Patterns of Gene Expression in *Schistosoma mansoni* larvae associated with Infection of the Mammalian Host

Sophia J Manuel

Submitted for

PhD

University of York

Department of Biology

February 2010

Abstract

Larval schistosomes infect the human host by penetration of unbroken skin, before gaining access to a blood vessel and beginning their intravascular life. The work in this thesis focuses on the mammalian infection process, as it would be the ideal point to interrupt the life cycle. A whole organism approach was taken, and the life cycle stages before, during and after skin penetration were studied, namely the intra-molluscan germ ball (embryonic cercaria), infective cercaria, and the *in vitro* cultured day 3 schistosomulum (equivalent to skin stage larva). Confocal microscopy was used for a morphological survey of these life cycle stages, establishing the timeline of cercarial embryogenesis and documenting the impressive changes they undergo. The temporal and spatial gene expression patterns underlying the changes were investigated using the first genome-wide microarray for *S. mansoni* for transcriptional profiling of the three stages. The known repertoire of molecules likely to be secreted during host entry was greatly expanded, particularly the proteases and venom allergen-like proteins. Genes involved in energy production and conservation were up-regulated in the cercaria, as were several genes encoding proteins deployed immediately on arrival in the skin. Additionally, micro exon genes (MEGs) encoding variant secreted proteins were highly up-regulated in the schistosomulum, emphasising their likely role after entry into the mammalian host. The transcription of many tegument and gut-associated genes was also increased in the schistosomulum; cathepsins were particularly notable, even in the cercaria, implying that the larval gut becomes active long before blood feeding begins. The first application of whole mount *in situ* hybridisation to germ balls confirmed localisation of invadolysin and VAL-10 to the nascent acetabular glands. However, SmKK7 and Sm16 were not expressed in these glands, questioning their putative roles in immunomodulation. Finally, three MEGs were revealed to be expressed in tissues at the host-parasite interface.

Table of Contents

List of Figures

Acknowledgements

I am very grateful to Alan for his excellent supervision, encouragement, clarity of thought and for constructive criticism. I am also thankful for his careful help with writing, and for teaching me what semicolons are for.

Thanks are also due to all members of the schistosome research group at York, past and present, for their encouragement, friendship, fun, and help with the work: Adrian, Adam, Ann, Bill, Fabienne, Gavin, Jenny, Joe, Pat, Pete, Rachel, Ross, Sarah, Shobana, Simon, Steph, and William. In particular, Gary taught me the techniques I needed in the laboratory, and was a great bench mate. In the Technology Facility: Graeme Park, Karen Chance and Meg Stark helped with microscopy. Celina Whalley and Naveed Aziz were always available with invaluable help and support. Also at York, I'm grateful to Professors Deborah Smith and Jenny Southgate for their advice and support as my training committee.

I am thankful to Dr Alasdair Ivens for all his work on the microarray project.

On a personal note I would like to thank my friends for keeping me sane, and sharing the good times and not-so-good times: Natalie Russell; Rachel, Sarah and Esther Hagger-Holt; Tom Selmes, Charlie Forman, Debbie Coldwell, Pete Morrison, Harry Saffin, Jerome Jackson, Silvia Nulissi, Francesca Vattari, Andreas Biternas, Charlotta Salmi, Mike Bright, Susanna Brown, Tom Bartlett and Clare Gladding, Anna Semlyen, Hannah and Norna, Susan Collier, and the lovely folks at St. Luke's Burton Stone Lane.

Last but not least I would like to thank my family, my fiancé Rich and his family for always cheering me on.

Thanks are due to the BBSRC, the NIH, and the Pathogen Sequencing Advisory Group at the Wellcome Trust Sanger Institute for funding.

Declaration

All of the following work is my own with the following exceptions:

Chapter 1:

Figure 1 was reproduced by kind permission of Prof RA Wilson

Chapter 3:

RNA was run on Bioanalyser gels by Mrs. Celina Whalley. The microarray design and statistical analysis was carried out by Dr A Ivens. Roche-NimbleGen synthesised the microarrays and carried out the hybridisations.

Chapter 4:

The clone of VAL-10 and antisera for SmKK7 and invadolysin were the kind gifts of Dr S Prasad at Cape Western University, Ohio, USA

Ms Jenny Middleton cloned SmKK7, and carried out WISH on larvae under my supervision as part of her undergraduate research project.

The negative control sense chorion probe used in WISH was kindly provided by Dr Gary P Dillon.

Electron micrograph of the anterior end of a cercaria kindly provided by Stephanie Hopkins and Prof Jim McKerrow, University of California, San Fransisco.

Perfusion of adult worms Prof Alan Wilson and Dr William C Borges

Electron microscopy of adult worms carried out by Ms Meg Stark

List of Abbreviations

Abbreviation	
20x SSC	0.3M trisodium citrate and 3.0M sodium chloride
2D	Two dimensional
3D	Three dimensional
5-HT	5-hydroxytryptamine
A	Acetabulum
AbD	Antibody diluent
AChE	Acetylcholine esterase
ADAM	A disintegrin and a metalloprotease
adjP	Adjusted P value
AFA	Fixative containing ethanol, formalin, and acetic acid
AG	Acetabular glands
ATP	Adenine triphosphate
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
C	cercaria
cDNA	Complementary deoxyribonucleic acid
CE	Cercarial elastase
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CM	Circular muscle
D	Acetabular gland duct bundle
D3	Day 3 *in vitro* cultured schistosomulum
DALY	Disability adjusted life year
DAPI	4',6' diamino-2-phenylindole
DIC	differential interference contrast image
DIG	digoxigenin
DM	Diagonal muscle
DNA	deoxyribonucleic acid
DPX	Mounting media containing distyrene, a plasticizer, and xylene
dscDNA	Double-stranded complementary deoxyribonucleic acid
dsRNA	Double-stranded RNA

EDTA	ethylenediaminetetraacetic acid
EFGP	Enhanced green fluorescent protein
EP	Excretory pore
ER	Endoplasmic reticulum
EST	Expressed sequence tag
f-actin	Filamentous actin (as opposed to globular)
GB	Germ ball
GFP	green fluorescent protein
GO	Gene ontology
GPCR	G-protein coupled receptor
HDL	High density lipoprotein
HG	Head gland
HSP	Heat shock protein
ID	identity
IHC	immunohistochemistry
ISH	*In situ* hybridisation
kDa	Kilo Daltons
LC	Langeron's carmine
LM	Longitudinal muscle
LMWP	Low molecular weight protein Smp_194860
LSM	Laser scanning microscope
M169	Parasite culture medium
MCM	Minichromosome maintenance proteins
MDR	Multidrug resistance protein
MEG	Micro-exon gene
mm	Mismatch
mRNA	Messenger ribonucleic acid
NG	Neural ganglia
ORESTES	Open reading frame expressed sequence tag - method for sequencing
PBS	Phosphate buffered saline
PBSAT	PBS and 0.1% Tween 20
PCNA	Proliferating cell nuclear antigen
PGN	Permeablising fluid (PBS with gelatin, saponin, and sodium azide)

PI	Propidium iodide
PIPLC	phosphatidylinositol-specific phospholipase C
pm	Perfect match
PZQ	Praziquantel
qPCR	Quantitative or real time PCR
RM	Radial muscle
RNA	ribonucleic acid
RNAi	RNA interference
RPMI	Roswell Park Memorial Institute culture medium
RT-PCR	Reverse transcriptase polymerase chain reaction
SCP	Sperm Coating Protein domain found in venom allergen-like proteins
SDS	Sodium dodecyl sulphate
SGTP1	*Schistosoma* glucose transporter protein 1
SGTP4	*Schistosoma* glucose transporter protein 4
Sm16	Smp_113760
SmKK7	Smp_194830
Smp	*S. mansoni* gene prediction identifier number from GeneDB
SmPepM8	*Schistosoma mansoni* metalloprotease eight (Smp_090100)
TBST	Tris buffered saline and tween
TIGR	The Institute for Genome Research
tRNA	Transfer RNA
TWIK	weakly inward rectifying potassium channels with a tandem of p domains
U	Undifferentiated cells
VAL	Venom allergen-like protein (containing sperm coating protein domain)
WHO	World Health Organisation
WISH	Whole mount *in situ* hybridisation
wk	Week(s) old
WTSI	Wellcome Trust Sanger Institute

1. Introduction

1.1. Part 1: Mammalian infection by larval schistosomes

1.1.1. Schistosomiasis

Schistosomiasis is a disease caused by blood flukes of the genus *Schistosoma*, affecting more than 200 million people in 74 countries [1]. An estimated 779 million people are 'at risk' of infection [2] and the annual number of deaths attributable to *S. mansoni* and *S. haematobium* is approximately 130,000 and 150,000 respectively, in Sub-Saharan Africa alone [3]. The symptoms are caused by the immune reaction to parasite eggs in the host tissue. The morbidity associated with infection by schistosomes was reassessed in 2005; it was suggested that the disability adjusted life years (DALYs) due to the disease should be increased from 0.5% (equivalent to a birth mark on the face, which is clearly absurd) to 2-15% [1]. Presently there is only one drug in use to control schistosomiasis: praziquantel (PZQ) and no vaccine available yet. It has been recognised, in the light of potential resistance, that novel drugs are needed, and should be used in combination therapy to extend the useful life of PZQ. New drug targets are being identified [4] and chemical libraries have been screened against schistosomes [5]. At present the WHO recommends focussing on morbidity reduction through mass chemotherapy [6]. Re-infection occurs rapidly after treatment, so a vaccine is an important goal of research. A combination of interventions, ideally including a vaccine, chemotherapy, safe water sources and improved sanitation, will be necessary to eliminate schistosomiasis [7]. As noted by Gryseels *et al* [8], schistosomiasis will be eradicated only when its underlying cause of poverty is no more.

1.1.2. Life cycle

Schistosomes have a two host life cycle, including a snail host and a mammalian host (Figure 1). Adult *S. mansoni* live paired in the blood vessels draining the intestines. The females lay up to 350 eggs per day [9], which either cross the intestine wall and exit in the faeces, or are washed back with the blood flow and become lodged in the liver. On contact with fresh water the egg hatches, releasing a miracidium larva; this will infect a snail of the genus *Biomphalaria.* There are two rounds of asexual reproduction which is important for life cycle maintenance. The miracidium penetrates the head-foot of the intermediate host and transforms into a mother sporocyst in which daughter sporocysts form. They leave the mother sporocyst

10-14 days after snail infection [10, 11] whereupon they enter the snail's circulatory system. The daughter sporocysts are carried by the flow of haemolymph via the heart to the digestive gland (hepatopancreas). On arrival here they grow quickly and lose their ability to migrate further. This is probably due to the availability of nutrients there. The daughter sporocysts settle and intertwine with the snail tissue. They can reproduce two ways: by sporocystogenesis producing a second generation of daughter sporocysts, or by cercariogenesis when cercarial embryos called germ balls grow and develop within the daughter sporocyst [12]. Mature cercariae leave the snail and enter fresh water. Here the larvae repeatedly swim to the surface and sink down again until they find a mammalian host, which they infect, by penetration of unbroken skin, using secretions from acetabular glands. On entry to the skin, the cercaria transforms into a schistosomulum larva, which exits the skin via a blood or lymph vessel and reaches the lungs four to six days after infection [13]. On arrival in the lungs the schistosomulum becomes longer and thinner to enable further migration through the narrow capillaries [13]. Migration through the lung takes an average of 30-35 hours in the mouse [14]. The schistosomulum leaves the lungs via the pulmonary vein and travels via the left side of the heart to the systemic organs [15]; it arrives at the liver between 6 and 21 days later [14]. Most schistosomula arrive at the hepatic portal system directly from the lungs and they remain there. However, if they are carried by the circulation to another site, they return to the lungs and complete further circuits until they arrive at the hepatic portal system [14]. It takes no more than 2-3 such circuits to recruit all the schistosomes that will mature successfully to the liver; those which do not arrive there will die in the other tissues [14].

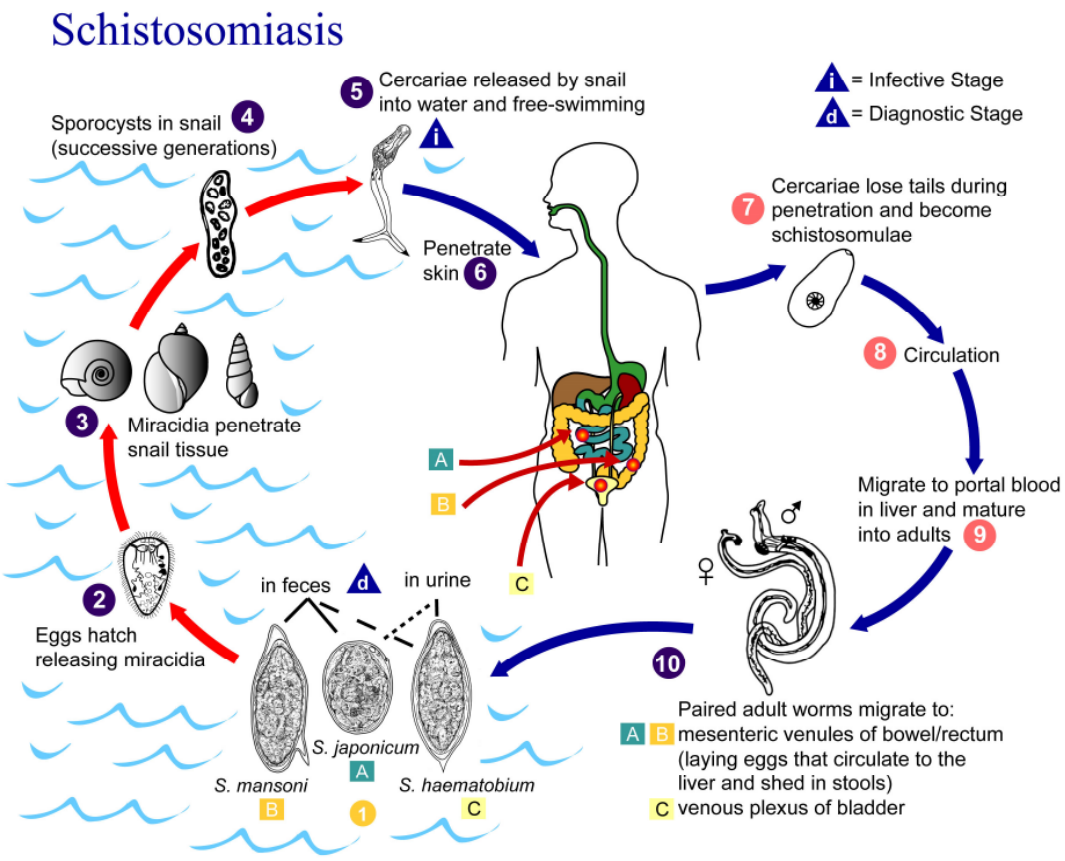


Figure 1-1 Life cycle of schistosomes.

Source: Centers for Disease Control, Division of Parasitic Diseases (http://www.dpd.cdc.gov/dpdx/HTML/Schistosomiasis.htm).

1.1.3. Gross anatomical features of adult schistosomes

Schistosomes are triploblastic acoelomates; they have no body cavities and three germ layers. Many cell types have a cell body with projections going elsewhere; examples include tegument cell bodies, nerve cells and muscle cells. Schistosomes have tissues and organs, but they cannot readily be dissected [16]. It has been suggested that laser microdissection could be used to cirvcumvent this problem [17]. However, cell bodies for various tissues can not be distinguished using light microscopy, and therefore would be difficult to target [17].

Tegument

The surface of schistosomes is an anucleate syncytium termed the tegument [18]. In intra-mammalian stages it is bounded apically by two lipid bilayers; the outer membranocalyx, and

the inner plasma membrane (see Figure 1-2). McLaren and Hockley noted in 1977 that this is a common feature of blood flukes, whereas flukes residing in the intestine or associated body cavities have a single outer membrane [19]. It was suggested that the membranocalyx may be useful for evading the immune system. The syncytium contains various inclusion bodies and is bounded basally by another membrane, below which lie the layers of circular and longitudinal muscle. The tegument also contains spines which protrude from the surface of the worm, but are covered apically by the membranocalyx. The spine's bases rest on the tegument basement membrane. Tegument cell bodies are situated underneath the two muscle layers and are joined to the syncytium by narrow microtubule-lined connections [20]. Everything that comprises the syncytial layer and its outer membrane originate in these cell bodies and is trafficked to the surface. The literature contains widely variable estimates of the turnover rate of the adult tegument, ranging from 20 minutes to 2 weeks [21]. Saunders *et al* carried out *in vivo* experiments and reported a half life of 5.4 days [22].

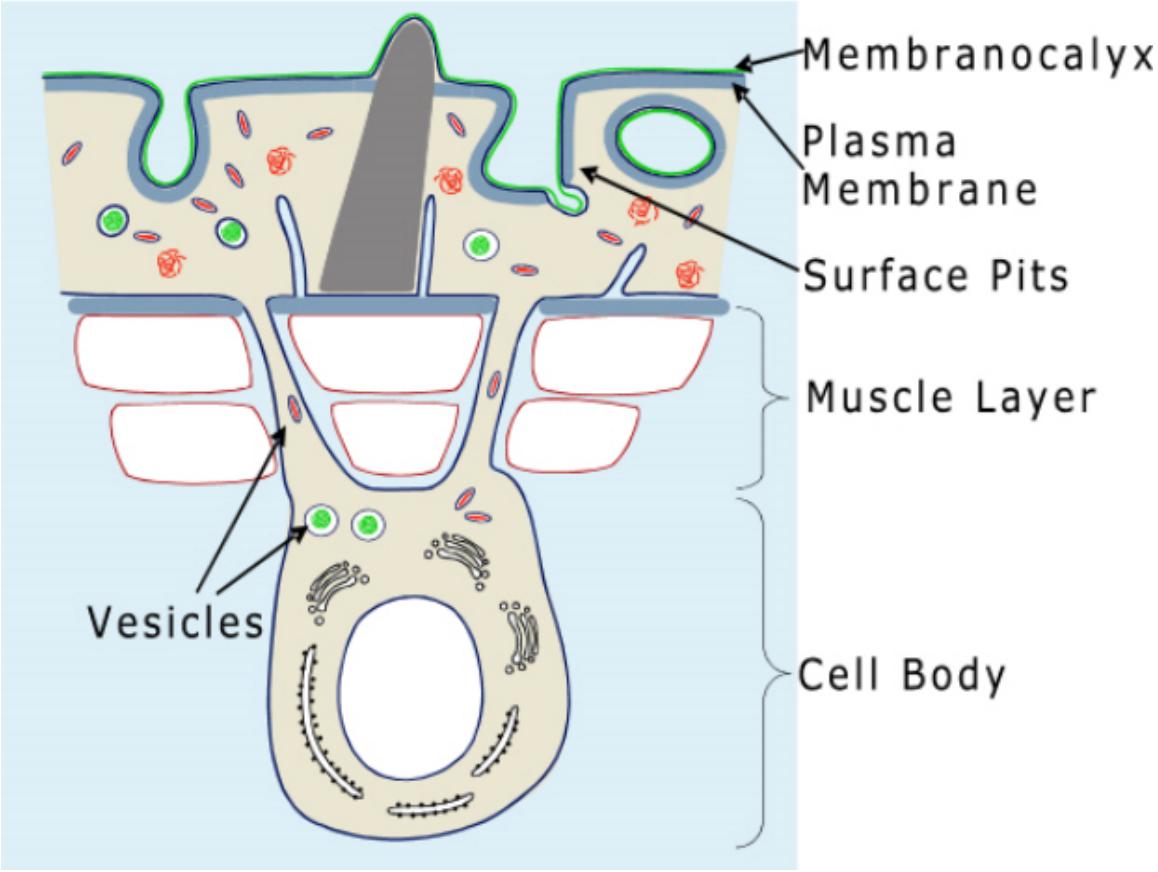


Figure 1-2 Diagram of the adult tegument.

Note the two apical membrane bilayers, the basement membrane (arrow), the muscle layers, and a tegument cell body with vesicles (reproduced courtesy of Prof RA Wilson).

Nerves

Schistosomes have an extensive nervous system. The brain [23] consists of a pair of anteriorly located central ganglia joined medially by a commisure. Longitudinal nerve trunks extend anterior and posterior from the brain. There is an extensive peripheral nerve net. The neurotransmitter 5-hydoxytryptamine (5-HT) has been shown in the peripheral nerve net by immunolocalisation, and in extensions to protrusions at the surface of adult males [24]. These protrusions are unicilliated sensory structures. A bulb filled with vesicles and containing nerve tissue is embedded in the tegument, but demarcated from it by a membrane [25]. The apical cilium is covered by the membranocalyx [25]. Unicillated sensory endings are said to be the only type of sensory ending in the adult [26]. The anterior ganglia, lateral nerve chords, and peripheral nerves are readily visualised by immunohistochemistry for the neuropeptide SALMFamide [27].

Protonephridial system

The protonephridial system consists of flame cells joined to collecting tubules; these lead to the bladder which is situated at the posterior of the worm and ends in the excretory pore [28]. The beating of the flagella gives the flame cell its flickering appearance, and also provides hydrostatic pressure to propel liquid down the tubules [29]. So called flame cells are made up of two cells, one flagella bearing cap cell which has a large nucleus, and second, a flattened cell which forms the tubule. The two cells are joined by interdigitations [28]. The ultrastructure of flame cells is the same in the cercaria, schistosomulum and adult worm [30]. The protonephridium is thought to function to remove organic metabolites from the deep tissues of the worm [28]. It has been suggested more recently that it may play a role in drug excretion, although the authors note that more experimental data are required to support this hypothesis [31].

Gut and oesophageal gland

In both male and female worms the mouth opens into the oesophagus which is lined by inturned tegument surrounded by circular and longitudinal muscle [32]. The gut, by contrast, is surrounded by circular muscle fibres only; these are thicker and more widely spaced than those of the oesophagus [32]. The oesophageal gland lies ventrally to the oesophagus at its

posterior end, just posterior to the ventral sucker and anterior to the gut [16]. The oesophagus is joined to the gut which splits into two lateral caecae which run either side of the genitalia, rejoining posterior to them and continuing to the posterior of the body. The gut is blind-ended and so excretion occurs by regurgitation.

Reproductive system

Adult male schistosomes have a gynaecophoric canal where the female resides; this begins just posterior to the ventral sucker. The genital pore is situated at the anterior end of the gynaecophoric canal and is connected to the four of five pairs of dorsal testes by a seminal vesicle and defferent duct [33] (p 54). Machado Silva propose that these are in fact a multi-lobed testis [34]. Female worms have a rather more complex reproductive system. The uterus opens via a genital pore just posterior to the ventral sucker and leads to the ootype or egg chamber. The ootype is connected both to the oviduct (leading to the centrally situated ovary) and the vitteline duct (leading to the vittelaria which continue to the posterior of the body) [33].

Musculature

Schistosomes are very muscular creatures. In adults there are three subtegumental muscle layers: outer circular, inner longitudinal, innermost are the diagonal fibres which are often paired [32].The oral and ventral suckers have radial fibres in addition to longitudinal and circular fibres [32]. Psuedo-striated muscle fibres have been described in the tail of cercariae [35]. The alimentary canal and female reproductive structures are also highly muscular as noted above.

1.1.4. Larval morphology

This thesis is concerned with mammalian infection by larval schistosomes. The larvae undergo considerable development as they progress from the snail, into fresh water and finally enter the mammalian host. Morphological studies have shed light on how each stage is adapted to its particular niche. The life cycle stages from daughter sporocyst to skin schistosomulum will now be described.

Daughter sporocyst morphology

Migratory daughter sporocysts are vermiform larvae approximately 250μm in length, and 15μm diameter [11]. They are covered in posterior facing spines at the anterior end, with fewer spines in the middle and none at the posterior end [36]. They are capable of moving at up to 2290μm per hour *in vitro* [10]. As they develop further, dilated zones appear, joined by narrow isthmuses [37]. Within these 'dilated zones' of mature daughter sporocysts, cercarial embryos (called germ balls) grow and develop [37].

Germ ball development

Germ ball development was investigated using light microscopy by Cheng and Bier in 1972 [38]. The authors designated stages from one single germinal cell, through a ball of cells, various stages with a stumpy tail, to the mature cercaria as illustrated in Table 1. The authors observed the development of germ balls at days 12, 17, 22, 32, 52 after snail infection. They report motile cercariae for the first time at day 32, but there is a large gap from the previous observation at day 22, when the most developed germ balls present were stage five, but the majority observed were stage one [38]. Meuleman and Holzman report that nearly mature cercariae are present at day 21 post infection [39].

Stage	Description
1	Single germinal cell
2	Embryo naked
3	Embryo covered by epithelium
4	Elongate embryo covered by epithelium
5	Developing cercaria – embryo with tail bud
6	Developing cercaria – longer tail with stem and bifurcation
7	Mature cercaria

Table 1: stages of germ ball development as designated by Cheng and Bier [38]

Ebrahimzadeh reported that the acetabular glands are the first 'organs' to develop [40]. Dorsey states that they can first be seen at the tail bud stage, before other organ systems are visible [41]. In contrast, Cheng and Bier observed them no earlier than stage six which is a

‘developing cercaria’, approximately 300μm long with the tail and body roughly equal in length [38]. The developing glands can be recognised by their size as they are larger than the other cells, and they are positioned in the centre towards the posterior end of the body [41]. The prominence of their nuclei is also noted. Dorsey uses the ratio of secretory granules to golgi apparatus and ER as a measure of maturity. Early on, machinery for protein synthesis is present, but secretory granules are yet to be formed; when the glands are mature the fundi and ducts are packed with vesicles containing the cercarial secretions [41]. In addition, he described the gland fundi and their ducts separately as it was not possible to reconstruct an image of the entire cells from the sections taken. Dorsey did not attempt a ‘strict chronological study’, as cercariae from different snails with the same post exposure date developed at different rates [41]. This is likely to be due to the different times taken by daughter sporocysts to migrate to the hepatopancreas, and any secondary daughter sporocysts arising from the first generation.

The development of the cercarial tegument has been observed using electron microscopy by Hockley [20] and Meuleman and Holzmann [39] and using conventional light microscopy by Cheng and Bier [38]. They each chose different time points to study; this fact, along with the different methods used and the lack of more general descriptions of the germ balls morphology means that these three studies are difficult to synthesise into one coherent description of germ ball development. The two electron microscopical studies describe a primitive epithelium that surrounds embryonic cercariae before their true tegument forms [20, 39]. Cheng and Bier did not observe this, hence their designation of stage two as a naked cell aggregate (Table 1) [38]. Hockley states that the primitive epithelium probably derives from the germ ball, and serves to protect it until the formation of the true tegument [20]. In contrast, Meuleman and Holzmann show evidence that it is formed from extensions of the daughter sporocyst’s tegument, and suggest that it may provide the developing embryo with nutrients from the snail haemolymph [39]. The two studies agree that the primitive epithelium is a nucleated syncytium, underneath which peripheral germ ball cells expand and coalesce to form the true tegument; the primitive epithelium is then shed [20, 39] and is not present 21 days post-infection [39]. It is also agreed that the subtegumental cell bodies arise independently of the tegument syncytium and form connections to it later, i.e. no evidence was found of nuclei from the syncytium sinking into the parenchyma [20, 39].

Morphology of the cercaria

Cercariae have been the subject of extensive ultrastructural studies. As a result much is known about cercarial anatomy [42]. The cercaria is a complex creature that comprises two main parts: the body and the tail. The tegument is covered in a single outer membrane, which is in turn enveloped/coated in a 1μm thick glycocalyx [20] which probably has a role in enabling this stage to be water-impermeable [43]. The ventral sucker or acetabulum is situated two thirds down the body. Schistosomes are very muscular, and the cercaria is no exception. There are two layers of muscle underlying the tegument surrounding the parasite; outermost is the circular layer, which encloses the inner longitudinal layer [35]. Approximately half to two thirds of the body volume is taken up by the ten large unicellular acetabular glands; two pairs of pre-acetabular glands lie anterior, and three pairs of post-acetabular glands posterior to the acetabulum (Figure 1.3). Post-acetabular glands contain homogenously granular secretory granules, and granules with electron dense bodies. Pre-acetabular glands contain granules with homogenous dense matrix and some with a less dense matrix containing electron lucid bodies [41]. When the glands are mature, the fundi and ducts are packed with vesicles containing the cercarial secretions, ready for action when they find a host. The contents of these gland cells are secreted during penetration of host skin via long ducts extending to the anterior tip of the cercaria. The ducts run in two lateral bundles, with two pre- and three post-acetabular in each. The bundles are surrounded by circular muscle fibres [44], and split as they reach the anterior end of the parasite so that each duct opens into the outside individually [44].

The anterior section of the body contains a thick muscular head capsule. The mouth lies on the ventral side of the head capsule, and opens into the oesophagus which is surrounded by two layers of muscle, outer circular and inner longitudinal [45]. The oesophagus runs through the posterior boundary of the head capsule, continues ventral to the brain [45], to the blind bifid caecum or gut, just anterior and dorsal to the acetabulum. The luminal surface of the gut bears many plates rather than villi. The gut is innervated by external neuromuscular junctions [45]. The gut is not active at this stage, the cercaria survives by using up a store of glycogen, more than half of which is in the tail [46].

The head capsule surrounds the head gland (Figure 1.3). This is a large purportedly unicellular gland which lies in the anterior area of the head capsule dorsal to the oesophagus and acetabular gland ducts [47]. There are many narrow microtubule-lined ducts which extend

from the fundus and open into the tegument – not to the exterior of the larva. There are three morphologically distinct types of vesicle present; in the cercaria these are found only in the fundus, not the ducts [47]. Dorsey suggests that the head gland secretions may repair the tegument at the anterior of the cercaria after damage occurs during penetration [47]. Crabtree later postulated that the head gland contents may be a source of 'lytic secretions' used by the schistosomulum to cross the epidermal basement membrane as the acetabular glands were empty at that stage [48].

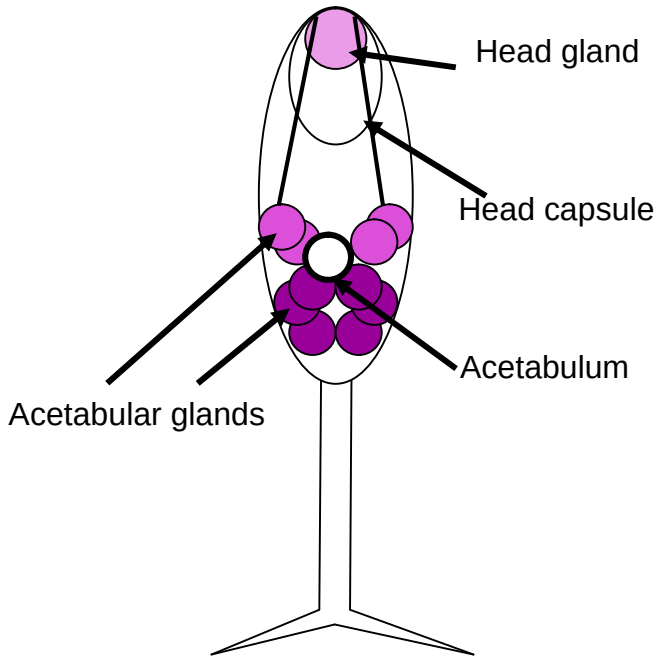


Figure 1-3 Diagram of a cercaria showing position of glands (not to scale).

The osmoregulatory or excretory system is made up of flame cells which are connected by tubules to the body tail junction. The tubules extend down the tail to the tip of each furcus where they end in excretory pores. There is confusion in the literature as to the number of flame cells in the cercaria. There are reports of six pairs in the body and one situated proximally in the tail. However, a diagram reproduced from Gordon *et al* by Stirewalt [49] shows four pairs in the body and one in the tail. This diagram also shows ciliated areas in the collecting tubules draining the flame cells. It is now acknowledged that cercariae have four pairs of flame cells in the body and one pair in the tail [42, 50].

The nervous system in the cercaria has been visualised by staining for acetylcholine esterase activity [51], neuropeptides [52], and by electron microscopy [53, 54]. The neurotransmitter 5-

HT was not detected by immunohistochemistry in the cercaria [52]. The ultrastructure of the internal nervous system has been described [54]. The bilobed brain measures 15μm at its widest point, and lies between the pre-acetabular gland fundi, and the head capsule. It lies dorsal to the oesophagus and acetabular gland duct bundles [54]. Sixteen nerve trunks extend from the brain, four pairs anterior, and four pairs posterior. In each direction there is one pair of central nerve trunks and three pairs peripheral: lateral, dorsal and ventral. The anterior central nerve trunks follow the acetabular gland ducts into the head capsule and continue to the anterior tip of the cercaria. The peripheral nerves also enter the head capsule and branch in order to innervate the body wall muscle [54]. The posterior nerves extend to the body tail junction. It has been reported by Cousin and Dorsey that they contribute to the dorsal and ventral nerve chords that extend the length of the tail stem [54]; however, Skuce *et al* report no connection between the nerves in the tail and those in the body [52]. When the cercaria was stained for AChE activity, a pair of 'anterior ganglia' was observed inside the head capsule, with nerves extending anteriorly to the tip and posteriorly to the brain [51]. Reactivity was also shown at the surface in structures which resembled "minute volcanic craters" [51]. These may correspond to the ciliated pits of Nuttman [53]. Just under the surface, a network made up of longitudinal and circular fibres, stained positive for AChE. The longitudinal fibres were spaced 3 – 4μm apart, and the circular fibres were arranged closer together, 1.5 – 2μm apart [51].

Three types of sensory ending have been described on the surface of the cercaria: 1) ciliated cavities, 2) unsheathed uniciliated bulbs, and 3) sheathed uniciliated bulbs [53]. Ciliated cavities are found on the cercarial body, adjacent to lateral nerve chords, but not on the tail. They form a bulb shaped cavity approximately 1.5μm wide which may be open to the environment. The pits contain five or six cilia and are packed with small dense and large clear vesicles which are extracellular [53]. It was suggested that these ciliated cavities could function as chemoreceptors as they are open to the environment [53].

Unsheathed uniciliated endings are found all over the cercaria including the tail and furcae [53]. They consist of a bulb embedded in the tegument, from which a cilium extends up to 7μm. The cilium is not covered by the tegument. The bulb is packed with large clear vesicles and microtubules which are continuous with those of the axon [53].

Sheathed uniciliated nerve endings are found only at the anterior tip of the cercaria [53, 55]. The cilium is shorter, only projecting 0.8- 2.5μm [55] or 1-2μm [53] and is surrounded by tegument apart from at its apex [53]. The basal bulb of these structures is raised from the tegument [55]. The acetabular gland duct openings are arranged laterally in two cresents. They are surrounded by tegumentary folds which vary in height from 0.5μm at the inner, concave edge, to 2.5μm at the outer, convex edge. The outer convex edge is reported to support five [53] or seven [55] sheathed unicilated nerve endings.

Skin stage morphology

Many important changes take place on entry to the mammalian host. These include loss of the tail, secretion of the acetabular glands contents and the swift transition to the double outer membrane. First the cercarial glycocalyx is shed, then the cercarial surface is lost by the formation and shedding of microvilli [56] and a new double outer membrane is formed. Skelly *et al* cloned, identified and sequenced 3 new *S. mansoni* glucose transporters (SGTP1, 2 and 4) and confirmed their function by expression studies in *Xenopus* oocytes [57]. Skelly and Shoemaker showed that SGTP4 protein is present in adults, and schistosomula (1hr post transformation), but not in eggs, cercariae, or the intramolluscan stages [58]. Jiang *et al* showed that SGTP4 is present in both the membranocalyx and plasma membrane, but not the tegument basement membrane of adult worms and schistosomula [59]. Thus SGTP4 can be used as a very effective marker for the outer tegumental membranes in intramammalian schistosomes. Indeed, Skelly and Shoemaker used immunostaining of SGTP4 to track the formation of the new tegument during transformation induced by incubation of cercariae in various different media [60]. SGTP4 appeared at the surface 30 minutes into transformation, and covered the surface by 3 hours [60]. These results show that the new tegument surface is formed rapidly on transformation.

As early as 30 minutes after penetration into mouse skin the head gland ducts have become wider and secretory vesicles can be seen in them and in the tegument [47]. It has been suggested that the material from the head gland, from the multilamellar vesicles in particular, is used to repair damage of the tegument sustained whilst migrating through mouse skin [47]. Dorsey suggests that the head gland is a specialised type of tegumental cell body [47]. Contrary to this, the authors of an investigation into penetration of the skin and entry into a

vein, using the hamster cheek pouch model, report that head gland secretions are likely to be involved in blood vessel penetration [48]. Vesicles of head gland type were observed in the host, outside the parasite [48]. Forty hours after initial penetration the contents of the acetabular glands had been secreted, whereas the head gland was still visible. Vesicles of the type from the head gland were seen in the tegument at the anterior of the head capsule as early as 30 minutes after entry to host skin [48]. Another study, which observed cercariae incubated in different media to encourage transformation to schistosomula, reported that although the acetabular glands were empty after 30 hours the head gland still contained secretory vesicles and was easier to see than in freshly emerged cercariae [40]. By day two the schistosomulum starts to elongate, and fewer spines are present between the mouth and the ventral sucker [61]. The cercarial ciliated papillae are lost; however, some sensory endings remain throughout migration [61]. By day three all the midbody spines had disappeared, but spines remain at the anterior and posterior ends, probably to be used as anchors during migration [61]. An investigation into schistosomula collected from lung 4 to 7 days after infection, showed that the head capsule musculature was no longer present.

The organisation of the flame cells in the parasite body does not change from cercaria to schistosomulum. The functional excretory pore is at the posterior end, where the body/tail junction had been, rather than at the tips of the tail furcae. There have been reports of the excretory pore serving as point for large molecules to enter the schistosomulum as it penetrates the mammalian host [62]. There is some discussion in the literature as to when the schistosome gut becomes active on entering the mammalian host [49]. It has been suggested that early schistosomula may ingest media [33]. However, in skin and lung stage schistosomula, electron microscopy shows that the oesophagus is too narrow to admit erythrocytes [48]. Crabtree noted that even if the oesophagus was large enough to accommodate cells, the mouth is pressed up against the blood vessel wall in migrating parasites, making ingestion of cells very difficult [63].

1.1.5. Physiological changes

Along with the morphological changes discussed above, there are several physiological changes that take place on entry to the mammalian host. The cercaria survives by using a store of glycogen as the substrate for aerobic metabolism. Lawson and Wilson noted that this would be used up after 24 hours *in vivo* so the schistosomulum must rely on the host as a source of

metabolites shortly after penetration [46]. This is borne out by the observation described earlier that the glucose transporter SGTP4 appears at the surface of the schistosomulum as early as 30 minutes after transformation [60].

Harrop and Wilson investigated the rate of protein synthesis and release by *in vitro* cultured schistosomula and found that much protein was released in the first 3 hours after transformation [64]. Hardly any translation occurred during the first 24 hours; after this point the rate steadily increased to a peak at day 8, then decreased for two days before rising again at day 11. An important finding of these authors was that in vitro cultured schistosomula would mature normally if transferred to a host intravenously [64]. In addition it was shown that schistosomula release proteins into the medium between 24 hours and 7 days after transformation. It was suggested that these proteins may originate in the head gland, as the acetabular glands were no longer present [64]. It is worth noting that the lack of translation in the first 24 hours is not due to a lack of mRNA. Blanton and Licate showed that post-transcriptional control is exerted during this period [65].

Metabolic function from cercaria to adult was investigated by Skelly and Shoemaker using genes involved in glucose uptake, phosphorylation, glycolysis, the Kreb's cycle, and the electron transport chain to probe northern and dot blots [66]. They discovered that cercarial tails have the highest level of transcripts involved in aerobic respiration, and schistosomula have the lowest. Adult worms had the highest levels of glucose transport transcripts, and the transcripts encoding aerobic respiration enzymes were increased, but not to the level of the cercarial tail [66]. It is worth noting that transcripts for SGTP1 and 4 were low and undetectable in cercariae and cercarial tails respectively [66]. This raises a question as to how SGTP4 protein arrives so quickly at the tegument of the schistosomulum.

1.2. Part 2: Gene expression

Having observed the morphological developments that take place during the infection process, the second part of this chapter deals with the molecular biology of schistosomes. First, the sequencing projects will be described, followed by the progress in and difficulties associated with genetic manipulation of schistosomes. The use of microarrays for gene expression profiling generally and in schistosomes specifically will be introduced. In addition to transcriptional profiling, the availability of sequence data has allowed proteomic studies to be

carried out. Our knowledge of proteins involved in the infection process will be presented. Finally, the importance of spatial expression patterns will be highlighted, and the methods available to elucidate them will be described.

1.2.1. The history of S. mansoni sequencing efforts

Simpson *et al* stated that understanding the genome and its expression would provide a new approach to discovering gene products of immunological or pharmalogical interest [67]. They estimated that the *S. mansoni* genome was 270 megabases in size in 1982 [67]. Ten years later the Schistosome Genome Project (including *S. mansoni* and *S. japonicum*) was launched by the World Health Organisation. The gene discovery programme began around the same time, using a directionally cloned adult worm cDNA library. The ends of these cDNAs were sequenced to produce expressed sequence tags (ESTs) [68]. This project resulted in 607 ESTs from 429 clones; 16% matched known S. mansoni genes, 22% matched genes from other organisms, and 33% had no significant matches to any sequences deposited in GenBank. Clustering revealed that the ESTs represented 169 genes, 154 of which were newly identified. Another study branched out from using adult worms only, and included cDNA libraries synthesised from eggs, cercariae and lung stage schistosomula [69]. The authors noted that larval stages can not be overlooked in the search for drug targets and vaccine candidates [69]. The resulting 1401 ESTs clustered together represented 466 unique genes, of which 8% matched known *S. mansoni* genes, 20% matched sequences from other organisms, but the vast majority 71.5% were new genes with no significant matches in GenBank [69]. This last category may represent schistosome specific genes, or simply genes which have not been sequenced from other organisms yet.

A new method called open reading frame expressed sequence tags or ORESTEs was developed to enable sequencing when starting material is limiting (ng of mRNA). It involves using arbitrary primers in conjunction with low-stringency RT-PCR. This method was reported to be three times more efficient than using directional cDNA libraries [70]. Other advantages of this method are firstly that it tends to sample the middle region of a cDNA rather than being biased to either end, and secondly, it normalises and samples rare transcripts very efficiently [71]. However, the direction of the sequences is unknown. In addition, the

relative expression of transcripts can not be estimated using this technique, as the most abundant transcripts are excluded before sequencing.

In 1999 Santos *et al* published a study of genes expressed by cercariae, having sequenced two cDNA libraries from this stage. Its purpose was to understand the biology of the host parasite relationship, and to identify novel gene products which could be useful as drug targets or vaccine candidates [72]. 64% of the ESTs discovered in this study had no matches at GenBank. The total 859 ESTs represented 453 genes. Those that could be putatively identified based on homology with other organisms were mostly involved in energy metabolism, gene expression, and regulation and signalling.

1.2.2. The Transcriptome

As a culmination of the sequencing efforts described above, a detailed and comprehensive analysis of the *S. mansoni* transcriptome was published in 2003 [73]. Sequences from a normalised adult worm cDNA library were added to ORESTES reads from eggs, miracidia, germ balls, cercariae, and schistosomula. Altogether, 163,000 ESTs were sequenced and compiled to form 31,000 assembled sequences (contigs). The authors estimated that this represented a 92% sampling of the probable 14,000 genes encoded by the genome [73]. A very high proportion of the sequences (77%) were new *S. mansoni* gene fragments, and 55% had no significant match to anything at GenBank. The presence of many genes encoding cell or tissue adhesion molecules was noted. Dicer and Piwi/Argonaute were also identified; these proteins form part of the RNA interference pathway. It was observed that no highly variable gene families were found, so evidence for antigenic variation was lacking at that time.

1.2.3. The Genome

The availability of extensive transcript data made gene finding possible, as the contig sequences could be used to train gene finding algorithms [74]. The genome of *S. mansoni* has been published recently alongside that of *S. japonicum*; these are the first Platyhelminthes to be fully sequenced [75]. In contrast to earlier estimates, it was found to be 363 mega bases (cf [67]), with at least 11, 092 genes. This is likely to be an underestimate, as more than 7,000 contigs are yet to be mapped to the genome [74]. Genomic DNA was extracted from cercariae

and libraries of varying sized inserts were sequenced randomly from either end using the Sanger method. Several gene-finding algorithms were trained using a set of 409 manually-curated genes before they were used on the sequence data. In addition to these automatically annotated genes, 958 genes were manually curated. All data are available at www.geneDB.org. Gene predictions are designated an identifier in the format Smp_xxxxxx; ESTs or contigs are named in the format Smxxxxx.

An exciting discovery reported by Berriman *et al* in the genome paper is a group of Micro-exon gene (MEG) families [75]. There were 14 families with 1 to 28 member genes in each; a total of 45 MEGs were reported. These genes have signal peptides, and, as their name suggests, very small exons and large introns. Not only are the exons very small, the transcripts display high splice variability. Indeed cDNA clones have been sequenced that reveal copies of MEG 1 with each of the exons missing [75]. The families are grouped together based on sequence similarity; there is no homology between families. It is also worth noting that these genes are represented in ESTs from intramammalian life cycle stages and germ balls, but were not found in miracidia [75].

The genome publication also illustrates that schistosomes lack certain genes involved in lipid metabolism and are reliant on the host for inositol [75].

The gene complement of *S. mansoni* was compared with that of the sea anemone *Nematostella victensis* to identify genes which may be needed for a parasitic life style, a third germ layer, and the formation of tissues into organs [75]. Tetraspannins, invadolysins and cathepsins were highlighted as important for parasitism. Genes encoding cadherins (cell-cell adhesion), Notch/Delta signalling (tissue patterning), and histone modification were all expanded compared to *Nematostella*. It was also noted that schistosomes have tools required for neurogenesis, axonguidance, and migration of neural cells. Many neuropeptides have been discovered in schistosomes, some of which appear to be platyhelminth-specific [75, 76]. Another area where schistosomes are different is in their complement of G protein-coupled receptors (GPCRs) and gated ion channels. No voltage-gated sodium channels were found, but many voltage-gated potassium channels are present [75].

Bioinformatic analyses were performed in order to identify potential new drug targets. One approach used was to build up a picture of the metabolic pathways and to highlight 'choke points' which could be targeted. A second approach was to search for schistosome orthologues of genes known to be targeted by drugs currently in use to treat human disease – either targeting a human protein, or a protein belonging to a human-infective pathogen. The rationale for this approach was to enable 'piggy-backing' of a current chemotherapy, potentially saving much time and money, as the current drugs have already passed the stringent safety testing required. However, this approach does not take into account the selective toxicity required to target a pathogen specifically, rather than the host.

The publication of the *S. mansoni* genome, 27 years after the first investigations into its size, is finally realising the aim stated by Simpson *et al* of providing a new approach to identify gene products of immunological or pharmacological interest [67]. Indeed, Caffrey *et al* carried out a comparative genomics analysis with *C. elegans* and *D. melanogaster* [4]. Genes which led to lethality, paralysis, or reduced motility when knocked out in the model organisms were identified and their orthologues in *S. mansoni* found, resulting in 72 possible leads. This set was filtered to exclude redundant genes, and those with no intramammalian expression pattern, leaving 57 genes. Of this set, 35 were found to be 'druggable' and the structure of 18 of these had been solved. The final 18 genes are the subject of further investigations by Caffrey *et al* [4].

1.2.4. Genetic manipulation of schistosomes

Progress has been made both in gene knockdown by RNA interference (RNAi), and in the application of transgenesis techniques. Genes encoding proteins in the RNAi pathway (Dicer and Piwi/argonaute) were discovered in the *S. mansoni* transcriptome [73]. Since then there have been several reports of successful, specific gene knock down in schistosomula [77, 78], eggs [79], and intra-molluscan stages [80]. On the basis of this, Brindley and Pearce state that "RNAi works powerfully in schistosomes and we see no reason why this approach should not become routine in laboratories studying gene function in schistosomes." [81]. It is worth noting that the genes for which a positive effect was observed are expressed either in the gut, or in tissues close to the surface. Both of these sites are readily accessible to the culture medium. Issues around delivery of dsRNA to deeper tissues must be addressed. A contrasting view was taken by Geldhof *et al* who reviewed the use of RNAi in parasitic helminths more

broadly [82]. They reported that non-specific effects, limited efficiency and reproducibility are problems shared by many species. They are more cautious than Brindley and Pearce, and say that the technique is in need of some development before it can be relied upon as a screen for functional genomics. Indeed, in 2009, de Moraes Maurão *et al* chose 32 genes known to be expressed by mother sporocysts, and carried out a phenotypic screen using RNAi [83]. Miracidia were allowed to transform *in vitro* into mother sporocysts in the presence of dsRNA. qPCR was carried out to measure transcript levels after seven days of culture, and any phenotypes were noted. Only 11 of the 32 treatments resulted in any phenotype – all of these were a reduction in parasite length. Transcript levels were reduced in 6/11 with decreased length, and 6/12 with no observed phenotype [83]. The authors note that gene specific effects must be controlled and experiments optimised carefully. Hence their suggestion that RNAi techniques should be developed further, before they can be used to screen for gene function [83].

1.2.5. Transgenesis

Heyers *et al* state that there are at least two obstacles for a reliable, tractable transgenesis system in schistosomes: 1) the development of constructs that allow stable integration into the genome, and 2) the demonstration that genetically modified parasites can complete the life cycle [84]. They report the use of particle bombardment to deliver a plasmid into miracidia and the subsequent infection of snails by transformed specimens. The plasmid used in this experiment encoded enhanced green fluorescent protein (EGFP), flanked by the *S. mansoni* heat shock protein 70 (HSP70) promoter and terminator. The construct was not designed to integrate into the genome. Transcript (but not protein) was demonstrated in infected snails several days after infection, and gold particles were evident inside germ balls within mother sporocysts. The development of mother sporocysts containing germ balls was taken as evidence that the miracidia were ‘completely vital’; snails were killed before the infection was patent, so there is no evidence as to whether cercariae would be produced from such an infection [84].

In 2002 Wippersteg *et al* used virions to deliver GFP under the control of the schistosome HSP70 promoter. They showed virions fusing with the worm tegument, and integration into the genome by southern blot [85]. However, this technique failed to result in delivery of the transgene to germ cells buried deep within the worm. They later used particle bombardment to

deliver GFP flanked by the promoter and termination regions of the cysteine protease ER60 [86]. This was carried out in order to determine the tissue specificity of ER60. RT-PCR and western blot confirmed expression of GFP. Although particles were distributed throughout the worm body, GFP was only detected in the excretory secretory system, confirming earlier results gained by conventional methods.

A promising result was described by Beckman *et al*, who used particle bombardment to deliver a plasmid encoding GFP under the control of the *S. mansoni* actin promoter to miracidia before allowing them to infect snails [87]. The resulting cercariae were collected and some from each snail were taken for molecular analysis, and the others used to infect mice. The resulting adult worms were also analysed by PCR. Confocal microscopy was carried out 48 hours post bombardment, and fluorescence was noted in a 'mosaic' like pattern on the tegument. This was taken as evidence that the DNA remained episomal rather than integrating into the genome [87]. Transcription levels of the transgene showed that the transgene was present in the first whole life cycle, but not in later rounds. Kines *et al* acheived incorporation of the firefly luciferase gene into the *S. mansoni* genome using a pseudotyped murine leukaemia virus [88]. Luciferase protein was detected by immunohistochemistry. However, it was not demonstrated that transduced schistosomula would mature and complete the life cycle if injected into mice [88].

Taken together, the techniques described above show great promise for the heritable introduction of a transgene into schistosomes; but substantial hurdles remain. If a construct for genome integration could be delivered to germ cells, by targeting an appropriate larval stage with particle bombardment, the transformed larvae could then be re-introduced to the life cycle and the resulting generations examined.

1.2.6. Global transcription profiling by microarray

Microarrays are a widely used tool for comparing transcription levels between different biological samples. This is achieved by hybridising fluorescently labelled cDNA from the samples under investigation to an array of cDNA or oligonucleotide probes immobilised on a solid substrate. After hybridisation, the array is washed under stringent conditions, and the fluorescence intensity level of each probe is recorded. The brighter the signal from a probe, the more cDNA was present in the test sample. It is important to note that information can

only be obtained for sequences which are represented on the array. In order to get statistically significant results, at least three biological replicates (i.e. material from separate organisms) must be carried out for each test.

Arrays can be made by printing cDNAs from a library onto glass slides, or by synthesising oligonucleotides *in situ* resulting in high-density formats. Many different platforms are commercially available. Affymetrix for example, use 20nt long probes perfect match (pm) paired with mis-matches (mm) where a single base is changed; 10 probes per gene are made, and the expression of a gene is calculated using the mismatch as a background control. Roche-NimbleGen use a digital micro-mirror array which enables them so synthesise 60mer probes, making mismatches unnecessary [89]. This technological advance allows very high density arrays to be synthesised. Whole genome arrays are available from a range of companies for various organisms, meaning that the expression of their entire genomes can be investigated under different conditions.

Data analysis from microarray experiments is no small task. First the fluorescence data must be quality assessed, and any aberrant arrays excluded. Background fluorescence must be accounted for, and the data must be normalised. Finally statistical analysis can be carried out to determine which genes are differentially expressed. It is usual to express the difference in gene expression in $\log_2$ fold change. Often a cut-off of fold change of 2 is applied. In addition, a statistical significance cut-off is used. The P-value is adjusted for multiple testing. Another statistic that is used is the B value, or log odds. This is the log (odds that a gene is differentially expressed). For example if B = 3, the odds that a gene is expressed is $e^3 = 20$, or 1 in 20, corresponding to a probability of 95%. If B = 0, the likelihood of differential expression is 50% [90].

There are many software packages available for microarray data analysis; one option is Bioconductor, an open source package of programmes in the statistical language R [91]. It can be used to perform a wide range of analyses from raw fluorescence data to meaningful biological information. Pair wise comparisons can be made using the limma package to produce top tables of the fold change in expression of each gene between samples [90].The Gene Ontology [92] is a controlled vocabulary of terms that describe gene products across species. There are three independent ontologies: molecular function, biological process and

cellular compartment. The ontologies are structured, so that root or parent terms are general and their descendants are more specific. One gene may be annotated to several GO terms, belonging to one or more ontologies. The GOstats package (http://bioconductor.org/packages/bioc/html/GOstats.html) can be used to determine whether particular groups of genes are enriched in one sample compared to another. Another possibility is to cluster the genes based on expression pattern. This can allow hypotheses to be made regarding the function of an unannotated gene if it clusters with genes whose products have been characterised.

Although microarrays offer a very sensitive method to characterise transcript levels in a sample, caution should be applied when inferring conclusions regarding the presence of a particular protein, as post-transcriptional control should be taken into account.

1.2.7. Previous applications of microarray technology to the study of schistosomes

The use of microarrays is a well-established tool for investigating gene expression patterns in schistosomes. Various different arrays have been designed and used to answer separate biological questions. As noted above, information about a particular sequence can only be gained if it is represented on the array. Microarrays have been used to address differences in gene expression between female and male adult worms [93, 94]; the rationale for this is that an insight into sexual development could lead to interventions to interfere with egg production, thereby reducing disease and transmission. Most recently this has been carried out using laser-capture micro-dissection of gastrodermis versus vitellaria [95]. Another popular approach is to measure gene expression between life cycle stages [96-99]. Here the focus is on identifying transcripts related to biological processes occurring in a particular life cycle stage. Even the effect of host gender on parasite gene expression profile has been investigated [100].

