Chapter 3

Transcriptome analysis of male and female *S. mansoni* during intra-mammalian development

3.1 Introduction

The aim of this Chapter is to explore the transcriptome changes that occur during the development of male and female schistosomes inside the mammalian host, with a particular emphasis on sexual maturation taking place after pairing. As discussed in detail in the introduction to this thesis, female maturation is dependent on stimulation by a male partner. Here, I examined the transcriptomes of female and male worms in the days leading up to pairing as well as the weeks following pairing. In particular, I wanted to explore how the male and female transcriptomes change over time and how the gene expression of worms from mixed sex (MS) infections differs from that of single sex (SS) infections.

In this Chapter, a variety of functional genomic tools and databases were used to investigate differentially expressed genes. Using a combination of different tools, such as GO terms, KEGG pathway and Pfam domain enrichment, I was able to get a more thorough understanding of the processes taking place in the developing worms.

While many previous studies on pairing-induced changes in gene expression used microarrays (Fitzpatrick & Hoffmann, 2006; Leutner *et al*., 2013), a recent study used RNA-Seq (Sun *et al.*, 2014). An analysis of female worms at 18 and 23 d.p.i. from both SS and MS infections was performed, to study female gene expression in the presence or absence of male worms, but also the changes that take place between the two time points. Although almost 3000 differentially expressed genes (DEGs) were identified in some of their comparisons, their

approach left room for improvement. Their method of RNA-Seq relied on sequencing tags, which are short, 17 base-pair sub-region of a whole cDNA sequence. As result the tags do not cover transcripts along their entire lengths which can make gene expression less quantifiable (Wang *et al.*, 2009). More importantly, the authors did not sequence any biological replicates for any of their samples for their conditions, although this has now become standard practice as it allows for a more robust statistical analysis by making it possible for the variation within treatment groups to be measured accurately (Schurch *et al.*, 2016). This is because variability between biological replicates is often higher than expected, so without biological replicates real differences in expression between treatment groups cannot be reliably distinguished from random fluctuations in expression (Schurch *et al.*, 2016). The method of determining differentially expressed genes is not described in detail in the methods (the authors refer to "a rigorous algorithm [that] was adopted to screen genes that were differentially expressed between samples"), making it difficult to replicate their analysis and their results (Sun *et al*., 2014). Finally, the authors performed a pathway analysis using th identified DEGs (Sun *et al.*, 2014). However, in their analysis Sun *et al.* (2014) did not evaluate the proportion of genes that were differentially expressed in each pathway and therefore did not examine the statistical significance of the up- or down-regulated pathways that they reported. For example, Sun *et al.* (2014) report that "pathways, such as cell adhesion molecules (CAMs), ribosome biogenesis in eukaryotes, the Notch signalling pathway" were up-regulated. However, it is not clear if what percentage of genes of a particular pathway were differentially expressed. A low proportion of DEGs

in a pathway could be the result of random chance, whereas a larger proportion could give greater conficdence in the result.

Another study examining the development of *S. mansoni* from cercaria to adult worm (Picard *et al.*, 2016). However, the results were limited by the fact that female specimens did not experience pairing and thus did not commence sexual maturation over the course of the experiment (Picard *et al.*, 2016). Most recently, Wang *et al*. (2017) published a RNA-seq time course of the intramammalian development of male and female *S. japonicum* from 14 to 28 d.p.i.. The results suggested that biogenic amines and insect-hormone-like receptors play crucial roles in male-female signalling and sexual maturation of ovaries and vitelline glands in female worms (see Chapter 1.6.2). Using RNAi targeted at the insect-hormone-like receptor, the authors provided evidence of a role of the insect-hormone-like receptor in the development of these reproductive tissues (Wang *et al.*, 2017). However, no RNAi results were reported for the AADC gene involved in the production of amines. The authors only analysed the transcriptomes of worms from MS infections. This makes it more difficult to distinguish the changes in expression linked to pairing from those that occur independently of pairing as the worms develop without the opposite sex. Notably, the development of *S. japonicum* as described by Wang *et al.* (2017) appears to be accelerated when compared to *S. mansoni* (see Chapter 1.2.1), as the worms were observed to pair and reach sexual maturity consistently earlier than in equivalent studies on *S. mansoni* (Biolchini *et al*, 2006). Differences in the speed of development, even between strains of the same species of schistosome, have been reported before (Hsü *et al.*, 1962).

In my experiment, male and female *S. mansoni* from MS and SS infections were collected and processed for RNA-Seq at six time points: days 18, 21, 28, 35, 38 and 49 d.p.i.. My analysis included not just the transcriptome changes taking place in female worms but also those of male worms. Robust statistical analyses were used to determine differentially expressed genes and pathway enrichment analysis. Additional bioinformatics tools were used, including Pfam domain analysis, GO terms and cluster analysis (see Chapter 2.4), to infer the potential roles of differentially expressed genes in the development and sexual maturation. Finally, I used whole mount *in situ* hybridisation and RNAi to characterised two genes found to be involved in female sexual maturity. This was done to explore the site of expression as well as the function of these genes in *S. mansoni* using respectively.

