up	849 bp	T7 polymerase	sense
sert	SERT_A_dw and SERT_2_up	1068 bp	SP6 polymerase	antisense
sert	SERT_A_dw and SERT_2_up	1068 bp	T7 polymerase	sense
tert	TRT_1_dw and TRT_2_up	1760 bp	T7 polymerase	antisense
tert	TRT_1_dw and TRT_2_up	1760 bp	SP6 polymerase	sense
tert	TERT-fw-for-probe and TERT-rev-for-probe	1353 bp	T7 polymerase	antisense
tert	TERT-fw-for-probe and TERT-rev-for-probe	1353 bp	SP6 polymerase	sense

10.9 Makro for cell identification in Fiji

```
macro "Cell identifier"{
    Dialog.create("Options for cell identification:");
    Dialog.addMessage("General requirements:\nDapi staining
    \nno Z-stack\n \nCAVE:\nResets results, roi manager and selections!");
    Dialog.setInsets(0, 20,0);
    Dialog.addRadioButtonGroup("The Dapi channel is", newArray("C1", "C2",
    "C3", "C4", "C5"),1,5, "C1");

    Dialog.show();
    dapi=Dialog.getRadioButton();

    waitForUser("Select the window you want to study (original)");
    run("Select None");
    roiManager("Reset");
    run("Clear Results");
    original=getTitle();

    //extracts nuclei from dapi stained cells
    //image processing
    run("Duplicate...", "title=[dup_"+original+"] duplicate");
    duplicate=getTitle();
    run("Split Channels");
    selectWindow(dapi+"-"+duplicate); //gets dapi channel
```

```
//extracts nuclei
run("Subtract Background...", "rolling=50");
run("Median...", "radius=5");
run("Enhance Contrast...", "saturated=0.5 normalize");
run("Auto Threshold", "method=Li white");
run("Watershed");
run("EDM Binary Operations", "iterations=3 operation=erode");
run("Analyze Particles...", "size=4-25 display exclude clear include summarize
add"); //only particles size 4-25

nCells=roiManager("Count");
if (isOpen("C1-"+duplicate)) {
	selectWindow("C1-"+duplicate);
	run("Close");
}
if (isOpen("C2-"+duplicate)) {
	selectWindow("C2-"+duplicate);
	run("Close");
}
if (isOpen("C3-"+duplicate)) {
	selectWindow("C3-"+duplicate);
	run("Close");
}
if (isOpen("C4-"+duplicate)) {
	selectWindow(C4+"-"+duplicate);
	run("Close");
}
if (isOpen("C5-"+duplicate)) {
	selectWindow(C5+"-"+duplicate);
	run("Close");
}
if (isOpen("Results")) {
	selectWindow("Results");
	run("Close");
}
if (isOpen("Summary")) {
	selectWindow("Summary");
```

```
        run("Close");
    }
    selectWindow(original);
    roiManager("Show All without labels");
    run("Grid...", "grid=Lines area=500 color=Cyan");

    print("The number of cells in \""+original+"\" is "+nCells+".");
}
```

10.10 Protocols

10.10.1 Preparation of c-DMEM-A (also known as A4 or A7-medium)

- Cultivate 1 million rat Reuber hepatoma cells (RH^--cells) in a 175 ml cell culture flask with 50 ml or 100 ml DMEM supplemented with 10 % FBS, penicillin (100 U/ml) and streptomycin (0,1 g/l) for 7 days at 37°C
- Sterile filtrate medium with a 0,02 µm filter
- Add fresh antibiotics to filtrated medium (penicillin (100 U/ml) and streptomycin (0,1 g/l))
- For use on metacestode vesicles add reducing agents (1 µl/ml L-cysteine (stock solution 100 mM in H_2O), 1 µl/ml β-Mercaptoethanol (stock solution 1:100, diluted in H_2O), 1 µl/mL bathocuproine disulfonic acid (stock solution: 10 mM in H_2O); all sterile filtrated using a 0,02 µm filter prior to use in cell culture)

10.10.2 Axenisation of *E. multilocularis* metacestode vesicles

- Transfer metacestode vesicles from cell culture flask in a 50 ml Falcon tube or in a beaker
- Discard as much medium as possible without loosing vesicles
- Wash 3 times with PBS
- Remove all broken and orange vesicles with a 5 ml pipette
- Remove PBS
- Transfer vesicles to a 175 ml cell culture flask and add c-DMEM-A with reducing agents
- Fill bottle with nitrogen an incubate at 37°C for 2-3 days until the medium is yellow and acidic

10.10.3 Primary cell isolation from *E. multilocularis* metacestode vesicles

Washing:

- Sieve metacestode vesicles from culture flask with a tea strainer
- Rinse vesicles with PBS
- Fill vesicles in a beaker glass and wash with PBS
- Remove leaky vesicles (red/yellow) with 5 ml pipette
- Optional: incubate 1 min in water to kill residual feeder cells by osmotic shock
- Wash with PBS 3-5 times to remove feeder cells
- Transfer vesicles in a 50 ml Falcon tube in the last washing step

Primary cell isolation:

- Remove PBS
- Destroy vesicles by pipetting up and down with 5 ml pipette (avoid bubbles, they can fill the vesicles and disturb the centrifugation process)
- Fill the 50 ml tube with PBS
- Centrifuge at 700 g for 10 min
- Discard the supernatant
- Destroy residual vesicles with 1 ml tip and refill with PBS
- Centrifuge at 700 g for 10 min
- Discard supernatant
- Refill tube with PBS, centrifuge at 400 g for 2 min, repeat washing step 3 times
- Add 6-8 volumes pre-warmed trypsin/ethylenediaminetetraacetic acid (EDTA) solution (0,05 %/0,02 % (w/v) respectively, in PBS) to 1 volume vesicle pellet
- Incubate at 37°C for 20-30 min, loosen pellet every 5 minutes by canting carefully
- Shake tube carefully to detach cells from the laminated layer
- Filter suspension through a 30 µm gauze filter
- Transfer residual vesicles from filter to a 50 ml tube and add some PBS (20 ml)
- Repeat shaking and filtering 4-7 times with incrementing shaking force (until liquid stays clear and laminated layers are clear and swim on top)
- Transfer the flow-through to 50 ml tubes and centrifuge at 80 g for 1 min to remove calcium bodies
- Transfer supernatant to new 50 ml tubes and centrifuge at 400 g for 10 min

- Discard supernatant
- Resuspend pellets in 500 µl PBS
- Filter cells through a 40 µm cell filter to remove bigger aggregations
- Wash tubes with a total of 500 µl PBS and filter that as well

Calculation of cell quantity:

- Add 12,5 µl cell suspension to 987,5 µl PBS
- Measure absorbance at λ=600 nm
- Calculate units

  λ=600 nm: 0,02 ≈ 1 Unit

  number of Units/12,5 µl = Units/µl