There are various different arrays in use by the community. They range from cDNA made exclusively from lung stage sequences [98, 101], a 44k oligo array [102], other oligo arrays based on ESTs available at TIGR [97] or WTSI [99], to an array with sequences from both *S. mansoni* and *S. japonicum* [103]. These different arrays have arisen as the sequence data available have increased over time, and/or to ensure that the genes relevant to the biological question under consideration were included.

Of particular relevance to this thesis are three published studies undertaking life cycle based experiments. The first, by Dillon *et al* reports the use of a custom made cDNA array comprising 6000 features representing ~3088 contigs and singlets from the lung stage larva [98, 101]. The authors state that this equates to approximately half of the 7,000 genes expected to be expressed by any one life cycle stage. This array was used to identify transcripts enriched at the lung stage compared to eggs, germ balls (dissected from snails 21 days post infection), day 2 and day 7 *in vitro* cultured schistosomula, and worms obtained from mice 21 days and 7 weeks post infection. The authors report a decrease in abundance of transcripts involved in aerobic metabolism and chromatin remodelling in schistosomula (both day 2 and day 7), compared to the reference sample. This supports the switch in larval metabolism described by Lawson and Wilson [104]. GOminer analysis revealed that genes involved in energy metabolism, cytoskeletal organisation, protease activity and chromosome remodelling showed the largest changes between lung schistosomula and adjacent life cycle stages. This array was designed specifically to show only lung stage transcripts, so it cannot answer questions regarding transcripts specific to the other life cycle stages.

The second, by Jolly *et al*, used an array comprising 12000 45-50mer probes representing contigs from the *S. mansoni* Genome Index at The Institute for Genome Research (TIGR). They compared daughter sporocysts (whole infected hepatopancreas six weeks post infection), cercariae and adult worms. Uninfected hepatopancreas was used as a control for the snail contribution to the daughter sporocyst sample. They report that, although the data set used to design the probes was of unknown orientation, >9,700 of the probes gave some signal, and therefore were likely to have been printed in the correct direction. Furthermore, only two probes gave a signal when hybridised with the snail only control. Jolly *et al* report that daughter sporocysts have high levels of transcripts encoding proteins involved in translation and quality control, as would be expected for this cercaria-producing stage. Transcripts involved in cell death and ubiquitination were also up-regulated; these may play an important part in morphological development. Genes highly expressed in cercariae were mainly those involved in mitochondrial function, enabling the energy production necessary for swimming. However, this stage was less transcriptionally active than the other stages studied. In contrast, adults produced a much wider range of transcripts; they were involved in egg production, immune evasion, and energy metabolism. Overall, 1,154 genes were differentially expressed in at least one contrast, most of these were enriched in the adult. 406 of the differentially

expressed genes had no known function. Interestingly, the hepatopancreas used was collected 6 weeks post infection, which the authors state is before cercariae were released. In our laboratory snails are shed routinely at five weeks post infection, so mature cercariae would have been present in the sample called daughter sporocyst. In spite of this, it was noted that the gene expression profiles of infected hepatopancreas and mature cercariae were different.

The third study was carried out by Fitzpatrick *et al*, using an array comprising 37,632 50mer probes [99]. The array was based on the sequence data available at GeneDB.org in 2005 [105]. They hybridised material from an impressive 15 separate life cycle stages: eggs; miracidia; mother sporocysts; daughter sporocysts (migratory stage obtained from snails 15 days after infection); cercariae; 3hr, 24hr, 3 day, 6 day *in vitro* cultured schistosomula; and worms obtained from mice at the following times post infection: 2wk, 3wk, 5wk, 7wk males, 7wk females. The authors analysed the data in several ways to get the most information from it. They carried out pair wise statistical analyses, a network analysis, GO category enrichment analysis, and focused on specific families of interest. Statistical analysis identified 973 non-redundant genes which were differentially expressed in at least one comparison, 448 transcripts exhibited a difference of more than 32 fold. Network analysis showed that 73% of the dataset were differentially expressed.

One finding Fitzpatrick *et al* highlighted was that the minichromosome maintenance (MCM) heterohexamer which is involved in licensing DNA replication is upregulated in the mother sporocyst and 3 week worm. This supports the observation by Clegg that cell division does not occur in migrating schistosomula, but remodelling is achieved by existing cells migrating [106]. Another discovery was that cercariae are not as transcriptionally inactive as previously thought. The authors focused on three gene families with may have value as intervention targets. Fucosyl transferases, tetraspanins, and G protein coupled receptors (GPCRs). The first group were examined as they are responsible for post translational modifications which may contribute to immunomodulation by the invading parasite. Tetraspanins are membrane proteins which are interesting not only because they have been identified on the surface of adult worms [107], but also because it has been shown that people with reduced susceptibility to re-infection by schistosomes have antibodies to them [108]. The array experiment showed that two tetraspanins (Sm04463 (best blast hit Smp_194970) and Smp_140000) were up-regulated in stages from cercaria to adult. Unfortunately only 19 of the 65 putative GPCRs

predicted in the genome were present on the array used by Fitzpartrick et al; however, they discovered that Smp_149770, Smp_152540 and Smp_127310 are enriched in the schistosomulum. Also highlighted as potential drug targets are Leishmanolysin, netrin and the netrin receptor.

Another approach Fitzpatrick *et al* took was to identify constitutively expressed genes which may be necessary for each life cycle stage. They take the view that platyhelminth-specific genes represent important genes to investigate for interventions – in contrast to the approach taken by Caffrey *et al* [4]. It could be a long and arduous task to discover more about the 294 platyhelminth-specific genes expressed in either cercaria to schistosomulum or later intra-mammalian stages and their products but one with rich rewards.

1.2.8. Molecules associated with the infection process

Long before microarrays were invented it was known that secretions from the acetabular glands of Schistosoma cercariae play an important role in host invasion. Gordon and Griffiths noted in 1951 that secretions from these glands separated layers of the stratum corneum and postulated that post-acetabular gland secretions may have a lytic effect on host cells [109]. In 1977, Dresden et al., demonstrated that the proteins from the cercaria had proteolytic activity against keratin and basement membrane proteins, but not collagen [110]. McKerrow *et al* showed that a 30kDa protease purified from cercarial secretions by chromatography could degrade type IV and VIII collagens as well as laminin and fibronectin [111]. Marikovsky *et al* immunolocalised a 28kDa protease to the pre- and post-acetabular glands, as well as the surface [112]. However, Salter *et al* stated that invasion is mediated by a single serine protease [113]. The first cercarial protease to be cloned was cercarial elastase (CE) in 1988 [114]. Newport *et al* demonstrated mRNA encoding CE in the post acetabular glands of developing cercariae by *in situ* hybridisation of a radiolabelled cDNA probe to sections of infected snail hepatopancreas. The transcript was not present in mature cercariae. Northern blots with the same probe showed the transcript in infected snails, but not in uninfected snails or adult schistosomes [114]. This work showed that there was transcriptional regulation both at the level of developmental stage and cell type. In 2002 Salter *et al* showed that there are five genes encoding cercarial elastase in the *S. mansoni* genome and stated that two of them account for more than 90% of the activity and protein released [115]. They also showed that

these genes are conserved between *S. mansoni* and *S. haematobium* and can also be identified in *S. douthitti*.

With the advent of both transcriptomic data and proteomic techniques, several studies of cercarial skin penetration have been carried out [116-119]. Curwen *et al* collected the proteins secreted by cercariae in the first three hours after mechanical transformation and separated them using 2D gel electrophoresis before identifying the spots by mass spectrometry. Cercarial elastase was found, along with SmPepM8 (a metalloprotease with homology to leishmanolysin), and a dipeptidyl peptidase IV. Curwen also reported several putative immunomodulatory proteins; these included 3 containing Sperm Coating Protein (SCP) domains (named SCP a, c, and d), a novel protein named SmKK7 with homology to BmKK7, (a potassium channel blocker found in scorpion venom), and Sm16 which had been identified by Rao and Ramaswamy [120]. This work suggested that the process of skin invasion by cercariae is more complex than had been previously acknowledged.

Proteins containing SCP domains are found throughout biology and have been identified in several helminths including hookworms *Ancylostoma duodenale* and *Necator americanus* [121], *Brugia malayi* [122] and *Onchocerca volvulus* [123]. Studies on the *Brugia* SCP (VAL-1) could not conclude whether it was secreted or not [122]. When jirds were vaccinated with Bm VAL-1 the average worm burden was not reduced [122]. An *Onchocerca* SCP has a role in promoting angiogenesis [123]. *Ancylostoma* secreted protein (ASP) contains the SCP domain; it is the largest constituent of the proteins secreted by infective L3 larvae. Anti-sera to the *N. americanus* ASP-2 inhibits larval migration *in-vitro.* Humans with anti-ASP-2 have lower average worm burdens [121]. This protein has undergone phase 1 clinical trails in naïve adults in non endemic region with some success; the vaccine was tolerated, and induced elevated antibody titres [124]. It remains to be seen whether there is any effect on worm burden.

A family of 28 *S. mansoni* genes encoding SCP domain containing proteins has been described recently by Chalmers *et al* [125]. They were named SmVAL (venom allergen like) 1-28. Expression of VALs 1-13 throughout the life cycle was investigated using qPCR. VALs 1, 4, 6, 7 and 10 were upregulated in cercariae relative to the other stages (Curwen's SCP a, c and d correspond to VALs 4, 10 and 18, respectively). VALs 8 and 12 were restricted to liver

stage worms, and VALs 2, 3, 5, and 9 were expressed by eggs, miracidia and mother sporocysts.

SmKK7 was named after the scorpion toxin BmKK7, the activity of which is dependent on the position of several key amino acids within a three-dimensional structure that is stabilized by three disulfide bridges [126]. The level of conservation between these important residues in the schistosome and scorpion proteins suggests a similar function for the schistosome molecule. Because T-cell activation is known to be regulated by potassium channels [127], it is plausible that the invading schistosome modulates the immune response of the host by this mechanism.

1.2.9. Localisation

A common theme in all of the *S. mansoni* sequencing projects to date has been the high proportion of 'schistosome unique' sequences obtained. Indeed, according to Wilson *et al*, informative identity may be as low as 26% [128]. The schistosome only genes are attractive as vaccine candidates and/or drug targets because of their lack of identity with host proteins [129]. Nevertheless, functional annotation of the remaining 74% of *S. mansoni* genes represents a large challenge. As discussed in Part 1, schistosomes have a complex body plan with tissues and organs. Therefore localisation of a specific gene can facilitate hypotheses as to its function. An example given by Dillon *et al* is the case of a protease: if localised to the gut it is likely to play a role in digestion, whereas if it is seen in the acetabular glands it is probably involved in host invasion [16]. An important fact to take into account when searching for vaccine candidates or drug targets is whether the target is accessible to the immune system or the potential drug. In the case of schistosomes, the host-parasite interface comprises the tegument, the gut, the secretory glands, and potentially the excretory system. Wilson and Coulson note that the most of the vaccine candidates tested in trials carried out by the WHO and other labs were cytosolic or cytoskeletal and therefore unlikely to be available to the host whilst the parasite is intact [130]. This fact led Curwen *et al* to state "it seems counterintuitive that such proteins would form the basis of an effective vaccine" [131], an opinion also held by DeMarco and Verjovski-Almeida [129], who also point out that *in silico* analyses can only be taken so far; a protein which is secreted may still be located deep within the body rather than exposed at the host-parasite interface [129]. There is no substitution for

experimental localisation data; this should be obtained early in the pipeline for both drug target and vaccine candidate validation.

Classically localisation studies are carried out by immunolocalisation of proteins. This is a long, difficult and costly process involving two particularly hard steps. Firstly the gene of interest must be cloned and expressed, or a peptide chosen and synthesised; secondly an animal is vaccinated with the recombinant protein to elicit an immune response. The serum is then collected and tested for reactivity to the protein in question. Finally the antiserum can be used to localise the protein, either in sectioned or whole mount worms, the latter is a long protocol in itself [32].

In order to allow higher throughput localisation studies, Dillon *et al* adapted whole-mount *in situ* hybridisation (WISH) for use in schistosomes. WISH has long been used in development to discover where specific transcripts are expressed during embryogenesis. This is the first application to a fully differentiated organism. A part of the gene of interest must be cloned (at least 250bp) and anti-sense digoxigenin (DIG)-labelled RNA probes made. These probes are hybridised to fixed, permeabilised whole mount parasites, and detected using an alkaline phosphatase-conjugated anti-DIG antibody in conjunction with a colourimetric or fluorimetric substrate. Not only is this technique higher through-put than immunohistochemistry (IHC), it also reduces the use of laboratory animals, and circumvents the often problematic step of protein expression. An example of this is antigen 10.3 (MEG 4.1) – a highly repetitive protein which was difficult to express *in vitro*. Dillon *et al* were able to synthesise probes and localised the transcript to the oesophageal gland of adult worms. The protocol described by these authors can be used for any of the intra-mammalian life cycle stages [16].

The availability of sequence data has enabled transcriptomic and proteomic experiments to be carried out, and so some of the molecules associated with infection by larval schistosomes have been identified. However, previous microarray studies have been hampered by the incomplete nature of the microarray platforms available and the exclusion of germ balls. A time line for germ ball development is lacking in the literature. Techniques for gene knock down and transgenesis are not yet routine in schistosomes. However, WISH can be used as a screen to discover which tissue(s) a gene is expressed in.

1.3. Aims of the thesis

The aim of this thesis is to investigate the gene expression patterns underlying infection of the mammalian host by larval schistosomes. This can be broken down further into two questions:
1) Which genes are differentially expressed from germ ball to cercaria to skin schistosomulum?
2) Where are these genes expressed?
In order to answer these questions a prerequisite is an in depth knowledge of the morphology of the life cycle stages involved. This is addressed in Chapter 2. The next part of the story was explored using a gene expression microarray experiment which is described in Chapter 3. Chapter 4 deals with localisation of some of the gene products highlighted in Chapter 3 and in the literature concerning the infection process.

2. Morphological study

2.1. Introduction

Schistosomes undergo two transitions in the part of the life cycle under investigation; firstly from the snail intermediate host to fresh water, and secondly, from fresh water to the mammalian definitive host. These three contrasting environments present different challenges for the parasite. Considerable morphological and physiological changes take place as the larva adapts to each niche. An understanding of these morphological changes will provide the context for investigations into the gene expression patterns associated with infection by larval schistosomes.

Morphological studies have been carried out on these life cycle stages. Cercariae have been studied more extensively than either germ balls or skin stage schistosomula; a situation which has not changed since a review published in 1974 [49]. The early work on schistosome larval morphology was carried out using conventional light [38], or electron microscopy [39, 41, 42, 44, 47, 132]. These methods enabled broad observations of body structure or detailed studies of ultrastructure, respectively, to be made. As a result, much has been revealed about the morphology of these important life cycle stages; however, several questions remain. No definitive timeline has been established for germ ball development. This is partly due to the asynchronous nature of development, but also to a lack of regular sampling [38] and the difficulty with reconstructing images from sections of infected snail hepatopancreas [41]. Another question which remains is the cellular make up and function of the head gland. In addition the skin stage schistosomulum has not been well studied.

More recently the advent of confocal microscopy has enabled detailed studies to be made of internal structures using whole mount worms. This technique allows pictures to be taken with a restricted focal depth, so that ‘optical sections’ may be imaged. A series of optical sections, called a z-stack, can be easily built up into a 3-D representation of the specimen using specialised software. This technique has several advantages over conventional and electron microscopy; the technically difficult process of sectioning small specimens is unnecessary, and the orientation is always known. This aids interpretation of the resulting images.

New staining methods have become available for use in conjunction with confocal microscopy. Langeron's carmine is a general protein stain which is visible by both light and fluorescence microscopy. It has been used successfully with confocal microscopy to provide novel insights into the anatomy of adult schistosomes [34]. It has not yet been applied to larvae. Phalloidin is a toxin from the 'Death Cap' mushroom (*Amanita phalloides*) which binds specifically to filamentous actin (f-actin). It can be conjugated to various fluorophores and used to examine musculature. Studies of adults [32], cercariae [35], miracidia and mother sporocysts [133] have been carried out using this method and have revealed that the outer muscle layers, oesophageal and gut musculature, acetabulum, and flame cells all stain well. This makes phalloidin an attractive stain for use with larval schistosomes. Nuclei may also be stained using propidium iodide (PI) and 4',6-diamidino-2-phenylindole (DAPI). These are aqueous stains which bind to DNA and fluoresce when excited by a 543nm (PI) or 405nm (DAPI) laser. These stains are ideal as a counterstain with phalloidin, but they cannot be used with the ethanolic Langeron's Carmine.

These techniques will be applied for the first time to daughter sporocysts, germ balls, and skin stage schistosomula. Although cercariae have been studied using phalloidin, this represents the first use of Langeron's carmine with this stage.

Aims

The aims of this study are:

- To discover when important cercarial structures, such as the secretory glands, are developing in the germ ball. The information gained will be used to ensure that germ balls are sampled appropriately for the gene expression study (Chapter 3).
- To describe the anatomy of the daughter sporocyst, cercaria, and skin stage schistosomulum. This knowledge will be necessary to interpret the data gained in the following chapters, particularly the localisation studies (Chapters 4).

2.2. Methods:

2.2.1. Parasite material

All parasite material was from a Puerto Rican isolate maintained at the University of York by passage through NMRI mice and albino *Biomphalaria glabrata* snails. To obtain daughter sporocysts and their developing germ balls, snails were infected with 40 miracidia each and dissected carefully in filter-sterilised 50% phosphate buffered saline (PBS; 3.2 mM Na2HPO4, 0.5 mM KH2PO4, 1.3 mM KCl, 135 mM NaCl, pH 7.4.) 23 – 27 days later. Obvious snail material was removed with a finely drawn-out Pasteur pipette. Daughter sporocysts and germ balls were harvested separately using a clean drawn out pipette and kept on ice until use.

Cercariae were collected from snails infected with 10 miracidia each. Five weeks after infection the snails were placed in the dark for two days and then placed under a bright light in approximately 10mls aerated tap water (pond water) for two hours to induce shedding of cercariae. The resulting cercariae were concentrated by cooling on ice for one hour to prevent swimming.

To obtain skin stage schistosomula, cercariae were obtained as above, and transformed mechanically before culturing for three days *in vitro* [134]. After concentration on ice and two brief washes in cold RPMI media (Invitrogen Gibco), the cercariae were transferred to 3mls M169 media (see below) and vortexed at full speed for 90 seconds to shear the bodies from their tails. They were then separated on a discontinous percoll gradient (70% and 45% percoll in RPMI). The bodies were recovered from the interface with a pastette, taking care to avoid the tails in the upper layer. Bodies were washed extensively in RPMI and then plated out into 24 well plates in M169 (BME (Eagle) liquid medium (Gibco) containing: 0.1% Lactalbumin hydrolysate, 0.1% Glucose, 0.5μM Hypoxanthine, 1μM Serotonin, 1μM Hydrocortisone, 0.2μM Triiodothyronine, 0.5% MEM vitamins, 5% Schneider's drosophila medium, 1% Penicillin/streptomycin) and incubated at 37°C with 5% CO_2 for three days. They were then recovered and washed twice more in RPMI [64].

2.2.2. Langeron's Carmine staining

Parasites were fixed in AFA (22% formalin, 33% 95% ethanol, 11.5% glacial acetic acid, 33.5% water) overnight, and then stored in 70% Ethanol until use (at least 10 minutes if they were used immediately). The ethanol was removed and replaced with a few drops of Langeron's Carmine (LC) and incubated at room temperature for 50 minutes. Excess stain was removed by washing the parasites several times with 70% ethanol. The parasites were then dehydrated through 90% ethanol and two changes of absolute ethanol (10 minute incubations at RTP). They were cleared for 15 minutes in Histoclear (Fisher Chemicals), mounted onto slides with DPX (BDH Chemicals), allowed to dry overnight and stored at room temperature.

2.2.3. Phalloidin and DAPI staining

Daughter sporocysts, germ balls, cercariae and skin stage schistosomula were stained with Alexafluor 488-conjugated phalloidin (Invitrogen, Molecular Probes) to visualize f-actin, and counterstained with DAPI according to Bahia *et al* [133]. Specimens were fixed for one hour in 4% paraformaldehyde in PBS on ice and then permeablised for an hour in PBS, 0.2% gelatin, 0.1% saponin, 0.1% NaN_3 (PGN) at room temperature, before addition of phalloidin (0.0165μM, 20 minutes) and/or DAPI (0.01μM, 30 minutes) in PGN at room temperature. Specimens were washed several times in PBS before they were viewed with the Zeiss invert confocal microscope.

2.2.4. Confocal microscopy

Confocal microscopy was carried out using a Zeiss LSM 510 meta on an Axioplan 2M (upright), or a Zeiss LSM 510 meta on an Axiovert 200M (invert).
Pictures were taken with a Plan-Aprochromat 63x/1.4 oil DIC under oil immersion.

Stain	**Laser**	**Beam splitter**	**Filter**
Langeron's Carmine	543nm	HFT 488/543	Long Pass 585nm
Phalloidin	488nm	405/488/543	Band Pass 505-530nm
DAPI	405nm	405/488/543	Band Pass 420-480nm
PI	543nm		Long Pass 585nm

Table 2.1 Confocal microscope settings

2.3. Results

2.3.1. Daughter Sporocyst

Migratory (Fig. 2.1) and mature daughter sporocysts (Fig. 2.3), were stained with phalloidin to visualise f-actin. Note coverage of anterior end with posterior-facing, actin-rich spines highlighted by phalloidin staining (Fig. 2.1). Spine coverage is sparse in the mid region of the sporocyst with no spines at the posterior end (Fig. 2.1). There is a protrusion at the anterior end (Fig. 2.1b). Circular and longitudinal muscle fibres can be seen, and are arranged in a pattern reminiscent of a loose weave material. Four flame cells are visible along with their collecting tubules (Fig. 2.1b).

Langeron's Carmine staining reveals the cellular make up of the migratory daughter sporocyst (Fig. 2.3). Most cells are irregularly shaped. The outer edges of the body comprise small cells (approximately 4-5μm in diameter) with fewer large cells (8-10μm diameter) comprising the internal core of the body. These core cells may be germ cells which will develop into germ balls. Extensive punctate staining in the nucleus of the majority of large cells provides evidence of cell division, the cytosol of these cells is full of protein (Fig. 2.2).

When the daughter sporocyst is further developed it is no longer vermiform (Fig. 2.3). Bulbous regions of the body are connected by actin-rich isthmuses. The surface spines are still in evidence, with the same distribution as in the younger specimens (Fig.s 2.1 and 2.2). The spines are most numerous on the very anterior bulb where they are paired; fewer can be seen on the second bulb. The PI staining reveals many cells packed into the bulbs. These are the 'brood chambers' where the germ balls will develop into cercariae. The outer muscle layers of these brood chambers form a 'loose weave'.

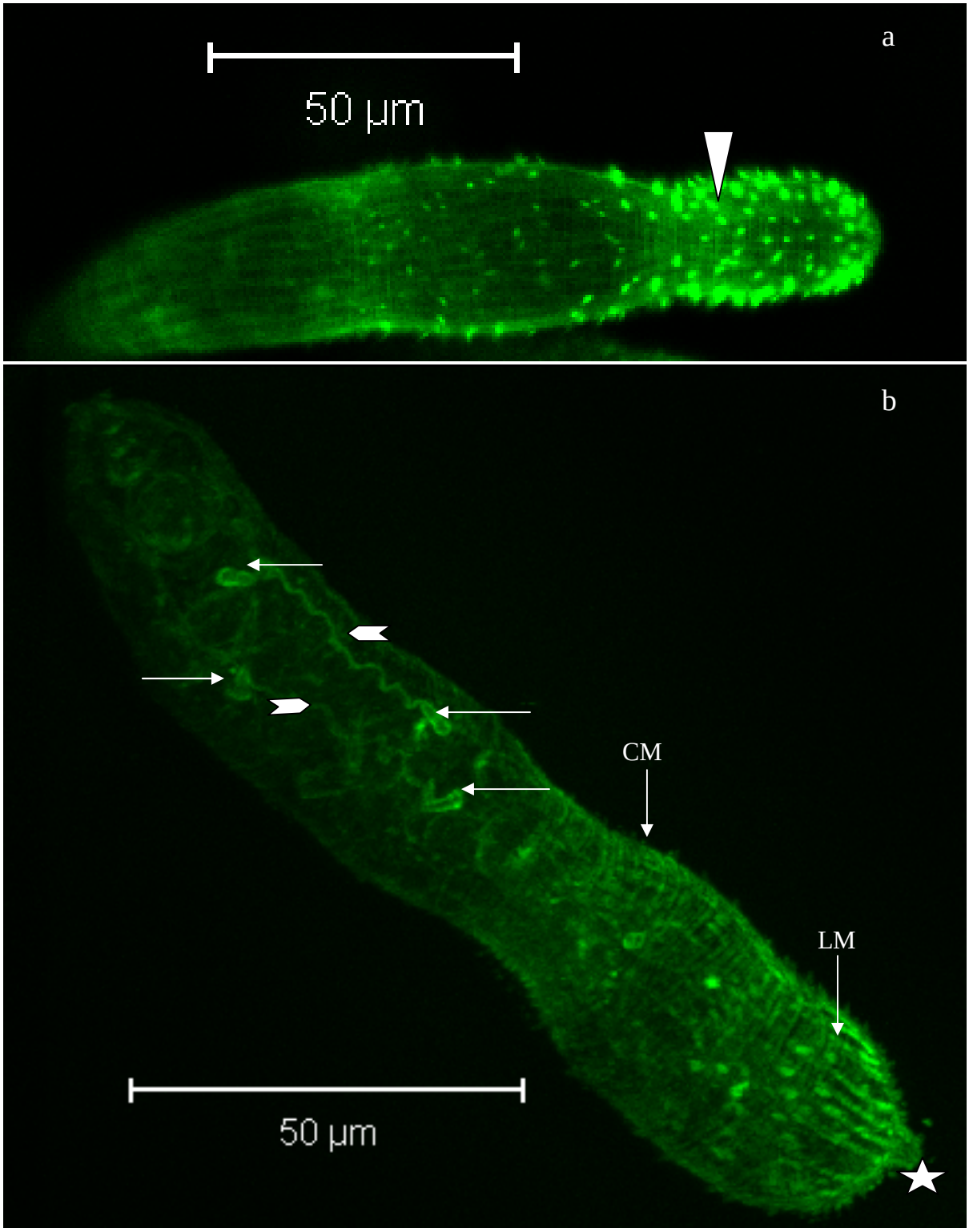


Figure 2-1 a and b Migratory daughter sporocysts stained with Alexafluor 488-conjugated phalloidin a) single slice b) 3D projection of 17.47µm deep z-stack. Note spines at anterior end (arrowhead), longitudinal (LM) and circular muscle (CM), and four flame cells (arrows) with the ducts (chevrons) connecting them. There is a protrusion at the anterior end (star).

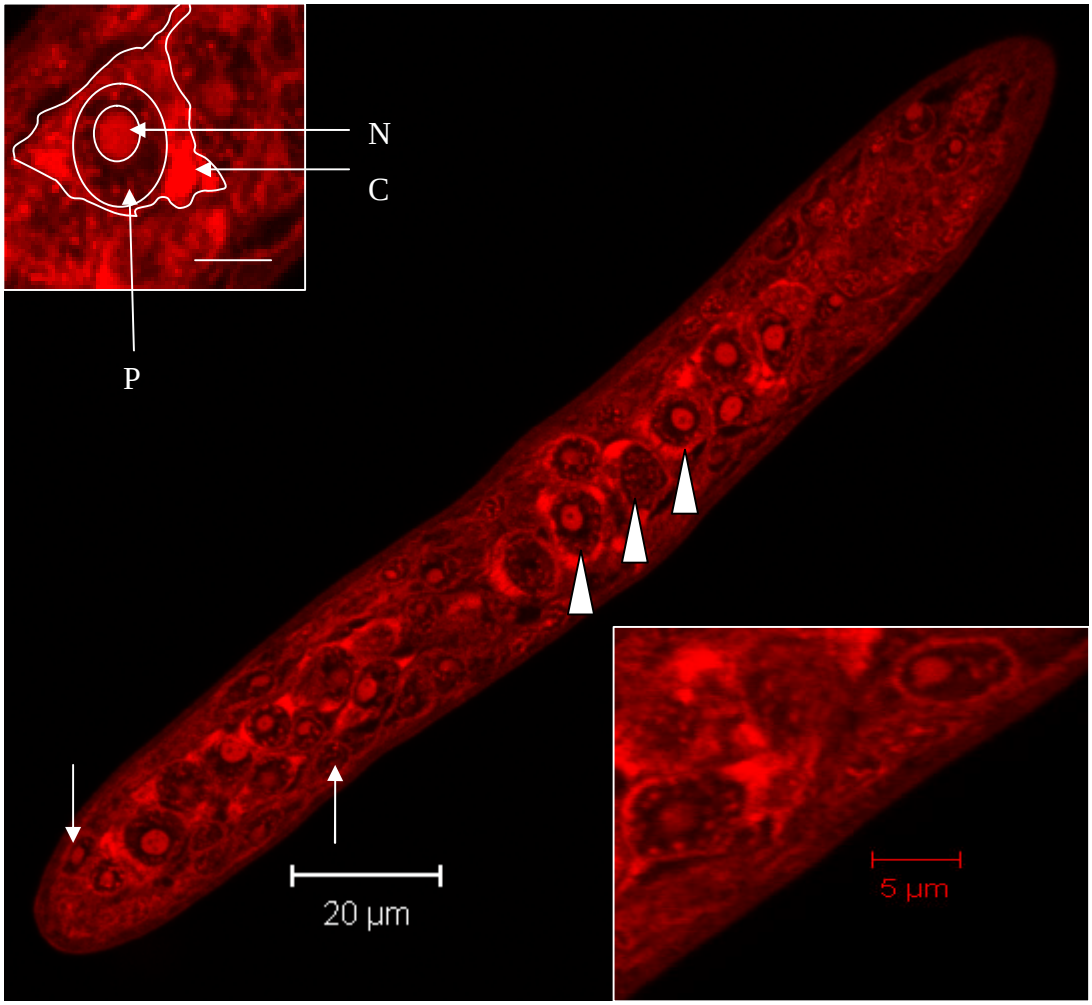


Figure 2-2 Migratory daughter sporocysts stained with Langeron's carmine.
Parasites were recovered from snails 21 days post-infection and fixed overnight in AFA and stored in 70% ethanol. They were and viewed using an upright confocal microscope using the 543nm laser. This specimen is ~180µm long. The optical section shown is 14µm into the 23µm z-stack. The core of the body consists of large cells (arrow heads) with smaller cells arranged along the periphery (arrows). Inset are enlargements of cells. Note prominent nucleolus (N), punctate staining in the nucleus (P), and staining throughout the cytosol (C) scale bar 5µm.

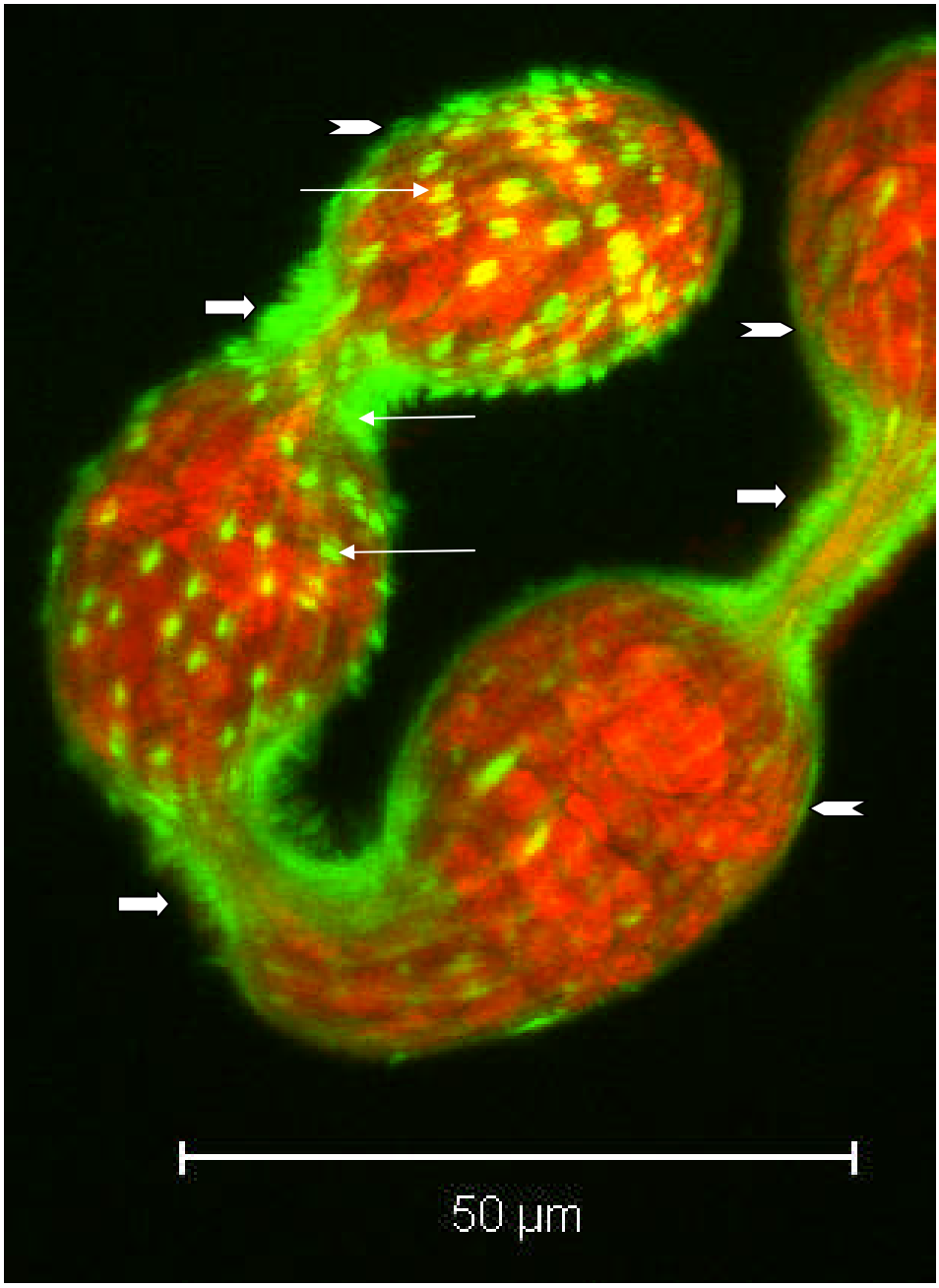


Figure 2-3 Projection of a 38µm deep z-stack of a daughter sporocyst

This specimen was stained with alexafluor 488-conjugated phalloidin (green) as above, with the addition of a 5 minute incubation with PI (red) before viewing by confocal microscope using 488nm and 543nm lasers. Note brood chambers (chevrons), narrow actin-rich constrictions (large arrows). Surface spines (arrows) are paired on the most anterior bulb and are distributed singly on the second bulb.

2.3.2. Germ ball development

The first germ ball documented is the single germ cell inside the daughter sporocyst. Next is a loosely connected ball of cells which are not uniform in size or shape (Fig. 2.4a).The cells

vary in diameter from 2-9μm, they have prominent nucleoli, and there is no staining in the cytoplasm. Some cells with two brightly staining spots can be seen indicating that they are about to divide (not shown). The next stage in development is that the germ ball elongates (Fig. 2.4b). At this point there are some prominent large (7μm diameter) cells at its edge. The remaining cells average 4-5μm in diameter. At this stage it is difficult to distinguish the anterior end from the posterior, and the germ ball does not have an outer epithelium. The cells are tightly packed together, in contrast to the earlier stage.

By the time the germ ball is ~50μm long, an epithelium surrounds the body (Fig 2.5). There are cells contiguous with this epithelium (Fig.s 2.5b and 2.7), which I will refer to as contributing cells. The posterior end of the germ ball is flattened (Fig. 2.5) and towards its anterior there is a group of cells arranged in an arc which may be the developing head capsule or the developing acetabular gland cells (Fig. 2.5a). The appearance of the majority of cells has changed. Unlike those in the early round germ ball (Fig. 2.4a), they stain positive for protein in the cytoplasm, indicative of protein synthesis.

When the germ ball is approximately 75μm in length (Fig. 2.6) the epithelium and some contributing cells are still clearly visible. In addition to this, bridges of protein stain are now evident, which may link internal cell bodies with the outer epithelium (Fig. 2.8). The head capsule area at the anterior end of the body is also more developed with cells elongated along the antero-posterior axis (Fig. 2.6). Phalloidin staining at this stage (Fig. 2.6c) shows that circular muscle is present, and although some strands of longitudinal muscle are evident, they are far fewer in number. There is a protrusion at the anterior end which may be developing sensory endings (Fig. 2.6b).

By the time the germ ball has a tail (body ~80μm, tail ~50μm, Fig. 2.9) its cells are well organised and some organ systems are recognisable. The body may be described in quarters from anterior to posterior. The first, anterior quarter, is made up of cells that will become the head capsule. They are well ordered, but there are no visible structures. The second quarter contains the neural ganglia. The third quarter comprises all 10 acetabular gland cells in the centre and three clearly separate bands or layers of cells on each side. The posterior quarter contains undifferentiated cells. The cells in the tail are neatly arranged in rows along the side and the tail is bifurcated from this early stage (Fig. 2.9a). A phalloidin stained germ ball of

approximately this age has two visible flame cells (Fig. 2.9B), and the outer muscle layers are denser in the region of the head capsule than across the rest of the body. Fibres of actin can be seen in squares on the tail which has two little knobs where the furcae are starting to grow.

Phalloidin staining of an older germ ball (Fig. 2.10) (tail detached during fixation) shows that the circular muscle layer is formed first, with the longitudinal muscles following. The latter are formed from bands of muscle cells in tiers around the body after the head capsule muscle boundary has formed. Diagonal muscle strands are also visible at this stage.

In order to discover the time line of development, snails were infected with 40 miracidia each and dissected daily starting 21 days later. At this early time point I noted that only small round germ balls were present (Fig. 2.4a – Cheng and Bier stage 2). Motile cercariae were observed for the first time at day 27 post- infection. At days 24 and 25, there was a range of germ balls from small round ones, through various stages with tails (Cheng and Bier stages 5 and 6) but not including motile cercariae.

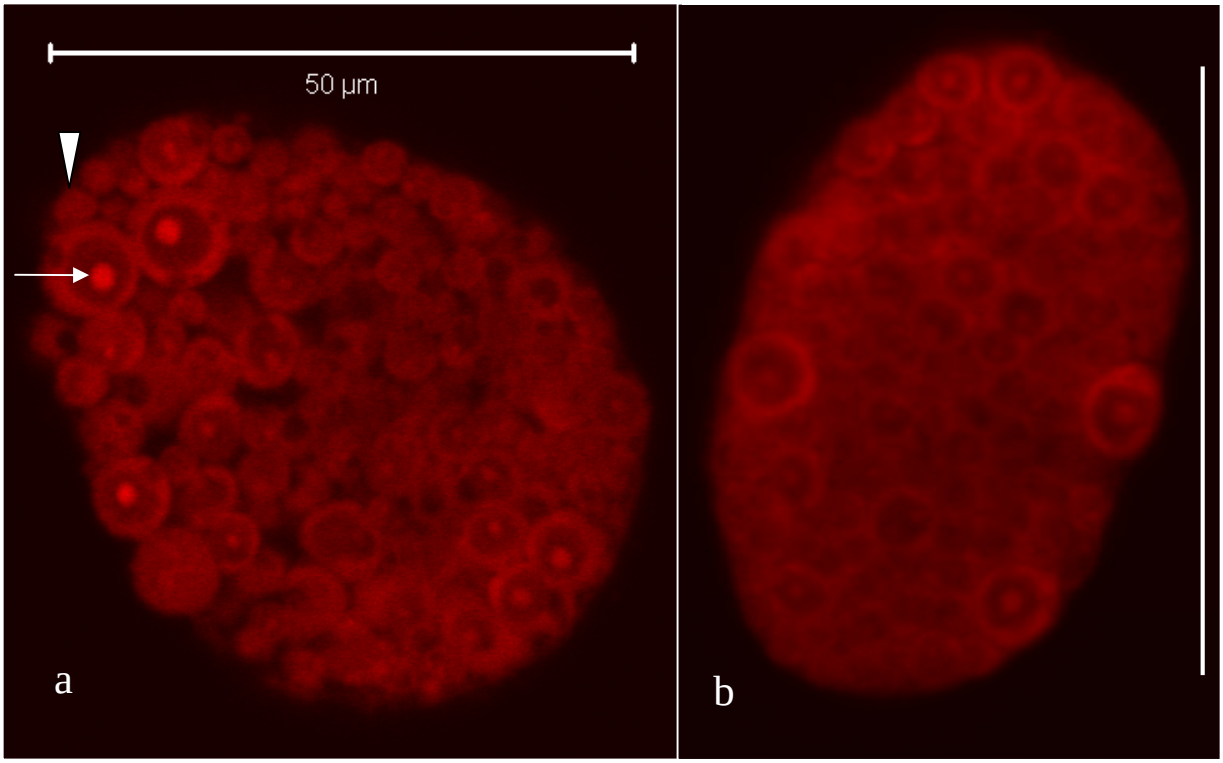


Figure 2-4 a and b. Optical sections of germ balls stained with Langeron's carmine.
Germ balls were recovered from snails at daily intervals from 21-27 days after infection with 40 miracidia and fixed in AFA. They were stained, dehydrated and viewed with a confocal microscope with a 543nm laser at 63x. Scale bars 50µm. a) cells vary widely in size and have prominent nucleoli (arrow). Protein staining is absent in the cytosol. Small cell highlighted by arrowhead. b) an elongated germ ball, anterior and posterior ends are indistinguishable. Cells are still variable sizes.

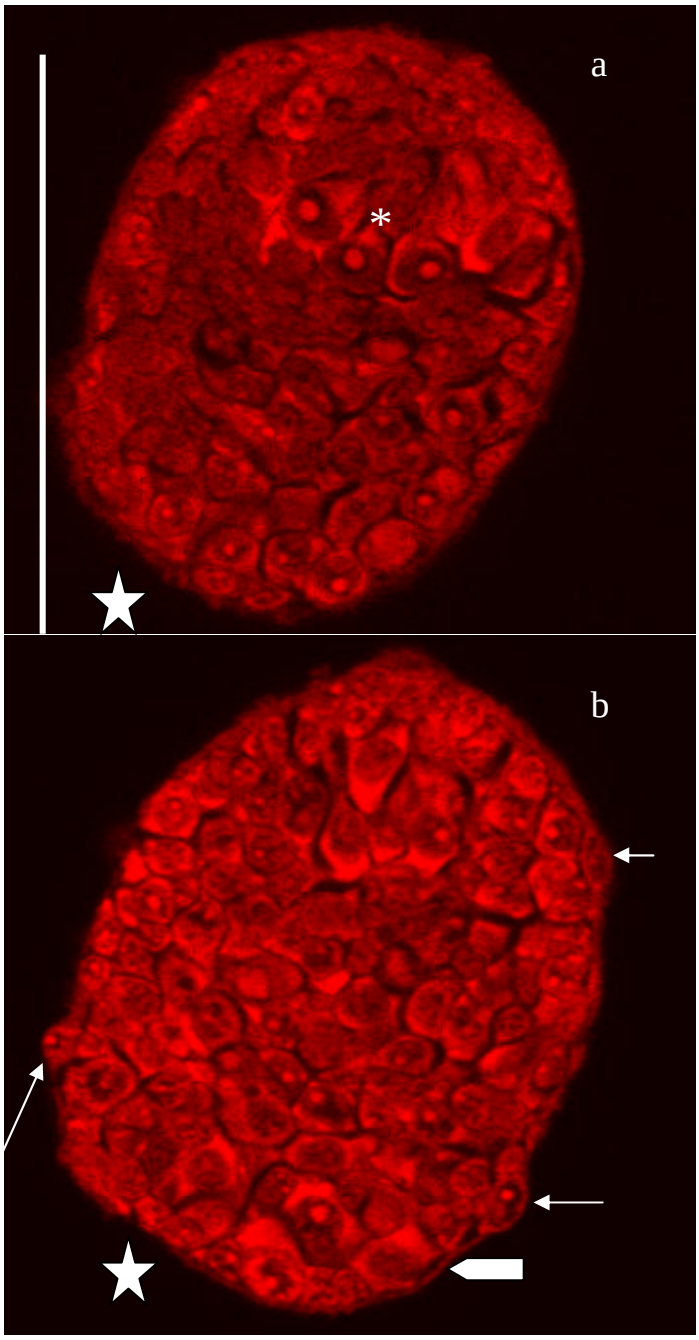


Figure 2-5 a and b. Young germ ball stained with Langeron's carmine

Selected optical sections, 23μm and 10μm through a 27μm deep z-stack, are shown. Note that the posterior end is flattened (stars). The body is more organised; towards the anterior end there is a group of cells arranged in an arc shape (asterix). Very thin Epithelium covers the body (block arrow), and cells contributing to it can be seen (thin arrows). Scale bar 50μm.

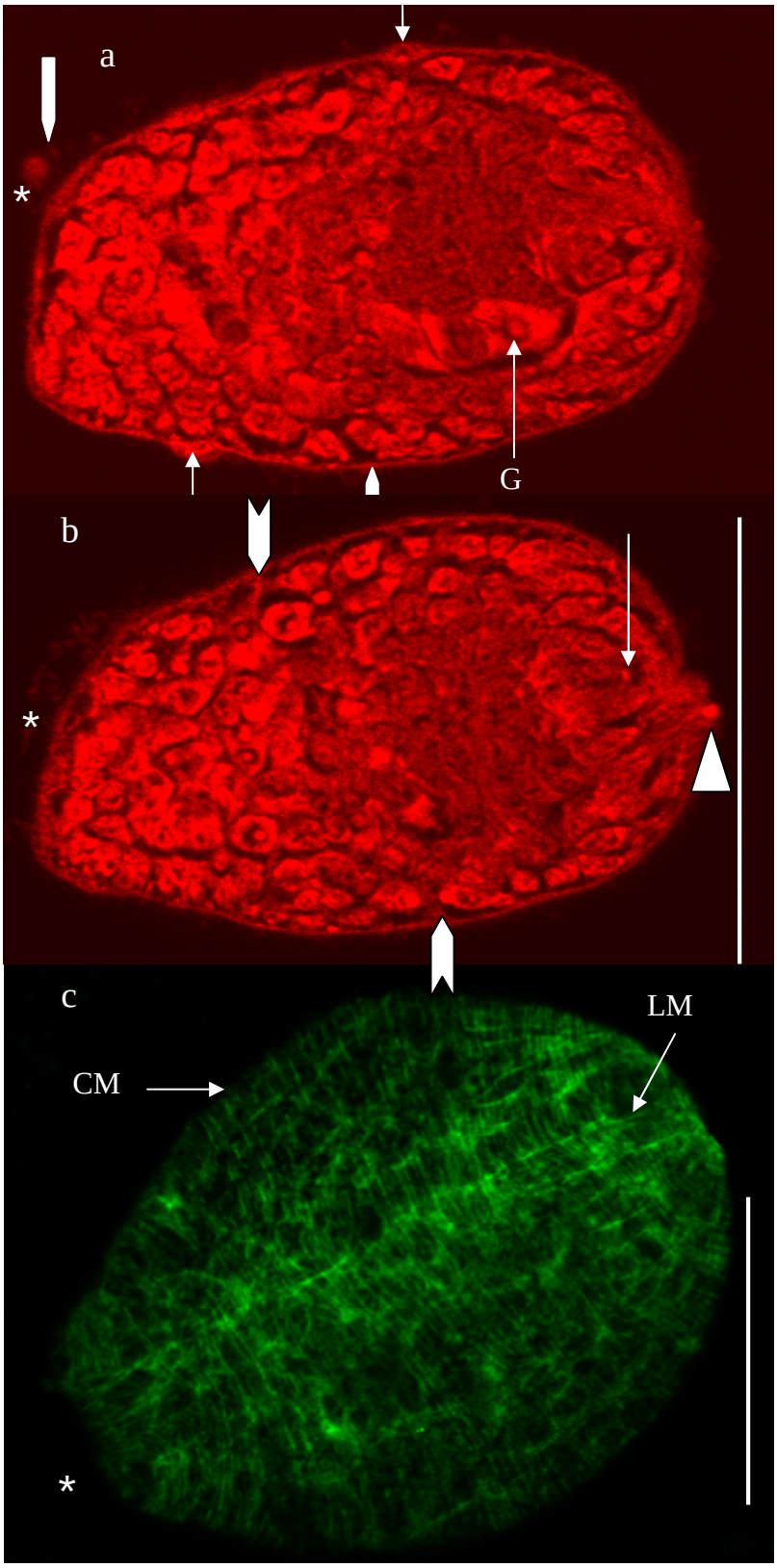


Figure 2-6 Developing germ ball.

Figure 2-6a) and b) optical sections, 24μm and 19μm through a 38μm deep z-stack, of a germ ball approximately 75μm long. Gland cells are appearing (G), the epithelium (block arrows) and contributing cells are still present (arrow). Connections between parechymal cells and the epithelium are evident (chevrons). The anterior end is organising into the head capsule cells are elongating along the anterioposterior axis (EC). There is a protrusion at the anterior end (arrow head). c) Germ ball of approximately the same developmental stage stained with phalloidin to visualize F-actin. Circular muscle is well developed (CM), and the longitudinal muscle is developing (LM), note that posterior end is flattened (asterix).

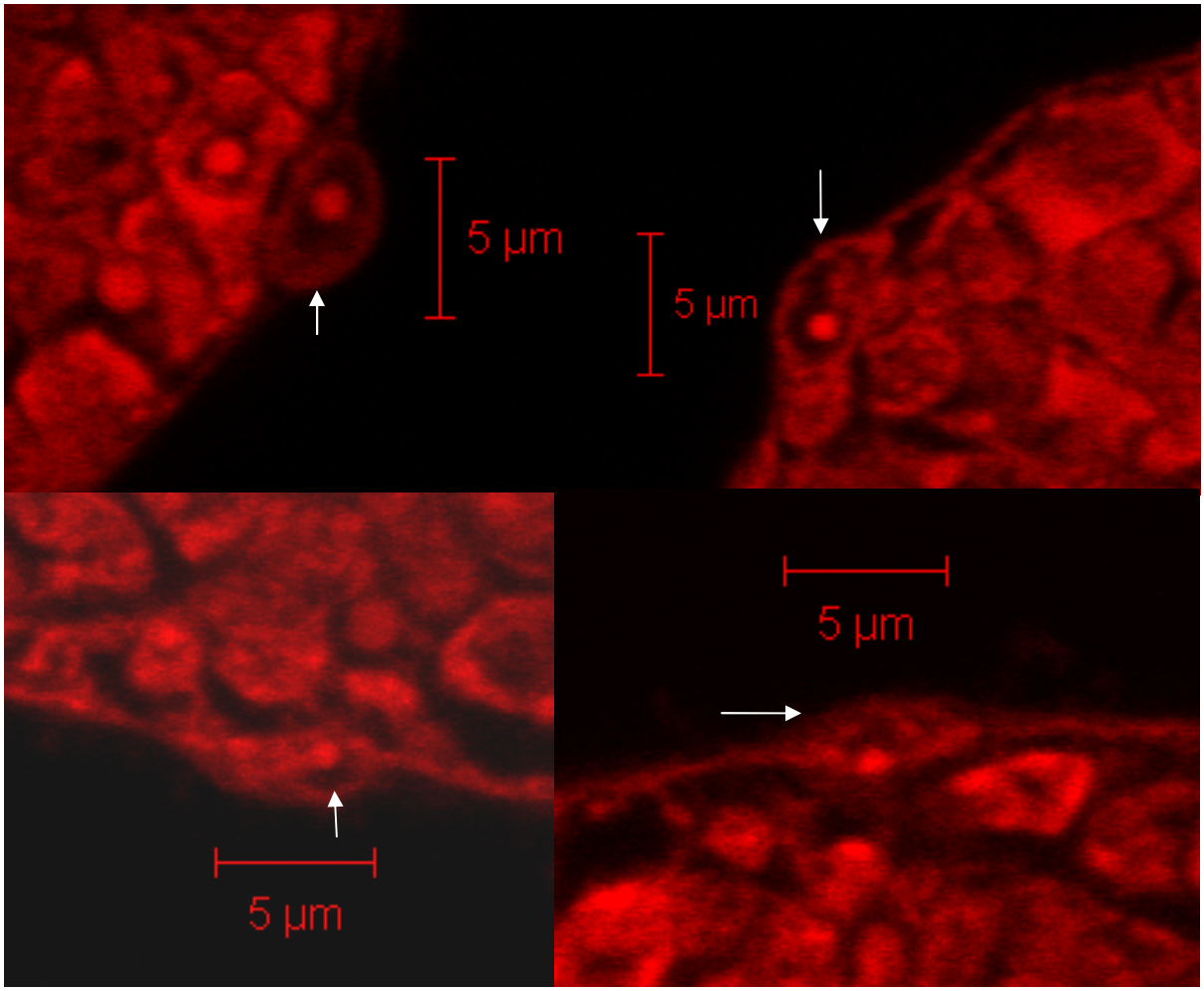


Figure 2-7 Outer epithelium of germ balls.

Areas cropped from germ balls stained with Langeron's carmine shown in figures 2.5 and 2.6. Cells forming the epithelium are shown (arrows).

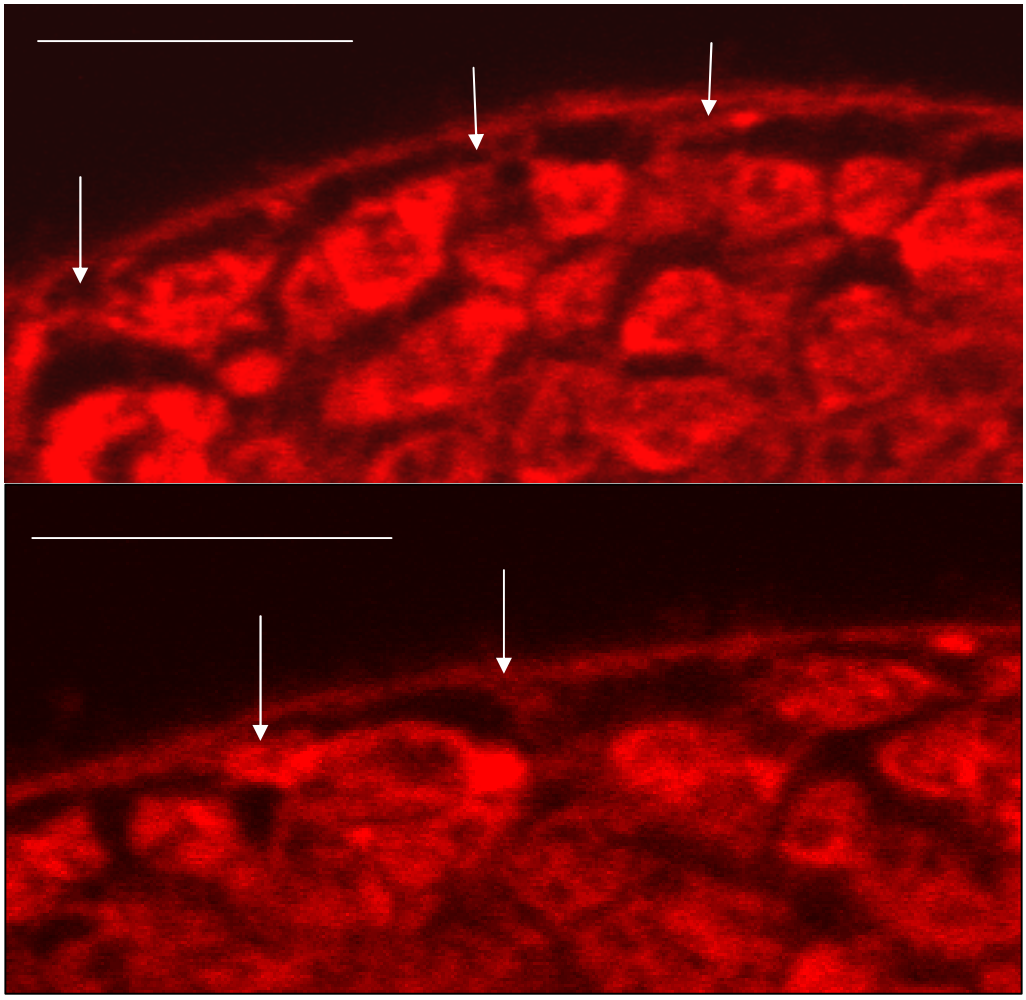


Figure 2-8 Possible subtegumental cells

Areas cropped from germ ball shown in Figure 2.6 highlighting protein ‘bridges’ which may connect cells inside the body to the epithelium (arrows). These cells may be the first sub-tegumental cell bodies. Scale bars 10µm.