3.2 Results

3.2.1 Egg laying

The livers of infected mice were examined for the presence of *S. mansoni* eggs (Figure 3.1), as described in Chapter 2.1.3. Schistosome eggs were found exclusively in livers of mice infected with MS parasites at 35 and 38 d.p.i. As expected, eggs were not observed in the livers of mice with the SS infections. Although paired worms were recovered at 28 d.p.i. from mice with MS infections, no eggs could be recovered from the livers of these mice. Since pairing is thought to take place from around 28 d.p.i. (Basch, 1991) most female worms would not

started egg laying at this time point. However, a very low egg count could have gone undetected.

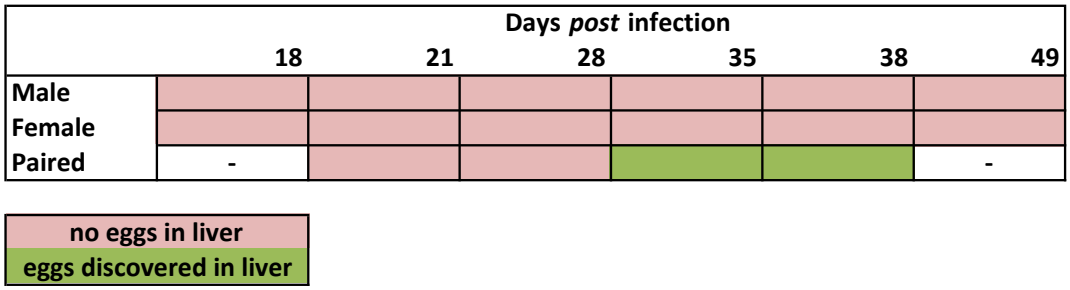


Figure 3.1: Eggs were only detected in the livers of mice with mixed sex infections. Single sex infections with male or female worms were found not to yield eggs at any time point, where as mixed sex infections were found to yield eggs from 35 d.p.i., but not 21 and 28 d.p.i. At 18 and 49 d.p.i. no mixed sex infection were available ("-").

3.2.2 Sequencing and sample clustering

The RNA-Seq data was prepared as described in Chapter 2.4, yielding over 20 million reads for each of the 70 samples (see Appendix A.1 for details). On average, 79% of the reads could be mapped to the *S. mansoni* genome (v5.2) (Appendix A.1), with the remainder mostly comprising PCR artefacts, especially primer dimers, created during the amplification of the sequencing library. Samples were then analysed using DESeq2. First, a principal component analysis (PCA) plot (Figure 3.2) was created of all samples. This gives a simplified overview of how developing male and female worms differ at the transcriptome level across the time course from 18 to 49 d.p.i. in MS and SS infections. While male worms from SS and MS infections develop at a similar rate, females from SS infections are known to become arrested at an early developmental stage (Popiel, 1986). This was reflected at the transcriptome level by the PCA plot as well as the following differential expression analysis (see Chapters 3.2.4 and

3.2.5). From 21 d.p.i. the gene expression of SS females remained almost unchanged, developmentally arrested, and most similar to MS females at 21 d.p.i. Females from MS infections became very distinct from all other samples from 28 to 38 d.p.i. at the transcriptome level.