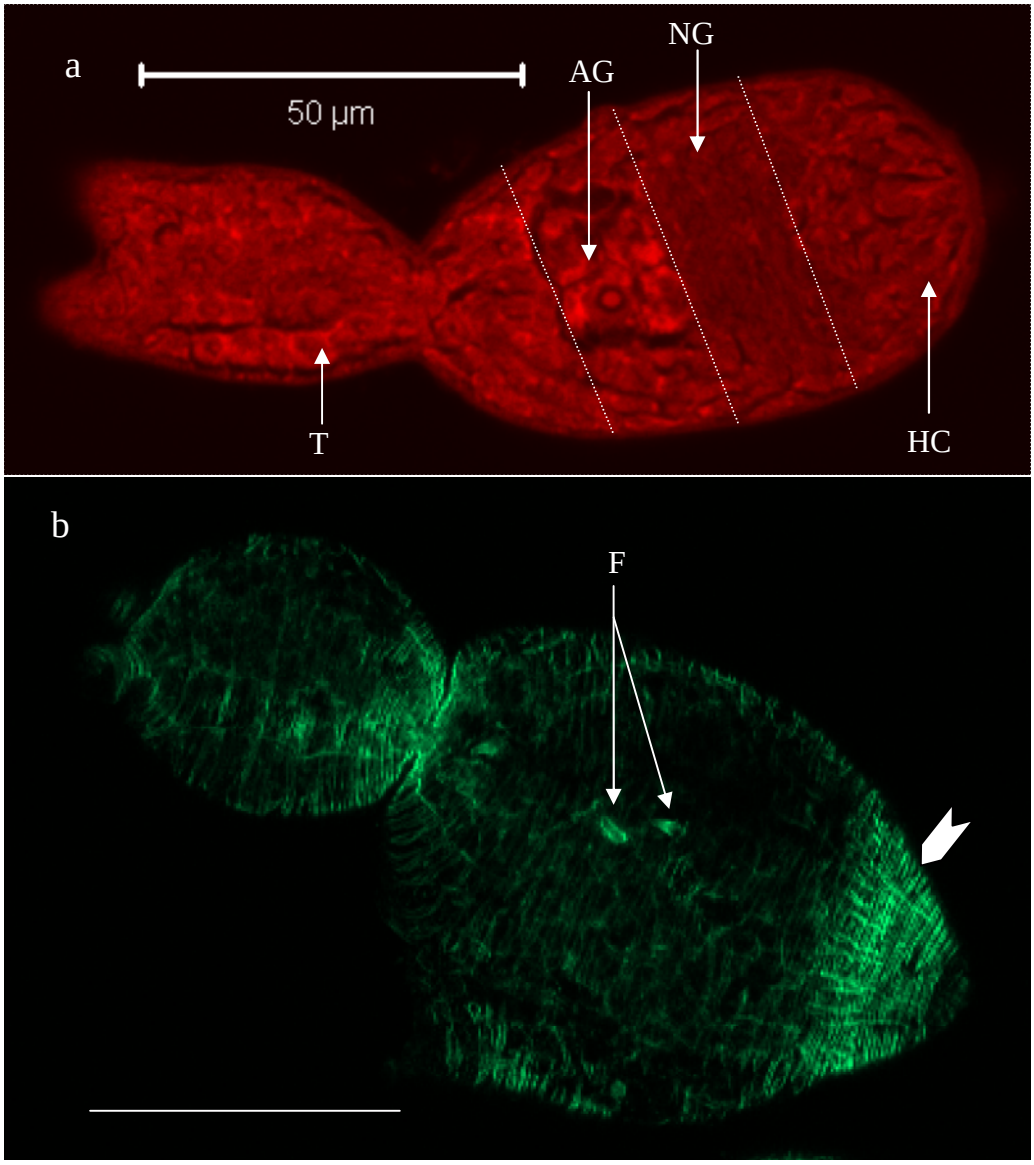


Figure 2-9 a and b Germ balls with stubby tails.

a) Optical section of a parasite stained with Langeron's carmine. Putative neural ganglia (NG), head capsule (HC), acetabular glands (AG) and tail (T) are visible. b) Optical section (1.46µm into a 10.2µm z-stack) of a germ ball stained with phalloidin, note two flame cells in the centre (F) and the well developed muscle at the anterior end (chevron). Scale bar 50µm

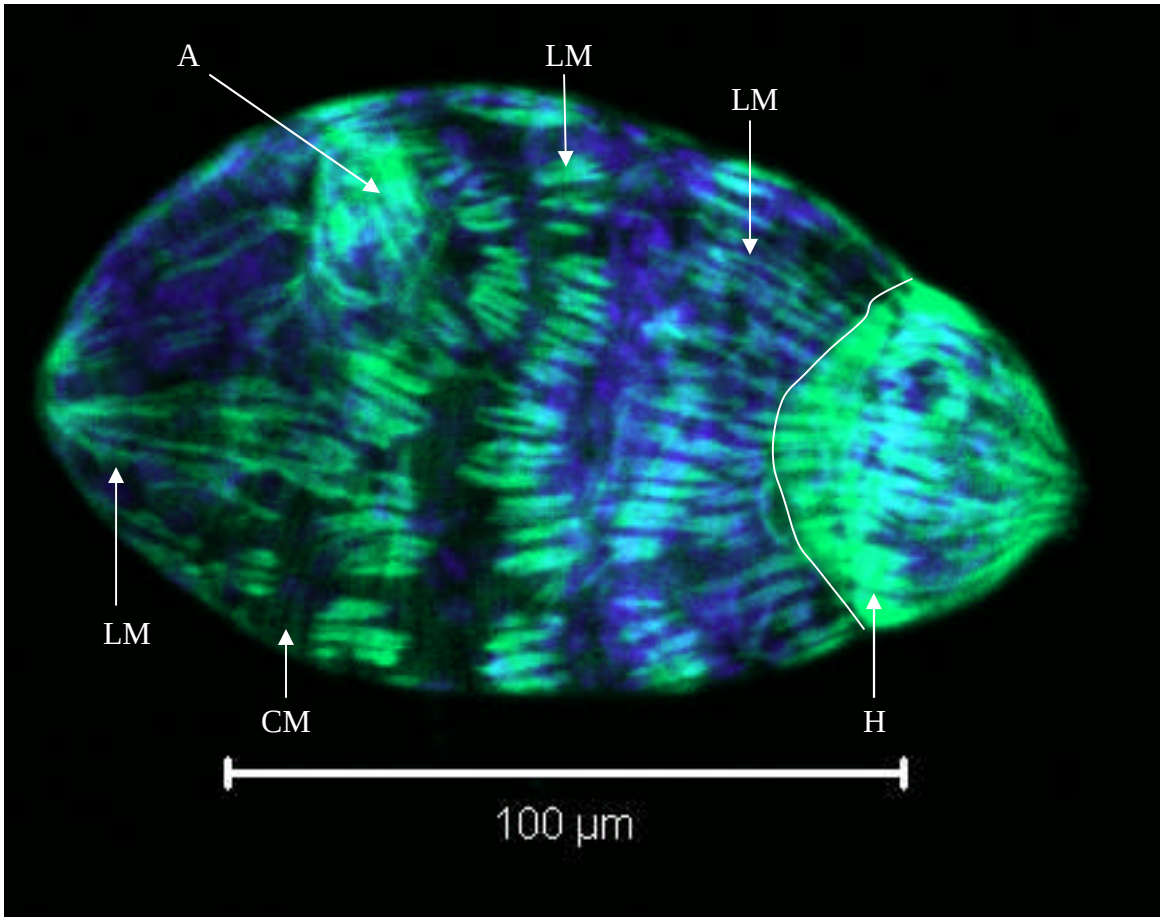


Figure 2-10 Muscle development in a maturing germ ball

Optical section of a more developed germ ball stained with phalloidin (green) and DAPI. The tail detached during fixation. Note acetabulum (A), the circular muscle is formed (CM), longitudinal muscle is developing (LM), head capsule musculature (H) more developed than that of the body.

2.3.3. The Cercaria

The fully developed cercarial body measures 190µm in length and 40µm in width. Posterior facing spines cover the body and tail. Most striking at this stage are the 10 acetabular gland cells which comprise approximately one third to half of the cercarial body volume (Fig. 2.11b). These cells have large nuclei and nucleoli (Fig. 2.13a and b). Their ducts form two bundles which are surrounded by muscle (Fig. 2.11a), and extend to the anterior of the head capsule, where they reach to the exterior of the larva through the tegument. There are protrusions from the surface in this area shown by LC staining (Figs 2.13a and b). The muscles surrounding the two bundles of acetabular gland ducts are stained by phalloidin, and can be seen on either side of the body just posterior to the head capsule boundary and also within the head capsule (Fig. 2.11a). This stucture extends posteriorly and is bounded by two layers of muscle: circular and longitudinal. At the anterior end of the body there are also

diagonal muscles, crossing from anterior to posterior (Fig. 2.11d). The head gland can be seen in Figs 2.13b and c, lying anterior and dorsal to the oesophagus, extending from the junction between the tegument and the head capsule muscle boundary, to the anterior end of the body. This gland is similar in appearance to the acetabular glands when stained with LC in that there is very little staining in the cytoplasm. There are two prominent nucleoli (Fig. 2.13b and c).

The acetabulum or ventral sucker is positioned two thirds down the length of the body. It is supported by eight lengths of muscle anchored dorso-laterally (Fig. 2.11a). The acetabulum is made up of many cells. This is highlighted by DAPI staining (Fig. 2.11a), whereas phalloidin staining shows concentric rings of actin fibres and strands of muscle extending radially from the centre to the edge (Fig. 2.12a). Pseudo-striated muscle can be seen in the tail (Fig. 2.13b). There are 8 flame cells in the cercarial body, and two at the proximal of the tail; they are arranged in pairs laterally. There are two pairs at the posterior of the body between the acetabulum and the body/tail joint, one pair lies dorsal, the other ventral (Figs 2.12 a and d, respectively). The middle pair lies in the same plane as the caecum, just posterior to it (Fig. 2.12b). The anteriormost pair is dorsally situated, between the caecum and the head capsule boundary (Fig. 2.12c). In the area between the acetabular glands and the head capsule, the cells are small and tightly packed. DAPI staining also shows a group of undifferentiated cells, approximately the same size as the acetabulum, on the ventral aspect, posterior to the acetabulum, between post -acetabular glands (Fig. 2.11a).

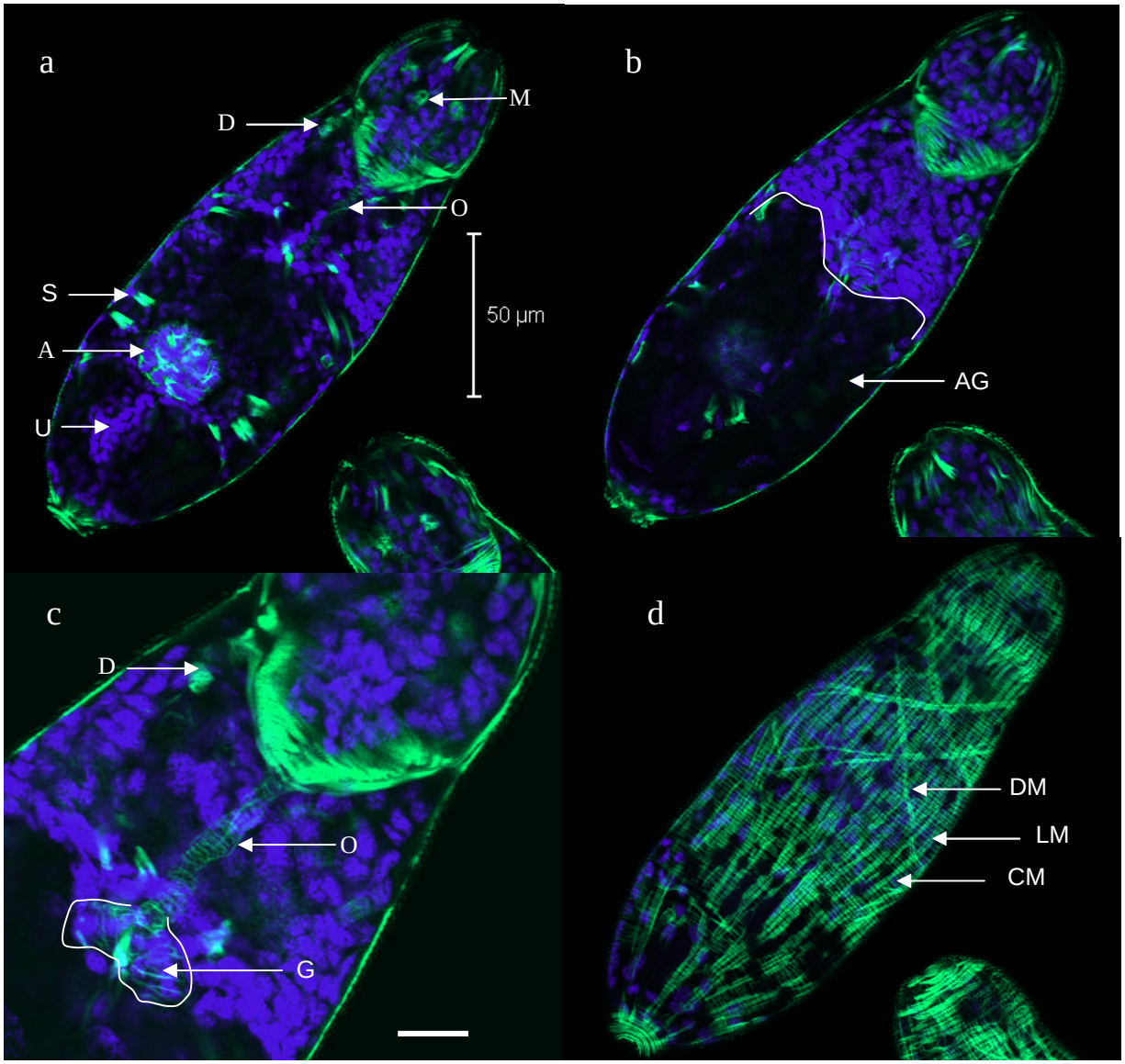


Figure 2-11 Cercaria stained with phalloidin (green) and DAPI (blue).

Optical sections shown are a) 3.78µm, b)4.91µm, and d) 6.23µm from the ventral surface. a) Note acetabulum (A) with supporting muscles (S), large acetabular glands with ducts (D), head capsule, oesophagus (O). A group of undifferentiated cells (U) can be seen at the posterior end. b) Area occupied by acetabular glands is shown. d) The outer muscle layers are visible: circular (CM) Longitudinal (LM) and diagonal (DM). c) area from a) enlarged to show oesophagus (O), position of the gut (G), and an acetabular gland duct bundle (D) in more detail. The oesophagus measures 4µm across, the gut branches are 7µm wide and 12µm long. Scale bar represents 10µm.

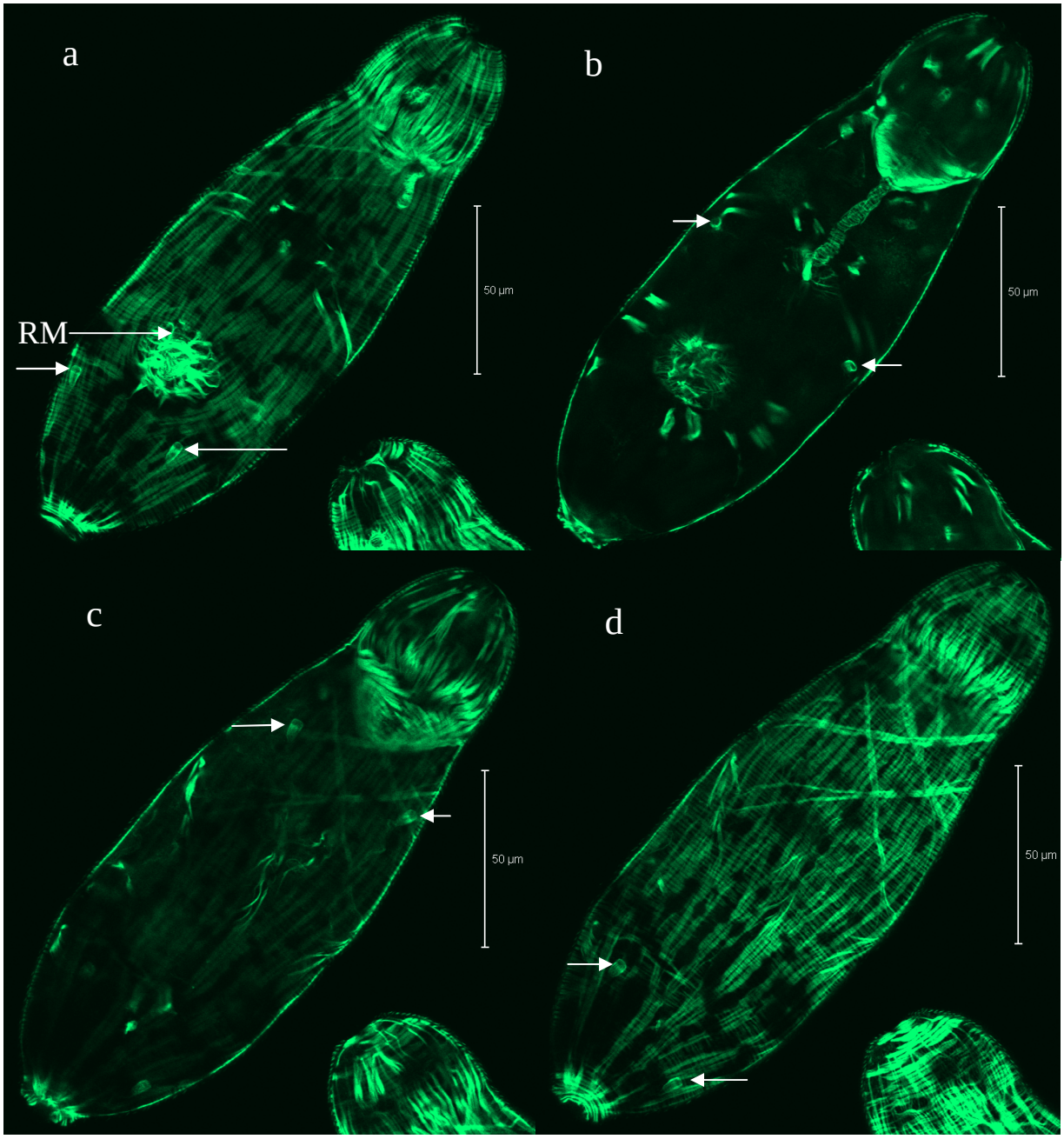


Figure 2-12 Cercaria stained with phalloidin

This is the same specimen as in figure 2.11, but Z sections, starting ventral (a) and moving dorsal (d), were selected to show the position of the flame cells (arrows), and DAPI is not shown. The two posterior pairs are shown: ventral (a; 2.64µm into the z-stack), and dorsal (d; 6.04µm). The middle pair (b; 4.15µm) are more ventral than the anterior pair (c; 5.48µm). There are no flame cells in the head capsule. Radial muscle fibres (RM) can be seen in the acetabulum.

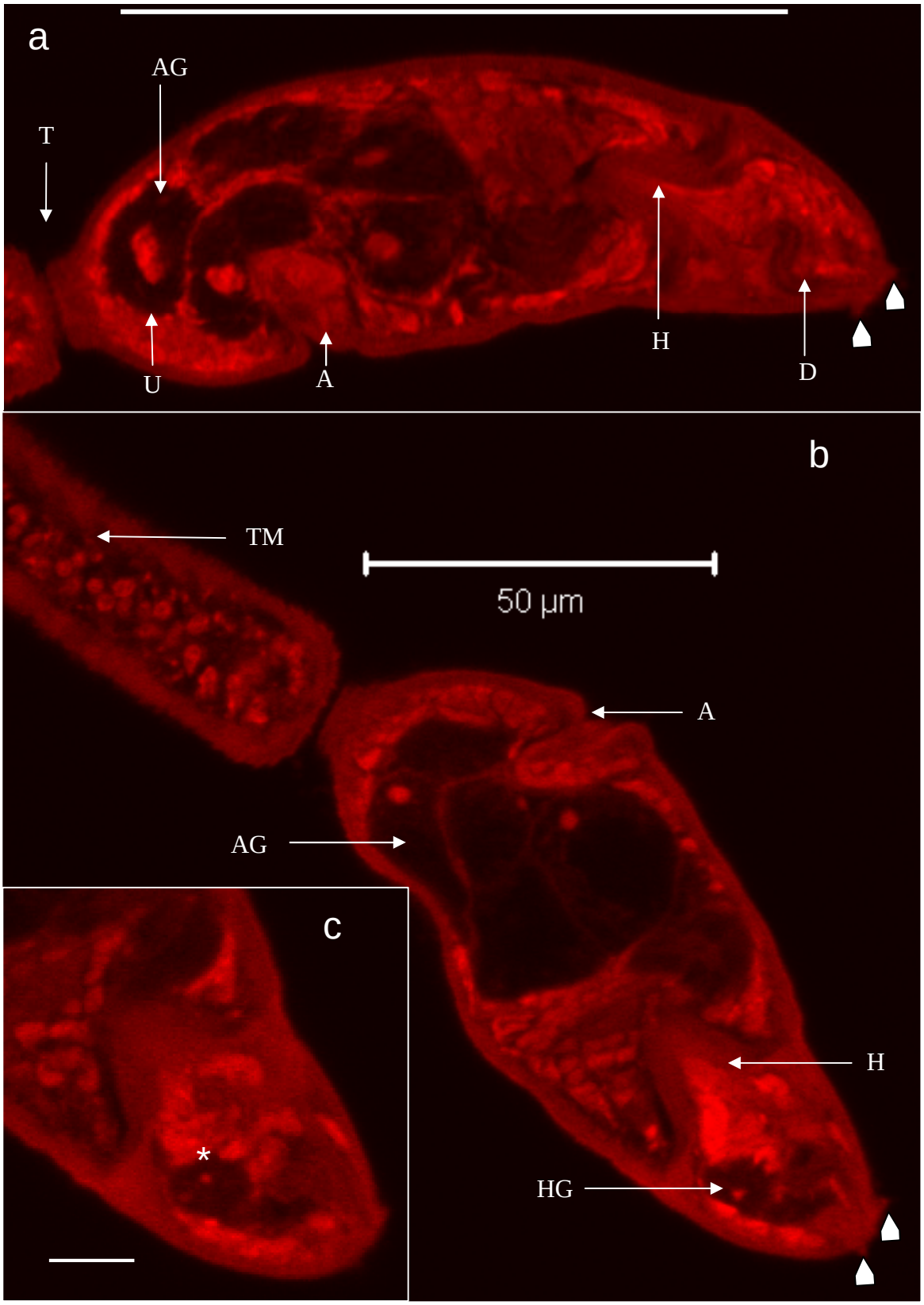


Figure 2-13 Cercariae stained with Langeron's carmine

Figure 2-13a) The body tail junction is visible (T) and the acetabular gland cells (AG) are prominent. The group of undifferentiated cells can be seen (U) just posterior to the acetabulum (A). Muscle boundary of the head capsule is stained (H), the acetabular gland ducts (D) wind through the head capsule, the putative sensory endings are indicated at the anterior (arrow heads). Scale bar represents 100μm. c) is a z section 2.59μm from b), showing a second brightly staining spot visible in the head gland (asterix). Scale bar 10μm**.**

2.3.4. Skin stage schistosomulum

The body of the skin schistosomulum is approximately 140μm long, and 55μm wide. Posterior oriented spines are prominent at either end, but almost absent from the middle third of the body (Fig. 2.15b). The two layers of muscle (circular and longitudinal), and the supporting diagonal strands are well developed and neatly arranged (Fig. 2.15b). Examination of the entire z stack reveals 8 flame cells, positioned as in the cercaria (Fig. 2.14). The acetabular glands and their ducts are no longer present; however, the head capsule with its strong muscle boundary is still intact (Fig. 2.14b) with the head gland still in evidence (Fig. 2.15a). It is different morphologically to the cercarial head gland in that it contains spherical vesicles (Fig. 2.15a), but as in the cercaria, there are many extensions which seem to be continuous with the tegument. The oesophagus is 3-4μm in diameter. The gut is larger and more prominent, and positioned dorsally just anterior to the acetabulum (Fig. 2.14b); the lumen stains positive for protein (Fig. 2.16). The neural ganglia are discernible by LC staining (Fig. 2.16), along with lateral nerve chords extending from them down the length of the body (Fig. 2.16). The acetabulum is prominent, as are the muscles supporting it (Fig. 2.14a). As in the cercaria, these are anchored dorso-laterally (Fig. 2.14a). The cells throughout the body are notably different morphologically from those in the cercaria and early germ ball, but similar to those of the later germ balls and daughter sporocyst (Fig. 2.2). There is distinct punctate staining in the nucleus of the vast majority of cells. The sensory endings that were present at the anterior end of the cercaria are no longer present.

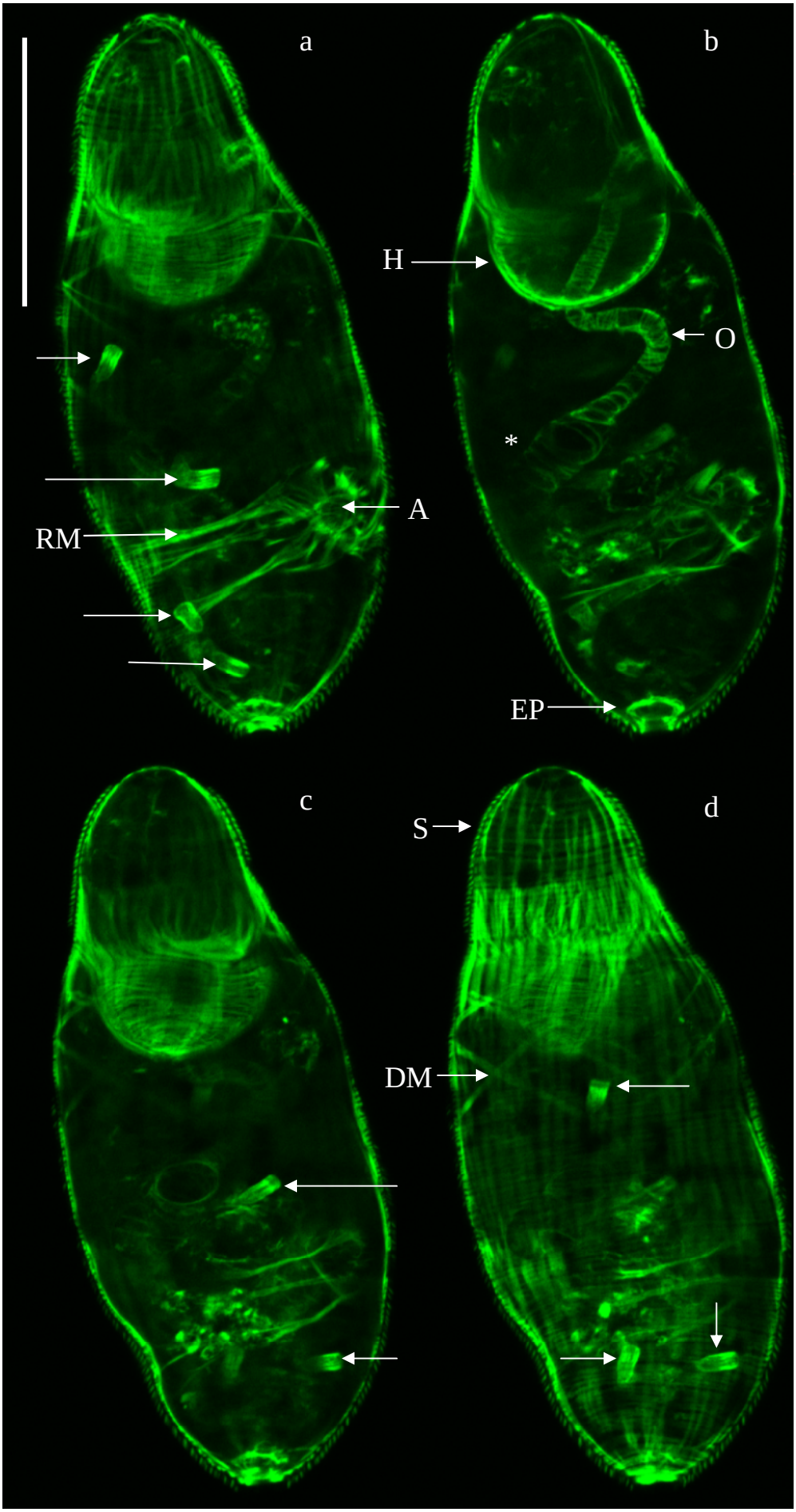

a
b
H
O
*
A
RM
EP
c
d
S
DM

Figure 2-14 Selected optical sections of a schistosomulum stained with phalloidin.
Sections shown in a) to d) are 4.46µm, 5.95µm, 7.43µm and 8.18µm, respectively, through an 11.52µm deep z-stack. a) Acetabulum (A) with its retractor muscles (RM) are highlighted, and flame cells are indicated (arrows). b) Head capsule muscle boundary is visible (H), with the oesophagus (O) passing through it, and leading to the dorsally situated blind gut (asterix). The oesophagus is 4µm wide, the gut sacs are 16µm long and 10µm wide. The excretory pore (EP) is seen at the posterior. c) ventral middle and posterior flame cells are shown (arrows). d) diagonal muscle (DM) and posterior dorsal and ventral flame cells (arrows) can be seen.

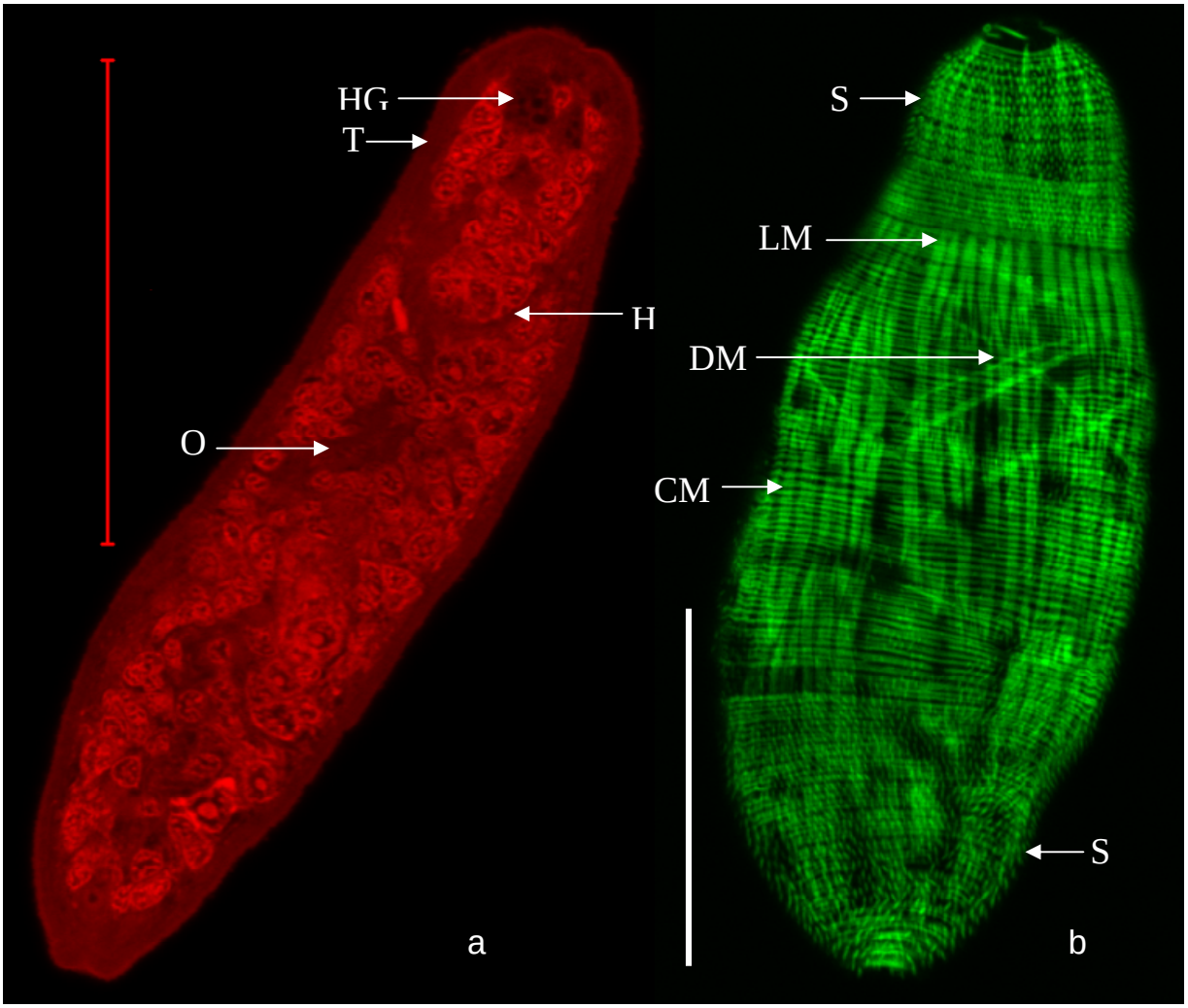


Figure 2-15 Skin stage schistosomulum

a) Optical section of a parasite stained with Langeron's carmine. Note head gland (HG) within head capsule (H), lack of acetabular glands and thick tegument (T). The oesophagus is also seen (O). b) Skin stage schistosomulum stained with phalloidin. Note anterior and posterior surface spines (S), and muscle layers: circular (CM), longitudinal (LM) and diagonal (DM). Scale bars 50µm.

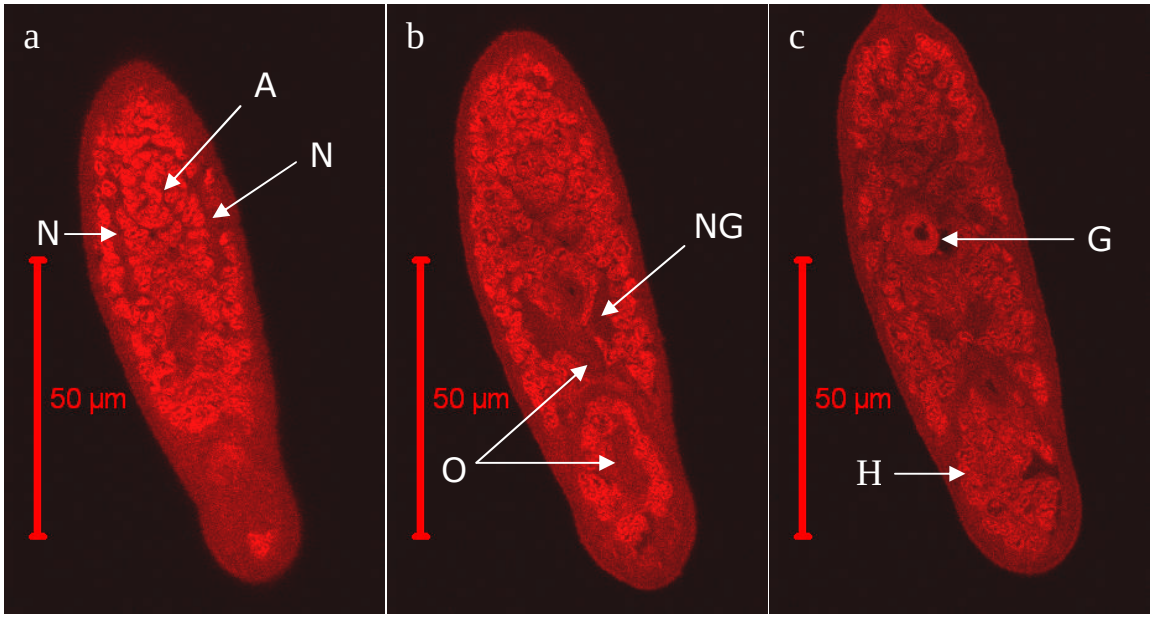


Figure 2-16 Schistosomulum stained with Langeron's carmine.

Selected optical sections (a) 2.92µm (ventral), (b) 5.85µm, (c) 11.69µm (central), taken from a 22.73µm deep z-stack. Note acetabulum (A), lateral nerve chords (N), oesophagus (O), neural ganglia (NG), protein staining is present in the gut lumen (G), head capsule (H) is still in evidence.

2.4. Discussion

Schistosomes undergo impressive and rapid morphological changes in a brief time, approximately 7 days from germ ball to cercaria, and a further 3 days to skin worm. Familiarity with the anatomy of these lifecycle stages will facilitate interpretation of results to come and put the infection process in context of considerable developmental changes for the schistosome. The advantages of using whole mount specimens are that the orientation of the animal is always known, and images do not have to be 'reconstituted' from serial sections. Both these facts make images easier to interpret.

2.4.1. Daughter sporocysts

To my knowledge this is the first application of confocal microscopy to daughter sporocysts. Previous studies have described the 'vermiform' nature of this stage, the spines, and the appearance of brood chambers [10, 36, 37]. Novel insights include the arrangement of muscle as revealed by phalloidin staining, and the proportion of the body core made up of large germ cells.

In both migratory and developed daughter sporocysts the outer muscle layers are arranged as in a loose basket weave conformation. This is particularly evident in the brood chambers. It is likely that this loose weave allows the brood chambers to expand in order to accommodate the growing germ balls. Phalloidin staining also enables visualisation of the protonephridia; four flame cells and their two connecting tubules can be seen, strongly suggesting the presence of actin in the tubule walls. This is the only stage in which the tubules are visible.

Langeron's carmine staining highlights the cellular make up of the daughter sporocyst. The majority of cells comprising the core of the body are large and may represent the germ cells which will develop into germ balls. Previous attempts at germ line transformation in schistosomes have focussed on larval stages such as miracidia or schistosomula which can easily be re-introduced to the lifecycle. It has been shown that daughter sporocysts can be successfully transplanted into snails and produce cercariae [135]. The prominence of germ cells and 'loose weave' muscle configuration revealed here, suggest that the migratory daughter sporocyst may be a prime target for transgenesis. DNA constructs could be delivered reliably to germ cells, and the transduced daughter sporocysts returned to snails to continue the life cycle.

2.4.2. Germ ball development

This is the first application of confocal microscopy to the study of germ ball development; previous studies have used light microscopy [38] or electron microscopy [41]. It is also the first time that this process has been observed daily from 21-27 days post-infection, thus filling the gap in the report of Cheng and Bier, who did not observe germ ball between days 22 and 32 post infection [38]. The findings are in agreement with both Cheng and Bier and Dorsey's observations that germ balls from snails infected on the same day, and indeed germ balls from the same daughter sporocyst develop at different rates. However, in contrast to Cheng and Bier's findings, only small round germ balls were present at day 21, germ balls with stubby tails were present at days 24 and 25. Motile cercariae were observed for the first time at day 27.

Germ balls corresponding to Cheng and Bier stages 2-7 were documented. Although these designated stages are helpful, there are some differences in the specimens observed presently. The authors show diagrams of germ balls with a spherical tail bud joined by a constriction to

the germ ball body. Novel observations in this study show that the posterior end of the germ ball first flattens and then extends, becoming bifurcated very early on. Cheng and Bier [38] show the acetabular gland ducts separately extending to the anterior of the germ ball. Later ultrastructural studies have shown that the ducts exist in a bundle surrounded by muscle fibres [44]. I have not observed the 'primitive epithelium' [20, 39]. However, this was reportedly lost at day 21 post infection [39], the time at which my observations began. The present data show that the tegument is joined by 'protein bridges' to cells in the body before the nuclei in the tegument are lost. This confirms the observation of Hockley, who reported that the tegument cell bodies arose independently of the tegument syncytium and that processes from these cells migrate towards the tegument and become continuous with it [20].

The circular muscle layer develops first, followed by the longitudinal layer, and then the diagonal strands appear last of all. The longitudinal fibres begin as short separate segments, which must coalesce to form the complete strand. It is possible that each cluster of fibres originates from an individual cell. The head capsule muscle is well developed by the time the tail is elongating; it can be seen as more developed than the body musculature at the 'stubby-tailed' stage. The acetabulum can be seen at the 'elongating-tail' stage. Flame cells can be seen first when there is a stubby tail. Nerve system development is not easy to follow with the stains chosen; however, the brain is discernable at the 'stubby-tailed' stage.

The contents of the cercarial glands are known to play an important role in host invasion. In order to include the transcripts for these proteins for the microarray experiment described in Chapter 3, it was necessary to sample germ balls whose glands (both acetabular and head) were developing. Although the development of cercarial glands has been described in some detail, no timescale was established. When the germ ball is 50μm long, the primordial acetabular gland cells can be seen towards the anterior of the body. The glands are prominent when the stubby tail is seen, and they develop further as the tail elongates. Tail length may be taken as a proxy for the appearance of acetabular glands. Ebrahimzadeh reported that the head gland develops later than the other cercarial glands [40]. The current results agree; I cannot discern this gland before the cercaria is mature. Taken together, these observations show that germ balls must be collected from the 'flat-bottomed' stage through the 'stubby-tailed' stage, but not including motile cercaria for the gene expression experiments. Thus, snails will be dissected between 22 and 26 days after infection.

2.4.3. Cercariae

Many features of cercariae have been examined in detail using electron microscopy. Phalloidin has also been used to describe the muscle structure. The present study used whole mount preparations, and the stains phalloidin, DAPI and Langeron's carmine in order to draw the anatomical information into an overall picture.

Acetabular glands are the most prominent feature of cercariae, taking up approximately ½ of the body volume. They failed to stain with either the general protein stain Langeron's Carmine, or the specific f-actin stain phalloidin. This indicates that the contents are not chiefly protein. Phalloidin staining shows clearly that the gut is found just anterior to the acetabular glands. The group of undifferentiated cells at the posterior of the cercaria have been described previously as 'genital anlagen', the implication being that these cells will in time develop into the adult reproductive system, which is the only anatomical system lacking in the cercaria. In agreement with Stirewalt *et al* [49], the current study shows four pairs of flame cells in the body and one pair located proximally in the tail. However, the ciliated tubules are not visible.

The other uncertainty in the literature is the cellular make up of the head gland. I have observed two nucleoli by LC staining, suggesting that this gland comprises two cells rather than one. This grasp of cercarial anatomy at the whole mount level will be essential for interpretation of localisation experiments in Chapter 4.

2.4.4. Skin stage schistosomula

Presented here is the first investigation into the morphology of skin stage schistosomula using confocal microscopy. The stains Langeron's carmine and phalloidin were used and show a detailed picture of the anatomy of this early intra-mammalian stage.

The developing schistosomulum is not larger than the cercaria, and has a similar range of sizes from 110μm when contracted to 170μm when elongated. The tegumental spines are retained at the anterior and posterior ends, and may help with migration, but there are no spines on the mid-body section. The acetabular glands have been emptied and are no longer present. However, the head capsule remains, indicating that this stage is indeed most similar to a 'skin

stage' schistosomulum as larvae recovered from lungs are reported to have lost the head capsule [13]. In addition, the head gland persists, but it is morphologically different to the cercarial head gland. Langeron's carmine staining shows spherical vesicles which may indicate that the gland is actively secreting at this stage. This finding is in agreement with Crabtree and Wilson who state that the head gland plays an important role in crossing the basal membrane of the epidermis to reach the dermis, and invasion of blood vessels once the larva is inside the skin [48].

The nervous system undergoes some remodelling between the cercaria and the skin stage larva, particularly the loss of longer sensory endings at the anterior end. The brain is more prominent and lateral nerve chords can be seen extending down the length of the body. These were not visible in the cercaria by Langeron's carmine staining which suggests that a degree of neural development occurs on entry to the host.

It has been shown that schistosomula are capable of taking in glucose through their tegument shortly after arrival in the skin. However, there is some discussion in the literature as to when the gut becomes active. The extra space available in the body once the acetabular glands have disappeared allows the gut to expand. The gut pouches had elongated and widened and were located more posteriorly, just anterior to the acetabulum. The oesophagus remains 4μm wide, but has elongated. Novel evidence that the gut is now 'active' is the protein staining in the lumen. This may be due to secretion of gut enzymes, in preparation for digestion, or the ingestion of media containing protein, or both. Although the oesophagus is still not large enough to admit erythrocytes, serum may well be consumed.

The excretory system is arranged much as in the cercaria; the flame cells are in the same positions. However, the bladder is more obvious along with a muscular pore which is surrounded by circular and radial fibres.

2.4.5. Final thoughts

The data presented in this chapter show the anatomical changes that occur during the mammalian infection process. This lays the ground work for understanding the underlying gene expression patterns which will be investigated in the coming chapters. The discovery that germ balls with developing glands can be recovered from snails 22 to 26 days after infection

will allow appropriate sampling for the microarray experiment. In addition, the broad picture of morphology gained will ensure well informed interpretation of localisation studies.

3. Microarray analysis of the germ ball, cercaria and day 3 schistosomulum

3.1. Introduction

The previous chapter described the morphological changes that occur in larval schistosomes during the mammalian infection process. The next task is to investigate the gene expression patterns underlying these developments. Several studies have used proteomics and transcriptomics to examine this process at the molecular level [97, 99, 116, 119].

Recently two microarray experiments have been published which address the transition from intermediate to definitive host, neither of which included all three of the life cycle stages analysed here (germ balls, cercariae, day 3 schistosomula). Jolly *et al* compared infected hepatopancreas with free cercariae and adult worms; schistosomula were not included [97]. Fitzpatrick *et al* omitted germ balls: their daughter sporocysts were migratory stage and did not contain developing cercariae [99]. It is also important to recognise that these studies were carried out using arrays designed before the full genome of *S. mansoni* was available. The array used by Jolly *et al* comprised 12,000 45-50mer probes representing contigs from the *S. mansoni* Genome Index at The Institute for Genome Research (TIGR) [97]. This dataset comprises the ORESTES-generated sequences described by Verjovsky-Almeida *et al* in 2003 [73]. The ORESTES method uses random primers with low stringency PCR, therefore direction of the resulting sequences is unknown [73]. Jolly *et al* report that, despite the unidentified orientation of the probes, >9,700 of them gave some signal, and therefore were likely to have been printed in the correct direction [97]. However, based on probability, it would be expected that approximately half of the probes would be in the wrong direction. Fitzpatrick *et al* used an array comprising 37,632 50mer probes, which was designed using the sequence data available at www.GeneDB.org in 2005 [99]. Although the authors went to great efforts to collect biological material from 15 separate life cycle stages, they were unable to carry out one of their chosen analyses (expression of G Protein-coupled receptors (GPCRs) throughout the lifecycle) as the array they used did not have unambiguous probes for the genes they were interested in [99].

Array construction technology has improved radically, such that extremely high-density arrays can be made. Combined with the availability of gene predictions encoded by the *S. mansoni* genome and cDNA libraries from the lifecycle stages involved in infection, this made it possible to design the first genome-wide microarray. This suggested that a new microarray

study of the mammalian infection process was due, as much new information should be gained by using a comprehensive array. In particular, knowledge about transformation and the genes expressed by skin stage schistosomula was lacking.

It has been shown that the transcripts encoding proteins secreted during cercarial transformation are present only in the intra-molluscan germ ball [116]; therefore, it was necessary to include this life-cycle stage in studies addressing infection by larval schistosomes. The germ balls used here were dissected from snails 22-26 days post infection, when their gland cells were developing (see Chapter 2). In order to discover the suite of genes required by the schistosomula to survive and thrive in the mammalian host, three-day *in vitro* cultured schistosomula were included. This stage corresponds to the skin stage larva (Chapter 2). Schistosomes transformed and cultured by the methods used in the current study are able to mature if transferred to the mouse as a definitive host [64], they are biologically comparable to *ex vivo* worms and can be produced in the large quantities needed to obtain enough RNA for hybridisation without amplification.

In this chapter, the design and use of the first genome-wide microarray platform for *S. mansoni* is described, and the insights gained into mammalian infection by larval schistosomes are presented.

Aims

The aims of this study are:

- To hybridise the novel *S. mansoni* microarray with cDNA from germ balls, cercariae and day 3 schistosomula
- Interrogate the resulting data to build a picture of the molecular processes underpinning infection by larval schistosomes
- Highlight novel genes involved in skin invasion and transformation

3.2. Methods

3.2.1. Array design

The array was designed by Dr A Ivens of Fios Genomics. All *S. mansoni* ESTs available at GeneDB.org whose direction was known were compiled using phrap (Phil Green at Genome Sciences Department, University of Washington, US; http://www.phrap.org/phredphrapconsed.html). These formed the input for the array design along with the predicted genes from version E of the *S. mansoni* genome (www.GeneDB.org). The input data were broken up into sequential 50mers and redundant sequences were removed using FAlite.pm (Ian Korf; http://homepage.mac.com/iankorf/) and associated Perl scripts. The unique sequences were mapped back to the genome assembly using exonerate (http://www.ebi.ac.uk/~guy/exonerate/). From a map ordered list, every 13th 50mer was chosen as a probe. No selection was made for the number of probes per predicted transcript. The design was sent to Roche-NimbleGen, who made some minor refinements for ease of synthesis and constructed the arrays using digital micromirror technology [89].

3.2.2. Biological material

Parasites were obtained as described in section 2.2.1. Germ balls were recovered from snails 22-26 days post-infection with 40 miracidia each. RNA was extracted from the larvae with TRIzol (Invitrogen, Paisley, UK) according to the manufacturer's instructions. Double stranded cDNA was synthesised using SuperScript II double-stranded cDNA synthesis kit (Invitrogen).

3.2.3. Hybridisation

NimbleGen were supplied with at least 2.7µg of double stranded cDNA for three biological replicates each from germ balls, cercariae, and day 3 schistosomula. Each replicate was necessarily made up of pooled cDNA from multiple syntheses. Care was taken to ensure that separate biological replicates were obtained i.e. no parasite homogenates were split across replicates. Each biological replicate of cDNA was labelled with Cy3 and hybridised to an array for 16 to 20 hours at 42°C. Slides were washed, and dried before fluorescence data was read using a NimbleGen MS 2000 Scanner with NimbleScan software.

3.2.4. Statistical analyses

NimbleGen supplied background corrected data. All subsequent statistical analysis was carried out by Dr A Ivens using programmes from the Bioconductor suite [91]. The data were quality-assessed by visual inspection of graphical representations of the raw probe level data. Box plots were drawn using the boxplot function from the graphics package. Correlation data were calculated using the cor function from the stats package and heatmaps were made by calling the heatmap.2 function in the gplots package. All arrays passed the quality assessment. Next the data were quantile normalised using the normalizeBetweenArrays function in the limma package. This ensures identical distributions of the data [136]. Following normalisation, the probe level data were summarised to yield 'gene level' data. The probes were re-mapped to the *S. mansoni* gene predictions at www.geneDB.org (version F) using exonerate. If a probe matched an Smp locus with an e value $<1^{-05}$ by both nucleotide and protein BLAST, the probe was annotated to that Smp locus. The intensity value for each locus was the mean intensity of all the probes by which it was represented. The resulting gene level data were the input for the differential expression analysis, which was carried out using the limma package. First a linear model fit was performed. This reduced the data for each gene to a mean value from each of the life cycle stages. Next, differential expression data were obtained by performing a contrast analysis. This compares the expression level of each gene in the following contrasts:

germ ball : cercaria
germ ball : day 3 schistosomulum
day 3 schistosomulum : cercaria

Multiple testing was corrected for using the eBayes function which employs the method of Benjamini and Hochberg [137]. This gives an adjusted P value (adjP). Genes which were differentially expressed with an adjP<0.0001 were chosen for further analysis. Finally a hypergeometric test was carried out using the Category package to discover whether particular GO terms were over-represented in the differentially expressed genes compared to the 'gene universe' of Smps on the array.

3.3. Results

The availability of sequence data for *S. mansoni* allowed the design of the first genome-wide array for this medically important species. The construction of the array will be described, followed by an overview of the contrasts made. The differentially expressed genes will be reported in groups based on functions.

3.3.1. Novel array

This is the most comprehensive array that has been constructed for *S. mansoni*. There are 385K features on the array comprising 377,598 sequences representing 16,222 loci from the *S. mansoni* compilation and 11,613 random sequences for hybridisation controls.

Of the loci represented on the array, 84.4% correspond to a gene prediction (Smp) or can be mapped to a gene prediction at GeneDB.org by BLAST, with an E value below a threshold of $1e^{-05}$. Gene predictions were classed as unannotated if they had no name, PFam ID, or Interpro ID. Loci with no annotation (ie Smp no annotation + EST with no matches to an Smp) make up 42.9% of the array.

Loci	Number	Percentage (of total)
locus designated by Smp	11,585	71.4%
• of which have an annotation	7,160	44.1%
• of which have no annotation	4,425	27.3%
locus designated by EST	4,637	28.6%
• of which hit Smp by BLAST nucleotide (n) or translated DNA versus protein (tblastx)	2,103	13%
• of which no BLAST hit to Smp	2,536	15.6%
Total loci	16,222	100%

Table 3.1 Summary of array features

Annotation is classed as having a protein name, Pfam ID or Interpro ID.

Of the 11,585 Smps, 4,425 had no annotation at all (38%).

3.3.2. Contrasts summary

There were large numbers of differentially regulated loci in each comparison made. The fewest differences were observed in the day 3 – cercaria comparison. The comparison with the highest number of differentially expressed genes was germ ball – cercaria. Before further analysis, two filters were applied; loci with B≥3 which corresponds to 95% confidence value (Chapter 1.2.6) and a fold change of ≥2.

Contrast	GB vs C	GB vs D3	D3 vs C
No. of loci with B≥3	1,731	1,431	1,066

Table 3.2 Summary of contrasts at gene level

3.3.3. GO analysis summary

GO analysis was carried out on the loci which matched Smps. Loci with no GO term annotation were automatically excluded from this analysis. Table 3.3.3 below shows a summary of the categories of genes enriched in each stage relative to the others.

Comparison	Some categories highlighted by GO analysis
GB > C	DNA replication, DNA repair, translation, Wnt signalling, proteolysis, macromolecule biosynthesis, metabolism
GB > D3	DNA replication, proteolysis, microtubule cytoskeketon, ATPase, biosynthesis, transferase activity, metabolism
C > GB	Neurotransmitter transport, proteolysis, carbohydrate biosynthesis, kinase activity, ion transmembrane transport, cytoskeleton
C > D3	Biosynthesis, cytoskeleton, catalytic activity, proteolysis, energy production
D3 > C	Membrane, pepsin A, lipid metabolism, transferase activity, calcium ion binding
D3 > GB	Proteolysis, enzyme regulation, neurotransmitter transport, ion transmembrane transport, membrane, transferase activity, lipid metabolism, hydrolase activity, calcium ion binding

Table 3.3.3 Examples of GO terms enriched at each stage relative to the others.

GO analysis reveals that DNA replication and proteolysis genes are up-regulated in the germ ball compared to the other two stages. The developmental patterning Wnt genes are also up-regulated in the germ ball along with genes involved in macromolecule biosynthesis and metabolism. Several categories are enriched in the cercaria including proteolysis, biosynthesis, cytoskeleton, neurotransmitter transport and energy production. Proteolysis and neurotransmitter transport are also enriched in the schistosomulum compared to the germ ball. Categories up-regulated exclusively in the schistosomulum include lipid metabolism and calcium ion binding.

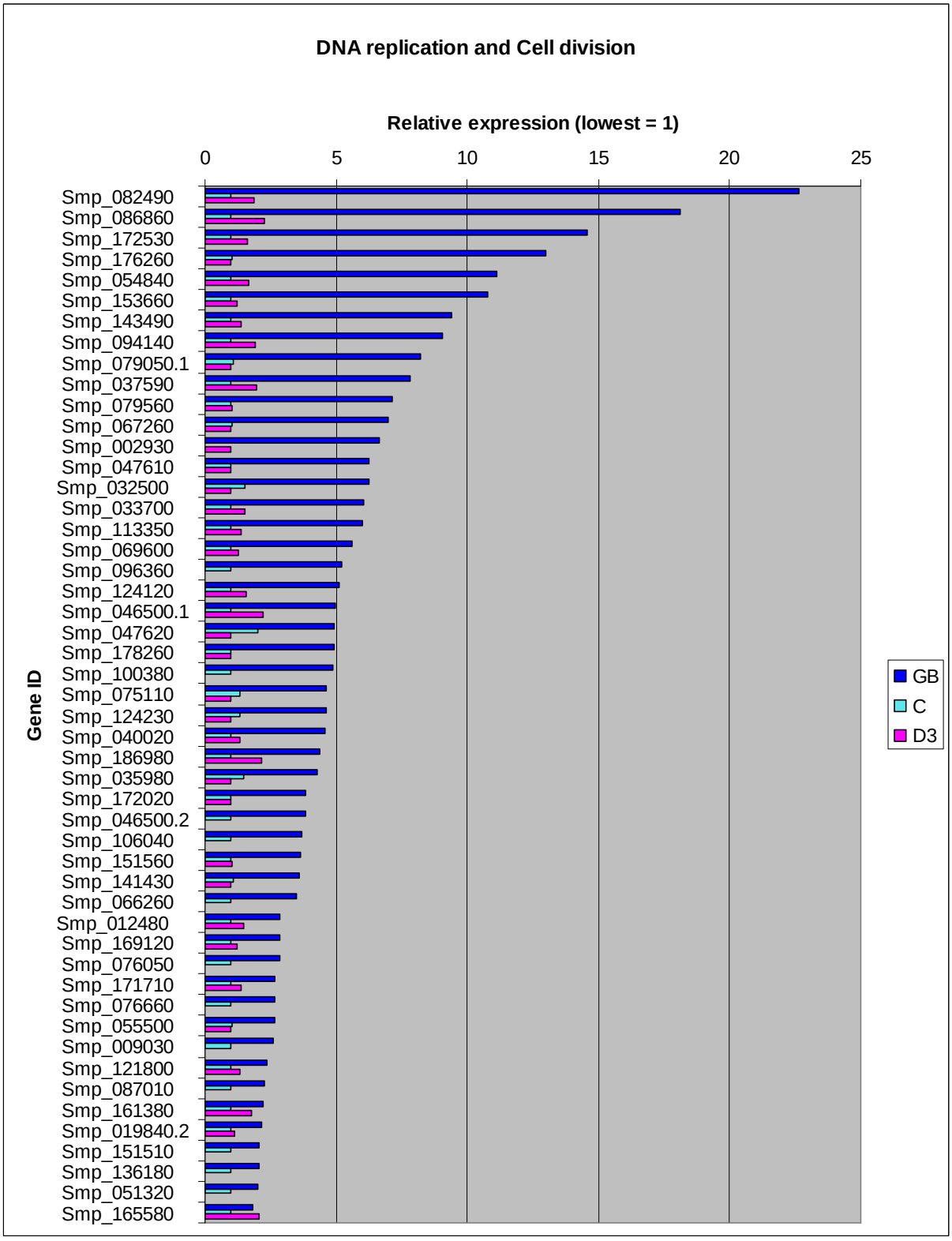
In addition to GO category analysis, categories based on those presented by Berriman *et al* in the genome paper [75], and custom groups based on proteomic and/or localisation evidence were also analysed. They are presented below in separate groups based on function. The bar charts show the relative fold change for each gene. Each bar chart has a key table showing the gene annotations in the order they appear in the chart. The baseline for each gene is the lowest level of expression for that gene in the life cycle stages studied; it is set to unity. Therefore, the actual baseline expression values may be different for each gene, and comparisons of absolute expression levels between genes, based on the charts shown here, are not valid. However, the patterns of gene expression may be compared between genes.

3.3.4. DNA replication and Cell division

DNA replication and cell division are important functions for an organism which is growing. The genes included in the chart below were highlighted by the GOstats analysis. All of the differentially regulated genes encoding proteins involved in DNA replication and cell division were up-regulated in the germ ball compared to the cercaria. The fold changes in this category ranged from 2 to 22 fold compared to the cercaria. The 22 fold change is exhibited by the mitotic cyclin B. This is followed by histone H2A, which is expressed 18 fold higher than in the cercaria. The DNA repair helicase, rad 25, is up-regulated approximately 2-fold in both the germ ball and the day 3 schistosomulum compared to the cercaria.

Seven DNA polymerase subunits are enriched in the germ ball along with DNA primase, DNA ligase, proliferating cell nuclear antigen (PCNA), and replication factors. Also noteworthy are the minichromosome maintenance (MCM) proteins with are thought to unwind DNA ahead of the replication fork; all six of the subunits which form the heterohexamer

(MCM2-7) are up-regulated in the germ ball with fold changes greater than five. Geminin (involved in loading MCMs onto DNA), GINS, and CDC45 also form part of the 'unwindosome' and are highly expressed in the germ ball.



ID	Annotation	ID	Annotation
Smp_082490	cyclin B	Smp_124230	DNA repair protein RAD51
Smp_086860	histone H2A	Smp_040020	centromere protein A
Smp_172530	MCM4	Smp_186980	flap endonuclease-1
Smp_176260.2	Geminin	Smp_035980	histone H2A
Smp_054840	MCM2	Smp_172020	conserved hypothetical protein
Smp_153660	cyclin d	Smp_046500.2	proliferating cell nuclear antigen
Smp_143490	MCM5	Smp_106040	DNA damage repair
Smp_094140	MCM6	Smp_151560	MCM1
Smp_079050.1	DNA primase large subunit	Smp_141430	hypothetical protein
Smp_037590	MCM3	Smp_066260	replication factor C
Smp_079560	MCM2	Smp_012480	exonuclease
Smp_067260	transforming growth factor-**β** receptor type I	Smp_169120	DNA mismatch repair
Smp_002930	histone H2A	Smp_076050	Telomere binding protein
Smp_047610	cell division control protein 45-related	Smp_171710	structural maintenance of chromosomes smc2
Smp_032500	MCM7	Smp_076660	DNA polymerase delta small subunit
Smp_033700	DNA photolyase	Smp_055500	DNA polymerase I
Smp_113350	MCM3	Smp_009030	ribonucleoside-diphosphate reductase, alpha subunit
Smp_069600	Ribosome biogenesis	Smp_121800	replication factor A 1, rfa1
Smp_096360	replication factor C	Smp_087010	DNA polymerase delta catalytic subunit
Smp_124120	DNA polymerase ε subunit B	Smp_161380	DNA polymerase gamma
Smp_046500.1	proliferating cell nuclear antigen	Smp_019840.2	DNA ligase I
Smp_047620	cyclin B3	Smp_151510	DNA repair nuclease
Smp_178260	DNA polymerase alpha catalytic subunit	Smp_136180	DNA polymerase eta
Smp_100380	GINS	Smp_051320	ap endonuclease
Smp_075110	rad1 DNA damage checkpoint protein	Smp_165580	rad25/xp-B DNA repair helicase

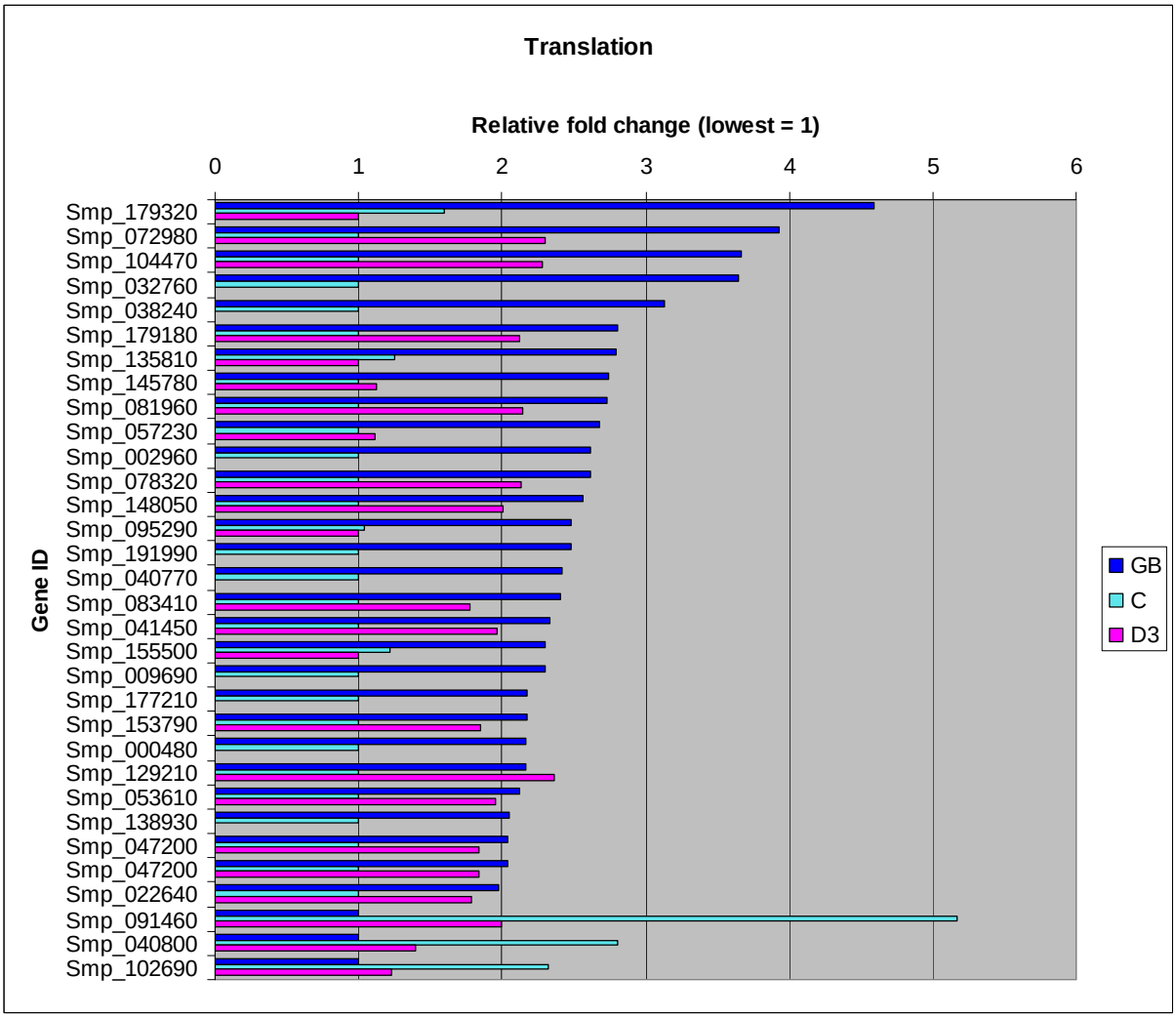
3.3.5. Translation

Translation or protein synthesis is a basic activity of cells, especially those which are dividing or differentiating. The GOStats analysis revealed that 32 genes of the 780 (4%) *S. mansoni* genes annotated to the GO term ‘translation’ are differentially regulated in the comparisons made. This includes genes encoding ribosomal proteins (six 60S, and one 28S, two S2, and one S11), ten tRNA synthases, five initiation factors and two elongation factors. None of the fold changes are exceptional; most are between 2-3 fold, the highest being 5.2. As figure 1 shows, the vast majority of these genes are most highly expressed in the germ ball. The exceptions are three genes which are up-regulated in the cercaria compared to the other two

life cycle stages. These are: glutamine synthetase, glycyl-tRNA synthetase, and eukaryotic translation initiation factor 2c.

There are five genes which are expressed at least three-fold higher in the germ ball compared to either the cercaria or the day 3 schistosomulum; these encode eukaryotic translation initiation factor 2c and elongation factor 2 kinase, lysyl-tRNA synthetase, ribosomal protein S11, and alanyl-tRNA synthetase.

Several genes are up-regulated more than 2 fold in both the germ ball and day 3 schistosomulum compared to the cercaria; these encode: eukaryotic translation elongation factor 2 kinase, lysyl-tRNA synthetase, tyrosyl-tRNA synthetase, ribosomal protein S2, and eukaryotic translation initiation factor 3 subunit 5.



Key

ID	Annotation	ID	Annotation
Smp_179320	translation initiation factor 2c	Smp_083410	elongation factor 1-gamma
Smp_072980	elongation factor 2 kinase	Smp_041450	aspartyl-tRNA synthetase
Smp_104470	lysyl-tRNA synthetase	Smp_155500	DNA polymerase theta
Smp_032760	ribosomal protein S11	Smp_009690	60S acidic ribosomal protein P0
Smp_038240	alanyl-tRNA synthetase	Smp_177210	60S ribosomal protein L7a
Smp_179180	ribosomal protein S2	Smp_153790	60S ribosomal protein L4
Smp_135810	expressed protein	Smp_000480	initiation factor 3 subunit 8
Smp_145780	28S ribosomal protein S5	Smp_129210	tyrosyl-tRNA synthetase
Smp_081960	ribosomal protein S2	Smp_053610	seryl-tRNA synthetase
Smp_057230	seryl-tRNA synthetase	Smp_138930	glutamyl-tRNA synthetase
Smp_002960	pelota	Smp_047200	60S ribosomal protein L3
Smp_078320	initiation factor 3 subunit 5	Smp_047200	60S ribosomal protein L3
Smp_148050	glutaminyl-tRNA synthetase	Smp_022640	60S ribosomal protein L13
Smp_095290	zinc finger protein	Smp_091460	glutamine synthetase 1, 2
Smp_191990	eukaryotic translation initiation factor 3 subunit (eif-3)	Smp_040800	glycyl-tRNA synthetase
Smp_040770	methionine-tRNA synthetase	Smp_102690	initiation factor 2c

3.3.6. Development

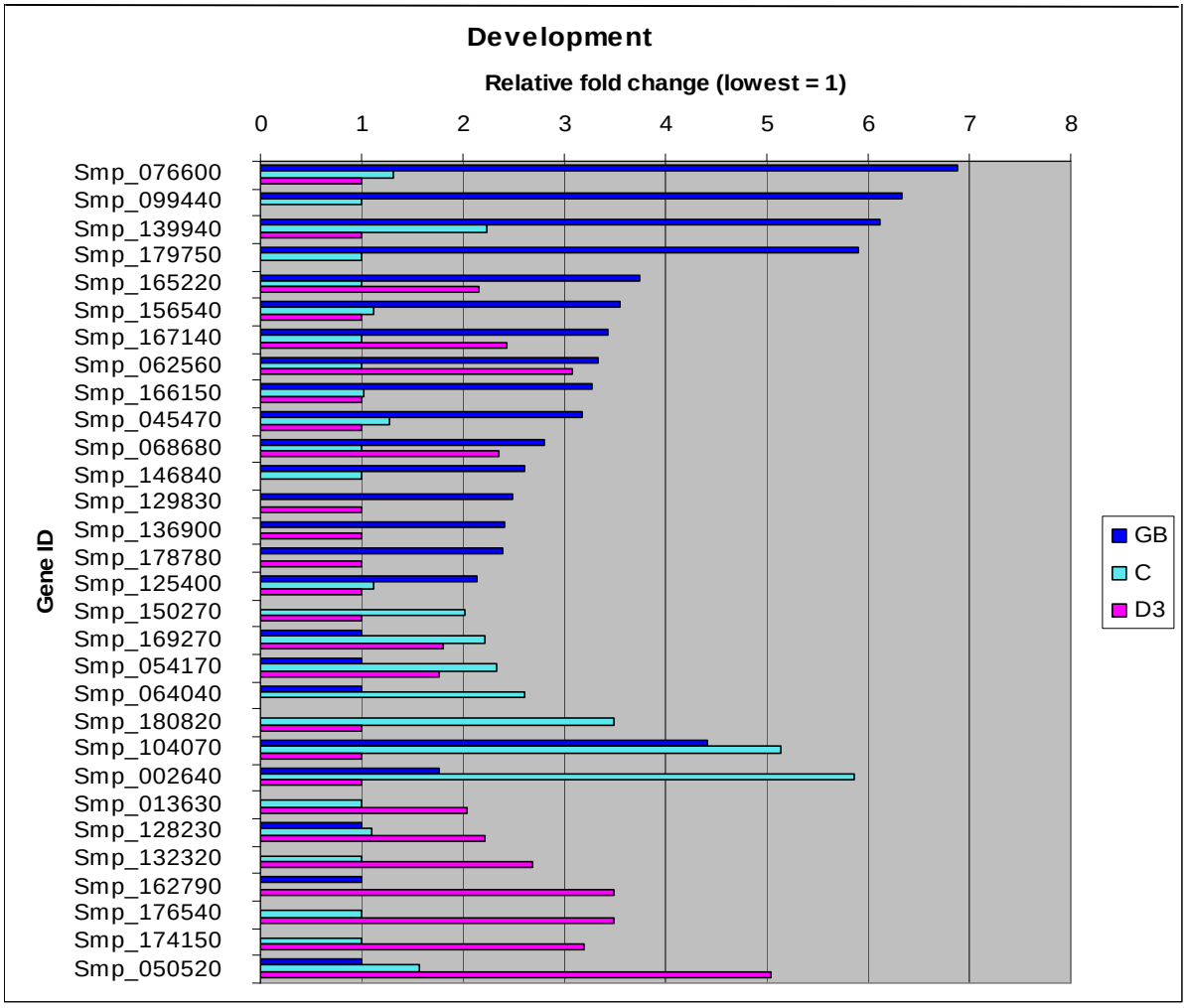
As has been noted in the previous two chapters, much development occurs during the transitions from the snail intermediate host to the mammalian definitive host. The expression patterns of the 62 genes described by Berriman *et al* [75] as involved in neural development were investigated. Nine of them (15%) were found to be differentially expressed. In addition, components of the Wnt signaling pathway, cadherins, homeobox proteins, and other transcription factors were highlighted by the GOstats analysis. The resulting 29 differentially expressed developmental genes are described below.

Whilst the majority of the genes in the 'development' category (16/29) were up-regulated in the germ ball, several are up-regulated in the cercaria (7/29) or day 3 schistosomulum (6/29). Of the nine differentially regulated neural development genes, four were up-regulated in the germ ball, with two expressed at the highest level in the cercaria and three in the day 3 schistosomulum. In the germ ball, the following early neural patterning genes are up-regulated: a 'sox-like' transcription factor, a maternal embryonic leucine zipper kinase, 'single-minded' and 'neurogenin'. The polarity complex component 'Mbt/PAK' is up-regulated in the cercaria; 'stardust' is enriched in the day 3 schistosomulum along with two 'notch' genes.

The four genes that are at least 2 fold up-regulated in both the germ ball and day 3 schistosomulum, compared to the cercaria, are an embryonic ectoderm development protein, a Wnt growth factor (Smp_167140) and two other genes involved in Wnt signaling (Smp_062560 and Smp_068680). Another Wnt gene is up-regulated solely in the germ ball (Smp_156540). A gene encoding Hox 8 is up-regulated in both the germ ball and cercaria compared to the day 3 schistosomulum.

Cadherins are proteins involved in cell-cell adhesion; 26 are present in the *S. mansoni* genome. Two cadherin transcripts are up-regulated in the germ ball, and three different cadherin transcripts are enriched in the day 3 schistosomulum. Different homeobox genes are up-regulated in each life-cycle stage.

The seven cercaria-enriched genes also include one which is involved in controlling neural differentiation and one which may play a role in photoreceptor cell morphogenesis.



ID	Annotation		
Smp_076600	sox transcription factor	Smp_178780	single-minded, meso-ectoderm gene expression control protein
Smp_099440	bone morphogenetic protein antagonist noggin	Smp_125400	neurogenin, bHLH transcription factor
Smp_139940	pou6f1/brn-5 transcription factor	Smp_150270	Cullin (ubiquitin ligase)
Smp_179750	cadherin-related	Smp_169270	homeobox protein six-related
Smp_165220	embryonic ectoderm development protein	Smp_054170	Neural proliferation differentiation control-1 protein (NPDC1)
Smp_156540	wnt related	Smp_064040.1	cell polarity protein, regulator of photoreceptor cell morphogenesis
Smp_167140	wnt related	Smp_180820	sox transcription factor
Smp_062560	wnt inhibitor frzb2	Smp_104070	SmHox8 (fragment)
Smp_166150	maternal embryonic leucine zipper kinase	Smp_002640	hox protein Smox1
		Smp_013630	cadherin-related
Smp_045470	homeobox protein prospero	Smp_128230	Delta/Serrate/lag-2 (DSL) protein
Smp_068680	secreted frizzled-related protein	Smp_132320	cadherin
Smp_146840	netrin	Smp_176540	cadherin
Smp_129830	cadherin	Smp_174150	homeobox protein meis
Smp_136900	homeobox protein distal-less dlx	Smp_050520	notch

3.3.7. Proteases

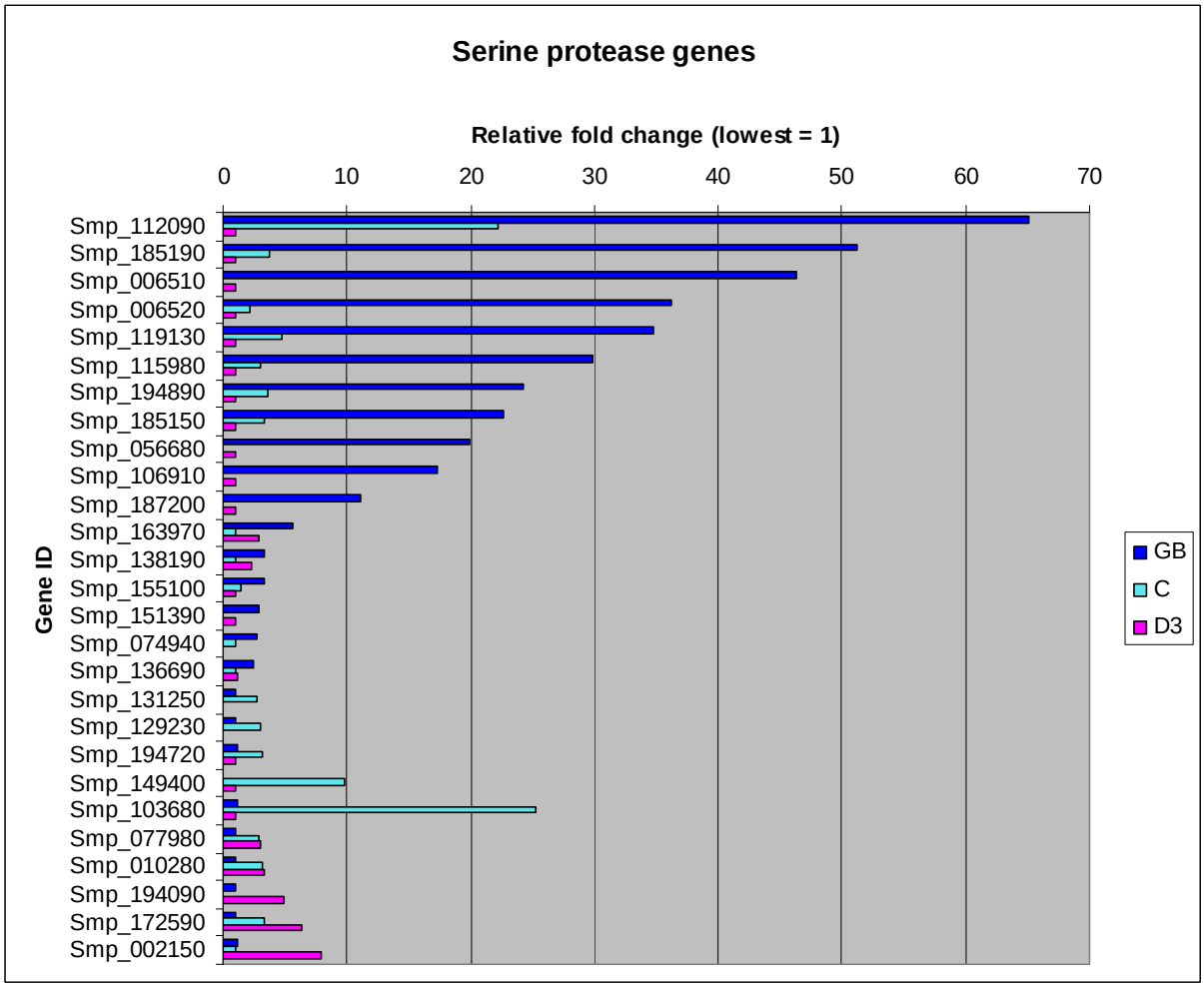
All of the 335 proteases annotated in the genome paper were investigated as it has long been known that these proteins play a large role in skin invasion by larval schistosomes [109]. They are described in four groups based on their functional classes as assigned by the MEROPS protease database [138]. No proteases which belong to the classes G (glutamic acid) or U (Unknown catalytic site) proteases and only two T (threonine) proteases are differentially regulated in the contrasts studied. One is a proteasome subunit (Smp_070930) which is up-regulated in the germ ball; the other (Smp_173480) is involved in degradation of glycoproteins and is up-regulated in both the germ ball and day 3 schistosomulum compared to the cercaria. The expression patterns of the differentially expressed proteases belonging to the remaining classes (serine, metallo, cysteine, and aspartic) will be reported in detail.

3.3.8. Serine proteases

Of the 78 serine proteases listed by Berriman et al [75], 27 (35%) were differentially expressed. This class of proteases encompasses the 11 cercarial elastase genes. Although the proteins they encode are secreted by cercariae, the transcripts are most highly expressed in the germ ball; in some cases they are between 10 and 65 fold higher in this stage compared to the day 3 schistosomulum. Also up-regulated in the germ ball are three family S09 esterases (non-peptidase homologues), one family S09 carboxypeptidase, gliotactin and a rhomboid protease (Smp_129230).

Five serine proteases are up-regulated in the cercaria; their fold changes are not as extreme as those of the germ ball-enriched transcripts (ranging from two to five fold). Two may have a role in cell differentiation (S33 family), one is a subtilase, and the remaining two belong to the S01 family (trypsin/chymotrypsin).

The remaining five differentially expressed serine proteases are up-regulated in the schistosomulum compared to the germ ball. This group includes two genes which are expressed at a similar level in the cercaria and day 3 schistosomulum; they belong to groups S08 (subtilase) and S09 (esterases). Two of the three which are up-regulated in the day 3 schistosomulum alone are trypsin/chymotrypsin-like (family S1A); the other is a carboxypeptidase (family S10).

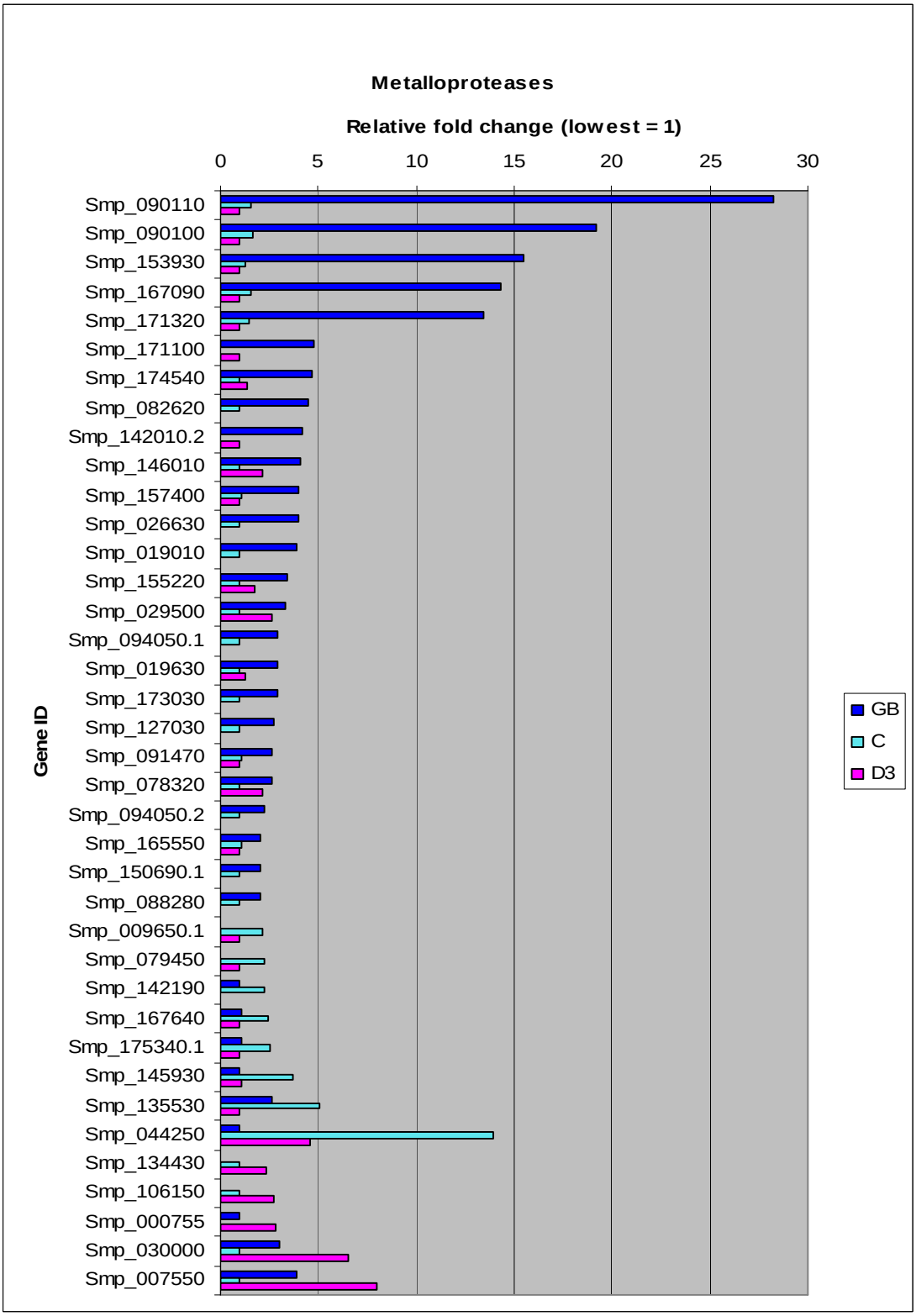


Gene ID	Annotation	Gene ID	Annotation
Smp_112090	cercarial elastase	Smp_155100	family S54, rhomboid
Smp_185190	cercarial elastase	Smp_151390	family S9 esterase
Smp_006510	cercarial elastase	Smp_074940	family S9 esterase
Smp_006520	cercarial elastase	Smp_136690	family S9 esterase
Smp_119130	cercarial elastase	Smp_131250	NDRG4 protein (S33 family)
Smp_115980	cercarial elastase	Smp_129230	subfamily S1A unassigned peptidase
Smp_194890	subfamily S1A non-peptidase homologue	Smp_194720	family S33 non-peptidase homologue
Smp_185150	cercarial elastase	Smp_149400	subfamily S8B unassigned peptidase
Smp_056680.2	cercarial elastase	Smp_103680	subfamily S1A unassigned peptidase
Smp_106910	cercarial elastase	Smp_077980	subfamily S8B non-peptidase homologue
Smp_187200	cercarial elastase	Smp_010280	family S9 unassigned peptidase (S09 family)
Smp_163970	family S10 non-peptidase homologue	Smp_194090	subfamily S1A unassigned peptidase
Smp_138190	gliotactin, septate junction protein, cholinesterase	Smp_172590	family S10 unassigned peptidase
		Smp_002150	subfamily S1A unassigned peptidase

3.3.9. Metalloproteases

Of the 117 metalloproteases described in the genome paper, 38 are differentially expressed during the germ ball to day 3 schistosomulum transitions (33%). The majority (25/38) of these are enriched in the germ ball. The five most highly expressed germ ball metalloproteases are invadolysins. A sixth invadolysin was expressed 2.2 fold higher in the germ ball compared to the cercaria, and another one was enriched in the cercaria compared to the other two stages (Smp_135530). The remaining germ ball-enriched metalloproteases are involved in translation and cell proliferation.

The metalloproteases up-regulated in the cercaria are two mitochondrial processing peptidases, two carboxypeptidases, an ADAM protease (a disintegrin and metalloprotease), a matrix metalloprotease and a proteasome regulatory subunit, which is expressed 14 fold higher than in the germ ball. Five metalloproteases are up-regulated in the day 3 schistosomulum. These include tolloid-like protease, a leucine aminopeptidase, which plays a role in protein turnover, and a glutamine hydrolase.


Metalloproteases
Relative fold change (lowest = 1)
0
5
10
15
20
25
30
Gene ID
Smp_090110
Smp_090100
Smp_153930
Smp_167090
Smp_171320
Smp_171100
Smp_174540
Smp_082620
Smp_142010.2
Smp_146010
Smp_157400
Smp_026630
Smp_019010
Smp_155220
Smp_029500
Smp_094050.1
Smp_019630
Smp_173030
Smp_127030
Smp_091470
Smp_078320
Smp_094050.2
Smp_165550
Smp_150690.1
Smp_088280
Smp_009650.1
Smp_079450
Smp_142190
Smp_167640
Smp_175340.1
Smp_145930
Smp_135530
Smp_044250
Smp_134430
Smp_106150
Smp_000755
Smp_030000
Smp_007550
GB
C
D3

Gene ID	Annotation	Gene ID	Annotation
Smp_090110	invadolysin (M08 family)	Smp_091470	cytosol alanyl aminopeptidase (M01 family)
Smp_090100	invadolysin (M08 family)	Smp_078320	eukaryotic translation initiation factor 3 subunit 5 (M67 family)
Smp_153930	invadolysin (M08 family)	Smp_094050.2	mitochondrial processing peptidase non-peptidase alpha subunit (M16 family)
Smp_167090	invadolysin (M08 family)	Smp_165550	Afg3-like protein 2 (M41 family)
Smp_171320	invadolysin (M08 family)	Smp_150690.1	proliferation-associated protein 2G4, 38kDa (M24 family)
Smp_171100	neprilysin-2 (M13 family)	Smp_088280	chromatin-specific transcription elongation factor 140 kDa subunit (M24 family)
Smp_174540	family M1 non-peptidase homologue (M01 family)	Smp_009650.1	mitochondrial processing peptidase beta-subunit (M16 family)
Smp_082620	farnesylated-protein converting enzyme 1 (M48 family)	Smp_079450	subfamily M16B non-peptidase homologue (M16 family)
Smp_142010.2	methionyl aminopeptidase 1 (M24 family)	Smp_142190	carboxypeptidase N (M14 family)
Smp_146010	insulysin unit 3 (M16 family)	Smp_167640	Zinc carboxypeptidase (M14 family)
Smp_157400	family M13 unassigned peptidase (M13 family)	Smp_175340.1	ADAM17 peptidase (M12 family)
Smp_026630	26S proteasome non-ATPase regulatory subunit 7 (M67 family)	Smp_145930	matrix metallopeptidase-7 (M10 family)
Smp_019010	dipeptidyl-peptidase III (M49 family)	Smp_135530	leishmanolysin-2 (M08 family)
Smp_155220	family M16 unassigned peptidase (M16 family)	Smp_044250	Proteasome regulatory subunit (M67 family)
Smp_029500	thimet oligopeptidase (M03 family)	Smp_134430	Tolloid-like peptidase (M12 family)
Smp_094050.1	mitochondrial processing peptidase non-peptidase alpha subunit (M16 family)	Smp_106150	aspartatecarbamoyltransferase, (glutamine-hydrolysing)
Smp_019630	aminoacylase (M20 family)	Smp_000755	family M13 non-peptidase homologue (M13 family)
Smp_173030	aminopeptidase A (M01 family)	Smp_030000	leucine aminopeptidase (M17 family)
Smp_127030	invadolysin (M08 family)	Smp_007550	leukotriene A4 hydrolase (M01 family)

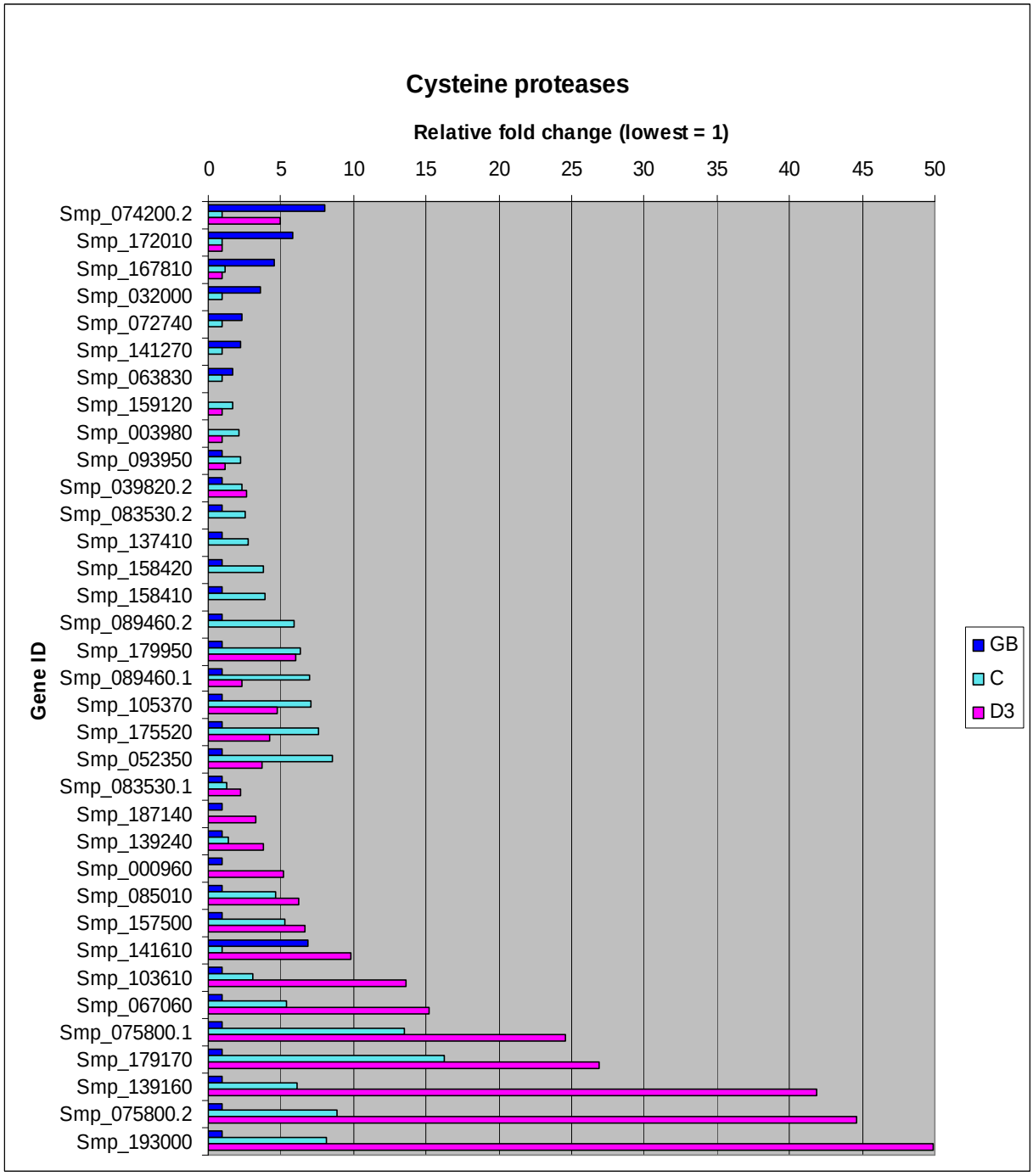
3.3.10. Cysteine proteases

Of the 97 cysteine proteases annotated in the genome paper, 35 were differentially expressed in the contrasts studied (36%). In contrast to the metallo- and serine- proteases, the cysteine proteases are mostly up-regulated in the cercaria and/or day 3 schistosomulum, the highest fold changes being in the day 3 schistosomulum. The fold changes of the seven cysteine proteases up-regulated in the germ ball are relatively low, levels ranging from 2 to 7.5 fold greater than the cercaria. Caspases account for three of the seven. These proteins are involved

in programmed cell death and are likely to be important during embryogenesis. A ubiquitin-specific protease involved in protein turnover is up-regulated in both the germ ball and the day 3 schistosomulum compared to the cercaria. The remaining germ ball-enriched transcripts encode a separase (caspase-like protein involved in cell division), GMP synthase, and a phytochelatin synthase which is involved in the synthesis of a heavy metal binding protein.

Fourteen genes in the Cysteine protease category are most highly expressed in the cercaria. Of these, five are calpains and four are cathepsin B-like. In addition, two deubiquitinating enzymes, two genes with homology to the *Drosophila* ovarian tumour gene and an autophagin are up-regulated at this stage.

With the exception of two calpains, all of the fourteen genes which are most highly expressed in the schistosomulum are annotated as cathepsin or asparaginyl endopeptidase. The increases are quite striking, ranging from 2 to 50 fold, with seven genes being up at least 10 fold. The majority are also up-regulated in the cercaria compared to the germ ball; their general pattern is one of increasing expression from germ ball to day 3 schistosomulum. However, one cathepsin B-like gene (Smp_141610) is up-regulated in both the germ ball and day 3 schistosomulum compared to the cercaria.

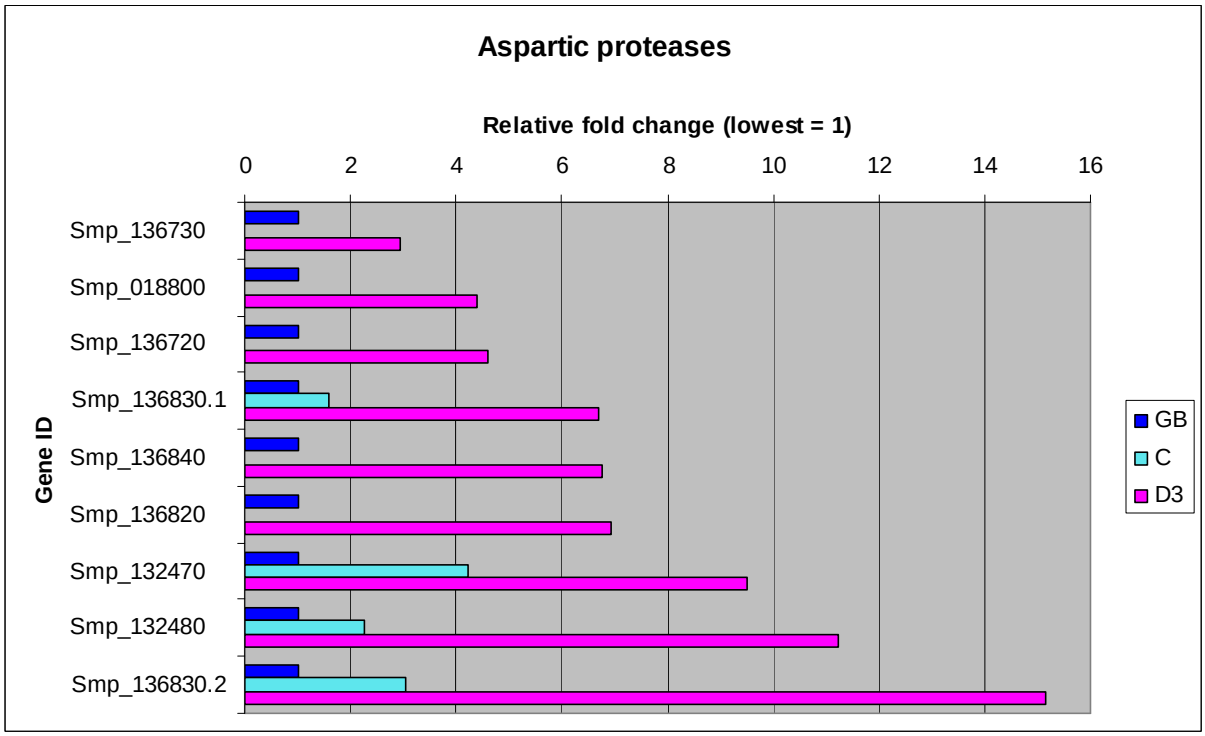

Cysteine proteases
Relative fold change (lowest = 1)
0
5
10
15
20
25
30
35
40
45
50
Gene ID
Smp_074200.2
Smp_172010
Smp_167810
Smp_032000
Smp_072740
Smp_141270
Smp_063830
Smp_159120
Smp_003980
Smp_093950
Smp_039820.2
Smp_083530.2
Smp_137410
Smp_158420
Smp_158410
Smp_089460.2
Smp_179950
Smp_089460.1
Smp_105370
Smp_175520
Smp_052350
Smp_083530.1
Smp_187140
Smp_139240
Smp_000960
Smp_085010
Smp_157500
Smp_141610
Smp_103610
Smp_067060
Smp_075800.1
Smp_179170
Smp_139160
Smp_075800.2
Smp_193000
GB
C
D3

Key

Gene ID	Annotation
Smp_074200.2	ubiquitin-specific peptidase 45 (C19 family)
Smp_172010	caspase-7 (C14 family)
Smp_167810	separase non-peptidase homologue (C50 family)
Smp_032000	caspase-7 (C14 family)
Smp_072740	Phytochelatin synthase (C83 family)
Smp_141270	Caspase (C14 family)
Smp_063830	GMP synthetase (C26 family)
Smp_159120	Deubiquitinating protease (C48 family)
Smp_003980	Calpain (C02 family)
Smp_093950	Ovarian tumour-like (C85 family)
Smp_039820	autophagin-1 (C54 family)
Smp_083530.2	calpain-7 (C02 family)
Smp_137410	calpain C (C02 family)
Smp_158420	cathepsin B-like peptidase (C01 family)
Smp_158410	cathepsin B-like peptidase (C01 family)
Smp_089460.2	Calpain (C02 family)
Smp_179950	cathepsin B-like peptidase (C01 family)
Smp_089460.1	Calpain (C02 family)
Smp_105370	cathepsin B-like peptidase (C01 family)
Smp_175520	ubiquitin-specific peptidase 2 (C19 family)
Smp_052350	Ovarian tumour-like (C64 family)
Smp_083530.1	calpain-7 (C02 family)
Smp_187140	papain (C01 family)
Smp_139240	cathepsin S (C01 family)
Smp_000960	cathepsin B-like peptidase (C01 family)
Smp_085010	cathepsin B-like peptidase (C01 family)
Smp_157500	calpain (C02 family)
Smp_141610	SmCB2 peptidase (C01 family)
Smp_103610	cathepsin B-like peptidase (C01 family)
Smp_067060	cathepsin B-like peptidase (C01 family)
Smp_075800.1	asparaginyl endopeptidase (C13 family)
Smp_179170	asparaginyl endopeptidase (C13 family)
Smp_139160	SmCL2-like peptidase (C01 family)
Smp_075800.2	asparaginyl endopeptidase (C13 family)
Smp_193000	SmCL2-like peptidase (C01 family)

3.3.11. Aspartic proteases

These proteins have homology to the acidic digestive enzymes Cathepsin D and pepsin. They are exclusively up-regulated in the day 3 schistosomulum compared to the other two stages with fold changes ranging from 2 to 15 fold compared to the germ ball. In this group,the expression of three genes increases from germ ball to cercaria and then again in the day 3 schistosomulum. These are Smp_132470, Smp_132480, and Smp_136830.2.



Key

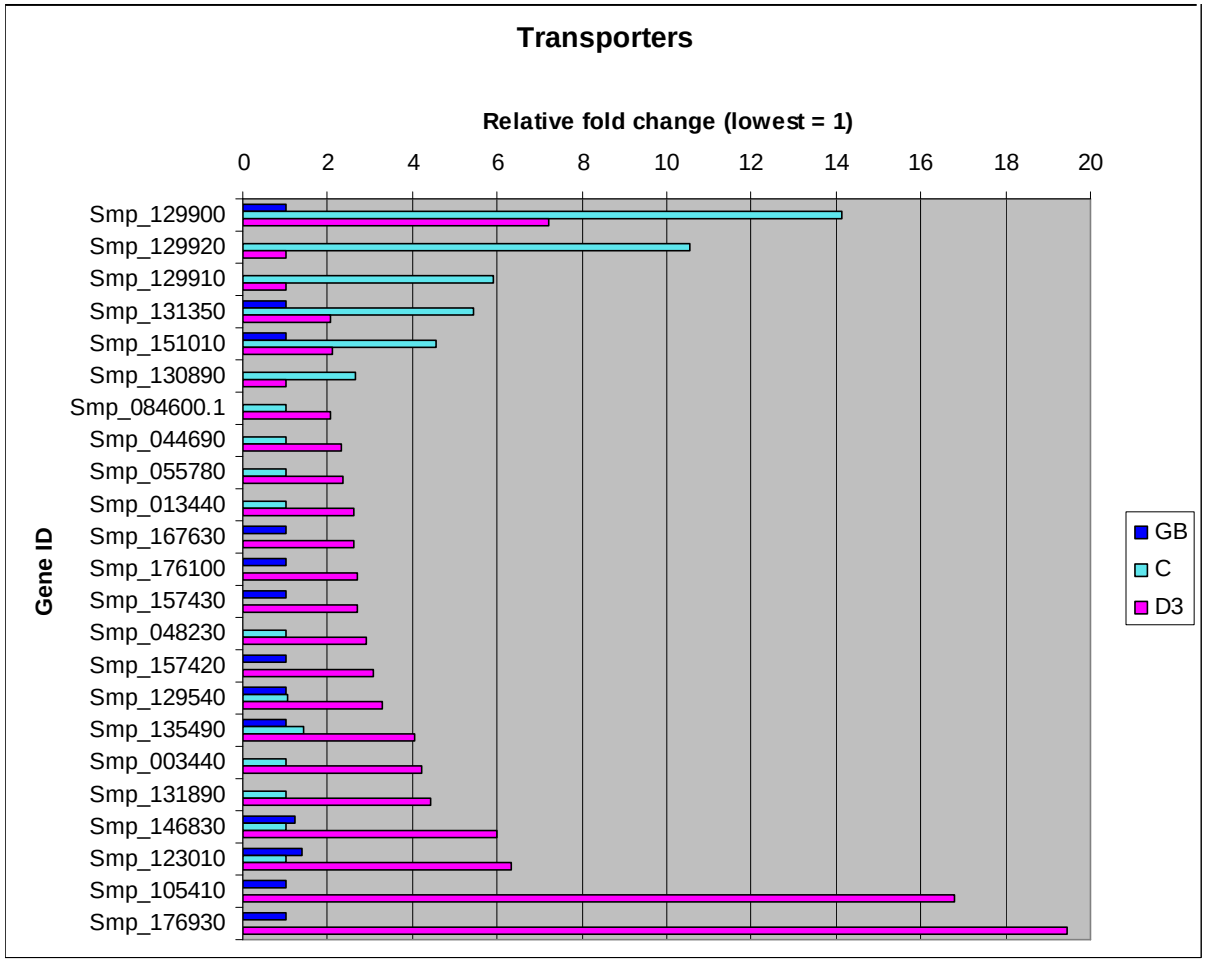
All of the genes in the above graph are annotated as subfamily A1A unassigned peptidase.

3.3.12. Membrane proteins

Transmembrane proteins (in particular, G –protein coupled receptors and channels) are the targets of many drugs currently in use. If expressed on cell surfaces or indeed on the tegument surface, they may represent novel intervention targets. As such, the differentially expressed membrane proteins are described in detail below. The GOstats analysis highlighted membrane proteins up-regulated in the cercariae and day 3 schistosomulum, but not in the germ ball. They are grouped together by putative function. The genes encoding membrane proteins that have been identified in the tegument are also included in the separate ‘tegument’ category.

3.3.13. Transporters

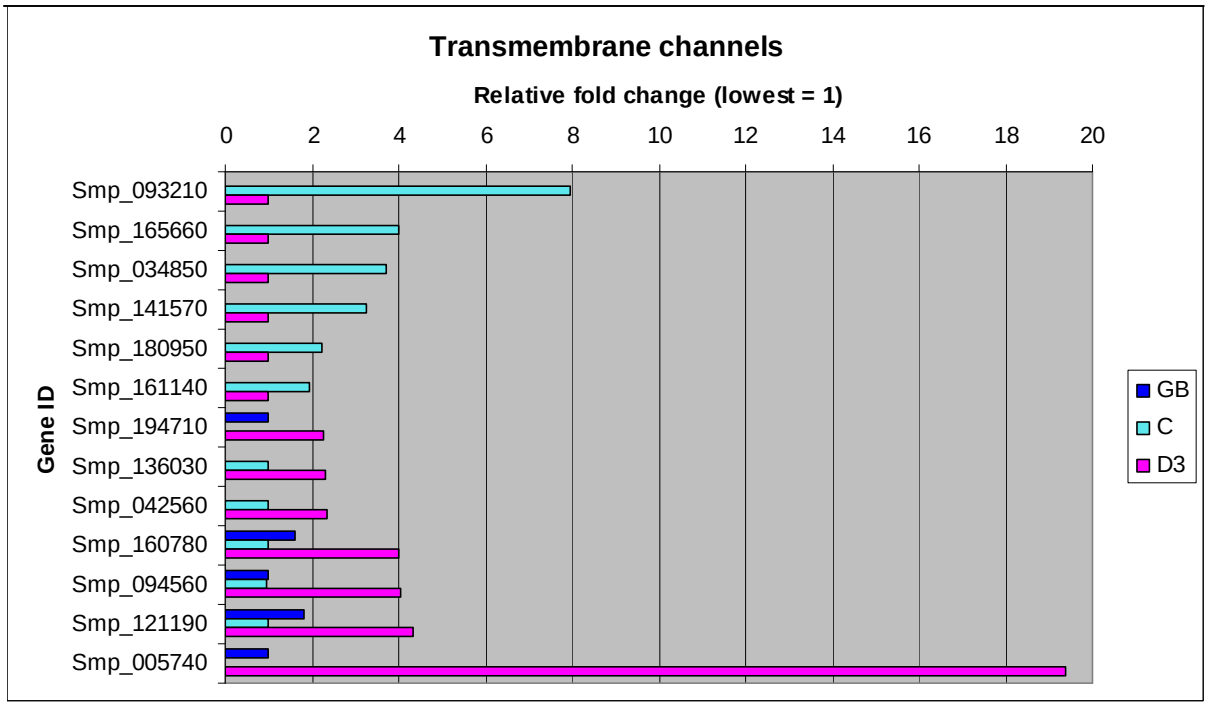
In the cercaria, three sodium-neurotransmitter symporters are up-regulated along with another sodium dependent transporter and a monocarboxylate transporter. The day 3 schistosomulum expresses a different suite of transporters; of these, a cationic amino acid transporter has the highest fold change (19 fold increase compared to the germ ball), followed closely by a glucose transporter (16 fold). Also enriched in the day 3 schistosomulum are two multidrug resistance pumps, three zinc transporters, two noradrenaline/adrenaline transporters, a copper uptake protein, a mitochondrial glutamate carrier and glycerol-3-phosphate transporter. A transcript encoding an intriguing protein called stomatin is also up-regulated in the schistosomulum; it was first described in the surface membrane of erythrocytes [139].



Gene ID	Annotation
Smp_129900	sodium-dependent neurotransmitter transporter
Smp_129920	sodium-dependent neurotransmitter transporter
Smp_129910	sodium-dependent neurotransmitter transporter
Smp_131350	Sodium-solute cotransporter related
Smp_151010	monocarboxylate transporter
Smp_130890	transporter
Smp_084600.1	solute carrier family 37 member 2 (glycerol-3-phosphate transporter
Smp_044690	mitochondrial glutamate carrier protein
Smp_055780	smdr2
Smp_013440	solute carrier protein
Smp_167630	solute carrier family
Smp_176100	cation efflux protein/ zinc transporter
Smp_157430	norepinephrine/norepinephrine transporter
Smp_048230	high-affinity copper uptake protein
Smp_157420	norepinephrine/norepinephrine transporter
Smp_129540	sugar transporter
Smp_135490	multidrug resistance pump
Smp_003440	stomatin-related
Smp_131890	sodium/chloride dependent transporter
Smp_146830	monocarboxylate porter
Smp_123010	cationic amino acid transporter
Smp_105410	glucose transport protein
Smp_176930	cationic amino acid transporter

3.3.14. Membrane Channels

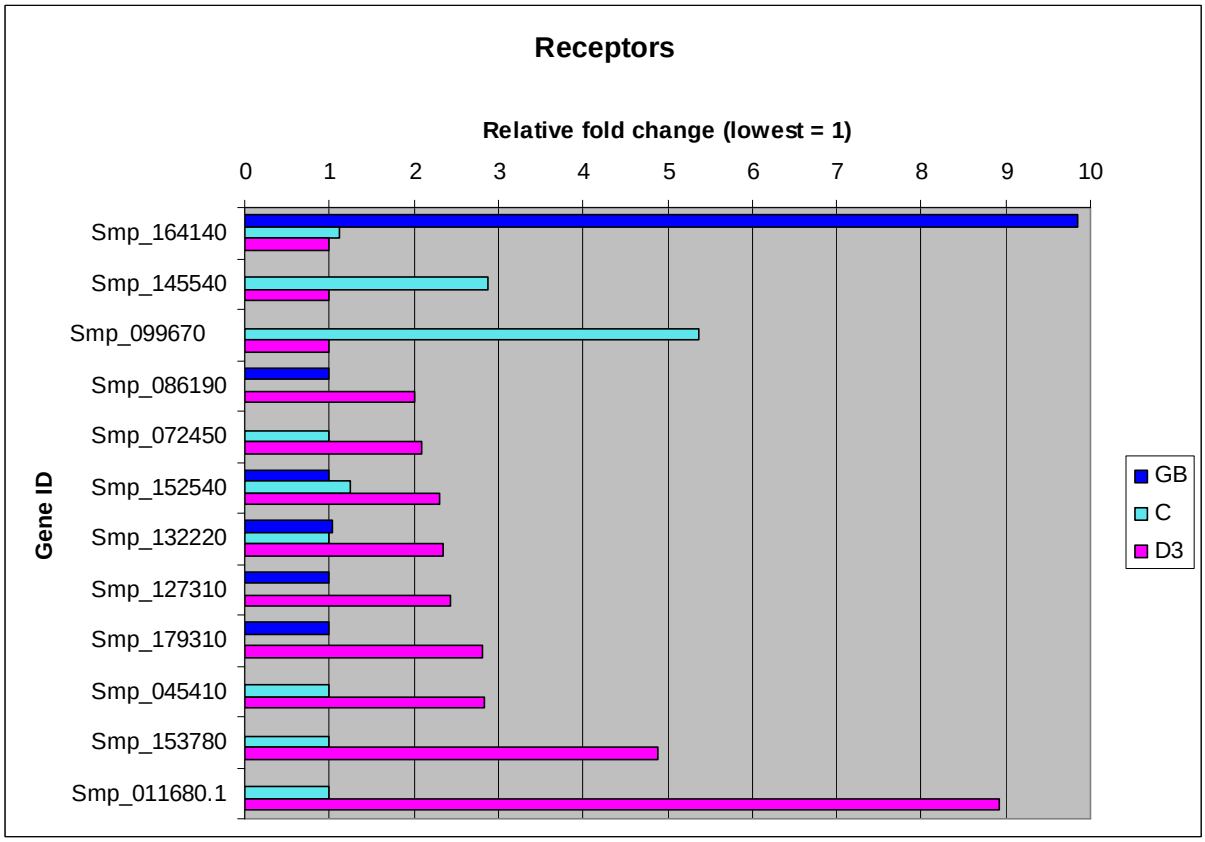
Channels are important for allowing substrates to pass into and out of cells and organelles, as such they may be expressed at the host-parasite interface. An amiloride sensitive sodium channel is expressed 8 fold higher in the cercaria; compared to the day 3 schistosomulum. Potassium channels are up-regulated in both the cercaria and the day 3 schistosomulum compared to the germ ball. Those enriched in cercariae are weakly inward rectifying potassium channels (TWIK family), which maintain membrane potentials. In contrast, voltage gated potassium channels are up-regulated in the day 3 schistosomulum. The most highly up-regulated channel in the day 3 schistosomulum is aquaporin; it is ~19 fold higher, compared to the germ ball. Also up-regulated in the day 3 schistosomulum is a porin (Smp_042560) which forms channels in the mitochondrial outer membrane.



Gene ID	Annotation
Smp_093210	amiloride-sensitive sodium channel-related
Smp_165660	Polycystin cation channel
Smp_034850	twik family of potassium channels-related
Smp_141570	twik family of potassium channels-related
Smp_180950	anion exchange protein
Smp_161140	voltage-gated potassium channel
Smp_194710	voltage-gated potassium channel
Smp_136030	anion exchange protein
Smp_042560	Porin3, Eukaryotic porin family
Smp_160780	voltage-gated potassium channel
Smp_094560	voltage-gated potassium channel
Smp_121190	voltage-gated potassium channel
Smp_005740	aquaporin-3

3.3.15. Receptors

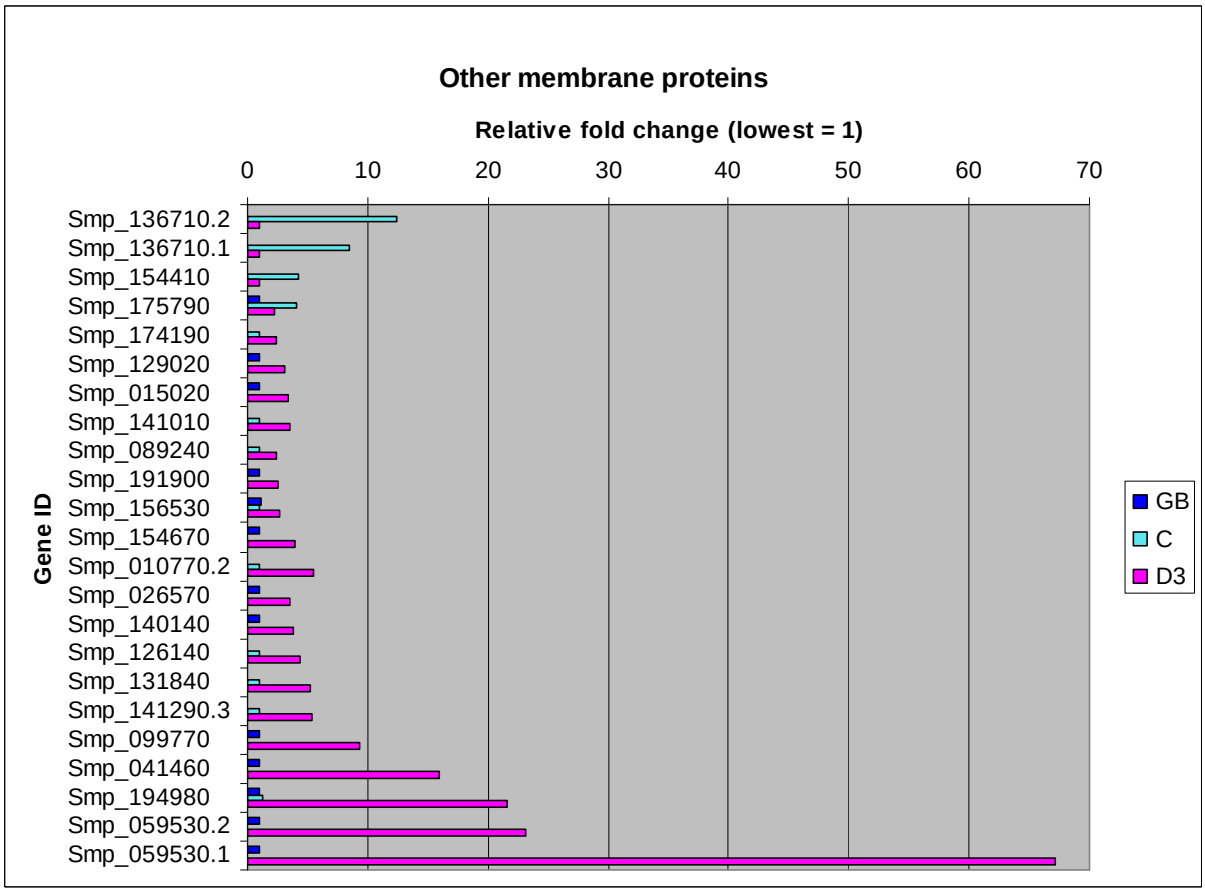
These genes are almost exclusively up-regulated in the day 3 schistosomulum. However, an opsin-like receptor is more than nine fold up-regulated in the germ ball compared to the day 3 schistosomulum. In addition, a muscarinic acetyl choline receptor and an orphan G-protein coupled receptor, are enriched in the cercaria. In the day 3 schistosomulum, three G-protein coupled receptors, an adiponectin receptor, a progestin/adipoQ receptor, a mitochondrial import receptor, along with receptors for biogenic amines, purines, and glutamate are up-regulated. The gene exhibiting the largest fold change in the day 3 schistosomulum is a CD-36 like scavenger receptor.



Gene ID	Annotation
Smp_164140	opsin-like receptor
Smp_145540	muscarinic acetylcholine (GAR) receptor
Smp_099670	GPCR
Smp_086190	progestin and adipoq receptor family member VI
Smp_072450	rhodopsin-like orphan GPCR
Smp_152540	G-protein coupled receptor fragment
Smp_132220	G-protein coupled receptor fragment
Smp_127310	biogenic amine (dopamine) receptor
Smp_179310	P2X purine receptor subunit
Smp_045410	adiponectin receptor
Smp_153780	glutamate receptor, kainate
Smp_011680.1	CD36-like class B scavenger receptor

3.3.16. Others

The majority of the remaining differentially expressed transcripts encoding membrane proteins are up-regulated in the schistosomulum. Exceptions to this are a fucosyl-transferase, and three ATPases (two of which are calcium-transporting sarcoplasmic reticulum type) which are up-regulated in the cercaria. A group of eight tetraspanins are most highly expressed in the schistosomulum along with three innexins and one integrin. Various transferases, a peptidyl-glycine alpha-amidating monooxygenase, an ATPase and a ferlin are also enriched in the day 3 schistosomulum.

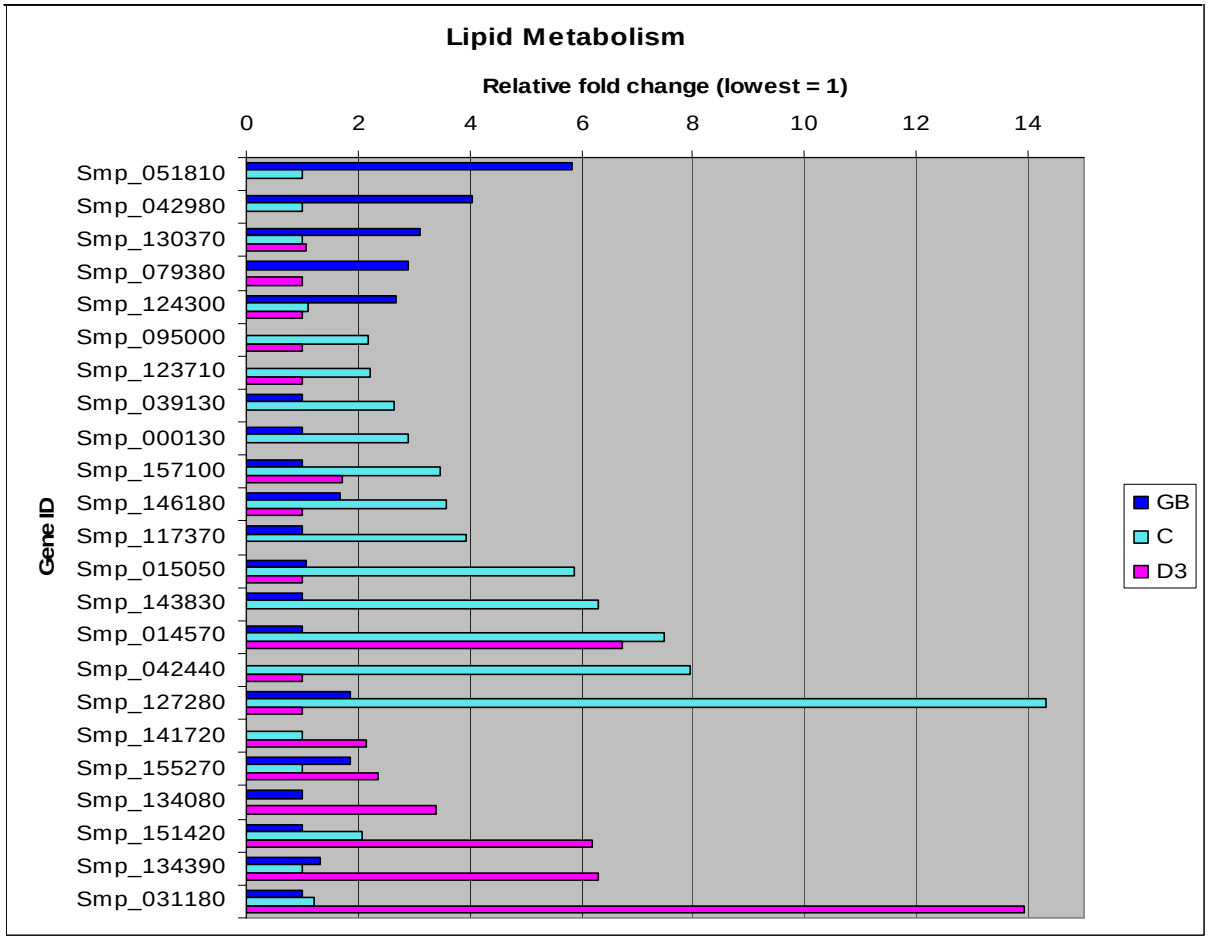


Gene ID	Annotation
Smp_136710.2	calcium-transporting atpase sarcoplasmic/endoplasmic reticulum type
Smp_136710.1	calcium-transporting atpase sarcoplasmic/endoplasmic reticulum type
Smp_154410	alpha(1,3)fucosyltransferase,putative
Smp_175790	atpase, class VI, type 11c
Smp_174190	Tetraspanin-18 (Tspan-18)
Smp_129020	innexin
Smp_015020	na+/k+ atpase alpha subunit
Smp_141010	ferlin-1-related
Smp_089240	beta-1,2-n-acetylglucosaminyltransferase II
Smp_191900	Signal recognition protein
Smp_156530	peptidyl-glycine alpha-amidating monooxygenase
Smp_154670	glycosyltransferase 14 family member
Smp_010770.2	fatty acid acyl transferase-related
Smp_026570	innexin
Smp_140140	tetraspanin
Smp_126140	integrin alpha
Smp_131840	tetraspanin
Smp_141290.3	innexin
Smp_099770	tetraspanin
Smp_041460	tetraspanin D76
Smp_194980	similar to tetraspanin TE736
Smp_059530.2	tetraspanin
Smp_059530.1	tetraspanin

3.3.17. Lipid metabolism

Lipid metabolism was singled out in the genome paper as being an area where schistosomes may be vulnerable to interventions [75]. Therefore, the expression patterns of those genes were investigated. Of the 102 genes described as playing a role in lipid synthesis, transport or degradation 23 were differentially expressed (22.5%). The five genes enriched in the germ ball are involved in lipid synthesis; they are up-regulated three to six fold compared to the cercaria. There is a switch in expression as genes involved in lipid uptake and breakdown are enriched in the cercaria and day 3 schistosomulum.

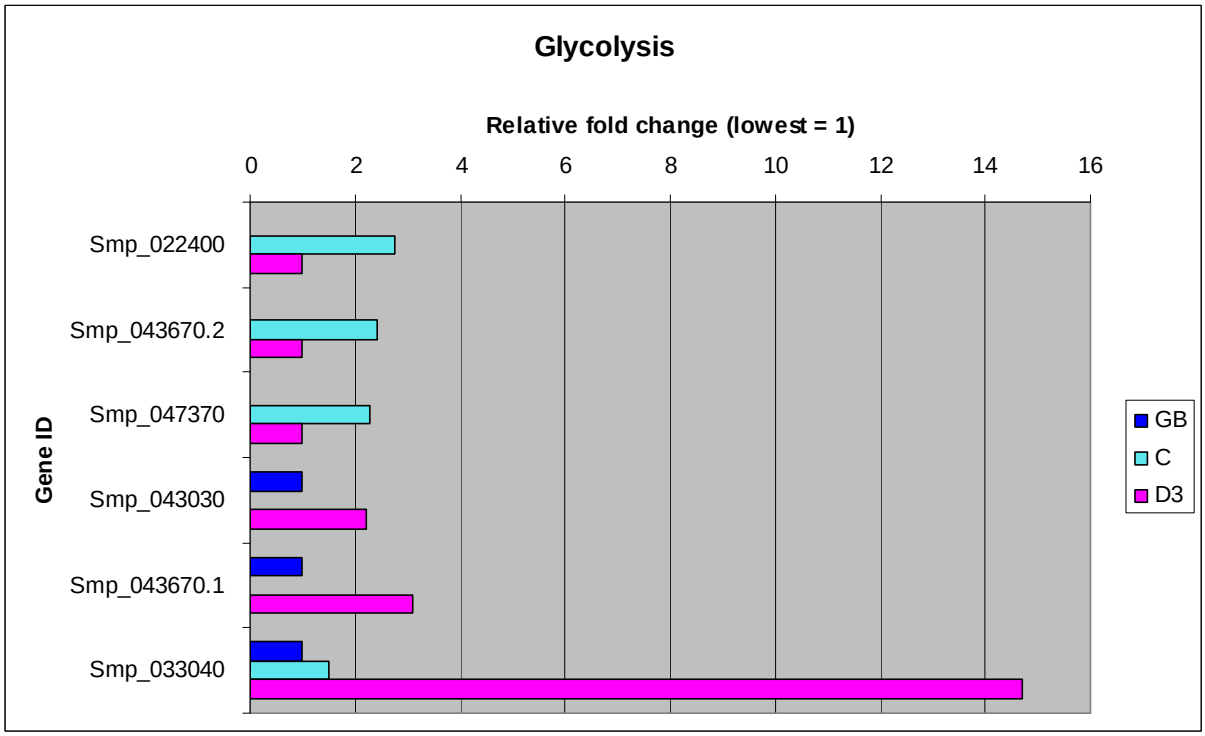
Twelve of the 23 lipid metabolism genes were up-regulated in the cercaria. The most highly expressed cercaria-enriched transcript encodes ceramide synthase, which is 14 fold up compared to the day 3 schistosomulum. The single gene which is up-regulated in both the cercaria and schistosomulum compared to the germ ball is a saposin. The genes most highly expressed in the day 3 schistosomulum include an inositol transporter and two phospholipases and a sterol O-acyltransferase.



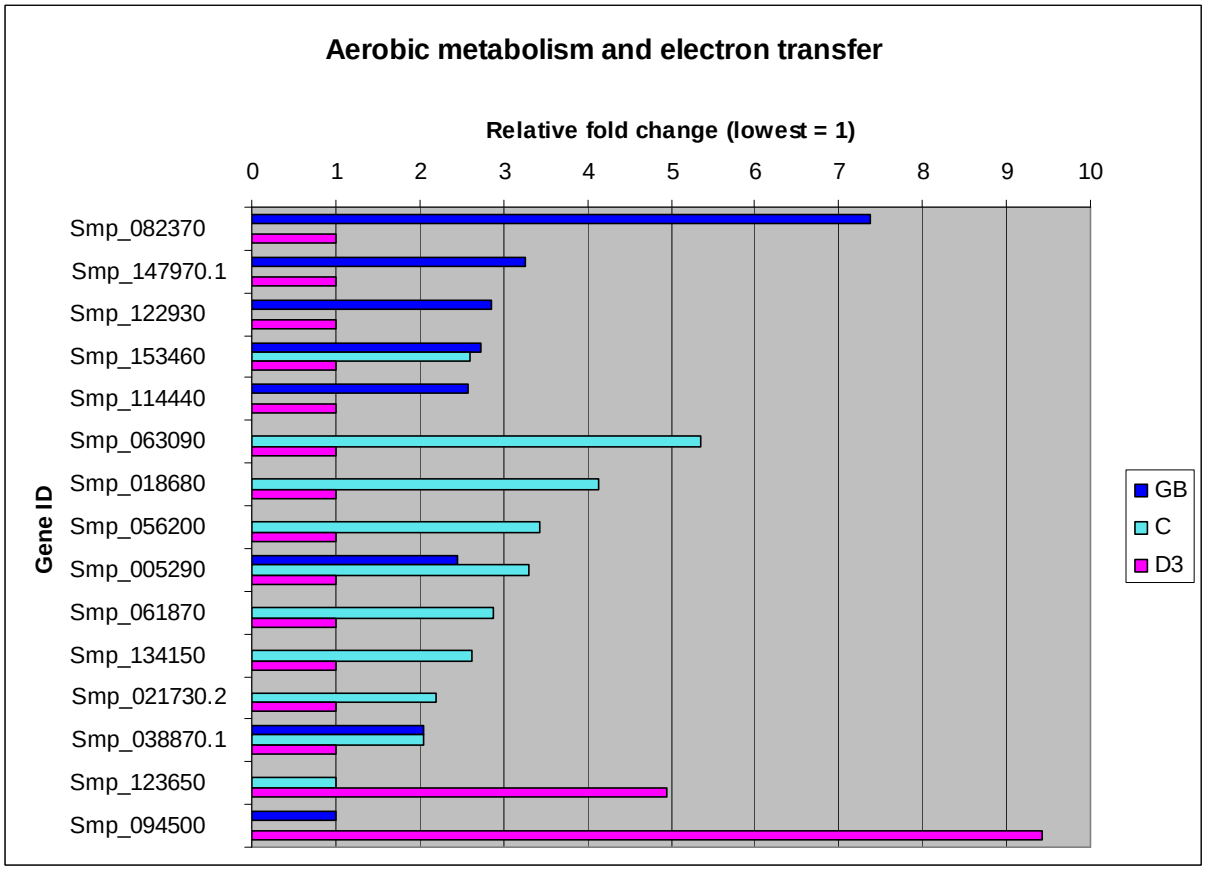
Gene ID	Annotation
Smp_051810	elongation of fatty acids protein 1
Smp_042980	mevalonate kinase
Smp_130370	elongation of fatty acids protein 1
Smp_079380	serine palmitoyltransferase 2
Smp_124300	ERG4/ERG24 ergosterol biosynthesis protein (ERG24)
Smp_095000	inositol monophosphatase
Smp_123710	acetyl-CoA carboxylase
Smp_039130	niemann-pick C1 (NPC1)
Smp_000130	hormone-sensitive lipase (S09 family)
Smp_157100	sphingoid long chain base kinase
Smp_146180	lipase 1,sterol esterase 1,sterol esterase 2,lysosomal acid lipase-related
Smp_117370	sphingosine kinase A, B
Smp_015050	choline kinase
Smp_143830	FFA transport protein
Smp_014570	saposin containing protein
Smp_042440	(dihydro)ceramide Synthase (LAG1)
Smp_127280	diacylglycerol O-acyltransferase 1
Smp_141720	3-keto-dihydrosphingosinereductase
Smp_155270	HMG-CoA synthase,hydroxymethylglutaryl-CoAsynthase
Smp_134080	inositol transporter
Smp_151420	phospholipase D
Smp_134390	sterol O-acyltransferase 1
Smp_031180	Phospholipase A

3.3.18. Energy metabolism

The life-cycle stages under investigation in the current study have different energy requirements and sources. Genes involved in these processes were highlighted by the GOstats analysis in each of the comparisons made. Of the 30 *S. mansoni* genes whose products are involved in glycolysis, six were found to be differentially expressed (20%). The highest fold change (14) was in the day 3 schistosomulum-enriched gene lactate dehydrogenase. The differentially expressed genes of the citric acid cycle (aerobic respiration) were all up-regulated in the cercaria. Likewise, a large proportion of 'electron transport'-related genes are enriched in the cercaria. All of these contribute to oxidative phosphorylation. The five genes involved in electron transport, which are up-regulated in the germ ball, are involved in mitochondrial transport and fatty acid metabolism, whereas choline dehydrogenase and peroxidasin are enriched in the schistosomulum.



Gene ID	Annotation
Smp_022400	glucose-6-phosphate isomerase
Smp_043670.2	6-phosphofructokinase
Smp_047370	malate dehydrogenase
Smp_043030	hexokinase
Smp_043670.1	6-phosphofructokinase
Smp_033040	L-lactate dehydrogenase



Gene ID	Annotation
Smp_082370	NADP transhydrogenase
Smp_147970.1	kif1
Smp_122930	acyl-CoA dehydrogenase
Smp_153460	monoxygenase
Smp_114440	acyl-CoA dehydrogenase
Smp_063090	aconitate hydratase
Smp_018680	isocitrate dehydrogenase
Smp_056200	Isocitrate dehydrogenase [NAD] subunit gamma, mitochondrial
Smp_005290	cytochrome C1
Smp_061870	ubiquinol--cytochrome C reductase
Smp_134150	glucose-methanol-choline (gmc) oxidoreductase
Smp_021730.2	cytochrome c oxidase
Smp_038870.1	NADH-ubiquinone oxidoreductase
Smp_123650	peroxidasin
Smp_094500	choline dehydrogenase

3.3.19. Custom categories

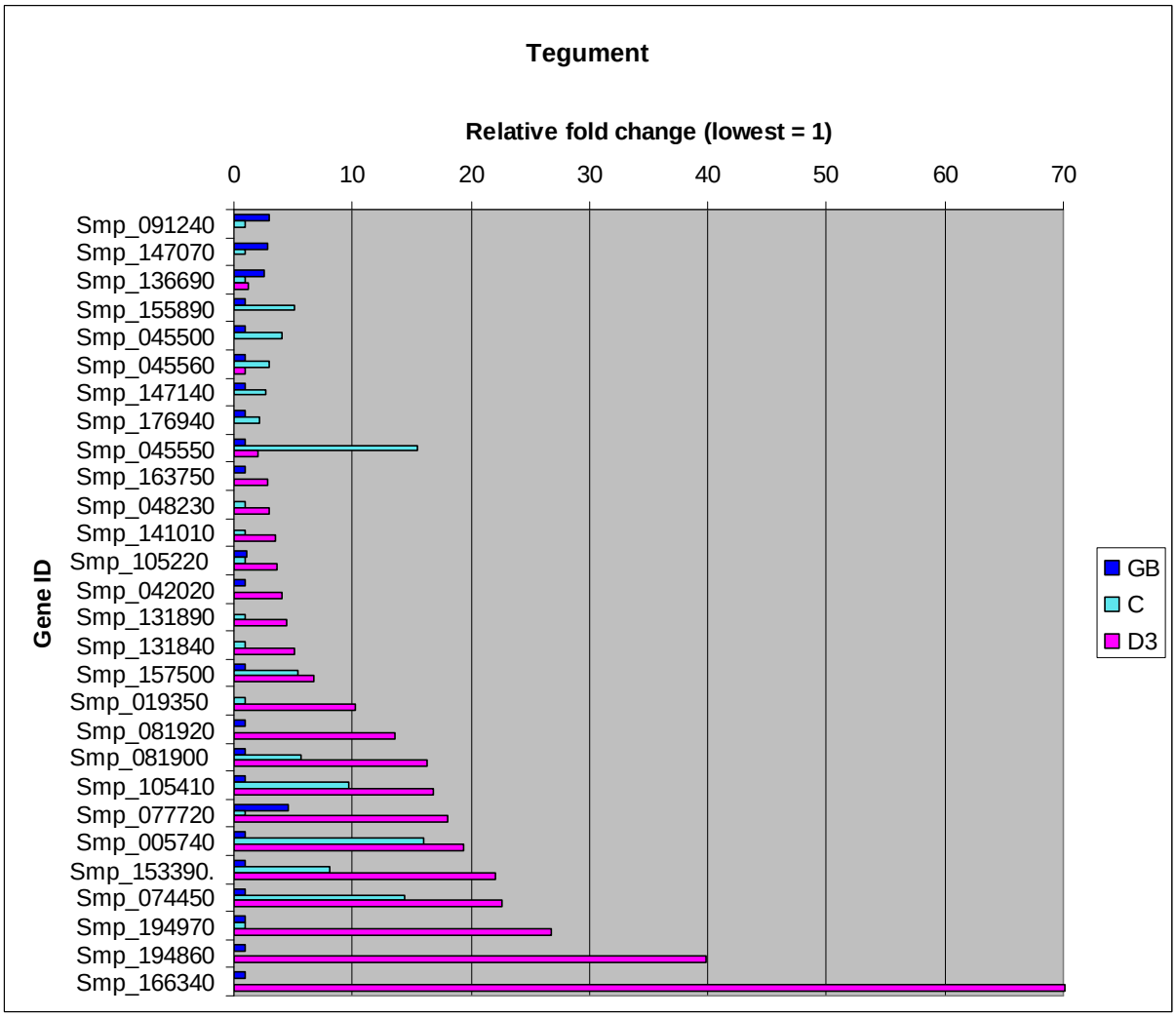
In addition to the GOstats analysis and categories based on genome annotations, the expression patterns of custom gene categories were investigated. These are based on proteomic or localisation data, and are linked with a tissue (tegument, gut) or condition such

as stress. Some of these genes are also included elsewhere (e.g. membrane and protease categories).

3.3.20. Tegument

The tegument undergoes substantial remodeling on entry to the mammalian host, and is a large component of the host-parasite interface. Of the 65 genes in this category, only three are most highly expressed in the germ ball. They are a voltage-dependent anion-selective channel, an amino acid transporter, and acetylcholinesterase.

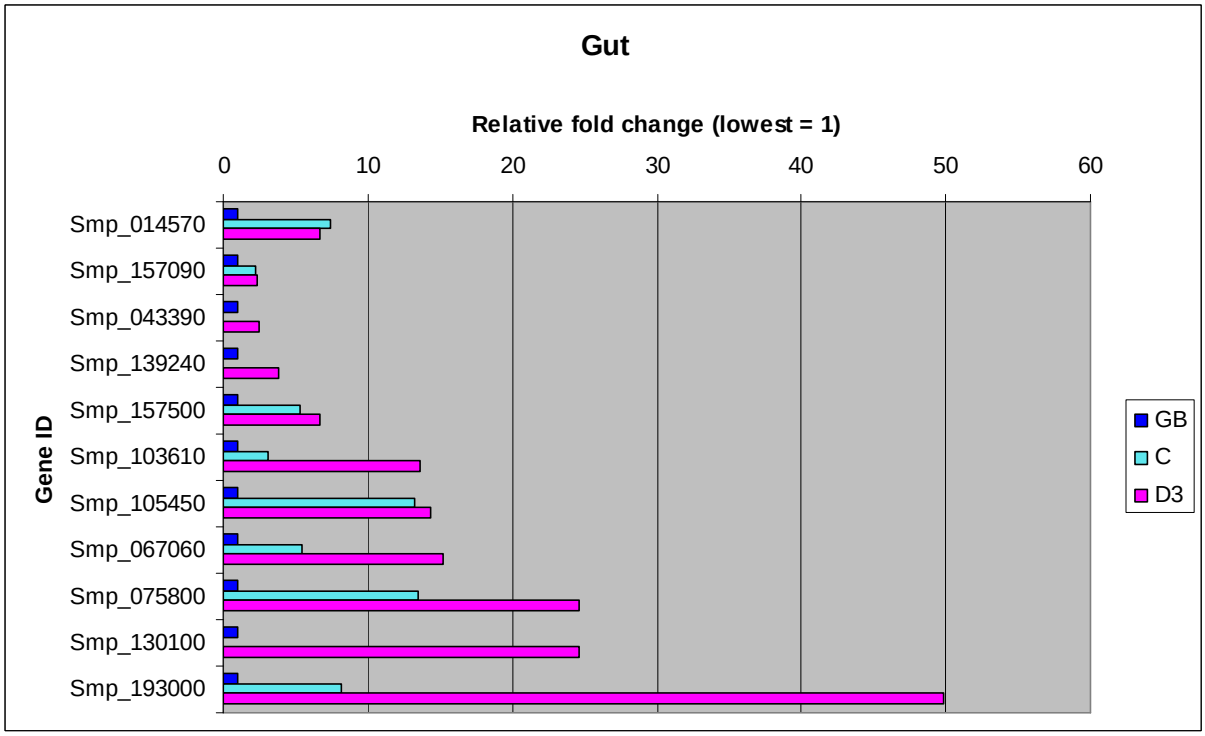
In the cercaria, three closely related annexins, alkaline phosphatase, a cation channel and a cationic amino acid transporter are most highly expressed. One of the annexins (Smp_045550) is expressed ~15 fold higher, compared to the germ ball. The majority of protein encoding-genes of the tegument are up-regulated in the day 3 schistosomulum compared to the germ ball. Five of these encode copies of a putative complement inhibitor CD59, one of which (Smp_166340) is ~70 fold higher than in the germ ball. Another of them (Smp_081900) is up-regulated in both the day 3 schistosomulum and the cercaria compared to the germ ball. Other genes that share this pattern are a calpain, the glucose transporter SGTP4, an aquaporin, an ectonucleotide-pyrophosphatase and a protein of unknown function, which was discovered on the schistosome surface by Braschi *et al* [140]. Transcripts which are up-regulated solely in the day 3 schistosomulum include otoferlin, ferlin, a copper transporter, and a sodium/chloride dependent transporter. Sm8.7 (low molecular weight secreted protein) is striking; it is up-regulated 40 fold compared to the germ ball.



Gene ID	Annotation	Gene ID	Annotation
Smp_091240	voltage-dependent anion-selective channel	Smp_131890	sodium/chloride dependent transporter
Smp_147070	amino acid transporter	Smp_131840	tetraspanin
Smp_136690	Acetylcholinesterase	Smp_157500	calpain (C02 family)
Smp_155890	alkaline phosphatase	Smp_019350	CD59
Smp_045500	annexin	Smp_081920	CD59
Smp_045560	annexin	Smp_081900	CD59
Smp_147140	transient receptor potential cation channel Braschi	Smp_105410	glucose transport protein SGTP4
		Smp_077720	annexin
Smp_176940	cationic amino acid transporter	Smp_005740	aquaporin-3
Smp_045550	annexin	Smp_153390	Ectonucleotide-pyrophosphatase
Smp_163750	otoferlin	Smp_074450	expressed protein (Braschi)
Smp_048230	high-affinity copper uptake protein	Smp_194970	similar to tetraspanin TE736
Smp_141010	ferlin-1-related	Smp_194860	Sm8.7 Low molecular weight protein
Smp_105220	CD59	Smp_166340	CD59

3.3.21. Gut-associated

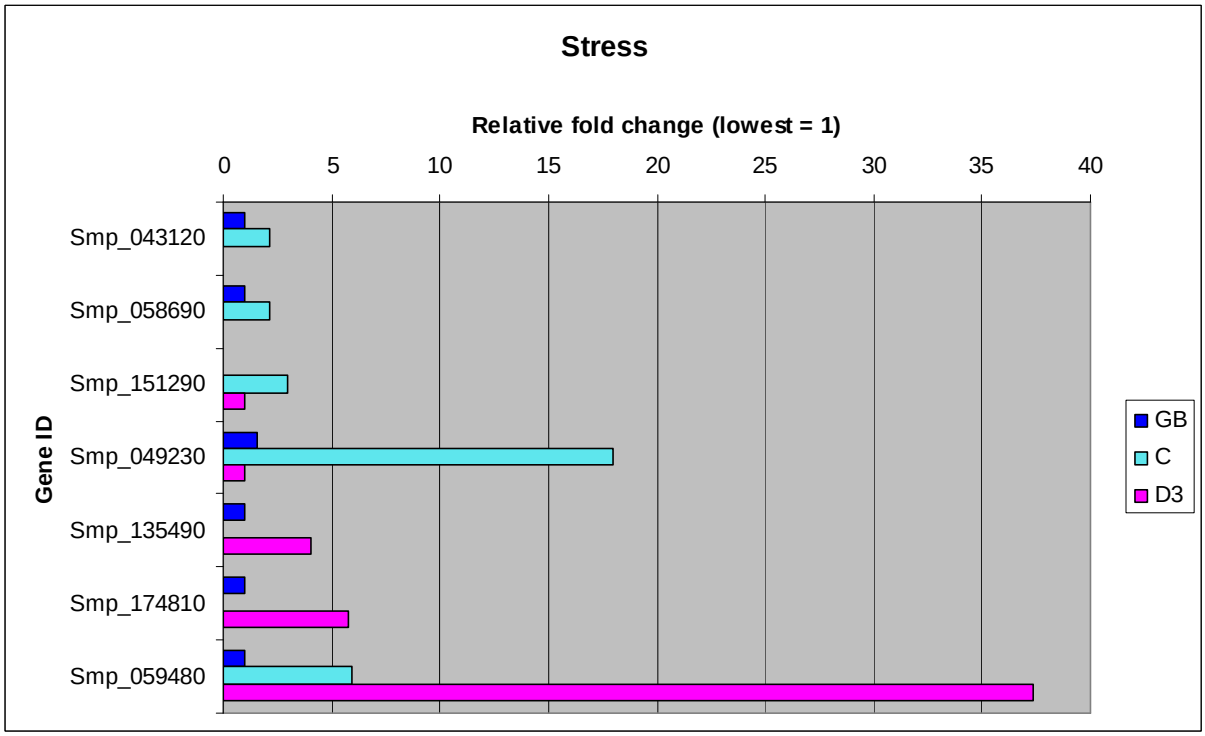
The expression patterns of genes encoding gut-associated proteins were investigated to shed light on the issue of when the larval gut becomes active. This category comprises genes encoding proteins that were identified in adult worm vomitus by Hall *et al* [141]. Of the 32 genes in this custom category, ten (31%) are differentially expressed. They include cathepsins, saposins, an asparaginyl endopeptidase, and a calpain. All of them are up-regulated in the day 3 schistosomulum compared to the germ ball. Most are also enriched in the cercaria; the exceptions are a glycosyl hydrolase, cathepsin S, and a saposin.



Gene ID	Annotation
Smp_014570	saposin containing protein
Smp_157090	SmCL3 (C01 family)
Smp_043390	Glycosyl hydrolase
Smp_139240	cathepsin S (C01 family)
Smp_157500	calpain (C02 family)
Smp_103610	cathepsin B-like peptidase (C01 family) Sm31
Smp_105450	saposin containing protein
Smp_067060	cathepsin B-like peptidase (C01 family)
Smp_075800	asparaginyl endopeptidase Sm32
Smp_130100	saposin containing protein
Smp_193000	SmCL2-like peptidase (C01 family)

3.3.22. Stress-related genes

To gauge the response to stress presumed to be experienced by larval schistosomes during the mammalian infection process; the expression patterns of stress-related genes were investigated. This custom category, based on literature surveys, comprises 29 genes, only six of which were differentially expressed (21%). None of the stress related genes were germ ball-enriched. Genes encoding universal stress protein, glutathione peroxidase and a multidrug resistance pump (MDR) are up-regulated in the cercaria. Notably up-regulated at this stage is a small heat shock protein, which is 18 fold higher than in the other two life cycle stages. A different MDR is up-regulated in the day 3 schistosomulum, along with a superoxide dismutase and a thioredoxin peroxidase. The latter gene shows the largest difference; it is 37 fold higher in the day 3 schistosomulum than in the germ ball.



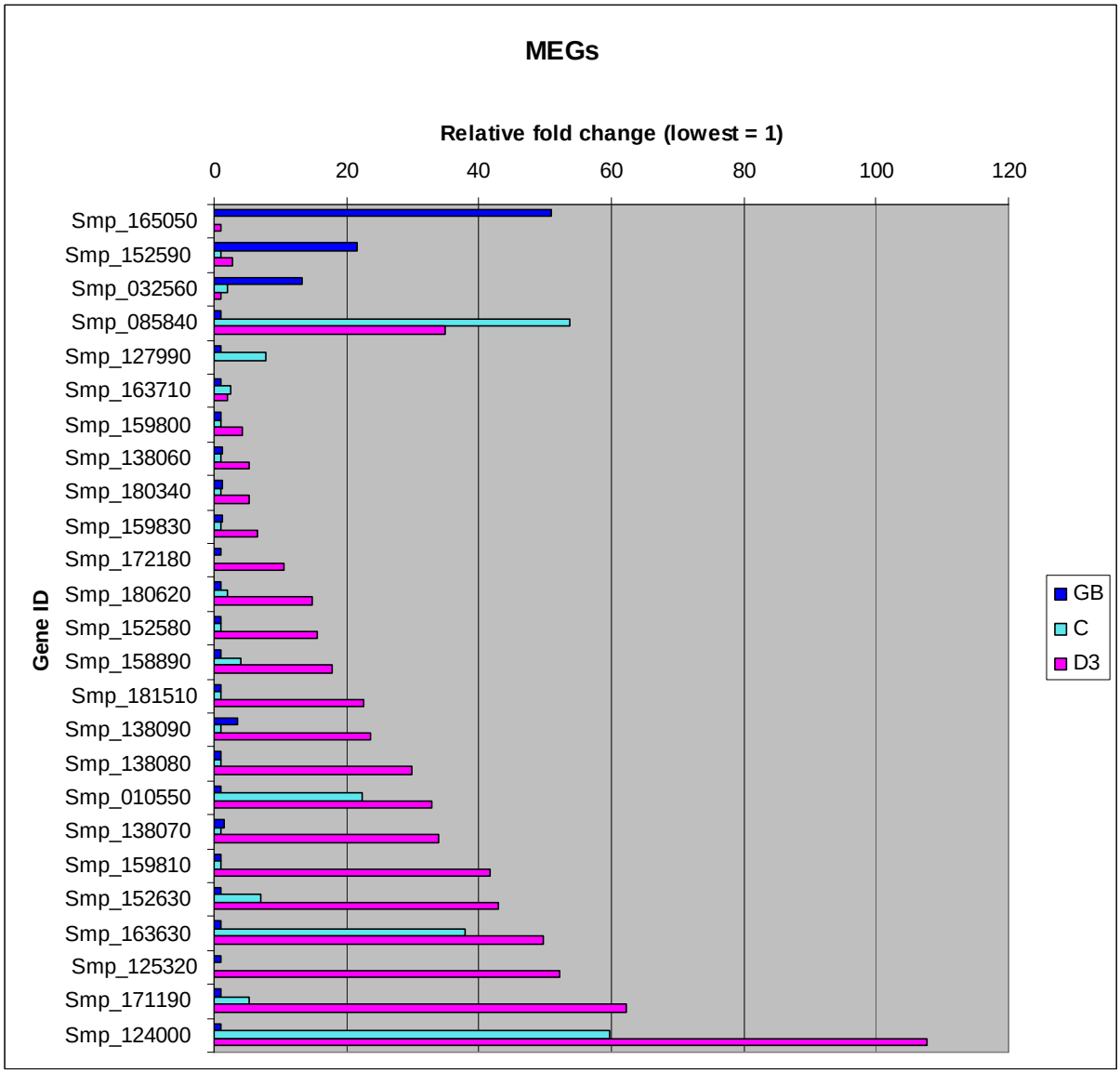
Gene ID	Annotation
Smp_043120	universal stress protein
Smp_058690	glutathione peroxidase
Smp_151290	multidrug resistance pump
Smp_049230	Sm-p40 (small heat shock protein)
Smp_135490	multidrug resistance pump
Smp_174810	cu/zn superoxide dismutase
Smp_059480	thioredoxin peroxidase

3.3.23. Micro exon genes

The micro exon genes were briefly described by Berriman *et al* [75] and a detailed account has been given by DeMarco *et al* [142]. They are hypothesised to represent a mechanism for generating variant proteins by schistosomes. The MEGs show the highest fold changes encountered in this entire study, with eight of the 25 having values greater than 40. Of the 14 MEG families described previously [75] all but three were found to be highly expressed in either the germ ball, cercaria or day 3 schistosomulum. Only MEGs 1 and 11 showed no differential expression during the transition studied. Fluorescence data across life cycle stages revealed that MEG 1 exhibits a consistent low level of expression, whereas MEG 6, which has small fold changes (up two fold in both cercaria and day 3 schistosomulum compared to germ ball) was highly expressed at each stage. No data could be gleaned for MEG 11.

Due to the very high fold changes exhibited by many of the MEGs, gene models of transcripts annotated as ‘hypothetical protein’ or ‘expressed protein’ that were highly expressed in one stage or another in my study were manually inspected. This led to the identification of four new MEGs (Smp_010550, Smp_158890, Smp_180620, Smp_032560). Further investigation revealed that the newly identified MEGs have no homology to previously identified MEG families, or to each other; they represent novel families, and have been named MEGs 15-18 respectively.

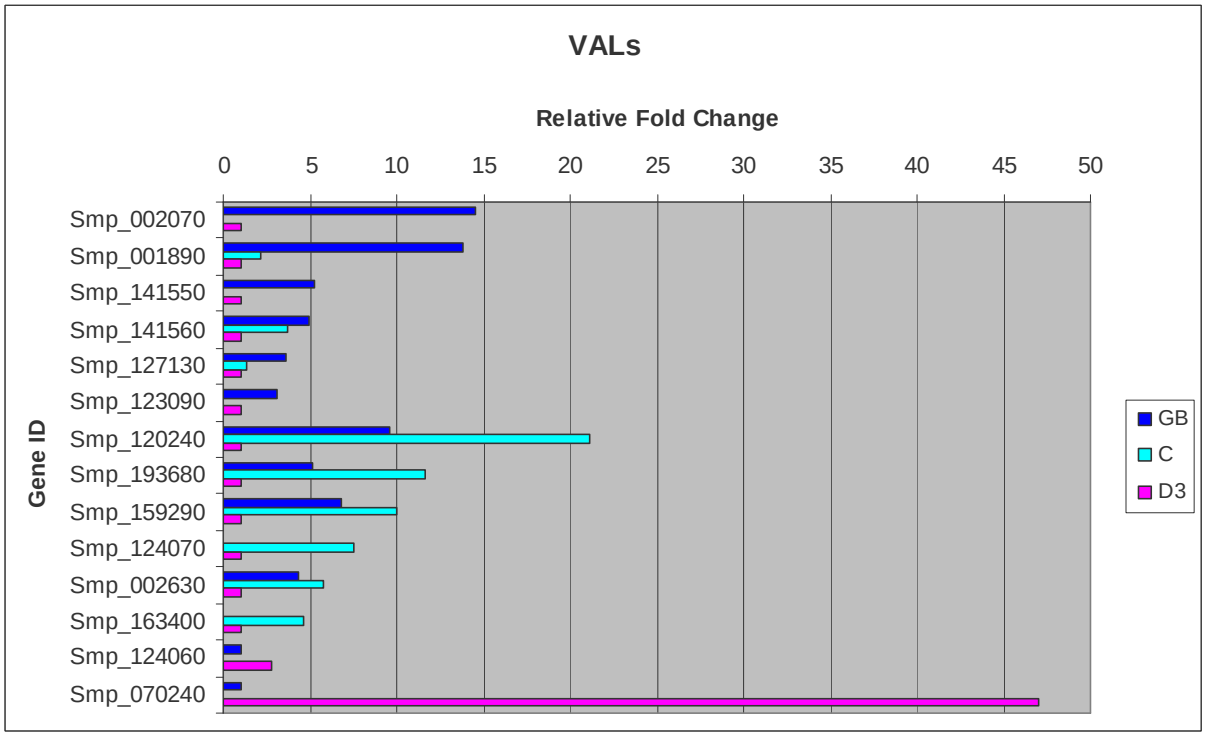
The majority of MEGs are up-regulated in the day 3 schistosomulum compared to the other two stages; however, MEGs 7, 10 and 18 are enriched in the germ ball, and MEGs 4.2, 13 and 6 are most highly expressed in the cercaria. Although MEGs 4.1, 8, 12, 14, 15, 16, and 17 are most highly expressed in the day 3 schistosomulum, they are also up-regulated in the cercaria compared to the germ ball. MEGs 2, 3, 5, 8 and 9 are up-regulated in the day 3 schistosomulum compared to both germ balls and cercariae.



Gene ID	Annotation	Gene ID	Annotation
Smp_165050	MEG 7	Smp_158890	MEG 16
Smp_152590	MEG 10	Smp_181510	MEG 2
Smp_032560	MEG 18	Smp_138090	MEG 3
Smp_085840	MEG 4	Smp_138080	MEG 3
Smp_127990	MEG 13	Smp_010550	MEG 15
Smp_163710	MEG 6	Smp_138070	MEG 3
Smp_159800	MEG 2	Smp_159810	MEG 2
Smp_138060	MEG 3	Smp_152630	MEG 12
Smp_180340	MEG 2	Smp_163630	MEG 4
Smp_159830	MEG 2	Smp_125320	MEG 9
Smp_172180	MEG 8	Smp_171190	MEG 8
Smp_180620	MEG 17	Smp_124000	MEG 14
Smp_152580	MEG 5		

3.3.24. Venom allergen-like proteins

VALs have been identified in secretions of schistosomes and parasitic nematodes. They are proposed to play a role in immune evasion [116]. Of the 28 VAL genes described by Chalmers *et al* [125], 13 are differentially expressed during the germ ball to day 3 schistosomulum transitions (46%). Most are enriched in the germ ball and/or cercaria compared to the day 3 schistosomulum. VALs 4, 18, 19, 20, 24 and 25 are most highly expressed in the germ ball compared to the day 3 schistosomulum. Of these, VALs 18 and 25 are also enriched in the cercaria. VALs 1, 2, 16, 17 and 21 are most highly expressed in the cercaria compared to the day 3 schistosomulum; of these, VALs 1, 2, and 21 are also up-regulated in the germ ball. Only two VALs (7 and 13) are most highly expressed in the schistosomulum. VAL 7 has the highest fold change of any of the VALs: it is ~47fold higher in the day 3 schistosomulum than the germ ball. This is noteworthy as none of the other VAL-encoding genes exhibit fold changes greater than 15.



Gene ID	Annotation
Smp_002070	VAL 04
Smp_001890	VAL 18
Smp_141550	VAL 24
Smp_141560	VAL 25
Smp_127130	VAL 20
Smp_123090	VAL 19
Smp_120240	VAL 01
Smp_193680	VAL 01
Smp_159290	VAL 21
Smp_124070	VAL 16
Smp_002630	VAL 02
Smp_163400	VAL 17
Smp_124060	VAL 13
Smp_070240	VAL 07

3.4. Discussion

3.4.1. Design and application of the first genome wide microarray for S. mansoni

The publication of the *S. mansoni* genome in 2009 [75] represents the culmination of nearly three decades of sequencing work (see Chapter 1.2). The availability of the gene predictions and many ESTs at GeneDB have enabled the design of the first genome wide array for *S. mansoni*. The ESTs were compiled with the gene predictions and 45mer probes were designed to the resulting contigs. Each locus is represented by 1-300 non-redundant probes ensuring both diversity and coverage. 15.6% of loci represented on the array do not have significant matches to gene predictions when blasted against the genome either by BLASTn or tBLASTx. There are two explanations for this: a) this sequence is not encoded by the *S. mansoni* genome b) the sequence is present in the *S. mansoni* genome, but has not been identified by gene finding software. The sequences which are not truly *S. mansoni* are likely to be contaminating sequences of host origin, (this is most likely for germ ball, or *ex vivo* intramammalian library sequences). Host sequences have been mistaken previously for novel schistosome genes [143]. The ESTs were included in the compilation for the array to reduce the chance of excluding bonafide schistosome genes that were not called by gene finding software. However, although some of the non-Smps represented on the array may be new uncalled schistosome-specific genes, they were filtered out before further analysis as there was no annotation available, and hypotheses would be difficult to make about their function.

An important caveat is that some of the gene models in GeneDB, on which the array is based, were faulty. For example, a gene encoding a secreted protein named LMWP (now Smp_194860) was concatenated with that for a histone (now Smp_194870), resulting in the erroneous annotation of a secreted histone (Smp_053310 in version 3.1). In the present study, such a gene model would be assigned an expression level based on the probes representing both the LMWP and histone portions.

Of the loci represented on the array, 42.9% have no annotation available. This figure is in keeping with previous sequencing projects using ESTs; analysis of the transcriptome showed that ~50% of the sequences available had no homology to anything at GenBank [73].

The expression patterns are presented as relative fold changes; for each gene, the baseline is the expression level of the lowest stage and is set to 1. Comparisons have not been made on an absolute basis. Low fold changes indicate only that the expression of a gene does not vary; it may be consistently low or consistently high across the stages. An example of the latter case is MEG 6 (see Micro-exon genes section in Results). Genes which were not differentially expressed were not reported, as the focus here is on the biological differences between the life cycle stages.

In the contrasts carried out here, the cercaria and day 3 schistosomulum were most similar to each other, and the germ ball and cercaria most different. The germ ball is an outlier. This can be explained by taking into account the environments and developments taking place. As we will see later, many of the genes required by the migrating schistosomulum are up-regulated in the cercariae in preparation for host entry. Both protein synthesis and cell division are processes highly up-regulated in the germ ball compared to the other stages. Furthermore, the genes needed specifically for intra-molluscan life and preparation for host entry are not transcribed at the later stages, a fact which also contributes to the high contrasts observed between this stage and the other two.

An overview of the functional categories whose member genes are particularly enriched in each life cycle stage will be given, followed by an in depth consideration of each category by named gene.

3.4.2. Germ ball

As we saw in Chapter 2, the term ‘germ ball’ encompasses embryonic cercariae at all stages of development, from a single cell, up to, but not including, the mature cercaria. The latter comprises approximately 1000 cells [42], which form tissues and organs. Although only 5% of the total *S. mansoni* genes involved in DNA replication are differentially regulated in the transitions studied here, it is noteworthy that all of them are up-regulated in the germ ball, reflecting the obvious occurrence of cell division. The majority of differentially expressed translation-associated genes are also up-regulated in the germ ball, but the fold changes are not as large; the maximum is 4.5. Both DNA replication and translation are processes necessary for the development of germ balls into cercariae. Likewise, several genes encoding morphogens such as homeobox and Wnt proteins are most highly expressed in the germ ball.

Again, the fold changes are not exceptional, the maximum being up-regulated seven fold compared to the day 3 schistosomulum. These genes are necessary for dorso-ventral and anterio-posterior patterning, essential processes in embryogenesis.

The expression of genes encoding serine and metallo-proteases was somewhat more dramatic, with 33 and 35% respectively of the *S. mansoni* gene complement differentially expressed. The vast majority in both cases were up-regulated in the germ ball with maximum fold changes of 65 and 27 respectively. In addition, several VAL-encoding genes were up-regulated in the germ ball. Proteases (cercarial elastases in particular) are known to be secreted during penetration of skin, enabling the parasite to gain access to the host [115, 116, 144]. VAL-proteins have been identified in cercarial secretions; the majority of transcripts encoding these putative immunomodulators are up-regulated in the germ ball and/or cercaria. These will be discussed in more detail in the relevant sections below.

3.4.3. Cercaria

The free-living cercaria swims actively in order to find a host, which it must do before it uses up its supply of glycogen. Thus, energy metabolism and neural functions are salient features of this highly motile life cycle stage. Genes that were most highly expressed in the cercaria belong to a wide range of groups. The most striking were those involved in lipid and energy metabolism. In each case, more than 50% of the differentially expressed genes are up-regulated in the cercaria; the highest fold changes in the two groups are 14 and 5, respectively. Genes encoding various membrane proteins involved in neural activity, including neurotransmitter transporters, are also up-regulated at this stage, with expression levels up to 14 fold that in the germ ball. In addition, many genes encoding proteins required by the schistosomulum immediately on entry to the host are already up-regulated in the cercaria compared to the germ ball. This includes gut-associated cysteine proteases and the tegumental aquaporin.

3.4.4. Day 3 schistosomulum

Inspection of the day 3 schistosomulum-enriched genes reveals a marked activation of genes across a wide range of categories on entry to the mammalian host. There is a switch from the serine and metalloproteases associated with skin penetration to aspartlyl and cysteine proteases involved in feeding. This is highlighted by the up-regulation of genes in the ‘gut’

category. The majority of genes encoding known tegument proteins were also up-regulated in the day 3 schistosomulum, reflecting the new surface that is required for survival in the mammalian host [145]. In particular, a CD59 gene was up-regulated 70 fold compared to the germ ball. In fact, the vast majority of the membrane proteins were most highly expressed in the day 3 schistosomulum, with increases up to 60 fold compared to the germ ball, emphasising the dramatic changes required for adaptation to the mammalian environment. However, the most striking category was the MEGs (de Marco et al, submitted). 76% of these genes were up-regulated in the day 3 schistosomulum, most being expressed at least 20 fold higher than the germ ball. MEG 14 exhibited the highest fold change in the entire study. It was expressed 110 fold higher compared to the germ ball. This indicates an important role for these genes in establishing the parasitic state.

3.4.5. Cell proliferation

All of the differentially expressed genes involved in cell proliferation are up-regulated solely in the germ ball. This includes DNA polymerases and Cyclins B and D, which control cell cycle progression. Pollok *et al* showed that CDC45 is rate limiting for DNA replication licensing in human cells; up-regulation of this important transcript in germ balls is in accordance with their rate of cell proliferation [146]. The results presented here are in keeping with those of earlier studies into larval maturation; although cell proliferation is necessary for cercariogenesis, Lawson and Wilson found no evidence that development occuring in the migrating parasite prior to arrival at the liver is due to cell proliferation; rather they suggest it is achieved through tissue reorganisation [104].

3.4.6. Protein synthesis genes are down-regulated in the cercaria

Only one RNA polymerase is differentially expressed in the transitions studied. Smp_004640 is up-regulated 2.66 fold in the germ ball compared to the cercaria. This suggests that each of the stages under investigation carry out transcription at a similar rate. However, transcripts encoding ribosomal proteins, tRNA synthetases, and translation initiation and elongation factors were all up-regulated in the germ ball compared to the cercaria and day 3 schistosomulum. Many of these genes were also up-regulated at day 3 compared to the cercaria. This pattern suggests the germ ball is the most translationally active, the cercaria the least. Blanton and Licate reported that protein synthesis occurs in the absence of a significant increase in transcription during the first 24 hours after transformation, indicating that post

transcriptional regulation occurs [65]. Jolly *et al* reported that genes encoding ribosomal proteins were up-regulated in infected hepatopancreas compared to cercaria and adult [97]. Harrop and Wilson studied protein synthesis rates in larval schistosomes using incorporation of radio-labelled methionine into proteins and noted a lag in the 24 hours after transformation before scqm incorporation occurred [64]. This may represent a lag phase in the uptake of amino acids from the media as well as a low level of translation. Three genes in this category were most highly expressed in the cercaria, one of which is glutamate synthase. As L-glutamate has been shown to contract *S. mansoni* nerve fibres [147] it may well be required by the cercaria for neuromuscular signalling during swimming.

3.4.7. Development

Only a small proportion of the genes implicated in neural development according to Berriman *et al* [75] are differentially regulated during the germ ball-day 3 schistosomulum transition. It may well be the case that they are required for other important phases of development during the lifecycle such as: maturation of the migrating larva to the adult stage on arrival at the liver, egg production, miracidial development, or sporocystogenesis. All differentially-regulated early neural patterning genes, including neurogenin, are enriched in the germ ball. This suggests that the basic pattern of the brain and nervous system is developed in the germ ball and remains unchanged until a later stage. In Chapter 2 we saw that the brain is visible at the stubby tailed stage of germ ball development. Although its extent increases, the position of the brain does not change throughout the intra-mammalian life.

It is logical that important body planning genes such as Hox and Wnt are enriched in the embryonic stage. Genes involved in cell polarity determination are up-regulated later – in the cercaria or day 3 schistosomulum. This reflects the morphological changes that are taking place. It has been proposed previously that the larvae spend some time ‘resting’ in the skin before onward migration. The results presented here show that developmental patterning genes are up-regulated at this point. This is in keeping with the observation that the acetabular glands and ducts and the cercarial sensory endings disappear during transformation.

Cadherins are up-regulated in the germ ball and the day 3 schistosomulum compared to the cercaria. They are membrane spanning proteins which mediate cell - cell adhesion. They are important for tissue formation and allow similar cells to aggregate. However, the repertoire of

cadherins expressed by the germ ball is different to that of the day 3 schistosomulum, suggesting that associations between different cell types form at these two distinct life cycle stages.

3.4.8. Proteases

The proteases will be discussed by catalytic type.

Serine proteases

The most striking group of serine proteases are the cercarial elastases (MEROPS class S01). Previously, Salter *et al* described five cercarial elastase genes, stating that two of them accounted for 90% of the protein present [115]. The data presented here show that the 11 genes encoding this enzyme are transcribed at high levels in the germ ball; all of them are expressed at least 10 fold higher than in the schistosomulum. This shows that skin invasion is more complex than had been thought previously. Also enriched in the germ ball is a rhomboid protease. Transmembrane proteases of this type have been implicated in cell invasion by *Toxoplasma* [148] and malaria parasites [149], immune evasion by *Entamoeba histolytica* [150], and in cell signalling in *Drosophila* [151]. Freeman notes that work is needed to clarify whether inhibition of rhomboid proteases prevents invasion by parasites, and furthermore whether they are suitable as drug targets [151]. The localisation of the germ ball-enriched rhomboid protease would help decipher its role in schistosomes, as both cell signalling and host interaction are plausible functions for this stage. The family S33 serine proteases up-regulated in the cercaria may play roles in development rather than skin invasion.

Metalloproteases

Most of the metalloproteases which are differentially expressed during the germ ball to day 3 schistosomulum transitions are up-regulated in the germ ball. The leishmanolysin identified in cercarial secretions by Curwen *et al* [116] (actually an invadolysin), was shown to be highly up-regulated at this stage, in addition to five other invadolysins. The number of these and their high expression (ranging from 2-28 fold higher in the germ ball than the other stages) indicates that they are all important during skin invasion. This along with the high expression of the previously unreported elastase genes greatly expands the known repertoire of proteases

involved in infection by larval schistosomes. The other germ ball-enriched metalloproteases include proliferation-associated protein (Smp_150690.1) and several peptidases involved in protein turnover, highlighting the importance of cell proliferation and protein synthesis for the embryonic cercaria.

An invadolysin, an ADAM protease, and a matrix metalloprotease are up-regulated in the cercaria. The two latter enzymes can be involved in tissue remodelling so they may play a role in parasite morphogenesis on entry to the host. However, ADAMs are found in snake venom where they inhibit platelet aggregation; this would be a useful function for a secreted protein to perform during blood vessel entry. Likewise, breakdown of the extracellular matrix in the skin would facilitate host invasion by schistosomes. Localisation of these proteases would increase our understanding of their function and aid decisions about whether they represent targets for intervention.

In the day 3 schistosomulum, a membrane metalloprotease and another M12 family protease (Smp_134430) are up-regulated. The latter has high homology to tolloid which is a dorso-ventral patterning protein in *Drosophila*. Together with the up-regulation of various cadherins, these may be the involved in the remodelling that occurs in this life cycle stage. Also up-regulated in the day 3 schistosomulum is a gene encoding leucine aminopeptidase which is a cytosolic enzyme involved in protein turnover. This protein has been localised to gastrodermal cells in adult worms, and is hypothesised to break down the peptides resulting from haemoglobin degradation in the gut to single amino acids [152].

Cysteine proteases

The gene expression patterns in this category reveal much about the biology of larval schistosomes. Whilst the majority of cysteine proteases are most highly expressed in the cercaria or day 3 schistosomulum, the up-regulation of caspases and a separase is consistent with cell death and division during cercariogenesis. Bos *et al* recently described *S. mansoni* proteases, and carried out a phylogeny of caspases [153]. They noted that Smp_172010 was in a clade with effector caspases from mice and humans, and Smp_141270 was closer to initiator caspases [153]. Both of these genes are up-regulated in the germ ball. It is tempting to

speculate that apoptosis is involved in the bifurcation of the tail, after the posterior of the germ ball has flattened and extended.

The cercaria expresses a de-ubiquitinating protease; this enzyme may serve to help avoid proteasomal degradation of proteins, which is costly in terms of ATP. In addition, two developmental proteases are up-regulated in the cercaria: an ovarian tumour-like protein, and autophagin. It is likely that these are both needed for transformation. The presence of autophagin, a protease involved in autophagy [154], indicates that this process may be used to destroy the acetabular glands once they are no longer needed. Calpains are also up-regulated in the cercaria. They are calcium-dependent proteases with diverse functions including signal transduction [155]. It is plausible that they are involved in neuromuscular signalling during swimming.

The remaining cysteine proteases are cathepsins and asparaginyl endopeptidases. Some of these are included in the 'gut' category as they have been localised to that tissue, or identified in adult worm vomit. It is likely that the additional cathepsins also have a role in digestion. The known digestive enzymes will be discussed in the 'gut' category.

Aspartyl proteases

Compared to the cathepsins, the function of these aspartyl proteases is poorly characterised in schistosomes. This is the first report on their expression. All of the differentially regulated enzymes in this class follow a pattern of increasing expression from germ ball to day 3 schistosomulum. They have homology to the acidic digestive enzymes pepsin and cathepsin D. The cathepsin D identified by Wong *et al* [156] is Smp_013040, which is not differentially expressed during the mammalian infection process. However, knock down studies using RNAi showed that this enzyme was required for schistosomula to grow [157]. Based on the similarity of their expression patterns with those of known gut proteases, it is likely that the aspartyl proteases included here are also involved in gut function. However, localisation and biochemical studies would be needed to confirm this hypothesis.

3.4.9. Energy production

With the exception of lactate dehydrogenase (up-regulated in the schistosomulum), the genes encoding glycolytic enzymes did not vary greatly in the germ ball to day 3 schistosomulum transitions. However, transcripts encoding citric acid cycle enzymes were consistently up-regulated in the cercaria. This is consistent with the findings of Skelly *et al* who showed that cercariae use aerobic metabolism, whereas schistosomula rely on glycolysis to supply energy [66, 158]. The cercaria requires energy to swim, and as it does not feed, it relies on a store of glycogen. It is logical that the cercaria uses aerobic metabolism to yield the maximum possible molecules of ATP from the limited resources available to it.

3.4.10. Lipid metabolism

Genes involved in lipid metabolism up-regulated in the day 3 schistosomulum include phospholipases; it is not clear in this context whether they are acting as digestive enzymes, or to generate signalling molecules. It is possible that both functions are undertaken, as both signalling and digestion are important for the day 3 schistosomulum as illustrated by the up-regulation of other gut-associated enzymes and neurotransmitter transporters. CD36 is also up-regulated in the day 3 schistosomulum. It is thought to scavenge oxidised low density lipoprotein and long chain fatty acids, and is also an adhesion molecule; in humans it is expressed on monocytes, macrophages, platelets, microvascular endothelial cells and adipose tissue. It suggests that the schistosomulum takes up lipid as well glucose from the host at this early larval stage. Verjovsky-Almeida *et al* included CD36 in their table of possible vaccine candidates [73]. Berriman *et al* noted that schistosomes must rely on the host for a source of inositol [75]. The current microarray data show that a putative inositol transporter (Smp_134080) is up-regulated in the day 3 larva compared to the germ ball. This should be investigated as an intervention target. The gene with the highest fold change at day 3 is lecithin:cholesterol acyltransferase. This is a secreted enzyme which metabolises plasma lipoproteins. Taken together these results imply that schistosomula are capable of ingesting and digesting plasma.

3.4.11. Membrane proteins

This category contains a wide range of proteins with various functions. Fitzpatrick *et al* note that GPCRs and ion channels are the target of a large proportion of drugs currently in use, and for that reason should be investigated thoroughly in schistosomes [99]. Efforts should be

concentrated on those which are up-regulated in the intra-mammalian stages. An important task is to narrow down the large repertoire of channels and receptors encoded in the *S. mansoni* genome, to those highly expressed in intra-mammalian stages. Therefore special attention will be given to the subset of membrane proteins enriched in the schistosomulum compared to the other two stages studied. Many of the membrane proteins have been identified as potential intervention targets. Those included in the 'tegument' category will be discussed later (section 3.4.12 below).

The increase in transcription of neurotransmitter-sodium symporters in the cercaria probably reflects the muscle activity required for swimming. Likewise, channels of the TWIK family (Tandem of P domains in a weakly inward rectifying K^+ channel) are enriched in the cercaria. TWIK channels are reported to be involved in maintaining the background membrane potential [159]. These proteins are also thought to be non-inactivating and so they may play an important role in maintaining the membrane potential in neuromuscular tissue in the cercarial tail enabling swimming. In contrast, voltage-gated potassium channels are up-regulated in the schistosomulum.

Only a small proportion of the predicted GPCRs are differentially regulated during the transition studied here. Although the GPCR up-regulated in the germ ball was assigned no specific annotation in the genome, BLASTp against NCBI shows highest similarity to opsin from *Branchiostoma belcheri* (37% identitiy, 52% positive), and to retinal pigment derived rhodopsin from house mouse (31% identity, 48% positive). It was highlighted in the transcriptome by Verjovsky-Almeida *et al* [73]. Although no eye spot-like morphological structure has yet been identified in the cercaria, this life cycle stage is known to be responsive to light. The opsin-like receptor, along with the cercaria-enriched photoreceptor cell development protein, is strongly indicative of pathways required for light detection.

Stomatin or erythrocyte band 7 protein was suggested by Verjovsky-Almeida *et al* as a gene that should be investigated as a vaccine candidate. It has been ascribed many roles including mechanoreception, HDL binding, phospholipid uptake, calcium transport, and interaction with antimalarial drugs [73]. Although its site of expression is not known, the current results reveal that it is expressed in the intra-mammalian larva, and as such shows promise as an intervention target.

3.4.12. Tegument

The dramatic change in tegument structure on entry to the mammalian host is one of the defining aspects of transformation. The genes included in this category are derived from proteomic studies and other published reports in which the gene product is localised to the tegument. Skelly and Shoemaker showed that the glucose transporter SGTP4 appears at the surface one hour into transformation induced by incubation in RPMI-rich media [60]. Thus the newly parasitic schistosomulum is able to acquire glucose from the host. This is important as the cercarial glycogen store is depleted after 24 hours *in vivo* [46]. In another study, Skelly and Shoemaker investigated the expression of SGTP4 using western blots: protein was not detectable in eggs, sporocysts, or cercariae, but was found in adult worms [58]. Furthermore, SGTP4 was not detected in protein extracted from schistosomula immediately after transformation, but was clearly visible 1 hour afterwards. In the present study, transcription of SGTP4 was lowest in the germ ball, increased in the cercaria, and reached a peak in the day 3 schistosomulum. The up-regulation of SGTP4 in the cercaria indicates that the transcript must be present and ready, allowing the schistosomulum to translate and traffic it to the surface immediately on entry to the host. As this protein has been shown on the apical surface of both schistosomula and adult worms, it represents a potential target for intervention. This would require either the design of a suitable inhibitor that binds with high specificity for the parasite channel, or an immune response to be elicited to an epitope not shared by the host homologue. Although these goals are not trivial, they may be possible, and would provide a treatment with potential against both larvae and adults.

The other known tegument surface proteins, aquaporin 3 [160], and nucleotide pyrophosphatase [161], share SGTP4's expression pattern. The aquaporin has been characterized recently by Skelly *et al* [160]. In addition to controlling osmoregulation, it has been ascribed an important role in allowing drugs to enter the parasite [160]. The nucleotide pyrophosphatase/ phosphodiesterase has been identified in tegument preparations by proteomics [107, 162, 163]. It has been demonstrated at the surface of adult worms by Rofatto *et al* using imunohistochemistry [161]. Western blots revealed that the protein was present in 7 day schistosomula. The authors established that the phosphodiesterase is active at the surface of live schistosomes and that the activity could be reduced with anti-sera raised against it [161].

CD59 is a complement inhibitory protein expressed on the surface of human cells [164]. It blocks incorporation of C9 into the C5b-9 complex, so that the membrane pore does not form, and cell lysis is avoided [165]. CD59 proteins have been discovered at the surface of adult schistosomes by proteomics (Braschi and Castro Borges *et al* unpublished). The current study shows that five genes encoding schistosome CD59 are up-regulated on entry to the mammalian host, indicating that this important defense strategy is employed throughout the intra-mammalian life. However, the most highly up-regulated CD59-encoding transcript has not been found by proteomics; this is the first evidence for its expression. It is possible that this is specific to the schistosomulum.

Tegument proteins should also be investigated as targets as they are accessible to the host, unlike cytosolic proteins [166]. It may be easier to achieve selective toxicity for schistosome specific tegument proteins rather than those with human orthologues, but these will need to be identified and localized first. Braschi *et al* used proteomic techniques to investigate the adult worm tegument and identified several *Schistosoma* specific proteins in the membrane fraction of the preparation [140]. One of these, Smp_074450, shares the expression pattern of SGTP4, implying that it may be present at the surface shortly after transformation, and remain throughout the intra-mammalian life. As such, this protein should be investigated as a vaccine target. Tran *et al* reported that two tetraspanins, found on the adult tegument, showed promising results when mice were vaccinated with the recombinant protein [167]. The genes encoding these particular proteins were not up-regulated in the schistosomulum in the current study. However, ten different tetraspanins were, and only two of these had previously been found on the surface on adult worms. Localisation studies would be required to discover whether the eight remaining schistosomulum-enriched tetraspanins are novel tegument proteins. The two day 3 schistosomulum-enriched tetraspanins that have been found in the adult tegument are worth investigating as they may be expressed on the surface of the worm throughout the intra-mammalian life, and as such may represent new vaccine targets.

Bhardwaj and Skelly propose that the expression of alkaline phosphatase and nucleotide pyrophosphatase at the tegument modulate the immune response by catabolizing free ATP which is a potent inflammatory mediator [168]. Taken with the expression of the anti-complement protein CD59 on the surface of adult worms (Castro Borges *et al*, unpublished),

this shows that the tegument has mechanisms for active defence; it is not simply an inert barrier.

Three of the genes included in the tegument category were shown to be up-regulated in the germ ball (a voltage-dependent anion-selective channel, an amino acid transporter, and acetylcholinesterase). This raises doubts regarding their expression in the intra-mammalian stages. As such these should not be considered as intervention targets.

3.4.13. Micro Exon Genes

MEGs are highly differentially regulated during the germ ball to day 3 schistosomulum transitions. The majority are up-regulated in the day 3 schistosomulum compared to the other two stages. Visual inspection of the gene models for some of the unannotated genes highly expressed at each stage enabled the identification of four new MEGs in addition to the 14 families reported in the genome paper. These represent new MEG families as they have no homology either to each other, or to the other 14 MEG families. These genes demonstrate considerable alternative splicing and exon skipping. Another important point to note is that they are almost all predicted to have signal peptides, but no transmembrane domains, so they are likely to be secreted [142]. No mechanism for antigenic variation had been found in schistosomes when the transcriptome was published [73]. The MEGs may represent a schistosome mechanism for protein variation [142]. Antigenic variation is an important feature of how other blood borne parasites survive (*Plasmodium* [169] and *Trypanosoma* [170]). In order to discover whether MEGs play a role in presenting variant proteins to the host, it will be necessary to discover whether these genes are expressed at the host parasite interface – either by secretion from the gut or glands, or expression on the tegument surface. Further study into the expression of proteins encoded by MEGs may shed light on the way that schistosomes are able to survive in the blood stream for decades. Arguably the mechanisms employed by schistosomes for immune evasion must be more effective than those used by the shorter lived unicellular intravascular parasites as schistosomes live in the bloodstream for up to 30 years.

3.4.14. Venom Allergen Like proteins

Another interesting group of genes that have been described recently are the VALs. These are not schistosome specific and have been implicated in angiogenesis [123] and skin invasion by parasitic nematodes [171]. Curwen *et al* identified VALs 4, 10 and 18 in cercarial secretions

by proteomics [116], the data here show that the transcripts for VAL 18 and VALs 1, 2, 19, 20, 21, 24, and 25 are all up-regulated in the germ ball. Chalmers *et al* investigated the expression of VALs 1-13 across various lifecycle stages using RT PCR and showed that VALs 1, 4, and 10 are up-regulated in the cercaria [125]. In addition to VAL 1, the present study showed VALs 2, 16, 17, 21 and 25 to be enriched at this stage compared to the day 3 schistosomulum.

Only two VALs were most highly expressed in the day 3, which is in contrast to the dominant expression pattern for MEGs. This suggests that these proteins play a role earlier in infection. Since other VALs are highly expressed in the intramolluscan stages, it may be the case that these proteins are important for germ balls and cercariae while they are in the intermediate host. However, expression of VAL 7 in the schistosomulum is more than 45 fold higher than that in the germ ball. A VAL in *Onchocerca* is reported to induce angiogenesis [123]. If SmVAL 7 protein were to interact with blood vessels to cause vasodilatation it would facilitate skin exit. The current study confirms that this family of proteins warrants further attention. In particular, VAL 7 should be prioritised as it is the most highly expressed VAL gene in the early intra-mammalian larva.

3.4.15. Stress

Only 21% of stress-related genes were differentially regulated in the contrasts studied here. This is somewhat surprising as the parasite must adjust to wildly different environments during this portion of the life cycle. Of those that were differentially expressed, none were enriched in the germ ball. This may be because the germ ball is shielded from the host tissue by the daughter sporocyst. Certain heat shock proteins are up-regulated in the cercaria. These proteins are chaperones which help keep proteins properly folded. Sm-p40 is noteworthy as Mathieson discovered that it is up-regulated in the miracidium [172]. He reported that this small heat shock protein acts to prevent irreversible aggregation of denatured proteins without using up precious ATP [172]. It is logical that the non-feeding cercaria expresses this gene as it too must conserve as much energy as possible. A thioredoxin peroxidise was up-regulated in the schistosomulum, suggesting involvement in defence against hydrogen peroxide. This gene was cloned and characterised by Kwatia *et al* in 2000, who found that the protein was active in adults [173]. This important defence protein is very likely to be expressed throughout the

intra-mammalian life-cycle stages, and should be investigated as an intervention target. Indeed, a new group of drug leads has recently been identified that inhibit thioredoxin glutathione reductase (Smp_048430) [174], highlighting the importance of investigating the genes in this category.

3.4.16. Gut

Proteins involved in the digestion of the blood meal have been studied extensively, as they are appealing targets for interventions (reviewed in [175-177]). Sm31 and Sm32 were the first to be discovered [178]. The former is cathepsin B, the latter was first named haemoglobinase. However, Dalton and Brindley noted that Sm32 has homology to legumain, an asparaginyl protease from legumes and suggested that it may activate cathepsins in the worm gut [179]. Sm32 has since been shown to self-activate at acidic pH [180] and to activate cathepsin B (Sm31) [181]. Given the large number of cathepsins in the schistosome gut, Dalton and Brindley suggest that it may be wiser to target Sm32; it is known to activate at least one of the cathepsins, and so it may be possible to target blood feeding by inhibiting that one enzyme. Suppression of cathepsin B by RNAi did not inhibit haematin formation, but led to stunting of treated worms [77].

In addition to the cathepsins, saposins are included in the ‘gut’ category. Sapsosins have been suggested to lyse blood cells on entry to the gut [182], and although Smp_105450 was localised to the gut of adult schistosomes, it was ineffective as a vaccine in the mouse model [182].

There has been some discussion in the literature as to when the larval gut becomes active. Crabtree states that lung schistosomula are unlikely to feed on plasma or cells during migration as their mouths are pressed up against blood vessel wall [63] and that the larval oesophagus is too narrow to admit red blood cells. Protein staining was observed in the larval gut in Chapter 2. The results presented in this chapter show that transcripts encoding many cathepsins are already up-regulated in the cercaria, with expression increased further at Day 3. This is consistent with the immunolocalisation of Sm31 and Sm32 to the cercarial gut reported by Skelly [183], and suggests that gut-specific proteases are translated and secreted early in

infection. In addition, Dillon *et al* localised cathepsin L mRNA to the gut of lung stage schistosomula [16]. Nevertheless, the question of what the larva ingests at this stage remains unanswered. The finding that there are important digestive proteases up-regulated throughout the intra-mammalian stages could mean that an intervention which targeted them would be effective against both larval and adult worms. This would be a useful feature of any new drug against schistosomes. The current treatment is only effective against adults; any migratory larvae present at the time of treatment mature unless follow-up treatment is administered.

4. Spatial Expression Patterns of Selected Genes

4.1. Introduction

The availability of genomic sequence data for *S. mansoni* has enabled transcriptomic and proteomic experiments to be carried out. Various approaches have been taken, some focusing on the differences between life cycle stages, others on the proteins expressed in certain tissues. Of particular relevance to the current study are those which tackle the infection process. Throughout all of the sequencing programmes for schistosomes, a common theme has been the large proportion of *S. mansoni* genes with no significant matches to any sequences available at GenBank. Indeed, Wilson *et al* noted that as few as 26% of *S. mansoni* sequences may have informative homology [128]. In order to assign functional annotation to the remaining 75%, an important early step is localisation of the gene product. Forming hypotheses regarding the function of a gene product is greatly aided by knowledge of where in the organism the gene is expressed. Localisation is also necessary when considering whether the gene under investigation could serve as a vaccine candidate, as such gene products must be available to the host. As several workers point out [129-131], cytosolic and cytoskeletal proteins are extremely unlikely to be of use as intervention targets and should be avoided in future. Localisation studies reveal whether a potential target is expressed at the host-parasite interface (tegument, gut, secretory glands, or excretory system); this enables researchers to disregard those which are found to be expressed internally with little likelihood of secretion.

A great step forward was achieved when Dillon *et al* adapted whole mount *in situ* hybridisation (WISH) for use in schistosomes, such that transcripts of interest could be localised in intra-mammalian larvae and adult worms [16]. The whole mount technique is particularly useful as small cells or tissues may be missed in the sectioning process for conventional *in situ* hybridisation (ISH) [16]. In addition, the life cycle stages under investigation in the present study are very small, making sectioning technically difficult and interpretation of images problematic. The method described by Dillon *et al* utilizes proteinase K to permeabilise the worms to allow the DIG-labelled RNA probes and detection antibodies to enter the parasite and reach their target mRNA. This treatment proved too harsh for germ balls to withstand. However, Dr Jean Illes of Sheffield University kindly provided an alternative protocol that uses methanol for permeabilisation [184]. Thus the first successful application of WISH to germ balls is described in this chapter.

I chose to localize transcripts encoding proteins previously identified in cercarial secretions and hypothesized by Curwen *et al* to play a role in infection, either by proteolysis or immunomodulation [116]. Of these, only one cercarial elastase had previously been localized. Newport *et al* carried out ISH on sections of infected snail hepatopancreas and showed that transcripts encoding cercarial elastase 1a (Smp_119130) were present in acetabular glands of intra-molluscan germ balls [114]. Thus Smp_119130 was used as a positive control in the current study. Expression of Curwen's SmPepM8 (invadolysin, Smp_090100), SCPc (VAL 10, Smp_002060), SmKK7 (Smp_194830) and Sm16 (Smp_113760) was localized in germ balls. Anti-sera for SmKK7 and invadolysin were also used to localise these proteins by immunohistochemistry.

As well as being identified in cercarial secretions, SmKK7 was discovered by mass spectrometry in the soluble fraction of the supernatant resulting from treatment of adult worms with trypsin and PIPLC, suggesting that it may be associated with the tegument (Dr W. C. Borges, personal communication). This led to the investigation of SmKK7 expression (both transcript and protein) in adult worms in addition to germ balls, cercariae, and schistosomula. This was carried out as adults are larger and better described, so it may be easier to determine which tissue SmKK7 was expressed in at that stage.

Sm16 was first identified by Ramaswamy *et al* in secretions from schistosomula, and was reported to have immunomodulatory properties [185]. Sm16 was cloned and shown not to be stage-specific [120]. More recently Sm16 protein has been expressed and its effect on human cells investigated [186, 187]. Transcript encoding this protein has been localized in the germ ball for the first time.

In addition, the microarray experiment described in the previous chapter revealed that MEGs are the most highly up-regulated genes in the early intra-mammalian larva. This coupled with the observation that MEGs encode highly variable secreted proteins with no homology to known proteins from any other organism, made them perfect targets for WISH. Three MEGS were chosen for localisation studies: MEG 4.2 (Smp_085840.2) was most highly expressed in the cercaria, whilst MEGs 3 (Smp_138070) and 14 (Smp_124000) were up-regulated in the D3 schistosomulum. MEG 3 protein was first identified in the 3hr to 3 day secretions of *in*

vitro cultured schistosomula [142]. These MEGs were localised to discover whether they are expressed at the host-parasite interface.

Aims:

- Confirm localisation of cercarial elastase transcript to acetabular glands using methanol protocol for WISH in germ balls
- Use WISH to localize chosen target transcripts in appropriate life cycle stages
- Interpret images and discuss implications for existing hypotheses regarding the function of the gene products under investigation.

4.2. Methods:

4.2.1. Biological material

Larval parasites were obtained as described in Chapter 2. In addition, day 10 schistosomula were obtained by transforming cercariae and were cultured using the same method as day 3 schistosomula, but for a further seven days. Adult worms were obtained by portal perfusion of mice up to seven weeks post-infection with 180 cercariae. Subsequently cercariae and schistosomula were fixed in Carnoy's fixative (Ethanol, chloroform, glacial acetic acid 6:3:1) for two hours on ice with shaking, followed by two five minute washes in 100% EtOH. Following this they were fixed for 1 hour in MEMFA (0.1M MOPS, 2mM EGTA, 1mMMgSO_4, 3.7% formaldehyde in water) at room temperature, then washed and stored in 100% ethanol at -20°C until use.

Germ balls were fixed in 4% paraformaldehyde at 4°C overnight with shaking, then washed twice for 5 minutes in PBS and stored in PBS at 4°C. To permeabilise the worms were dehydrated through 10 minute washes in a graded series of PBSAT plus 25%, 50%, and 75% MeOH at room temperature, before two washes in 100% MeOH. Parasites were stored in 100% MeOH at 20°C until use, or for at least 30 minutes if used immediately.

4.2.2. PCR

RNA was extracted from germ balls as described above. A 2µg sample of RNA was reverse transcribed in a 20µl reaction using Omniscript RT (Qiagen) with an oligo dT primer (Promega). Genes chosen for WISH were amplified by PCR using the primers in Table 4.1 and germ ball cDNA. Primers were designed using sequences available at www.geneDB.org. VAL 10 was subcloned from a clone kindly provided by Dr S Prasad of the Department of Molecular Medicine at Cape Western Reserve University, Cleveland, Ohio, USA, using the pENTR primer pair. SmKK7 was cloned by Ms Jenny Middleton.

Primer	Gene name and ID	Sequence
Sm09202 3F	Cercarial elastase 1a	CTG TCA TCG CAT TCT TAA CGA C
Sm09202 3R	Smp_119130	CGT TAT CAT CCC TTC CAT AAC C
Smp_090100 1 F	Invadolysin (SmPepM8)	ATG ATA CCC TGT TCA AGA AAT CTC TT
Smp_090100 1 R	Smp_090100	CCT ATC AGT TGT AGG ATG CAT TTC
pENTR 1 F	VAL 10 (for subcloning)	GCC TTG TTT AAC TTT AAG AAG GAG C
pENTR 1 R	Smp_002060	ATA ATG ACT TTG TAC AAG AAA GC
Sm16 1 F	Sm16	ATG ACA TTG ATC ACA GCT ACA ACG
Sm16 1 R	Smp_113760	CAT CAT CTT ATC CAG TTT CTT CGC
KK7 F	SmKK7	GAA TTC AAA CCT GGC CGA GTC AAG TGC AGCG
KK7 R	Smp_194830	TCT AGA TCA TGC AAT TTA TGT TCA TCA TAG G
Smp_124000 F	MEG 14	ATG AAT AGG TTC TTT TGG ACT GTC A
Smp_124000 R	Smp_124000	TAC GAT AGG GAC AGC CGC
Smp_085840.2 F	MEG 4.2	ATG AAT TTC TTG ACA CTT TAC GTA ACT
Smp_085840.2 R	Smp_085840.2	TGA AGT AAT ATG ATA TAG CTC TTG GAA
Sm12949 F	MEG 3.2	TTAATTATAGTTAACAAACAGCCAAGA
Sm12949 R	Smp_138070	TCGACTGTGTATTCACAGCTCG

Table 4.1 Primers used to amplify gene portions for WISH

4.2.3. Cloning

PCR products were ligated into PGEM-T easy (Promega) according to the manufacturer's instructions. The reaction was allowed to proceed for 1hr at room temperature followed by overnight incubation at 4°C. A 2µl aliquot of the resulting ligation was transformed into chemically competent DH5α *E. coli* (Invitrogen). The bacteria were spread onto LB agar plates with 0.1mg/ml Ampicillin and 40µg/ml X-gal for blue/white selection. White colonies were picked and grown in 10ml LB cultures with 40µg/ml ampicillin overnight at 37°C with shaking. Plasmid DNA was extracted using Qiagen miniprep kit according to the manufacturer's instructions. Plasmid was sequenced by the in house Technology Facility using

T7 primer in order to check the quality and direction of the insert. Sequences were checked by BLAST analysis against the *S. mansoni* genome at www.genedb.org using the default settings.

4.2.4. Probe synthesis

DIG-labelled probes were synthesised using the method described by Dillon *et al* [16]. Briefly, plasmids containing schistosome gene portions were linearised, by cutting at the 5' end of the insert with an appropriate restriction enzyme (NcoI or NdeI, New England Biolabs). *In vitro* transcription was carried out using T7 or S6 polymerase (Promega) incorporating DIG-labelled dUTP (Roche). Probes longer than 600bp were hydrolysed to yield probes between 250 and 600bp as described by Dillon *et al* [16].

4.2.5. Whole mount in situ *hybridisation*

Cercariae, schistosomula and adults were treated according to the method described by Dillon *et al* [16]. The more delicate germ balls do not withstand proteinase K treatment, and so they were permeablised in methanol according a protocol kindly supplied by Dr J Illes of Sheffield University. Germ balls were rehydrated through 10 minute washes at room temperature in a graded series 100%, 75%, 50%, and 25% MeOH in phosphate buffered saline with 0.1% Tween 20 (PBSAT), followed by two washes in PBSAT. Following rehydration, germ balls were incubated in pre-hybridisation solution (50% formamide, 5X SSC (pH7), 2% blocking powder (Boeringer Mannheim),1% Triton X-100, 0.5% CHAPS, 100µg Yeast RNA, 50µM EDTA, and 50µg/ml heparin) for 1 hour at 65°C. After this step the prehybrisation buffer was replaced with fresh hybridisation solution containing 2µl/ml digoxygenin (DIG)-labelled RNA probe and the germ balls were incubated overnight at 65°C with constant rocking.

Hybridisation solution was removed and germ balls were washed twice 30 minutes each in pre-warmed Solution 1 (50% formamide, 5X SSC pH 4.5, 1%SDS) at 65°C, followed by two 30 minute washes in Solution 2 (50% formamide, 2X SSC pH4.5, 1% tween 20) at 65°C. Next the germ balls were washed three times in tris-buffered saline and tween (TBST; 0.14M NaCl. 2.7mM KCl, 25mM Tris HCl pH7.5, 0.1% tween) for five minutes each at room temperature, this step was followed by a blocking step: 90 minute incubation at room temperature in TBST with 10% heat inactivated sheep serum. Germ balls were incubated in fresh blocking solution containing alkaline phosphatase-conjugated anti-DIG FAb fragments (Roche) (1:2000

dilution) overnight at 4°C. The following morning, antibody solution was removed and the germ balls were subjected to three 5 minute washes, and four 1 hour washes in TBST at room temperature. A final TBST wash was applied overnight at 4°C. Colour was developed by addition of BM purple (Roche) or Fast Red fluorescent substrate (Sigma). When signal had developed sufficiently, germ balls were washed twice in PBSAT and stored in 10% formalin in PBSAT. Negative controls were always included, using DIG-labelled sense probe for the *S. mansoni* chorion gene. This probe was the kind gift of Dr GP Dillon, University of York. When Fast Red substrate was used, the worms were counterstained with DAPI as in Chapter 2. Parasites were imaged using a Leica DM2500 microscope attached to an 18.2 colour mosaic camera (Diagnostic Instruments) with SPOT Advanced software (SPOT Imaging Solutions). Confocal microscopy was carried out using a Zeiss LSM 510 meta on an Axiovert 200M with a 543nm laser and a 560-615nm filter. To visualise DAPI, a 405nm laser was used with a 420-480nm band pass filter.

4.2.6. Immunohistochemistry (IHC)

Antisera for SmKK7 and Leishmanolysin were a kind gift of Dr S Prasad at Cape Western University, Ohio, USA. IHC was carried out on adult worms according to the protocol described by Mair *et al* [32]. Larvae were fixed overnight in 4% paraformaldehyde at 4°C with shaking. The following morning they were washed twice for at least 5 minutes in PBS and stored at 4°C until use. All of the following steps were carried out at 4°C with shaking. Bodies were washed three times in permeablising fluid (PBS, 1% triton x-100, 0.1% SDS, 10% naïve sera from secondary Ab species, 0.1% NaN_3), 30 minutes. This fluid was replaced with Antibody Diluent (AbD) PBS with 0.5 % Triton X-100, 0.1%BSA, 10% naïve sera from secondary Ab species, and 0.1% NaN_3 and 1:200 antiserum to recombinant antigen. Parasites were incubated for 4 days at 4°C with shaking. Non – specifically bound antisera were removed by 3 washes in AbD over 24hrs at 4°C with shaking. Secondary goat anti-rat antibody was added at 1:200 in AbD and incubated for 48 hours. In adults, the secondary antibody was conjugated to Alexafluor 488, but for larvae the fluorophore was Alexafluor 647 (Invitrogen). Excess secondary antibody was removed by 3 – 4 washes in AbD over 24 hours. Larvae were counterstained with Alexafluor 488-conjugated phalliodin by incubating for 2-3 days in AbD. After two brief washes in AbD, larvae were mounted on slides and viewed with a confocal microscope. Lasers used were 488nm and 633nm with long pass filters of 505 and 650nm, respectively.

4.3. Results

4.3.1. Cercarial Elastase

Newport's elastase 1a (Smp_119130) was used as a positive control for germ ball WISH as it had been shown previously in the acetabular glands of embryonic cercariae [114]. This is the first time that transcript encoding this important protease has been shown in whole mount germ balls. Figure 4.1 shows germ balls of various stages of development. Only the young round germ ball did not stain positive for elastase 1a mRNA (Fig.4.1A). Figure 4.1B shows positive staining in a young germ ball, approximately 110μm in length with a flattened posterior end; elastase 1a transcript is clearly discernable in the cytoplasm of acetabular gland cells towards the anterior of the body (Fig. 4.1B). In a slightly older germ ball, with an extended posterior, elastase 1a transcript is present in several cells – some are more posterior and some more anteriorly located (Fig. 4.1C). The early 'stubby-tailed' stage (Fig.4.1D) exhibits the strongest staining. Here transcript is seen in a band across the middle of the body. The same pattern is seen in a ~100μm long germ ball at the 'elongating-tail' stage (Fig.4.1F). By the time the germ ball is approaching maturity (body ~125μm long, Fig.4.1E), staining is less bright and can be seen in a narrow band across the body at the anterior of the acetabular gland fundi.

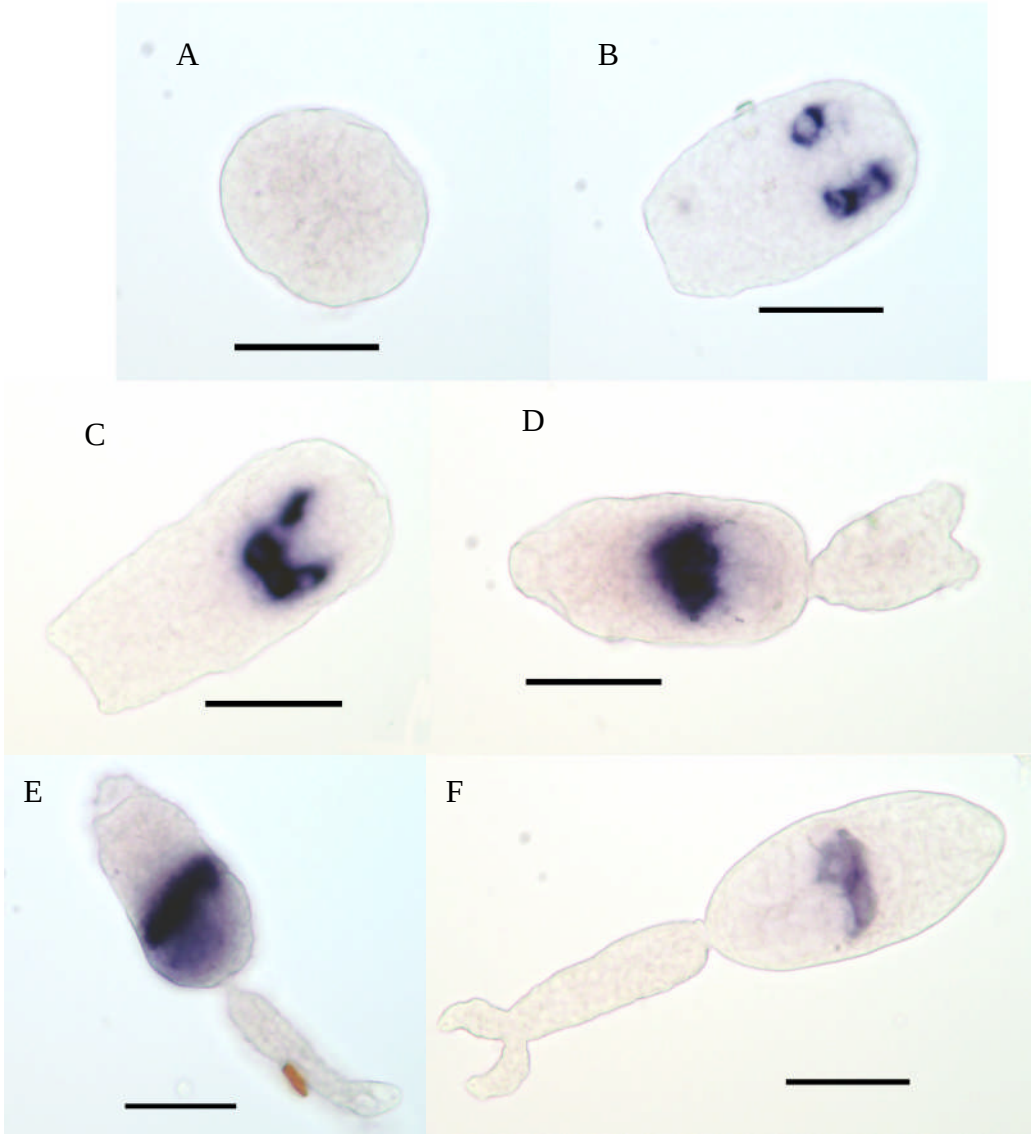


Figure 4-1 Transcript for cercarial elastase 1a localised in germ balls

Staining can be seen clearly in the acetabular glands of developing germ balls. It is not transcribed in young round germ balls (A) It is first visible in young germ balls with a flattened posterior (B), and is expressed continually until the germ ball is almost mature (C, D, E, F). Scale bars 50μm.

4.3.2. Invadolysin

Invadolysin (Smp_090100) was localized in germ balls. Figure 4.2 shows germ balls at various stages of development from a small round germ ball (Fig.4.2A), one at the 'flattened-posterior' stage (Fig.4.2B), a stubby tailed specimen (Fig.4.2C), and two with elongating tails (Fig.4.2D and E). Positive staining is only seen in the most mature germ ball (Fig.4.2F). Invadolysin transcript is detected at the anterior and lateral periphery of the acetabular gland cells, which are well developed by this stage.

Invadolysin protein was visualized in the cercaria by IHC. Figure 4.3 shows two optical sections of a cercaria showing presence of invadolysin in the acetabular gland fundi (Fig. 4.3A and B) and ducts, extending into the head capsule (Fig. 4.3A). It is unclear whether expression of this protein is limited to the pre-acetabular glands, or whether the staining visible in Fig. 4.3B is in the post-acetabular glands.

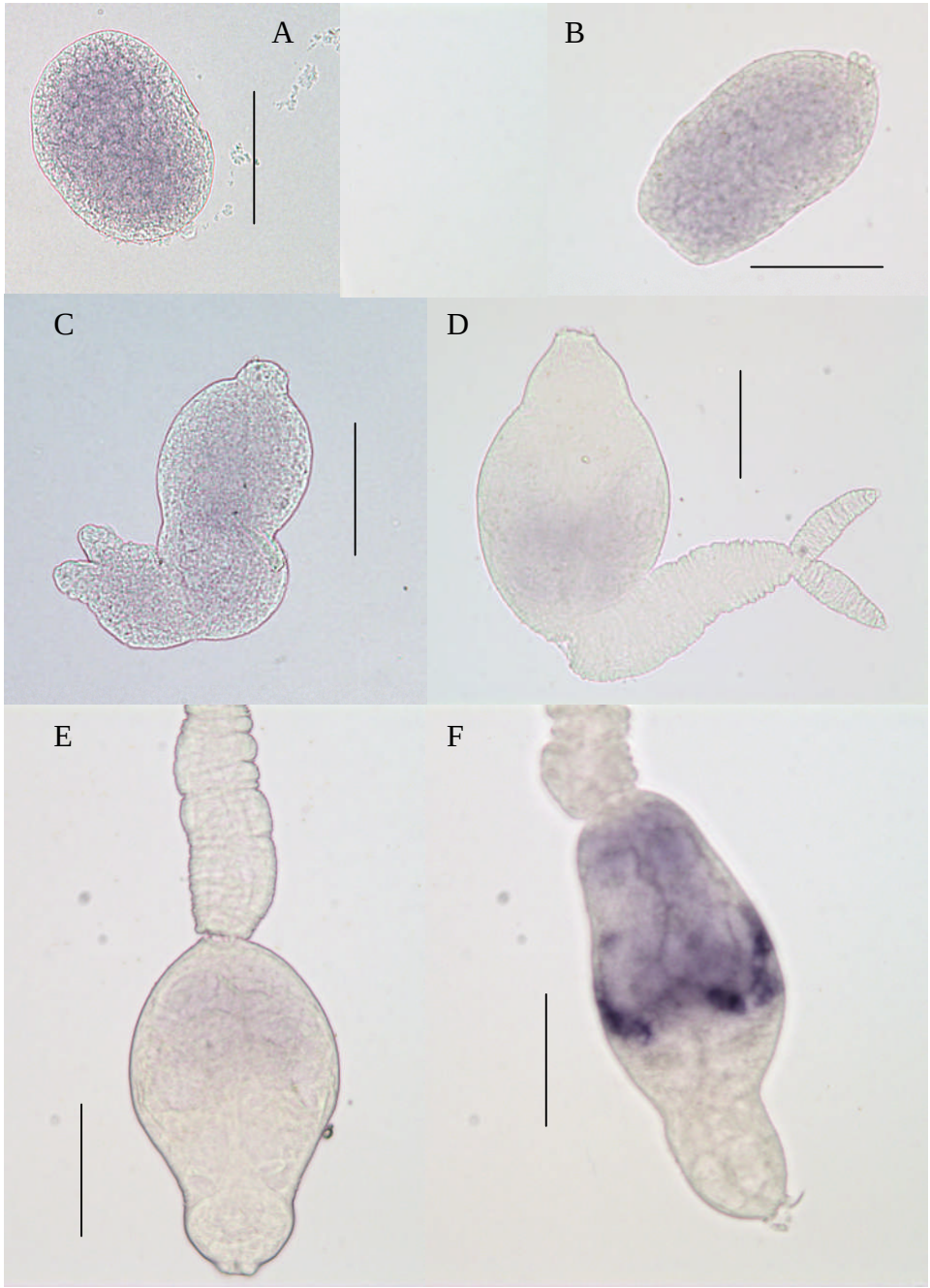


Figure 4-2 Invadolysin transcript localized in germ balls by WISH.
Transcript is not visible in young round (A), 'flattened-posterior' (B), 'stubby-tailed' (C), or young 'elongating-tail' (D and E) germ balls. A germ ball nearing maturity (F) stains positive for invadolysin mRNA at the periphery of the acetabular gland fundi. Scale bar 50μm

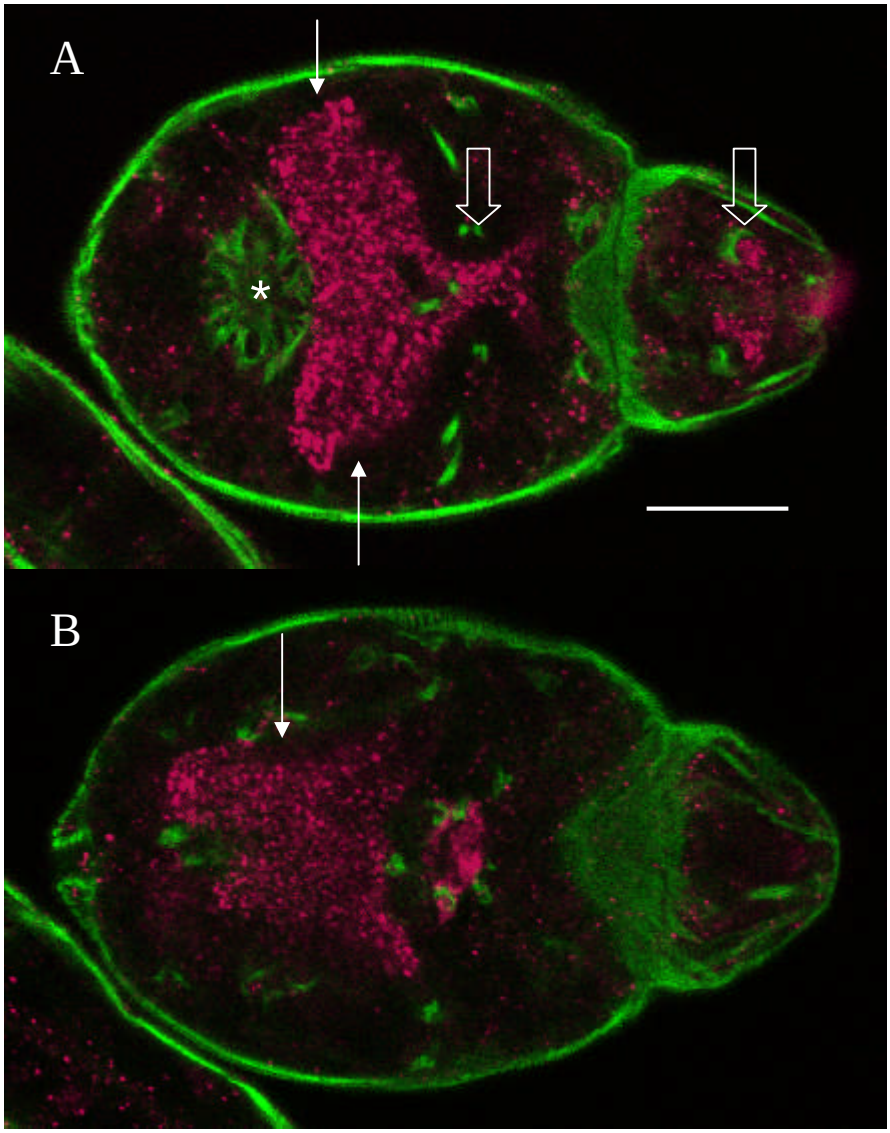


Figure 4-3 Invadolysin protein localizes to the acetabular glands

Selected optical sections of a cercaria after immunohistochemistry for detection of invadolysin (Smp_090100). Note acetabulum (asterisk), invadolysin protein immunolocalised to the acetabular glands (arrows) and their ducts (block arrows). A) section 10.47μm from the ventral surface B) 17.47μm from the ventral surface. Scale bar 20μm.

4.3.3. VAL-10

Transcript encoding VAL-10 was localised in germ balls. Expression of this gene was not apparent in small round (Fig. 4.4A) or 'stubby-tailed' germ balls (Fig. 4.4B and C). Staining can be seen only at the elongating-tail stage (Fig. 4.4 D) and in nearly mature germ balls (Fig. 4.4 E and F). The expression pattern of VAL-10 is similar to that of invadolysin described

above. It is visible in a band across the middle of the body along the anterior and lateral edges of the acetabular gland cell fundi (Fig. 4.4D).

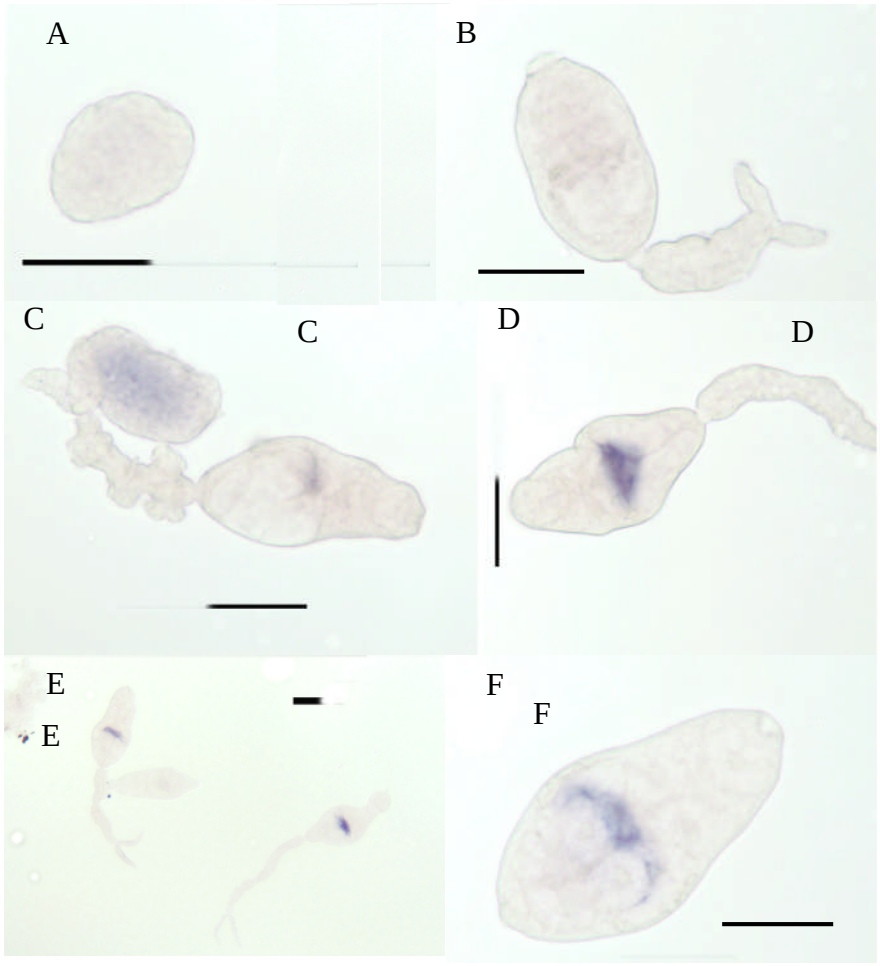


Figure 4-4 VAL 10 transcript localised in germ balls by WISH.

Transcript is present at the anterior of the acetabular glands. No staining is visible in younger germ balls (A, B, C); expression is restricted to germ balls nearing maturity (D, E, F). Scale bars 50μm

4.3.4. Sm16

In the present study Sm16 mRNA was localised in germ balls. A selection of images is shown in figure 4.5. In stark contrast to the circumscribed patterns reported above for elastase 1a, invadolysin and VAL-10, Sm16 is distributed throughout the body. It is first transcribed in germ balls with flattened posteriors (Fig. 4.5A). Images of stubby-tailed germ balls reveal that Sm16 expression is not limited to any particular tissue; staining is seen throughout the body including the tail (Fig. 4.5 B and C).However, there are patches that lack staining (Fig. 4.5 A, C and D). These correspond to the location of the acetabular gland fundi and ducts. This is

most clearly evident in the ‘elongating-tail’ stage (Fig. 4.5D), but is also apparent in the ‘flattened-posterior’ and ‘stubby-tailed’ stages (Fig. 4.4A and C).

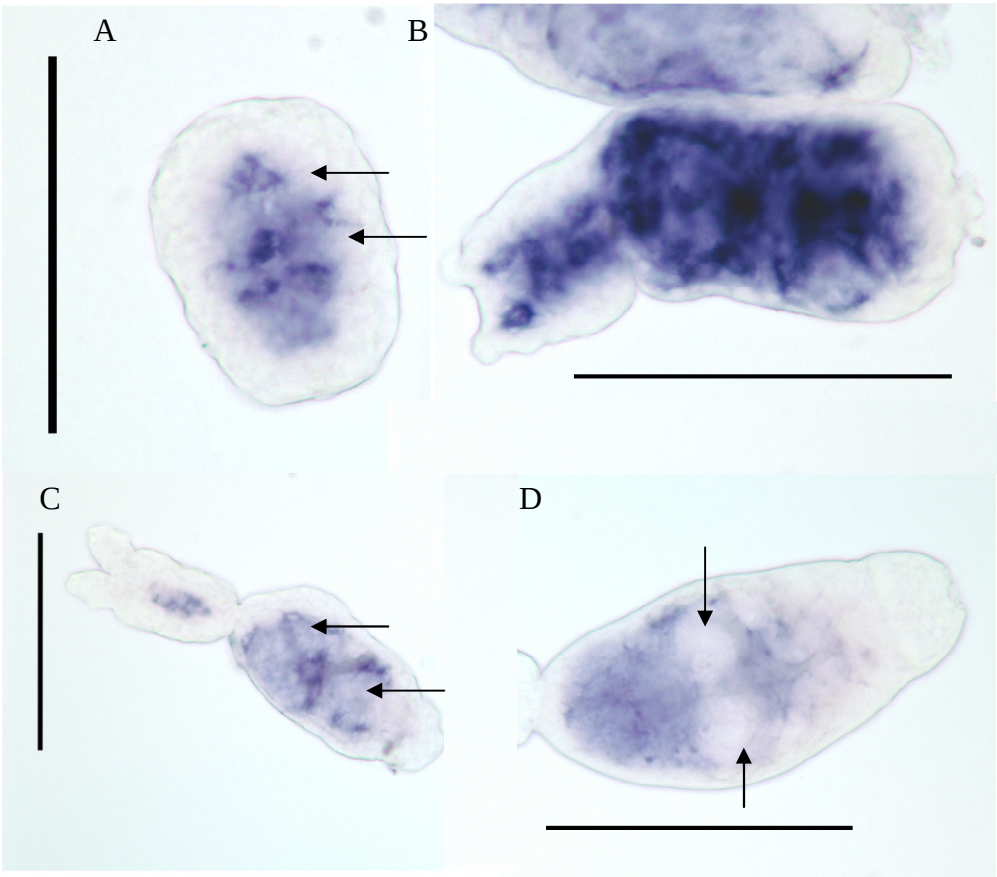


Figure 4-5 Expression of Sm16 was localised in germ balls by WISH.
Note the presence of transcript in the tail (B and C), and clear spots corresponding to the acetabular glands and ducts (A, C, D arrows). Scale bars 50μm

4.3.5. SmKK7 expression in larvae

SmKK7 transcript was localised in the germ ball, cercaria and schistosomulum (Fig. 4.6). In the cercaria, SmKK7 is transcribed in six cells in the body (Fig.4.6 B) and two towards the posterior of the tail, just anterior to the bifurcation (Fig. 4.5A). The same pattern is seen in the schistosomulum (Fig. 4.5C). The KK7-producing cells are arranged in pairs laterally; the anterior pair is within the head capsule, one pair is situated in the middle of the body, and the third pair is located towards the posterior of the body. WISH in germ balls revealed that the middle and posterior cells transcribe SmKK7 at the ‘elongating-tail’ stage (Fig. 4.6 D). This is followed by the appearance of the anterior (Fig. 4.6E) and tail pairs (Fig.4.6F) as the germ ball develops further.

Immunohistochemistry revealed the presence of KK7 in a network of fine processes within the head capsule and throughout the body (Fig. 4.7A), and on the external surface of the cercarial head capsule (Fig.4.7 B). There are also KK7-containing knobs at the very anterior of the cercaria, external to the phalloidin staining. Examination of the entire z stack shows these protuberances are ventral, and arranged in three pairs laterally (Fig.4.8). Figure 4.9 shows a 3D projection of a cercaria stained for SmKK7 protein (A), with an electron micrograph of the anterior end of a cercaria kindly provided by Stephanie Hopkins and Prof Jim McKerrow, University of California, San Fransisco (C). The 'buckyball'-like protrusions are the contents of the acetabular glands exiting the ducts. Figure 4.9 B is a diagrammatic representation of the sensory endings, illustrating the three different types observed. The nerve endings are arranged in pairs laterally. The ventral-most pair has a raised base that is wider in diameter than the thin cilium that protrudes from it. The next pair also has a raised base, but the cilium is wider. The most striking pair is the middle pair. The base is comprised of two concentric, raised tegumentary folds; the cilium is thin and much longer than the others (~3.5μm). Dorsal to the pair of sensory endings with long cilia is another pair with a raised base and a short thin cilium. Finally, the most dorsal pair is similar to the second ventral pair with a raised base and a short, wide cilium. This shows that although there are clearly more sensory endings than SmKK7 positive protrusions (Fig. 4.9), the KK7 staining pattern is consistent with that of the nerve endings. It is possible that the sensory structures with the short, wide cilium are those that stain positive for SmKK7 protein.

The pattern revealed by immunohistochemistry in the schistosomulum is striking (Fig. 4.10). The same six cells which were identified by WISH are highlighted, and the protein is distributed in a cyton network. Phalloidin counter-staining reveals that the middle pair of SmKK7-positive cells lies just posterior and dorsal to the gut sacs (Fig. 4.10C). Two processes from each of these cells extends to the anterior, the medial pair meet and join dorsal to the oesophagus (Figure 4.10C). There are six ventral strands (Fig.4.10A) and six dorsal strands (Fig.4.10 D) extending from the posterior of the worm towards the anterior, they connect with lateral processes just posterior to the acetabulum. SmKK7 staining is not discernible on the surface of the schistosomulum, and the anterior knobs are no longer visible.

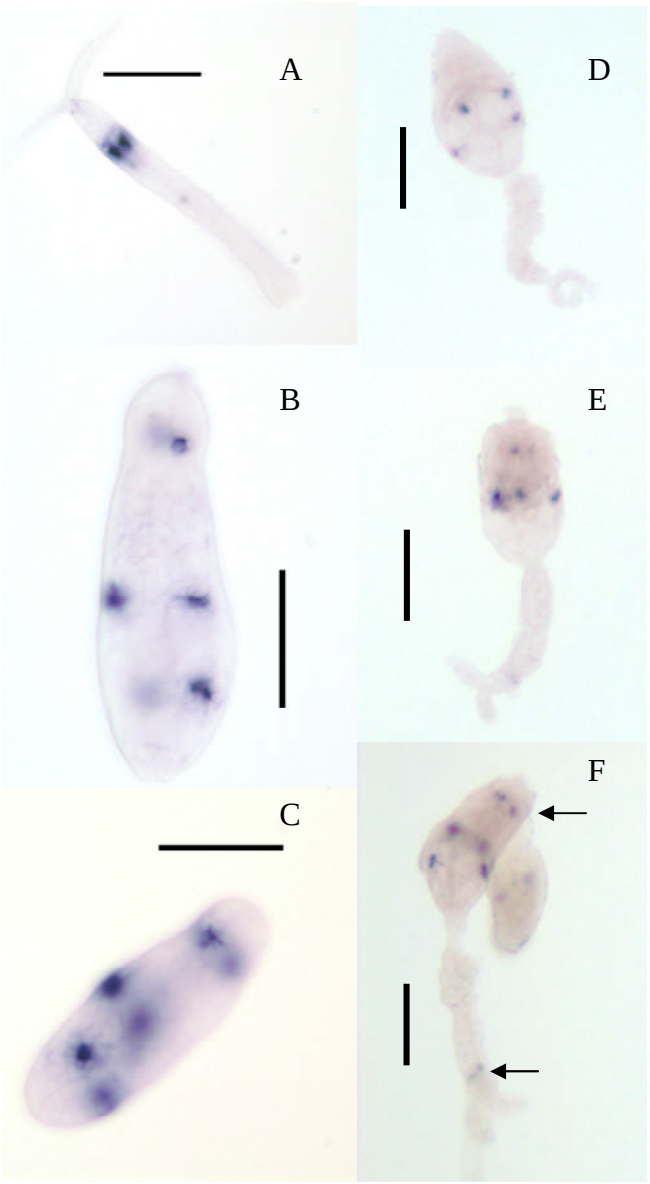


Figure 4-6 SmKK7 transcript was localised in larvae by WISH

KK7 transcript is localised in two cells at the posterior of the cercarial tail stem (A), six distinct cells in the cercarial body (B) and schistosomulum (C). SmKK7 transcript can be seen in four cells in the body at first (D), appearing in the head (E) and tail as the larva develops (F, arrows). Scale bars 50μm.

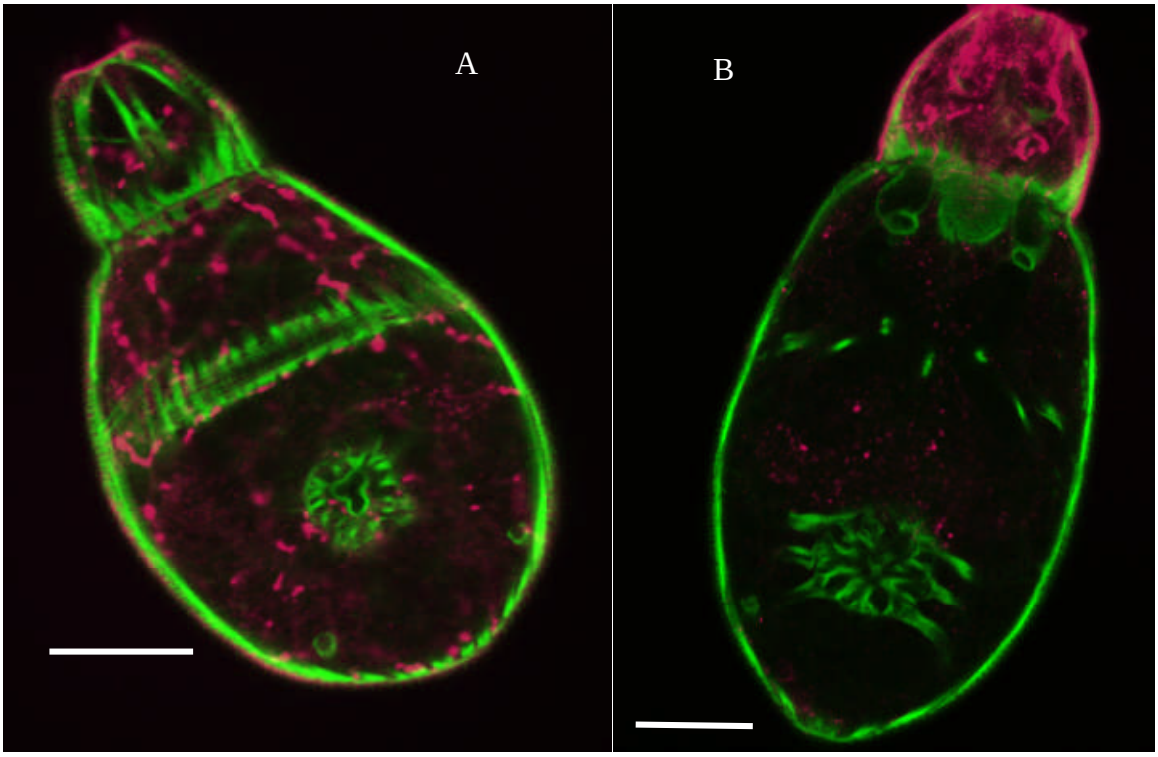


Figure 4-7 SmKK7 protein distribution in cercariae

SmKK7 (pink) was localised by immunohistochemistry using Alexafluor 488-phalloidin (green) as a counterstain. Optical sections of two typical specimens are shown. Note internal network staining (A). There is also extensive staining external to the the head capsule musculature and within the head capsule (B). Scale bars 20μm.

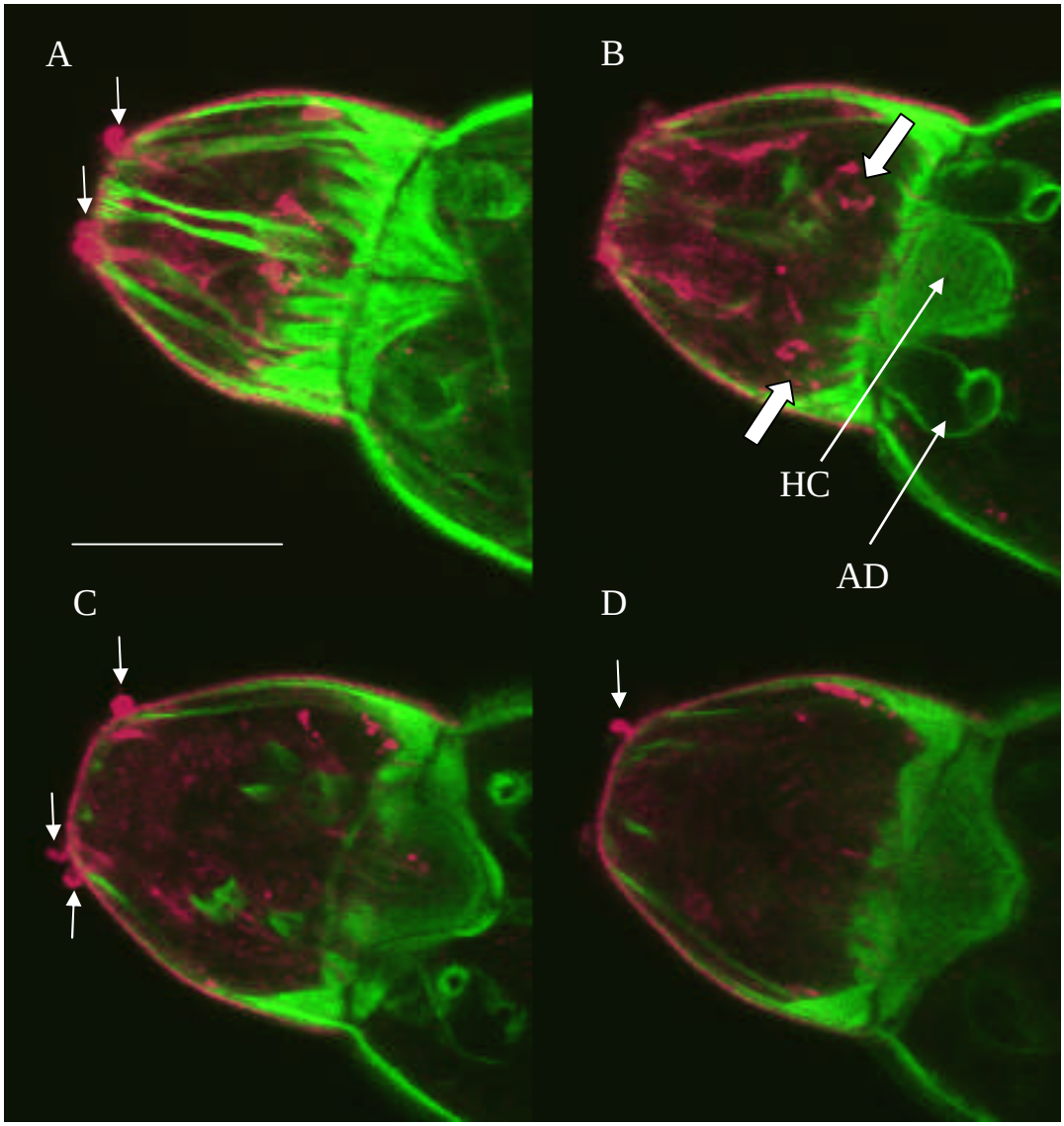
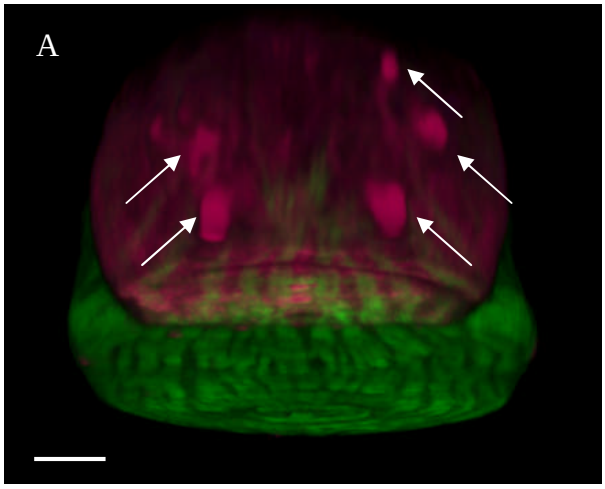
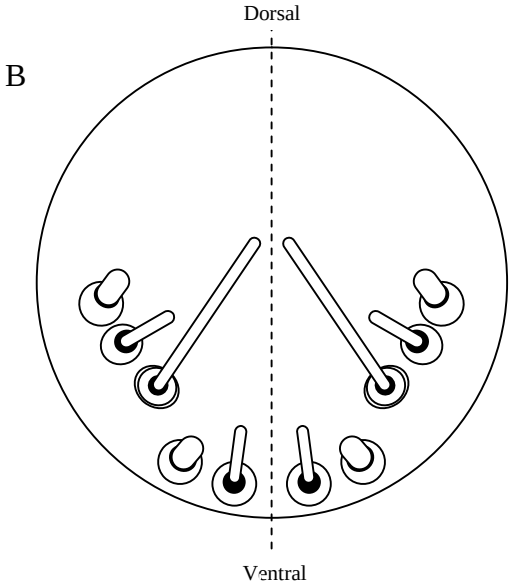


Figure 4-8 SmKK7 positive protrusions at anterior of cercaria

Selected optical sections of the head capsule of a cercaria stained with anti-SmKK7 antisera (pink) counterstained with phalloidin (green). Sections are presented ventral (A) to dorsal (D) and are 3.35µm, 5.95µm, 8.18µm and 11.53µm through a 19.71µm deep z-stack. Note positively staining protrusions from anterior end of head capsule (arrows), and cell bodies (block arrows). Head capsule muscle boundary (HC) and acetabular gland ducts (AD) are highlighted by phalloidin staining. Scale bar 20µm.


A


Dorsal
B
Ventral

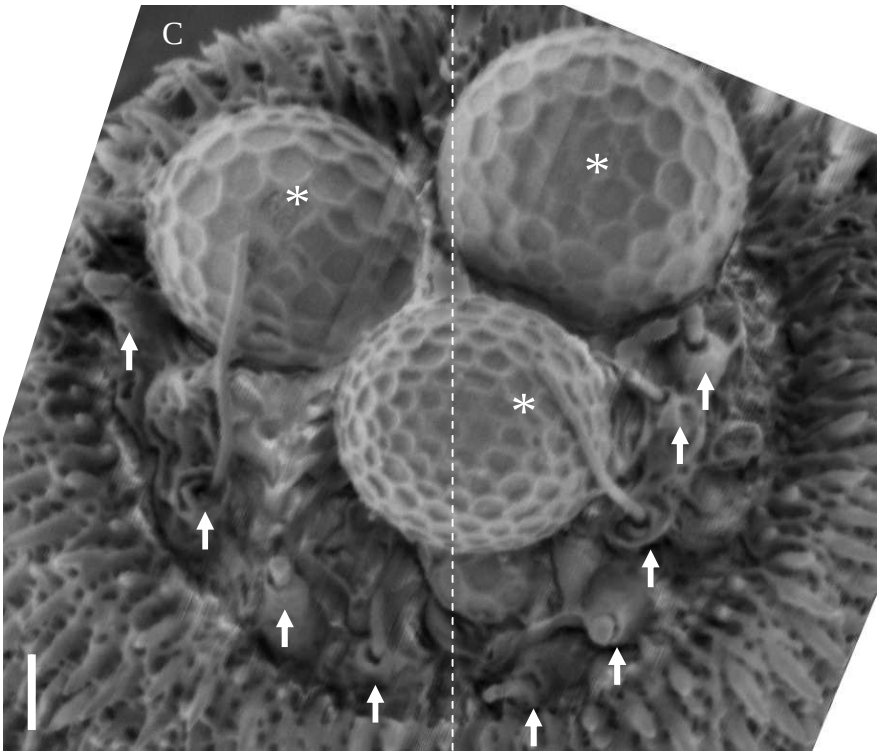

C

Figure 4-9 SmKK7 protein in a cercaria alongside an electron micrograph of the anterior
A) 3D projection of a cercaria (the specimen shown in Figure 4.8) highlighting the position of SmKK7 positive protrusions (arrows). Scale bar represents 10μm. B) Diagram showing position of anterior cercarial sensory endings as revealed by electron microscopy (C; arrows). Secretions can be seen exiting the acetabular glands (asterisks). Electron micrograph kindly provided by Stephanie Hopkins and Prof Jim McKerrow, University of California, San Fransisco. Dorso-ventral axis is shown (dotted lines), and sensory endings are indicated (arrows). Scale bar 1μm.

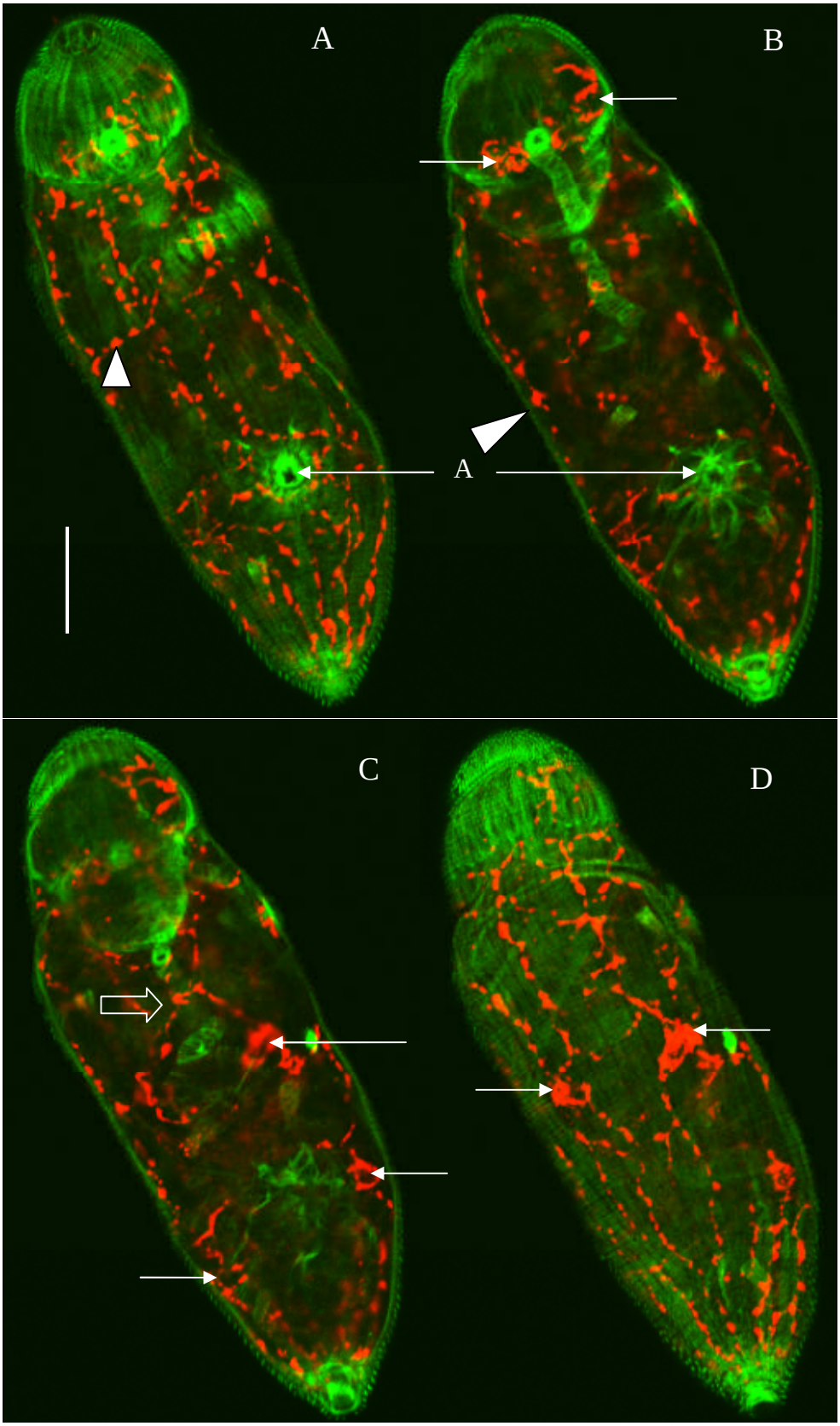


Figure 4-10 Distribution of SmKK7 protein in schistosomula

Figure 4-10 Selected optical sections of day 3 schistosomulum after IHC for SmKK7 (red) counterstained with Alexafluor 488-phalliodin (green). The sections shown are 3.71μm, 4.83μm, 5.94μm, and 7.43μm (A to D, respectively) through a 10.03μm deep z-stack. Note ventral (A) and dorsal (D) network, and lateral chords (arrow head, B) stained positive for KK7. The middle pair of cells is most prominent (D). A pair of processes meets dorsal to the oesophagus (block arrow, C).

4.3.6. KK7 transcript localises to numerous cell bodies in adults

In adults, KK7 was localised using a fluorescent substrate for WISH allowing visualisation by confocal microscopy. Many more KK7-producing cells were seen in this stage. Particularly striking were those around the oral sucker and near the surface of the body (Fig. 4.11). Following the positive staining for transcript in adults, IHC was carried out on cryo-sections of adults. Figure 4.12 shows electron micrographs of the tegument of an adult male (A, B) alongside a projection of a z-stack taken of a section of an adult male worm after IHC (Fig.4.12 C). SmKK7 protein is visible in a distinct and complex network, which protrudes into selected dorsal tubercles. In contrast, no staining is seen in the tegument on the ventral side (in contact with the female). The electron micrographs reveal the presence of sensory endings on dorsal tubercles. Figures 4.12 A and B show the same area of a male worm at different magnifications. Tubercles covered in spines alternate with tubercles with many fewer spines (Fig 4.12A). The sensory structures are present on the tubercles with fewer spines (Fig.4.12 B).

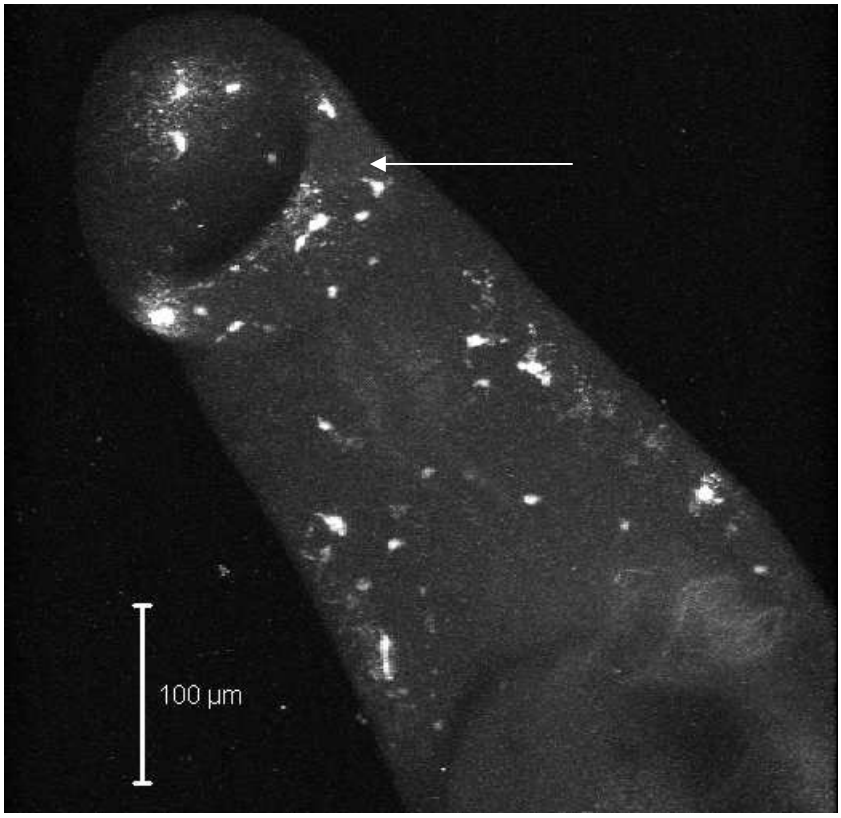


Figure 4-11 SmKK7 transcript localised in whole mount adult worms.

Projection of a z-stack of adult male anterior after WISH. Fast Red was the substrate. Cell bodies expressing SmKK7 (white) are located around the oral sucker (arrow) and laterally on the body.

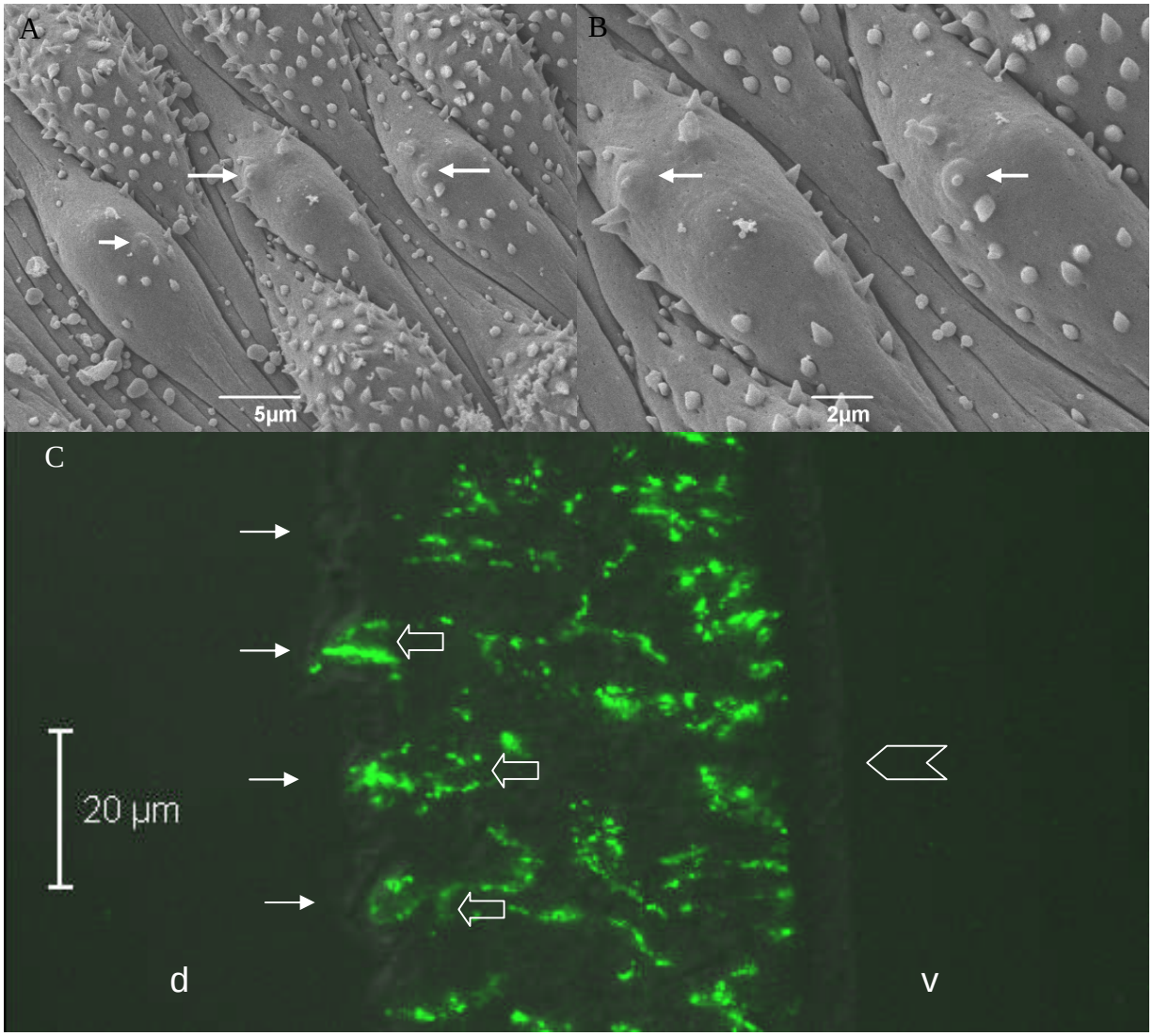


Figure 4-12 Distribution of SmKK7 protein in adult worms

A) Electron micrograph showing sensory endings (arrows) on the tubercles at the dorsal surface of an adult male worm dorsal. B) the same specimen as A, shown at higher magnification. Note the sensory structures are present on tubercles with many fewer spines. C) immunolocalisation of KK7 protein in a section of an adult male. Network containing SmKK7 is visible (green) including loops in the dorsal (d) tubercles (block arrows). Not all tubercles stain positive for SmKK7 (arrows). Note lack of staining in tegument on ventral (v) side (chevron).

4.3.7. Microexon genes

Three microexon genes were chosen for localisation studies using WISH. Figure 4.13 shows MEGs 4.2 and 14 associated with the gut primordium of the cercaria (Fig 4.13 A and B). Phalloidin staining reveals the position of this organ in the cercaria (Fig. 4.13 C). It is centrally located, posterior to the head capsule. In addition, transcripts encoding MEGs 4.2 and 14 are revealed here to be expressed in the gut of the day 3 schistosomulum (Fig. 4.13 D and E). The positively staining area is posterior to the head capsule, anterior to the acetabulum. Expression

is increased at this stage as evidenced by the darker staining. Visualisation of f-actin shows minimal expansion of the oesophagus and gut at this stage (Fig. 4.13F). Expression of MEG 3.2 was shown to be sharply demarcated in the head gland of the schistosomulum (Figure 4.14). It is the first transcript to be localised to this gland. The positive staining is seen at the very anterior of the worm, within the head capsule. The morphology of the larva is highlighted by counterstaining with DAPI and simultaneous collection of a differential interference contrast image. This image highlights the absence of nuclei in the tegument.

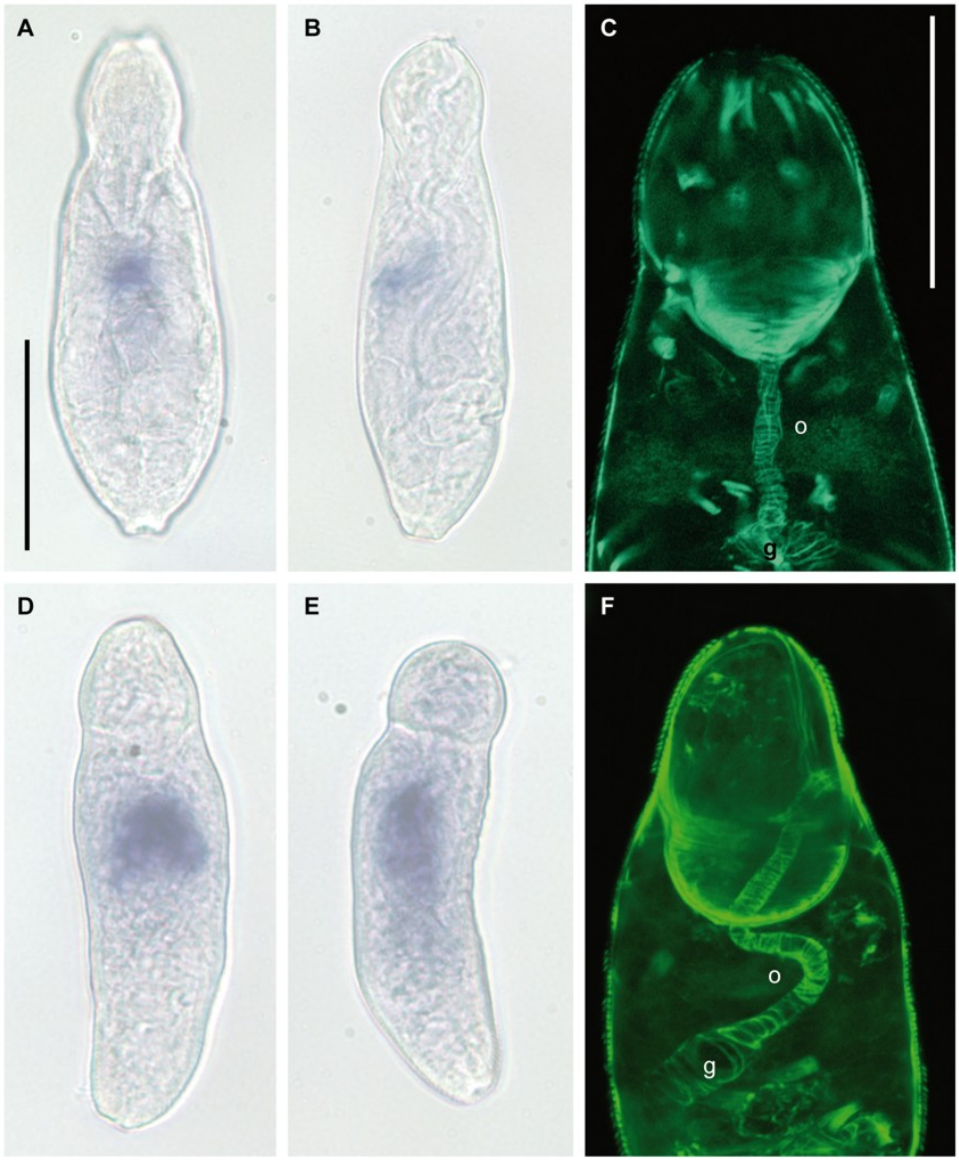


Figure 4-13 Distribution of MEG transcripts in larvae

WISH reveals transcripts encoding MEGs 4.2 (A) and MEG 14 (B) in the cercarial gut. Gut (g) and oesophagus (o) revealed by phalloidin staining of actin in surrounding muscle cells in the cercaria (C) and day 3 schistosomulum (F). MEG 4.2 (D) and MEG 14 (E) transcript is also expressed in the gut in the day 3 schistosomulum. Scale bars A, B, D, E 20μm; C and F 50μm.

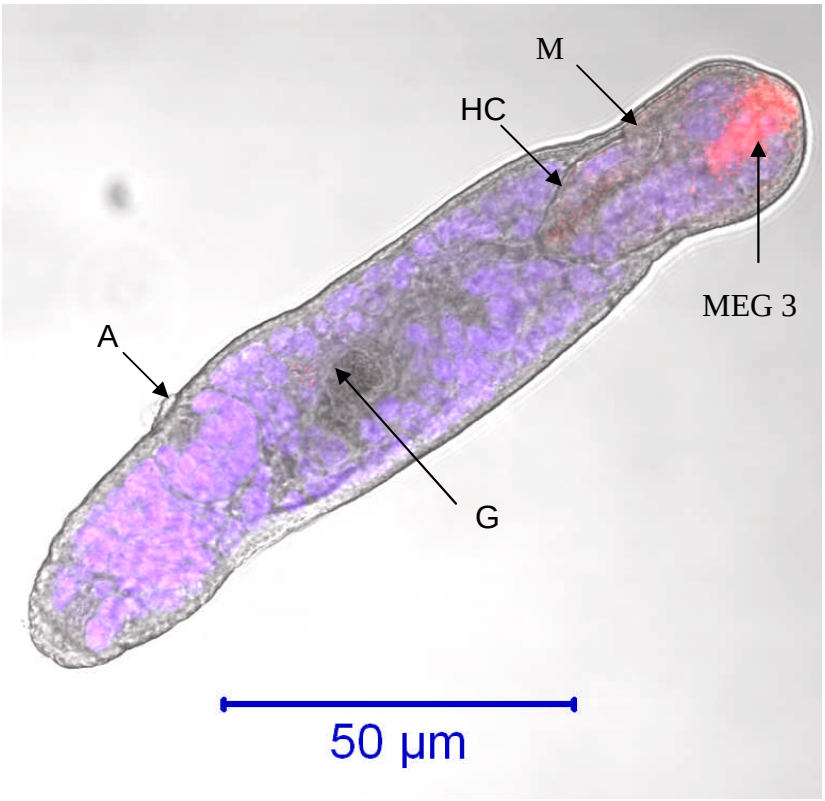


Figure 4-14 MEG3.2 (Smp_138070) transcript locaslied in a Day 10 schistosomulum
Fast Red (Sigma) was the substrate for alkaline phosphatase, thus MEG 3 transcript is highlighted by the red stain in the head gland. Parasites were counterstained with DAPI to visualise nuclei (blue). The mouth (M), head capsule muscle boundary (HC), gut (G), and acetabulum (A) are visible due to counterstaining with DAPI and collection of a non-confocal differential interference contrast image (grey).

4.4. Discussion

4.4.1. Methods

The use of methanol for permeabilisation, instead of proteinase Ktreatment, has resulted in the first successful application of WISH to germ balls. As we saw in Chapter two, germ balls are fragile, it seems they cannot withstand proteinase Ktreatment. However, this harsher permeablisation method is necessary for cercariae and all other lifecycle stages [16]. DNA is preserved during WISH as the technique was developed to localise nucleic acids. This makes DAPI an ideal counterstain for WISH when Fast Red is used as the alkaline phosphate substrate for visualisation by confocal microscopy.

4.4.2. Cercarial elastase 1a

Cercarial elastase 1a transcript was chosen as a positive control both for WISH in germ balls generally, and for localisation to the acetabular glands in particular, as Newport *et al* showed that cercarial elastase 1a was produced in these cells [114]. The results of the current study reveal that cercarial elastase mRNA is abundant in young germ balls at the flattened posterior stage, and expression continues until the germ ball is nearly mature, although the staining is not so intense. This pattern allows the morphogenesis of the acetabular glands to be tracked throughout germ ball development. It confirms that the large prominent cells observed towards the anterior of young germ balls in Chapter 2 are the primordial acetabular glands. This is in keeping with statements by Ebrahimzadeh and Dorsey that the acetabular glands are the first tissues that are recognisable in very young germ balls [40, 41]. In addition to identifying novel proteins in cercarial secretions, Curwen *et al* investigated the temporal expression pattern using pulse labelling with radio-labelled methionine [116]. Infected snails were incubated with [35S] methionine, and the post transformation secretions were collected from cercariae shed three, five, seven and nine days later. The results showed that the bulk of cercarial elastase was translated seven days prior to shedding [116].This is consistent with the present observation that transcript encoding this important enzyme is transcribed in very young germ balls.

4.4.3. Invadolysin

Following identification of invadolysin (SmPepM8) in cercarial secretions, it was hypothesised by Curwen *et al* to be secreted during host invasion, and to play a role in allowing the parasite to travel through the extracellular matrix in the skin and across the basement membrane [116]. In the current study, the presence of invadolysin has been verified in the acetabular glands by both WISH and IHC. This confirms the earlier observation that it is secreted during entry into the skin. The authors also showed that the bulk of this enzyme was synthesised three days prior to exit from the snail. The results detailed here also show that invadolysin is expressed later than cercarial elastase, as transcript was detected only in more mature germ balls.

4.4.4. VAL-10

VALs 4, 10 and 18 were identified in cercarial secretions by Curwen *et al* [116], and although the family of *S. mansoni* VAL-protein encoding genes was described by Chalmers *et al* recently [125], no function has yet been ascribed to these proteins. We saw in Chapter 3 that just under half of the VAL genes are differentially expressed during the infection process. The majority of these are up-regulated in the germ ball or the cercaria. In this chapter, transcript encoding VAL-10 was localised to the acetabular glands. This confirms the likely secretion of VAL-10 protein in the early stages of infection. The peripheral nature of the staining pattern exhibited by acetabular gland associated transcripts in well-developed germ balls is in keeping with the observation of Dorsey that the protein synthesis machinery is pushed to the edge of the acetabular gland fundi as these cells develop and fill with secretory vesicles [41].

4.4.5. Sm16

Sm16 has been variously described as an immunomodulator and stathmin-like. Stathmin is a protein with microtubule regulating activity [188]. It binds to free tubulin monomers and inhibits microtubule formation. Valle *et al* reported that Sm16 was stage specific, it was found in sporocysts but was not detectable in schistosomula 48 hours post infection [189]. The authors suggested that Sm16 plays a role in regulating development at specific points in the life cycle. Sm16 was found in the proteins released by cercariae in the first three hours after mechanical transformation [116] and reported to be translated by both cercariae and schistosomula. Rao *et al* cloned and expressed Sm16 and suggested that it is secreted by schistosomula whilst they are in the skin to down regulate the host's inflammatory response [120].

Subsequent work has failed to confirm either that Sm16 is stathmin-like, or that it has a potent immunomodulatory effect. Holmfeldt *et al* showed that recombinant Sm16 did not inhibit microtubule formation *in vitro* [186]. They found that the presence of Sm16 in the cytosol of human cells led to apoptosis. However, this was achieved by expressing Sm16 in human cells. It was found that when Sm16 is added to human cells, it binds to the surface, is endocytosed and degraded. In this case, the cells do not apoptose [186]. Later, Brannstrom *et al* expressed Sm16 in *Pichia pastoris* and discovered that it had not cause T-lymphocytes to proliferate; neither did it lead to a change in the cytokine production of whole human blood or monocytes

[187]. However, Sm16 did inhibit the response of various cells to lipopolysaccharide and poly (I:C) [187]. In an attempt to better understand this controversial protein, mRNA encoding Sm16 was localised in germ balls.

WISH results presented here show that the Sm16 transcript has an expression pattern quite distinct from those of elastase 1a, invadolysin and VAL-10. This indicates that Sm16 is not expressed in cells normally expected to secrete their contents into the host during infection. Due to the widespread expression of Sm16 throughout the tissues of the germ ball, including the tail (but excluding the acetabular glands), it is plausible that Sm16 would be released as the tail detaches, rather than as a result of holocrine secretion from the acetabular glands. However, the current results are in agreement with the timescale of expression detailed by Curwen *et al*. The authors state that Sm16 was translated between seven to three days prior to exit from the snail [116]; here Sm16 transcript is evident in germ balls from the flattened-posterior to the nearly mature stage.

The discovery that Sm16 is not expressed in secretory glands underlines the importance of localisation studies for confirmation that a particular gene is expressed at the host parasite interface. Such studies should be carried out prior to further studies aimed at discovering how the protein of interest interacts with the host.

4.4.6. SmKK7

SmKK7 was first identified in cercarial secretions. However, it has also been discovered in the soluble fractions of supernatants when the tegument of adult worms is perturbed. In the light of these findings, it is perhaps not surprising that SmKK7 is not expressed in the acetabular glands. SmKK7 was annotated because of homology to the potassium channel blocker BmKK7 from venom of the scorpion *Mesobuthus martensii* [116]. The two proteins are 30% identical, with conserved cysteine residues. It was hypothesised by Curwen *et al* to play a part in immunomodulation by interacting with T cells [116].

The transcript for SmKK7 was localised to eight cells in the germ ball and cercaria, and six cells in the day 3 schistosomulum. At first glance these cells look similar in distribution to the flame cells; however, more careful observation reveals that they are different. In the tail, flame

cells are proximal, whereas the KK7 producing cells are distally located, just before the tail bifurcates. In addition, the anterior KK7-positive cells are situated in the head capsule, whereas the most anterior flame cells are posterior to the head capsule muscle boundary (see phalloidin staining of larvae in section 2.3.2). Confirmation that the KK7-producing cells are not flame cells is the lack of colocalistaion of KK7 staining with the flame cells as highlighted by phalloidin counterstaining with IHC [133].

It is rather impressive that the extensive network stained positive for KK7 protein arises solely from the six cells identified by WISH. The anterior pair of cells may be the pair with extensions to the surface, that was highlighted by Bruckner and Voge as staining for acetylcholine esterase activity [51]. The varied morphology of the sensory structures at the anterior of the cercaria suggests that they serve different functions. The shape of the KK7-positive protrusions indicates that they may be the sensory endings that have a short wide cilium, rather than either of the types with a narrow cilium. Robson and Erasmus studied the anterior of the cercaria using electron microscopy. They showed the five sensory structures detailed in the present study with the addition of two at the dorsal side of the crescents that surround the acetabular duct openings [55]. Nuttman showed that the axon and raised bulb (but not the cilium) of the anterior sensory endings were filled with vesicles [190]. Based on the current observations, it is possible that SmKK7 is within these vesicles. However, immunohistochemistry in conjunction with electron microscopy would have to be carried out to confirm this.

The expression of SmKK7 has also been investigated in adult worms. The KK7-producing cells are far more numerous in the adult, and visualisation of the protein reveals an extensive net there, with processes entering tubercles on the adult male dorsal surface. Staining was restricted to the area under the two muscle layers on the ventral side. Neuropeptide S1 (KYSALMF-NH_2) has previously been localised with a pattern similar to KK7; it was present in the peripheral nerve net including occasional dorsal tubercles [27]. In order to demonstrate that KK7 enters the tubercles with fewer spines, counterstaining with an appropriately labelled phalloidin would be necessary. It seems likely that these are ‘touch’ receptors, as the lack of spines would aid perception. SmKK7 staining could be used to track the development of the schistosome peripheral nervous system from comprising only six cells in the larvae to the fully developed state in the adult.

The neural location of SmKK7, revealed here, seems at odds with its appearance in the 0-3 hour released proteins collected when cercariae are mechanically transformed. The simplest explanation is that KK7 is expressed in the peripheral nerves in the worm where it functions as a modulator of the action potential through inhibition of potassium channels. In this scenario, SmKK7 is not expected to come into contact with the host under normal circumstances and so the presence of KK7 in the released proteins is likely to be an artefact of mechanical transformation (incubation on ice followed by vortexing), or in the case of the adult by deliberate damage to the tegument. Conversely, it is known that the anterior cercarial nerve endings disappear during transformation. It may be the case that SmKK7 is released into the host as part of that process, whereupon it acts on host potassium channels. To investigate the possible effect of KK7 on host blood vessels, the protein was expressed in *Pichia pastoris* by Dr GP Dillon, and applied to mouse blood vessels. However, no effect was observed. SmKK7 has also been the subject of a knock-down study using RNAi in schistosomula (Dr P McVeigh, personal communication). Again, no phenotype was observed. This is perhaps not surprising as it may take a long time for all of the existing KK7 protein to be turned over. This would have to happen before the knockdown of the mRNA would have an effect.

Co-localisation with known neural markers may give extra information of the type of nerves in which KK7 is expressed, and shed some light on its function. Although many questions remain regarding its role in schistosome biology, it is clear from the current study that it expression is not limited to the cercaria.

4.4.7. MEGs

The remarkable up-regulation of MEG expression is the day 3 schistosomulum documented in chapter three, leads to the conclusion that whatever their function, the proteins encoded by these genes are important for the parasite during early intra-mammalian life. Although genes containing single micro-exons have been described in other species, DeMarco *et al* report that MEGs (comprising mostly micro-exons) are only found in schistosomes [142]. The authors state that MEGs seem to have appeared and expanded rapidly in number in an ancient schistosome ancestor. It is likely that this was an adaption to parasitism, specifically to life in the bloodstream [142].

Three MEGs have been localised here for the first time in cercariae and/or schistosomula. They are without exception expressed in tissues that form part of the host-parasite interface. MEGs 4.2 and 14 are transcribed in the larval gut; Dillon *et al* have previously shown that MEG 4.1 is expressed in the oesophageal gland of adult worms [16]. It is difficult to identify this tissue in cercariae and schistosomula. Nevertheless, the localisation of these transcripts to the alimentary tract confirms that they could come into contact with the host; schistosomes have a blind ending gut therefore waste products are ejected by vomit. DeMarco *et al* postulate that protein variation arising from MEGs is likely to be part of the immune evasion strategy employed by schistosomes, but exactly how this is of benefit to the parasite is under investigation [142]. It is also more evidence, taken with the up-regulation of gut proteases, that the gut is transcriptionally active at this stage. MEG 3.2 stands out as it is the first gene known to be expressed in the head gland. This is noteworthy as the contents of this gland are secreted during skin invasion [48].

Summary

The localisation of novel schistosome gene products reveals unexpected patterns, and highlights the importance of WISH both as a screen for selection of genes expressed at the host-parasite interface, and in its role in hypothesis forming for functional annotation.

5. Final Discussion

The work presented in this thesis focuses on the *S. mansoni* larval stages before, during, and after infection of the mammalian host. This encompasses many varied developmental changes for the parasite: embryogenesis within the daughter sporocyst, exit from the snail, and survival in fresh water with only a store of glycogen for nourishment. Once the cercaria has found a suitable host, it enters by penetration of unbroken skin. The newly parasitic schistosomulum undergoes rapid morphological changes that allow it, not only successfully to evade the host immune system, but also to thrive. The next challenges the parasite must overcome are two barriers: the epidermal basement membrane, and the epithelium of a blood vessel. The schistosome persists in the vascular system for the rest of its life, which could be up to 30 years.

An intervention directed at larvae is attractive as it would interrupt the life cycle prior to the onset of egg deposition. This is important for two reasons: firstly, pathology would be prevented as it is the host response to the egg that causes it; secondly, transmission would be blocked as this requires eggs to escape in the faeces, allowing miracidia to hatch and infect a snail intermediate host. Therefore, a more thorough understanding of the infection process was sought, with the aim of discovering new ways to target larvae early during infection.

The current study took a whole organism approach. First, the morphology of the relevant life cycle stages was characterised. This was followed by a whole genome microarray experiment to discover the gene expression underlying the impressive morphological and physiological changes that take place. Lastly, the expression of selected genes was localised, putting them into the context of the whole organism. The new methods used in this study will be described. I will then draw together the novel findings discovered through the application of these techniques and discuss how our knowledge of each life cycle stage and the infection process has been furthered.

The cercaria had previously been studied extensively using electron microscopy but the development of germ balls and schistosomula had been less well characterised. A large hurdle for the study of germ balls was the use of sectioned hepatopancreas specimens. This made interpretation of images difficult, as the orientation was hard to determine, and images had to be reconstructed from a series of sections. The first application of confocal microscopy to whole mount daughter sporocysts, germ balls, and schistosomula, presented here, allowed a

detailed study of the morphological changes that take place throughout the infection process. This enabled the information gained in the later chapters to be related to the big picture.

The data that can be gleaned from a microarray experiment are limited to the sequences represented on the array. When undertaking such an experiment, one must ensure that genes of interest are included. This has led to the design and use of a variety of different microarray platforms for *S. mansoni*, each of which was tailor-made to address the biological questions under consideration at the time. All of these previous arrays have included only part of the transcriptome; many comprised sequences derived from one life cycle stage only. With the availability of a draft genome sequence and EST libraries from many life cycle stages, an opportunity arose to design the first comprehensive microarray for this medically important species. This was carried out in collaboration with the Wellcome Trust Sanger Institute and the array was synthesised by Roche-NimbleGen. The design will be made freely available, and should be useful for the research community, as it could be applied to any gene expression question with any life cycle stage(s). It could also be re-annotated as the genome assembly improves and the number of un-annotated genes is reduced.

The other novel aspect of the microarray experiment detailed in this thesis is the hybridisation of germ balls to an array comprising sequences expressed at that stage. It has been shown that the proteins secreted by cercariae during skin penetration are synthesised whilst germ balls are in the snail. Therefore, gaining knowledge of this life cycle stage is an integral part of investigating the mammalian infection process. The only other investigators to hybridise germ ball material to an array were Dillon *et al* who used a microarray based entirely on lung schistosomulum sequences to identify transcripts preferentially expressed in the schistosomulum [98]. However, germ balls have not been included before in microarray studies aimed at investigating the infection process. Other investigators have included daughter sporocysts or infected hepatopancreas instead.

In order to relate the temporal expression pattern to the biology of schistosomes, a small group of genes was chosen for localisation studies. The whole mount *in situ* hybridisation technique developed by Dillon *et al* was applied to cercariae, schistosomula and adult worms. In order to successfully localise transcripts in germ balls, an alternative permeablisation method was

employed. This allowed genes of interest to be localised in each of the life cycle stages under investigation. It was the first application of WISH to germ balls.

The morphological study revealed that migratory daughter sporocysts have large germ cells beneath a relatively thin layer of muscles. These cells could be readily targeted by particle bombardment, and may solve the problem of delivering DNA to germ cells. It is worth noting that daughter sporocysts can be delivered into snails where they continue the life cycle and could produce transgenic cercariae.

In order for a germ ball to become a cercaria, much cell division, translation, and organisation of cells into tissues must take place. Daily sampling in the morphological analysis revealed that germ balls develop between 22-26 days after infection, and mature cercariae are present at day 27. This was an important finding, as it allowed germ balls of all stages to be sampled in the absence of cercariae, so that their gene expression profile could be compared with cercariae and day 3 schistosomula. I have confirmed that the germ ball tegument forms as cells at the surface coalesce, and that connections are made between these contributing cells and cells further inside the body before the nuclei in the tegument are lost. Thus the epithelial origins of the adult tegument syncytium are confirmed in the germ ball.

Results from the morphological study indicated that the acetabular glands begin to develop very early. This was confirmed by the first application of WISH to germ balls; it was apparent, following the localisation of cercarial elastase mRNA, that the acetabular glands take longer to mature than any other tissue. They begin developing as soon as the anterior-posterior axis is discernable by the flattened posterior, and transcripts encoding secreted proteins are still present in nearly mature specimens. The duration of expression noted for cercarial elastase is in stark contrast to the brief appearance of transcripts encoding both SmPepM8 and VAL-10, which are only expressed in nearly mature germ balls.

The microarray study detailed in Chapter 3 shed light on the gene expression patterns underlying cercarial embryogenesis. As well as DNA polymerase and genes whose products are required for cell cycle progression, it was revealed that a subset of cadherins, caspases, homeobox and Wnt genes are up-regulated at this stage. Cadherins are adhesion molecules which regulate cell-cell interactions. It is likely that the four cadherins expressed in the germ

ball are involved in sorting tissues. Caspases are involved in programmed cell death, which is also important in embryogenesis. Although much remains to be learnt regarding the complex process of cercariogenesis, WISH may be used to localise genes shown in this study to be enriched in germ balls, enabling hypotheses to be generated regarding their precise functions.

Five cercarial elastase genes had been described previously by Salter *et al* [115]. The current thesis provides the first evidence that 11 cercarial elastase genes are up-regulated in the germ ball and five invadolysins additional to Curwen's SmPepM8 are also germ ball-enriched. The results reveal that the infection process is more complex than had been thought. It is certainly not the case that cercarial entry into the host is mediated by a single protease class. This is the first report of VAL expression in germ balls. Again, the known repertoire was expanded; six VALs were shown to be up-regulated at this stage. The up-regulation of VALs in germ balls and/or cercariae suggests that their functions are related to the early stages of infection. Transcripts encoding VAL 10 and invadolsysin (SmPepM8) were localised to the acetabular glands, confirming their importance as secreted proteins involved in host invasion.

More surprising results were obtained when Sm16 and SmKK7 were localised. Both of these proteins, Sm16 in particular, had been suggested to function as immunomodulators during early infection. A prerequisite for such a function is that the protein must come into contact with the host. However, transcripts encoding both of these proteins were localised to tissues not expected to secrete their contents to the exterior of the worm. This could be taken as a salutary tale; caution should be taken when assigning putative functions to novel proteins. Even their presence in preparations enriched for secretions is not conclusive proof that such proteins are, in fact, secreted by the worm. Localisation studies can be used to confirm hypotheses prior to further research. In this case, the putative functions of both SmKK7 and Sm16 can now be reappraised in the light of their localisation patterns. SmKK7 is present in the peripheral nerves. The simplest explanation is that SmKK7 is involved in neural function within the worm. Many questions remain. Assuming it is a potassium channel blocker, which potassium channels does it affect and which signals does it respond to? It is very difficult to make any hypotheses regarding the function of Sm16 based on its expression throughout the body. However, it is not transcribed in the acetabular glands and as such, despite several reports that it is found in preparations enriched for secreted proteins; it does not seem to be secreted 'on purpose' by the worm.

The cercaria has been thought of as quite transcriptionally inactive. It is a complex organism adapted for its short stint as a free-living, non-feeding, aquatic creature. As previous studies have shown, the cercaria uses aerobic metabolism in order to make the most of its glycogen store. The free-living, snail-infective miracidium also uses aerobic metabolism. Another similarity with the miracidium is the up-regulation of the small heat shock protein Sm-p40, which conserves energy by avoiding proteasomal degradation of denatured proteins. A novel finding of the current study is that the cercaria transcribes genes whose products are needed on entry to the host. Indeed, 'planning ahead' seems to be a salient feature of schistosome biology. The contents of the acetabular glands are synthesised even before the germ ball has a tail and tegument- and gut-associated protein encoding genes are up-regulated in the cercaria which does not feed, and whose tegument has no use for glucose transporters or complement inhibitors.

Cercariae are known to emerge from snails when it is light. Although the morphological study did not pin-point the light sensitive organ, the microarray experiment revealed that an opsin-like GPCR is up-regulated in the germ ball. There is also a light sensitive cell development protein whose transcript is up-regulated in the cercaria. WISH could be used to localise the expression of either of these genes in the germ ball, potentially leading to the identification of the cercarial light sensing organ. The cercaria also relies on neural signalling for swimming and to find a host. This is reflected in the up-regulation of neurotransmitter transporters at this life cycle stage and the conspicuous nerve endings at the anterior, some of which contain large amounts of the novel neural protein SmKK7.

The head gland, whilst known to be important in the infection process is an enigmatic organ. It was not known whether it was unicellular like the acetabular glands; it was shown in Chapter 2 that it contains two nuclei. It was also revealed in this study that MEG 3 is transcribed in this gland.

Considerable remodelling occurs in the three days after transformation. It was noted in Chapter 2 that the anterior sensory endings and acetabular glands and ducts are no longer visible in the day 3 schistosomulum, and that the lateral nerve chords can be distinguished at this stage. Cadherins and various other morphogens are implicated in this developmental

phase by virtue of their increased expression in the day 3 schistosomulum. It is worth noting that the cadherins expressed at this stage are different to those up-regulated in the germ ball.

The morphological study of the schistosomula revealed the presence of protein in the larval gut and the microarray study demonstrated that gut-related proteases and transport proteins are most highly expressed at this stage compared to the other two. It remains unclear whether the protein observed in the gut is synthesised or ingested by the schistosomulum. This question could be addressed by incubating schistosomula in medium containing fluorescently labelled proteins and subsequently viewing them by confocal microscopy. It seems likely that early larvae do ingest plasma, as nutrition is necessary for the onward migration. However, at this stage the oesophagus is still too narrow to admit red blood cells.

Another finding of the current study was the marked up-regulation of certain tegument protein-encoding genes in the day 3 schistosomulum. The tegument is the ideal target for a vaccine, as it is accessible to the host. Those proteins expressed at the surface throughout intra-mammalian life may prove to be successful candidates, as they are by implication necessary for survival in the host, and may enable the host to target the parasite before it reaches maturity and causes pathology. Tegument proteins linked to immune evasion by adult worms are up-regulated in the day 3 schistosomulum. Of particular note are a group of CD59-encoding genes. In humans, CD59 is a complement inhibitor. The schistosome genome contains seven of these paralogues, four of which have been identified on the surface of adult worms. This is the first report showing their up-regulation in the schistosomulum. Functional studies are yet to be carried out on schistosome CD59. These would be needed as well as localisation studies in the larva to determine its role in complement inhibition and immune evasion.

An intriguing part of the infection process is the appearance of micro exon genes in the early intra-mammalian larva. Four new MEGs were identified in the current study. The highest fold change observed in the study was exhibited by a MEG. Not only are the majority of MEGs up-regulated in the schistosomulum, three of them have been localised to tissues at the host parasite interface. The extreme splice-variability of these transcripts, coupled with their expression at the host parasite interface begs the question: what is their primary function? It seems likely that the variants are generated to bamboozle the host immune response. Whether

that is in order to mask an important function carried out by a specific variant, or rather to allow the parasite as a whole to avoid immune attack remains to be discovered.

A large proportion of the genes in the *S. mansoni* genome have no homology to any other known sequences. Due to time constraints, these have not been investigated. However, the discovery of four novel MEGs among the un-annotated genes points likely presence of other novel schistosome genes that may shed light on how these worms gain entry to the host and remain there.

6. List of References

1. King CH, Dickman K, Tisch DJ: **Reassessment of the cost of chronic helmintic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis**. *Lancet* 2005, **365**(9470):1561-1569.
2. Steinmann P, Keiser J, Bos R, Tanner M, Utzinger J: **Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk**. *Lancet Infect Dis* 2006, **6**(7):411-425.
3. van der Werf MJ, de Vlas SJ, Brooker S, Looman CW, Nagelkerke NJ, Habbema JD, Engels D: **Quantification of clinical morbidity associated with schistosome infection in sub-Saharan Africa**. *Acta Trop* 2003, **86**(2-3):125-139.
4. Caffrey CR, Rohwer A, Oellien F, Marhofer RJ, Braschi S, Oliveira G, McKerrow JH, Selzer PM: **A comparative chemogenomics strategy to predict potential drug targets in the metazoan pathogen, Schistosoma mansoni**. *PLoS ONE* 2009, **4**(2):e4413.
5. Abdulla MH, Ruelas DS, Wolff B, Snedecor J, Lim KC, Xu F, Renslo AR, Williams J, McKerrow JH, Caffrey CR: **Drug discovery for schistosomiasis: hit and lead compounds identified in a library of known drugs by medium-throughput phenotypic screening**. *PLoS Negl Trop Dis* 2009, **3**(7):e478.
6. **Prevention and control of schistosomiasis and soil-transmitted helminthiasis**. *World Health Organ Tech Rep Ser* 2002, **912**:i-vi, 1-57, back cover.
7. Bethony JM, Diemert DJ, Oliveira SC, Loukas A: **Can schistosomiasis really be consigned to history without a vaccine?** *Vaccine* 2008, **26**(27-28):3373-3376.
8. Gryseels B, Polman K, Clerinx J, Kestens L: **Human schistosomiasis**. *Lancet* 2006, **368**(9541):1106-1118.
9. Cheever AW, Macedonia JG, Mosimann JE, Cheever EA: **Kinetics of egg production and egg excretion by Schistosoma mansoni and S. japonicum in mice infected with a single pair of worms**. *Am J Trop Med Hyg* 1994, **50**(3):281-295.
10. Becker W: **[Studies on the migrating daughter sporocysts of Schistosoma mansoni. II. The migration to the hepatopancreas]**. *Z Parasitenkd* 1970, **34**(3):226-241.
11. Meuleman EA, Holzmann PJ, Peet RC: **Development Of Daughter Sporocysts Inside The Mother Sporocyst Of Schistosoma-Mansoni With Special Reference To The Ultrastructure Of The Body Wall**. *Zeitschrift Fur Parasitenkunde-Parasitology Research* 1980, **61**(3):201-212.
12. Jourdane J, Theron A, Combes C: **Demonstration of several sporocysts generations as a normal pattern of reproduction of Schistosoma mansoni**. *Acta Trop* 1980, **37**(2):177-182.
13. Crabtree JE, Wilson RA: **Schistosoma mansoni: an ultrastructural examination of pulmonary migration**. *Parasitology* 1986, **92 (Pt 2)**:343-354.
14. Wilson RA, Coulson PS: **Schistosoma mansoni: dynamics of migration through the vascular system of the mouse**. *Parasitology* 1986, **92 (Pt 1)**:83-100.
15. Miller P, Wilson RA: **Migration of the schistosomula of Schistosoma mansoni from the lungs to the hepatic portal system**. *Parasitology* 1980, **80**(2):267-288.
16. Dillon GP, Illes JC, Isaacs HV, Wilson RA: **Patterns of gene expression in schistosomes: localization by whole mount in situ hybridization**. *Parasitology* 2007, **134**(Pt 11):1589-1597.

17. Jones MK, Higgins T, Stenzel DJ, Gobert GN: **Towards tissue specific transcriptomics and expression pattern analysis in schistosomes using laser microdissection microscopy**. *Exp Parasitol* 2007, **117**(3):259-266.
18. Morris GP, Threadgold LT: **Ultrastructure of the tegument of adult Schistosoma mansoni**. *J Parasitol* 1968, **54**(1):15-27.
19. McLaren DJ, Hockley DJ: **Blood flukes have a double outer membrane**. *Nature* 1977, **269**(5624):147-149.
20. Hockley DJ: **Schistosoma mansoni: the development of the cercarial tegument**. *Parasitology* 1972, **64**(2):245-252.
21. Skelly PJ, Alan Wilson R: **Making sense of the schistosome surface**. *Adv Parasitol* 2006, **63**:185-284.
22. Saunders N, Wilson RA, Coulson PS: **The outer bilayer of the adult schistosome tegument surface has a low turnover rate in vitro and in vivo**. *Mol Biochem Parasitol* 1987, **25**(2):123-131.
23. Halton DW, Gustafsson, M. K. S.: **Functional morphology of the platyhelminth nervous system**. *Parasitology* 1996, **113**:S47-S72.
24. Gustafsson MK: **Immunocytochemical demonstration of neuropeptides and serotonin in the nervous systems of adult Schistosoma mansoni**. *Parasitol Res* 1987, **74**(2):168-174.
25. Silk MH, Spence IM: **Ultrastructural studies of the blood fluke--schistosoma mansoni. 3. The nerve tissue and sensory structures**. *S Afr J Med Sci* 1969, **34**(4):93-104.
26. Hockley DJ: **Ultrastructure of the tegument of Schistosoma**. *Adv Parasitol* 1973, **11**(0):233-305.
27. Brownlee DJ, Fairweather I, Johnston CF, Thorndyke MC, Skuce PJ: **Immunocytochemical demonstration of a SALMFamide-like neuropeptide in the nervous system of adult and larval stages of the human blood fluke, Schistosoma mansoni**. *Parasitology* 1995, **110 (Pt 2)**:143-153.
28. Wilson RA, Webster LA: **Protonephridia**. *Biol Rev Camb Philos Soc* 1974, **49**(2):127-160.
29. Wilson RA: **A physiological study of the development of the egg of Fasciola hepatica L., the common liver fluke**. *Comp Biochem Physiol* 1967, **21**(2):307-320.
30. Bogers JJ, Nibbeling HA, Van Marck EA, Deelder AM: **Immunoelectron microscopical localization of a circulating antigen in the excretory system of Schistosoma mansoni. Ultrastructural localization studies of the excretory system of S. mansoni**. *Parasitol Res* 1995, **81**(5):375-381.
31. Kusel JR, McVeigh P, Thornhill JA: **The schistosome excretory system: a key to regulation of metabolism, drug excretion and host interaction**. *Trends Parasitol* 2009, **25**(8):353-358.
32. Mair GR, Maule AG, Day TA, Halton DW: **A confocal microscopical study of the musculature of adult Schistosoma mansoni**. *Parasitology* 2000, **121 (Pt 2)**:163-170.
33. **The biology of schistsosomes from genes to latrines**.
34. Machado-Silva JR, Pelajo-Machado M, Lenzi HL, Gomes DC: **Morphological study of adult male worms of Schistosoma mansoni Sambon, 1907 by confocal laser scanning microscopy**. *Mem Inst Oswaldo Cruz* 1998, **93 Suppl 1**:303-307.
35. Mair GR, Maule AG, Fried B, Day TA, Halton DW: **Organization of the musculature of schistosome cercariae**. *J Parasitol* 2003, **89**(3):623-625.
36. Hansen E, Perez-Mendez G: **Scanning electron microscopy of Schistosoma mansoni daughter sporocysts**. *Int J Parasitol* 1972, **2**(1):174.

37. Fournier A, Theron A: **[Morpho-anatomic and functional sectorization of Schistosoma mansoni daughter sporocysts]**. *Z Parasitenkd* 1985, **71**(3):325-336.
38. Cheng TC, Bier JW: **Studies on molluscan schistosomiasis: an analysis of the development of the cercaria of Schistosoma mansoni**. *Parasitology* 1972, **64**(1):129-141.
39. Meuleman EA, Holzmann PJ: **Development Of Primitive Epithelium And True Tegument In Cercaria Of Schistosoma-Mansoni**. *Zeitschrift Fur Parasitenkunde-Parasitology Research* 1975, **45**(4):307-318.
40. Ebrahimzadeh A: **[Development, histology and histochemistry of the glandular system of the cercaria of Schistosoma mansoni Sambon (1907)]**. *Z Parasitenkd* 1970, **34**(4):319-342.
41. Dorsey CH: **Schistosoma mansoni: development of acetabular glands of cercaria at ultrastructural level**. *Exp Parasitol* 1975, **37**(1):37-59.
42. Dorsey CH, Cousin CE, Lewis FA, Stirewalt MA: **Ultrastructure of the Schistosoma mansoni cercaria**. *Micron* 2002, **33**(3):279-323.
43. Stirewalt MA: **Cercaria vs. schistosomule (Schistosoma mansoni): absence of the pericercarial envelope in vivo and the early physiological and histological metamorphosis of the parasite**. *Exp Parasitol* 1963, **13**:395-406.
44. Dorsey CH, Stirewalt MA: **Schistosoma mansoni: fine structure of cercarial acetabular glands**. *Exp Parasitol* 1971, **30**(2):199-214.
45. Ebrahimzadeh A, Kraft M: **[Ultrastructural studies on the morphology of the cercariae Schistosoma mansoni. I. The alimentary tract]**. *Z Parasitenkd* 1969, **32**(2):157-175.
46. Lawson JR, Wilson RA: **The survival of the cercariae of Schistosoma mansoni in relation to water temperature and glycogen utilization**. *Parasitology* 1980, **81**(2):337-348.
47. Dorsey CH: **Schistosoma mansoni: description of the head gland of cercariae and schistosomules at the ultrastructural level**. *Exp Parasitol* 1976, **39**(3):444-459.
48. Crabtree JE, Wilson RA: **Schistosoma mansoni: an ultrastructural examination of skin migration in the hamster cheek pouch**. *Parasitology* 1985, **91 (Pt 1)**:111-120.
49. Stirewalt MA: **Schistosoma mansoni: cercaria to schistosomule**. *Adv Parasitol* 1974, **12**:115-182.
50. Mecozzi B, Rossi A, Lazzaretti P, Kady M, Kaiser S, Valle C, Cioli D, Klinkert MQ: **Molecular cloning of Schistosoma mansoni calcineurin subunits and immunolocalization to the excretory system**. *Mol Biochem Parasitol* 2000, **110**(2):333-343.
51. Bruckner DA, Voge M: **The nervous system of larval Schistosoma mansoni as revealed by acetylcholinesterase staining**. *J Parasitol* 1974, **60**(3):437-446.
52. Skuce PJ, Johnston CF, Fairweather I, Halton DW, Shaw C: **A confocal scanning laser microscope study of the peptidergic and serotoninergic components of the nervous system in larval Schistosoma mansoni**. *Parasitology* 1990, **101 Pt 2**:227-234.
53. Nuttman CJ: **The fine structure of ciliated nerve endings in the cercariae of Schistosoma mansoni**. *The Journal of Parasitology* 1971, **57**):855-859.
54. Cousin CE, Dorsey CH: **Nervous system of Schistosoma mansoni cercaria: organization and fine structure**. *Parasitol Res* 1991, **77**(2):132-141.
55. Robson RT, Erasmus DA: **The ultrastructure, based on stereoscan observations, of the oral sucker of the cercaria of Schistosoma mansoni with special reference to penetration**. *Z Parasitenkd* 1970, **35**(1):76-86.

56. McLaren DJ, Hockley DJ: **Schistosoma mansoni: the occurrence of microvilli on the surface of the tegument during transformation from cercaria to schistosomulum**. *Parasitology* 1976, **73**(2):169-187.
57. Skelly PJ, Kim JW, Cunningham J, Shoemaker CB: **Cloning, characterization, and functional expression of cDNAs encoding glucose transporter proteins from the human parasite Schistosoma mansoni**. *J Biol Chem* 1994, **269**(6):4247-4253.
58. Skelly PJ, Shoemaker CB: **Rapid appearance and asymmetric distribution of glucose transporter SGTP4 at the apical surface of intramammalian-stage Schistosoma mansoni**. *Proc Natl Acad Sci U S A* 1996, **93**(8):3642-3646.
59. Jiang J, Skelly PJ, Shoemaker CB, Caulfield JP: **Schistosoma mansoni: the glucose transport protein SGTP4 is present in tegumental multilamellar bodies, discoid bodies, and the surface lipid bilayers**. *Exp Parasitol* 1996, **82**(2):201-210.
60. Skelly PJ, Shoemaker CB: **Induction cues for tegument formation during the transformation of Schistosoma mansoni cercariae**. *Int J Parasitol* 2000, **30**(5):625-631.
61. Crabtree JE, Wilson RA: **Schistosoma mansoni: a scanning electron microscope study of the developing schistosomulum**. *Parasitology* 1980, **81**(Pt 3):553-564.
62. Thornhill J, Coelho PM, McVeigh P, Maule A, Jurberg AD, Kusel JR: **Schistosoma mansoni cercariae experience influx of macromolecules during skin penetration**. *Parasitology* 2009, **136**(11):1257-1267.
63. Crabtree JE: **An ultrastructural study of the migrating schistosomulum of *Schistosoma mansoni***. *D Phil thesis, University of York* 1982.
64. Harrop R, Wilson RA: **Protein synthesis and release by cultured schistosomula of Schistosoma mansoni**. *Parasitology* 1993, **107 (Pt 3)**:265-274.
65. Blanton RE, Licate LS: **Developmental regulation of protein synthesis in schistosomes**. *Mol Biochem Parasitol* 1992, **51**(2):201-208.
66. Skelly PJ, Shoemaker CB: **A molecular genetic study of the variations in metabolic function during schistosome development**. *Mem Inst Oswaldo Cruz* 1995, **90**(2):281-284.
67. Simpson AJ, Sher A, McCutchan TF: **The genome of Schistosoma mansoni: isolation of DNA, its size, bases and repetitive sequences**. *Mol Biochem Parasitol* 1982, **6**(2):125-137.
68. Franco GR, Adams MD, Soares MB, Simpson AJ, Venter JC, Pena SD: **Identification of new Schistosoma mansoni genes by the EST strategy using a directional cDNA library**. *Gene* 1995, **152**(2):141-147.
69. Franco GR, Rabelo EM, Azevedo V, Pena HB, Ortega JM, Santos TM, Meira WS, Rodrigues NA, Dias CM, Harrop R *et al*: **Evaluation of cDNA libraries from different developmental stages of Schistosoma mansoni for production of expressed sequence tags (ESTs)**. *DNA Res* 1997, **4**(3):231-240.
70. Dias Neto E, Harrop R, Correa-Oliveira R, Wilson RA, Pena SD, Simpson AJ: **Minilibraries constructed from cDNA generated by arbitrarily primed RT-PCR: an alternative to normalized libraries for the generation of ESTs from nanogram quantities of mRNA**. *Gene* 1997, **186**(1):135-142.
71. Dias Neto E, Correa RG, Verjovski-Almeida S, Briones MR, Nagai MA, da Silva W, Jr., Zago MA, Bordin S, Costa FF, Goldman GH *et al*: **Shotgun sequencing of the human transcriptome with ORF expressed sequence tags**. *Proc Natl Acad Sci U S A* 2000, **97**(7):3491-3496.
72. Santos TM, Johnston DA, Azevedo V, Ridgers IL, Martinez MF, Marotta GB, Santos RL, Fonseca SJ, Ortega JM, Rabelo EM *et al*: **Analysis of the gene expression profile**

of Schistosoma mansoni cercariae using the expressed sequence tag approach. *Mol Biochem Parasitol* 1999, **103**(1):79-97.
73. Verjovski-Almeida S, DeMarco R, Martins EA, Guimaraes PE, Ojopi EP, Paquola AC, Piazza JP, Nishiyama MY, Jr., Kitajima JP, Adamson RE *et al*: **Transcriptome analysis of the acoelomate human parasite Schistosoma mansoni**. *Nat Genet* 2003, **35**(2):148-157.
74. Haas BJ, Berriman M, Hirai H, Cerqueira GG, Loverde PT, El-Sayed NM: **Schistosoma mansoni genome: closing in on a final gene set**. *Exp Parasitol* 2007, **117**(3):225-228.
75. Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC, Mashiyama ST, Al-Lazikani B, Andrade LF, Ashton PD *et al*: **The genome of the blood fluke Schistosoma mansoni**. *Nature* 2009, **460**(7253):352-358.
76. McVeigh P, Mair GR, Atkinson L, Ladurner P, Zamanian M, Novozhilova E, Marks NJ, Day TA, Maule AG: **Discovery of multiple neuropeptide families in the phylum Platyhelminthes**. *Int J Parasitol* 2009.
77. Correnti JM, Brindley PJ, Pearce EJ: **Long-term suppression of cathepsin B levels by RNA interference retards schistosome growth**. *Mol Biochem Parasitol* 2005, **143**(2):209-215.
78. Skelly PJ, Da'dara A, Harn DA: **Suppression of cathepsin B expression in Schistosoma mansoni by RNA interference**. *Int J Parasitol* 2003, **33**(4):363-369.
79. Rinaldi G, Morales ME, Alrefaei YN, Cancela M, Castillo E, Dalton JP, Tort JF, Brindley PJ: **RNA interference targeting leucine aminopeptidase blocks hatching of Schistosoma mansoni eggs**. *Mol Biochem Parasitol* 2009, **167**(2):118-126.
80. Boyle JP, Wu XJ, Shoemaker CB, Yoshino TP: **Using RNA interference to manipulate endogenous gene expression in Schistosoma mansoni sporocysts**. *Mol Biochem Parasitol* 2003, **128**(2):205-215.
81. Brindley PJ, Pearce EJ: **Genetic manipulation of schistosomes**. *Int J Parasitol* 2007, **37**(5):465-473.
82. Geldhof P, Visser A, Clark D, Saunders G, Britton C, Gilleard J, Berriman M, Knox D: **RNA interference in parasitic helminths: current situation, potential pitfalls and future prospects**. *Parasitology* 2007, **134**(Pt 5):609-619.
83. de Moraes Mourao M, Dinguirard N, Franco GR, Yoshino TP: **Phenotypic Screen of Early-Developing Larvae of the Blood Fluke, Schistosoma mansoni, using RNA Interference**. *PLoS Negl Trop Dis* 2009, **3**(8):e502.
84. Heyers O, Walduck AK, Brindley PJ, Bleiss W, Lucius R, Dorbic T, Wittig B, Kalinna BH: **Schistosoma mansoni miracidia transformed by particle bombardment infect Biomphalaria glabrata snails and develop into transgenic sporocysts**. *Exp Parasitol* 2003, **105**(2):174-178.
85. Wippersteg V, Kapp K, Kunz W, Jackstadt WP, Zahner H, Grevelding CG: **HSP70-controlled GFP expression in transiently transformed schistosomes**. *Mol Biochem Parasitol* 2002, **120**(1):141-150.
86. Wippersteg V, Kapp K, Kunz W, Grevelding CG: **Characterisation of the cysteine protease ER60 in transgenic Schistosoma mansoni larvae**. *Int J Parasitol* 2002, **32**(10):1219-1224.
87. Beckmann S, Wippersteg V, El-Bahay A, Hirzmann J, Oliveira G, Grevelding CG: **Schistosoma mansoni: germ-line transformation approaches and actin-promoter analysis**. *Exp Parasitol* 2007, **117**(3):292-303.
88. Kines KJ, Morales ME, Mann VH, Gobert GN, Brindley PJ: **Integration of reporter transgenes into Schistosoma mansoni chromosomes mediated by pseudotyped murine leukemia virus**. *Faseb J* 2008, **22**(8):2936-2948.

89. Singh-Gasson S, Green RD, Yue Y, Nelson C, Blattner F, Sussman MR, Cerrina F: **Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array**. *Nat Biotechnol* 1999, **17**(10):974-978.
90. Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments**. *Stat Appl Genet Mol Biol* 2004, **3**:Article3.
91. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J *et al*: **Bioconductor: open software development for computational biology and bioinformatics**. *Genome Biol* 2004, **5**(10):R80.
92. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT *et al*: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium**. *Nat Genet* 2000, **25**(1):25-29.
93. Fitzpatrick JM, Johnston DA, Williams GW, Williams DJ, Freeman TC, Dunne DW, Hoffmann KF: **An oligonucleotide microarray for transcriptome analysis of Schistosoma mansoni and its application/use to investigate gender-associated gene expression**. *Mol Biochem Parasitol* 2005, **141**(1):1-13.
94. Hoffmann KF, Johnston DA, Dunne DW: **Identification of Schistosoma mansoni gender-associated gene transcripts by cDNA microarray profiling**. *Genome Biol* 2002, **3**(8):RESEARCH0041.
95. Gobert GN, McManus DP, Nawaratna S, Moertel L, Mulvenna J, Jones MK: **Tissue Specific Profiling of Females of Schistosoma japonicum by Integrated Laser Microdissection Microscopy and Microarray Analysis**. *PLoS Negl Trop Dis* 2009, **3**(6):e469.
96. Vermeire JJ, Taft AS, Hoffmann KF, Fitzpatrick JM, Yoshino TP: **Schistosoma mansoni: DNA microarray gene expression profiling during the miracidium-to-mother sporocyst transformation**. *Mol Biochem Parasitol* 2006, **147**(1):39-47.
97. Jolly ER, Chin CS, Miller S, Bahgat MM, Lim KC, DeRisi J, McKerrow JH: **Gene expression patterns during adaptation of a helminth parasite to different environmental niches**. *Genome Biol* 2007, **8**(4):R65.
98. Dillon GP, Feltwell T, Skelton JP, Ashton PD, Coulson PS, Quail MA, Nikolaidou-Katsaridou N, Wilson RA, Ivens AC: **Microarray analysis identifies genes preferentially expressed in the lung schistosomulum of Schistosoma mansoni**. *Int J Parasitol* 2006, **36**(1):1-8.
99. Fitzpatrick JM, Peak E, Perally S, Chalmers IW, Barrett J, Yoshino TP, Ivens AC, Hoffmann KF: **Anti-schistosomal Intervention Targets Identified by Lifecycle Transcriptomic Analyses**. *PLoS Negl Trop Dis* 2009, **3**(11):e543.
100. Waisberg M, Lobo FP, Cerqueira GC, Passos LK, Carvalho OS, El-Sayed NM, Franco GR: **Schistosoma mansoni: Microarray analysis of gene expression induced by host sex**. *Exp Parasitol* 2008, **120**(4):357-363.
101. Dillon GP, Feltwell T, Skelton J, Coulson PS, Wilson RA, Ivens AC: **Altered patterns of gene expression underlying the enhanced immunogenicity of radiation-attenuated schistosomes**. *PLoS Negl Trop Dis* 2008, **2**(5):e240.
102. Verjovski-Almeida S, Venancio TM, Oliveira KC, Almeida GT, DeMarco R: **Use of a 44k oligoarray to explore the transcriptome of Schistosoma mansoni adult worms**. *Exp Parasitol* 2007, **117**(3):236-245.
103. Gobert GN, McInnes R, Moertel L, Nelson C, Jones MK, Hu W, McManus DP: **Transcriptomics tool for the human Schistosoma blood flukes using microarray gene expression profiling**. *Exp Parasitol* 2006, **114**(3):160-172.
104. Lawson JR, Wilson RA: **Metabolic changes associated with the migration of the schistosomulum of Schistosoma mansoni in the mammal host**. *Parasitology* 1980, **81**(2):325-336.

105. Fitzpatrick JM, Protasio AV, McArdle AJ, Williams GA, Johnston DA, Hoffmann KF: **Use of Genomic DNA as an Indirect Reference for Identifying Gender-Associated Transcripts in Morphologically Identical, but Chromosomally Distinct, Schistosoma mansoni Cercariae**. *PLoS Negl Trop Dis* 2008, **2**(10):e323.
106. Clegg JA: **In Vitro Cultivation Of Schistosoma Mansoni**. *Exp Parasitol* 1965, **16**:133-147.
107. Braschi S, Wilson RA: **Proteins exposed at the adult schistosome surface revealed by biotinylation**. *Mol Cell Proteomics* 2006, **5**(2):347-356.
108. Webster M, Fulford AJ, Braun G, Ouma JH, Kariuki HC, Havercroft JC, Gachuhi K, Sturrock RF, Butterworth AE, Dunne DW: **Human immunoglobulin E responses to a recombinant 22.6-kilodalton antigen from Schistosoma mansoni adult worms are associated with low intensities of reinfection after treatment**. *Infect Immun* 1996, **64**(10):4042-4046.
109. Gordon RM, Griffiths RB: **Observations on the means by which the cercariae of Schistosoma mansoni penetrate mammalian skin, together with an account of certain morphological changes observed in the newly penetrated larvae**. *Ann Trop Med Parasitol* 1951, **45**(3-4):227-243.
110. Dresden MH, Lewis JC, Krisko I: **Proteolytic action of Schistosoma mansoni cercarial proteases on keratin and basement membrane proteins**. *J Parasitol* 1977, **63**(5):941-943.
111. McKerrow JH, Jones P, Sage H, Pino-Heiss S: **Proteinases from invasive larvae of the trematode parasite Schistosoma mansoni degrade connective-tissue and basement-membrane macromolecules**. *Biochem J* 1985, **231**(1):47-51.
112. Marikovsky M, Arnon R, Fishelson Z: **Schistosoma mansoni: localization of the 28 kDa secreted protease in cercaria**. *Parasite Immunol* 1990, **12**(4):389-401.
113. Salter JP, Lim KC, Hansell E, Hsieh I, McKerrow JH: **Schistosome invasion of human skin and degradation of dermal elastin are mediated by a single serine protease**. *J Biol Chem* 2000, **275**(49):38667-38673.
114. Newport GR, McKerrow JH, Hedstrom R, Petitt M, McGarrigle L, Barr PJ, Agabian N: **Cloning of the proteinase that facilitates infection by schistosome parasites**. *J Biol Chem* 1988, **263**(26):13179-13184.
115. Salter JP, Choe Y, Albrecht H, Franklin C, Lim KC, Craik CS, McKerrow JH: **Cercarial elastase is encoded by a functionally conserved gene family across multiple species of schistosomes**. *J Biol Chem* 2002, **277**(27):24618-24624.
116. Curwen RS, Ashton PD, Sundaralingam S, Wilson RA: **Identification of novel proteases and immunomodulators in the secretions of schistosome cercariae that facilitate host entry**. *Mol Cell Proteomics* 2006, **5**(5):835-844.
117. Dvorak J, Mashiyama ST, Braschi S, Sajid M, Knudsen GM, Hansell E, Lim KC, Hsieh I, Bahgat M, Mackenzie B *et al*: **Differential use of protease families for invasion by schistosome cercariae**. *Biochimie* 2008, **90**(2):345-358.
118. Hansell E, Braschi S, Medzihradszky KF, Sajid M, Debnath M, Ingram J, Lim KC, McKerrow JH: **Proteomic analysis of skin invasion by blood fluke larvae**. *PLoS Negl Trop Dis* 2008, **2**(7):e262.
119. Knudsen GM, Medzihradszky KF, Lim KC, Hansell E, McKerrow JH: **Proteomic analysis of Schistosoma mansoni cercarial secretions**. *Mol Cell Proteomics* 2005, **4**(12):1862-1875.
120. Rao KV, Ramaswamy K: **Cloning and expression of a gene encoding Sm16, an anti-inflammatory protein from Schistosoma mansoni**. *Mol Biochem Parasitol* 2000, **108**(1):101-108.

121. Asojo OA, Goud G, Dhar K, Loukas A, Zhan B, Deumic V, Liu S, Borgstahl GE, Hotez PJ: **X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, Necator americanus, and a vaccine antigen for human hookworm infection**. *J Mol Biol* 2005, **346**(3):801-814.
122. Murray J, Gregory WF, Gomez-Escobar N, Atmadja AK, Maizels RM: **Expression and immune recognition of Brugia malayi VAL-1, a homologue of vespid venom allergens and Ancylostoma secreted proteins**. *Mol Biochem Parasitol* 2001, **118**(1):89-96.
123. Tawe W, Pearlman E, Unnasch TR, Lustigman S: **Angiogenic activity of Onchocerca volvulus recombinant proteins similar to vespid venom antigen 5**. *Mol Biochem Parasitol* 2000, **109**(2):91-99.
124. Bethony JM, Simon G, Diemert DJ, Parenti D, Desrosiers A, Schuck S, Fujiwara R, Santiago H, Hotez PJ: **Randomized, placebo-controlled, double-blind trial of the Na-ASP-2 hookworm vaccine in unexposed adults**. *Vaccine* 2008, **26**(19):2408-2417.
125. Chalmers IW, McArdle AJ, Coulson RM, Wagner MA, Schmid R, Hirai H, Hoffmann KF: **Developmentally regulated expression, alternative splicing and distinct sub-groupings in members of the Schistosoma mansoni venom allergen-like (SmVAL) gene family**. *BMC Genomics* 2008, **9**:89.
126. Korolkova YV, Bocharov EV, Angelo K, Maslennikov IV, Grinenko OV, Lipkin AV, Nosyreva ED, Pluzhnikov KA, Olesen SP, Arseniev AS *et al*: **New binding site on common molecular scaffold provides HERG channel specificity of scorpion toxin BeKm-1**. *J Biol Chem* 2002, **277**(45):43104-43109.
127. Panyi G, Varga Z, Gaspar R: **Ion channels and lymphocyte activation**. *Immunol Lett* 2004, **92**(1-2):55-66.
128. Wilson RA, Ashton PD, Braschi S, Dillon GP, Berriman M, Ivens A: **'Oming in on schistosomes: prospects and limitations for post-genomics**. *Trends Parasitol* 2007, **23**(1):14-20.
129. DeMarco R, Verjovski-Almeida S: **Schistosomes--proteomics studies for potential novel vaccines and drug targets**. *Drug Discov Today* 2009, **14**(9-10):472-478.
130. Wilson RA, Coulson PS: **Schistosome vaccines: a critical appraisal**. *Mem Inst Oswaldo Cruz* 2006, **101 Suppl 1**:13-20.
131. Curwen RS, Ashton PD, Johnston DA, Wilson RA: **The Schistosoma mansoni soluble proteome: a comparison across four life-cycle stages**. *Mol Biochem Parasitol* 2004, **138**(1):57-66.
132. Meuleman EA, Holzmann PJ, Peet RC: **The development of daughter sporocysts inside the mother sporocyst of Schistosoma mansoni with special reference to the ultrastructure of the body wall**. *Z Parasitenkd* 1980, **61**(3):201-212.
133. Bahia D, Avelar LG, Vigorosi F, Cioli D, Oliveira GC, Mortara RA: **The distribution of motor proteins in the muscles and flame cells of the Schistosoma mansoni miracidium and primary sporocyst**. *Parasitology* 2006, **133**(Pt 3):321-329.
134. Ramalho-Pinto FJ, Gazzinelli G, Howells RE, Mota-Santos TA, Figueiredo EA, Pellegrino J: **Schistosoma mansoni: defined system for stepwise transformation of cercaria to schistosomule in vitro**. *Exp Parasitol* 1974, **36**(3):360-372.
135. Jourdane J, Theron A: **Schistosoma mansoni: cloning by microsurgical transplantation of sporocysts**. *Exp Parasitol* 1980, **50**(3):349-357.
136. Bolstad BM, Irizarry RA, Astrand M, Speed TP: **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias**. *Bioinformatics* 2003, **19**(2):185-193.

137. Hochberg Y, Benjamini Y: **More powerful procedures for multiple significance testing**. *Stat Med* 1990, **9**(7):811-818.
138. Rawlings ND, Morton FR, Kok CY, Kong J, Barrett AJ: **MEROPS: the peptidase database**. *Nucleic Acids Res* 2008, **36**(Database issue):D320-325.
139. Stewart GW, Hepworth-Jones BE, Keen JN, Dash BC, Argent AC, Casimir CM: **Isolation of cDNA coding for an ubiquitous membrane protein deficient in high Na+, low K+ stomatocytic erythrocytes**. *Blood* 1992, **79**(6):1593-1601.
140. Braschi S, Curwen RS, Ashton PD, Verjovski-Almeida S, Wilson A: **The tegument surface membranes of the human blood parasite Schistosoma mansoni: a proteomic analysis after differential extraction**. *Proteomics* 2006, **6**(5):1471-1482.
141. Hall SL: **The use of proteomic techniques to identify antibody targets in schistosomiasis**. University of York; 2005.
142. DeMarco R, Mathieson W, Manuel SJ, Dillon GP, Curwen RS, Ashton PD, Ivens A, Berriman M, Verjovski-Almeida S, Wilson RA: **Protein variation in blood-dwelling schistosome worms generated by differential splicing of micro-exon gene transcripts.** *Submitted* 2009.
143. DeMarco R, Mathieson W, Dillon GP, Wilson RA: **Schistosome albumin is of host, not parasite, origin**. *Int J Parasitol* 2007, **37**(11):1201-1208.
144. Newport G, McKerrow J, Hedstrom R, Culpepper J, McGarrigle L, Agabian N: **Schistosome elastases: biological importance, structure, function and stage-specific expression**. *Biochem Soc Symp* 1987, **53**:115-121.
145. Hockley DJ, McLaren DJ: **Schistosoma mansoni: changes in the outer membrane of the tegument during development from cercaria to adult worm**. *Int J Parasitol* 1973, **3**(1):13-25.
146. Pollok S, Bauerschmidt C, Sanger J, Nasheuer HP, Grosse F: **Human Cdc45 is a proliferation-associated antigen**. *Febs J* 2007, **274**(14):3669-3684.
147. Miller CL, Day TA, Bennett JL, Pax RA: **Schistosoma mansoni: L-glutamate-induced contractions in isolated muscle fibers; evidence for a glutamate transporter**. *Exp Parasitol* 1996, **84**(3):410-419.
148. Brossier F, Starnes GL, Beatty WL, Sibley LD: **Microneme rhomboid protease TgROM1 is required for efficient intracellular growth of Toxoplasma gondii**. *Eukaryot Cell* 2008, **7**(4):664-674.
149. Dowse TJ, Pascall JC, Brown KD, Soldati D: **Apicomplexan rhomboids have a potential role in microneme protein cleavage during host cell invasion**. *Int J Parasitol* 2005, **35**(7):747-756.
150. Baxt LA, Baker RP, Singh U, Urban S: **An Entamoeba histolytica rhomboid protease with atypical specificity cleaves a surface lectin involved in phagocytosis and immune evasion**. *Genes Dev* 2008, **22**(12):1636-1646.
151. Freeman M: **Rhomboids: 7 years of a new protease family**. *Semin Cell Dev Biol* 2009, **20**(2):231-239.
152. McCarthy E, Stack C, Donnelly SM, Doyle S, Mann VH, Brindley PJ, Stewart M, Day TA, Maule AG, Dalton JP: **Leucine aminopeptidase of the human blood flukes, Schistosoma mansoni and Schistosoma japonicum**. *Int J Parasitol* 2004, **34**(6):703-714.
153. Bos DH, Mayfield C, Minchella DJ: **Analysis of regulatory protease sequences identified through bioinformatic data mining of the Schistosoma mansoni genome**. *BMC Genomics* 2009, **10**:488.
154. Marino G, Uria JA, Puente XS, Quesada V, Bordallo J, Lopez-Otin C: **Human autophagins, a family of cysteine proteinases potentially implicated in cell degradation by autophagy**. *J Biol Chem* 2003, **278**(6):3671-3678.

155. Sorimachi H, Ishiura S, Suzuki K: **Structure and physiological function of calpains**. *Biochem J* 1997, **328 (Pt 3)**:721-732.
156. Wong JY, Harrop SA, Day SR, Brindley PJ: **Schistosomes express two forms of cathepsin D**. *Biochim Biophys Acta* 1997, **1338**(2):156-160.
157. Morales ME, Rinaldi G, Gobert GN, Kines KJ, Tort JF, Brindley PJ: **RNA interference of Schistosoma mansoni cathepsin D, the apical enzyme of the hemoglobin proteolysis cascade**. *Mol Biochem Parasitol* 2008, **157**(2):160-168.
158. Skelly PJ, Stein LD, Shoemaker CB: **Expression of Schistosoma mansoni genes involved in anaerobic and oxidative glucose metabolism during the cercaria to adult transformation**. *Mol Biochem Parasitol* 1993, **60**(1):93-104.
159. Lesage F, Guillemare E, Fink M, Duprat F, Lazdunski M, Romey G, Barhanin J: **TWIK-1, a ubiquitous human weakly inward rectifying K+ channel with a novel structure**. *Embo J* 1996, **15**(5):1004-1011.
160. Faghiri Z, Skelly PJ: **The role of tegumental aquaporin from the human parasitic worm, Schistosoma mansoni, in osmoregulation and drug uptake**. *Faseb J* 2009, **23**(8):2780-2789.
161. Rofatto HK, Tararam CA, Borges WC, Wilson RA, Leite LC, Farias LP: **Characterization of phosphodiesterase-5 as a surface protein in the tegument of Schistosoma mansoni**. *Mol Biochem Parasitol* 2009, **166**(1):32-41.
162. Braschi S, Borges WC, Wilson RA: **Proteomic analysis of the schistosome tegument and its surface membranes**. *Mem Inst Oswaldo Cruz* 2006, **101 Suppl 1**:205-212.
163. van Balkom BW, van Gestel RA, Brouwers JF, Krijgsveld J, Tielens AG, Heck AJ, van Hellemond JJ: **Mass spectrometric analysis of the Schistosoma mansoni tegumental sub-proteome**. *J Proteome Res* 2005, **4**(3):958-966.
164. Davies A, Simmons DL, Hale G, Harrison RA, Tighe H, Lachmann PJ, Waldmann H: **CD59, an LY-6-like protein expressed in human lymphoid cells, regulates the action of the complement membrane attack complex on homologous cells**. *J Exp Med* 1989, **170**(3):637-654.
165. Rollins SA, Sims PJ: **The complement-inhibitory activity of CD59 resides in its capacity to block incorporation of C9 into membrane C5b-9**. *J Immunol* 1990, **144**(9):3478-3483.
166. Loukas A, Tran M, Pearson MS: **Schistosome membrane proteins as vaccines**. *Int J Parasitol* 2007, **37**(3-4):257-263.
167. Tran MH, Pearson MS, Bethony JM, Smyth DJ, Jones MK, Duke M, Don TA, McManus DP, Correa-Oliveira R, Loukas A: **Tetraspanins on the surface of Schistosoma mansoni are protective antigens against schistosomiasis**. *Nat Med* 2006, **12**(7):835-840.
168. Bhardwaj R, Skelly PJ: **Purinergic signaling and immune modulation at the schistosome surface?** *Trends Parasitol* 2009, **25**(6):256-260.
169. Kyes SA, Kraemer SM, Smith JD: **Antigenic variation in Plasmodium falciparum: gene organization and regulation of the var multigene family**. *Eukaryot Cell* 2007, **6**(9):1511-1520.
170. Taylor JE, Rudenko G: **Switching trypanosome coats: what's in the wardrobe?** *Trends Genet* 2006, **22**(11):614-620.
171. Hawdon JM, Jones BF, Hoffman DR, Hotez PJ: **Cloning and characterization of Ancylostoma-secreted protein. A novel protein associated with the transition to parasitism by infective hookworm larvae**. *J Biol Chem* 1996, **271**(12):6672-6678.
172. Mathieson W, Wilson RA: **A comparative proteomic study of the undeveloped and developed Schistosoma mansoni egg and its contents: The miracidium, hatch fluid and secretions**. *Int J Parasitol* 2009.

173. Kwatia MA, Botkin DJ, Williams DL: **Molecular and enzymatic characterization of Schistosoma mansoni thioredoxin peroxidase**. *J Parasitol* 2000, **86**(5):908-915.
174. Sayed AA, Simeonov A, Thomas CJ, Inglese J, Austin CP, Williams DL: **Identification of oxadiazoles as new drug leads for the control of schistosomiasis**. *Nat Med* 2008, **14**(4):407-412.
175. Dalton JP, Smith AM, Clough KA, Brindley PJ: **Digestion of haemoglobin by schistosomes: 35 years on**. *Parasitol Today* 1995, **11**(8):299-303.
176. Brindley PJ, Kalinna BH, Dalton JP, Day SR, Wong JY, Smythe ML, McManus DP: **Proteolytic degradation of host hemoglobin by schistosomes**. *Mol Biochem Parasitol* 1997, **89**(1):1-9.
177. Caffrey CR, McKerrow JH, Salter JP, Sajid M: **Blood 'n' guts: an update on schistosome digestive peptidases**. *Trends Parasitol* 2004, **20**(5):241-248.
178. Ruppel A, Diesfeld HJ, Rother U: **Immunoblot analysis of Schistosoma mansoni antigens with sera of schistosomiasis patients: diagnostic potential of an adult schistosome polypeptide**. *Clin Exp Immunol* 1985, **62**(3):499-506.
179. Dalton JP, Brindley PJ: **Schistosome asparaginyl endopeptidase SM32 in hemoglobin digestion**. *Parasitol Today* 1996, **12**(3):125.
180. Caffrey CR, Mathieu MA, Gaffney AM, Salter JP, Sajid M, Lucas KD, Franklin C, Bogyo M, McKerrow JH: **Identification of a cDNA encoding an active asparaginyl endopeptidase of Schistosoma mansoni and its expression in Pichia pastoris**. *FEBS Lett* 2000, **466**(2-3):244-248.
181. Sajid M, McKerrow JH, Hansell E, Mathieu MA, Lucas KD, Hsieh I, Greenbaum D, Bogyo M, Salter JP, Lim KC *et al*: **Functional expression and characterization of Schistosoma mansoni cathepsin B and its trans-activation by an endogenous asparaginyl endopeptidase**. *Mol Biochem Parasitol* 2003, **131**(1):65-75.
182. Don TA, Bethony JM, Loukas A: **Saposin-like proteins are expressed in the gastrodermis of Schistosoma mansoni and are immunogenic in natural infections**. *Int J Infect Dis* 2008, **12**(6):e39-47.
183. Skelly PJ, Shoemaker CB: **Schistosoma mansoni proteases Sm31 (cathepsin B) and Sm32 (legumain) are expressed in the cecum and protonephridia of cercariae**. *J Parasitol* 2001, **87**(5):1218-1221.
184. Vesque C, Ellis S, Lee A, Szabo M, Thomas P, Beddington R, Placzek M: **Development of chick axial mesoderm: specification of prechordal mesoderm by anterior endoderm-derived TGFbeta family signalling**. *Development* 2000, **127**(13):2795-2809.
185. Ramaswamy K, Salafsky B, Potluri S, He YX, Li JW, Shibuya T: **Secretion of an anti-inflammatory, immunomodulatory factor by Schistosomulae of Schistosoma mansoni**. *J Inflamm* 1995, **46**(1):13-22.
186. Holmfeldt P, Brannstrom K, Sellin ME, Segerman B, Carlsson SR, Gullberg M: **The Schistosoma mansoni protein Sm16/SmSLP/SmSPO-1 is a membrane-binding protein that lacks the proposed microtubule-regulatory activity**. *Mol Biochem Parasitol* 2007, **156**(2):225-234.
187. Brannstrom K, Sellin ME, Holmfeldt P, Brattsand M, Gullberg M: **The Schistosoma mansoni protein Sm16/SmSLP/SmSPO-1 assembles into a nine-subunit oligomer with potential To inhibit Toll-like receptor signaling**. *Infect Immun* 2009, **77**(3):1144-1154.
188. Cassimeris L: **The oncoprotein 18/stathmin family of microtubule destabilizers**. *Curr Opin Cell Biol* 2002, **14**(1):18-24.

189. Valle C, Festucci A, Calogero A, Macri P, Mecozzi B, Liberti P, Cioli D: **Stage-specific expression of a Schistosoma mansoni polypeptide similar to the vertebrate regulatory protein stathmin**. *J Biol Chem* 1999, **274**(48):33869-33874.
190. Nuttman CJ: **The stucture and behaviour of the cercaria of Schistosoma mansoni**. University of York; 1975.