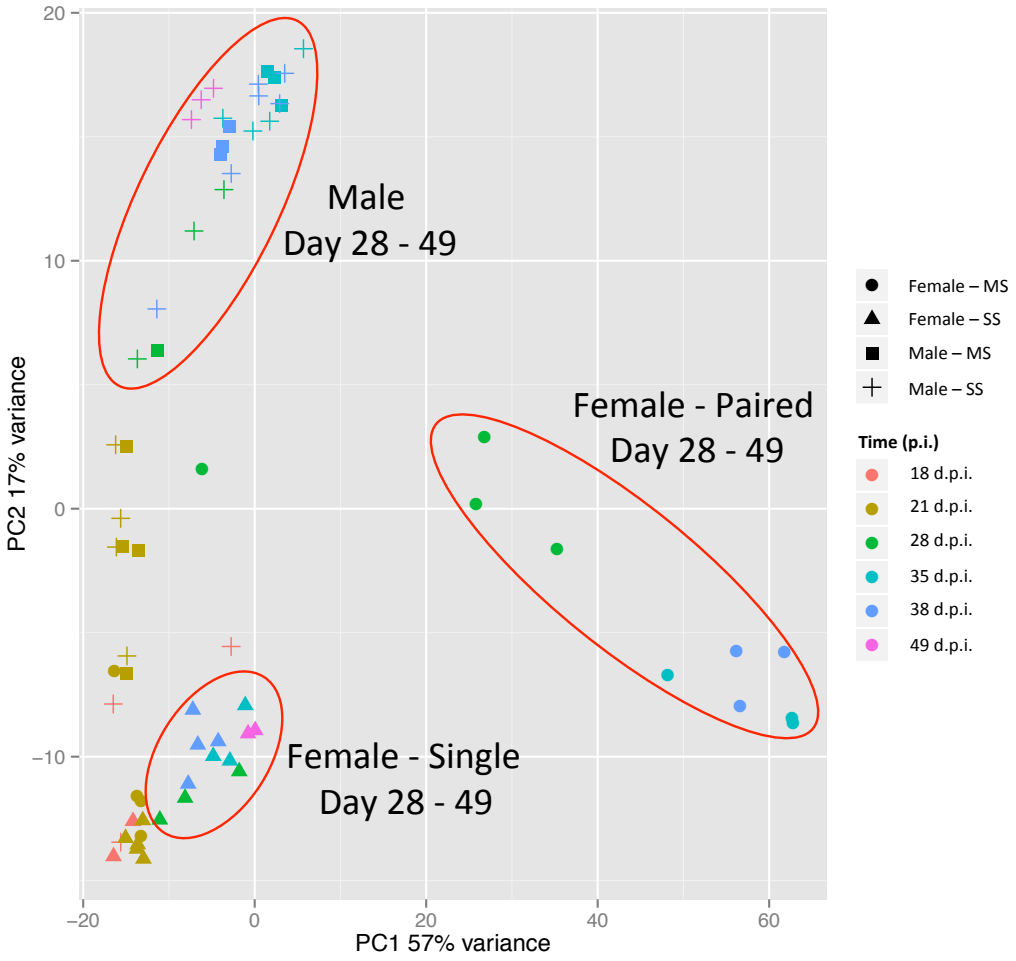


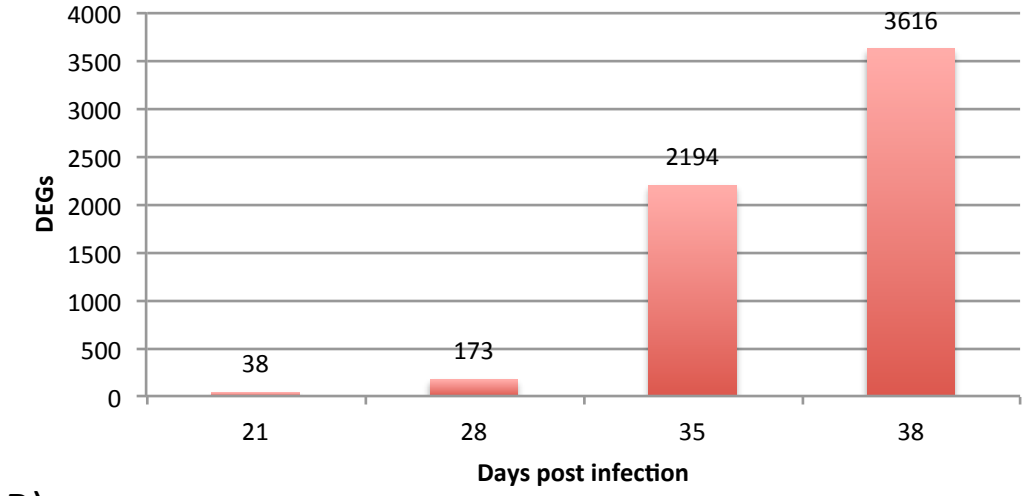
Figure 3.2: Samples group into male, as well as mixed and single sex female clusters. Principle Component Analysis (PCA) plot of gene expression in developing male and female *S.mansoni* from single sex (SS) and mixed sex (MS) infections at 18 – 49 days *post* infection (d.p.i.) PC1 explains 57% of the variance observed in the data set and broadly correlates with female sexual maturity over time, whereas PC2 explains only 17% of the variance and broadly correlates to male maturity. Together 74% of variance across the data set is captured by the first two principal components.

These findings reflected that female sexual development depends on pairing with a male whereas male maturation is largely independent of the female, as has been noted in the literature (Basch, 1991). In this context, it was also interesting to note that female worms from MS and SS infections were transcriptomically very similar at 21 d.p.i., but diverged significantly after 28 d.p.i., when pairing is thought to take place (Basch, 1991). It should also be noted that male and female samples from MS infections separate very well at all available time points (it should be noted that the males and females from SS infections at 18 d.p.i. did not separate). This shows that males and females were correctly identified even at 21 d.p.i. using morphological differences outlined by Basch (1991).

3.2.3 Gender-specific gene expression

During their intra-mammalian development, male and female worms grow from morphologically similar immature stages to distinct male and female worms that are usually reproductively active by 35 d.p.i. (Biolchini *et al.*, 2006). Pearce and Huang (2015) suggested that the sexual dimorphism of male and female worms can be explained by adaptations for different biological roles. This allows for a unique co-dependency that maximises egg production. The male, for instance, is responsible for moving the pair to the site of oviposition (Basch, 1991).

A)


Males compared to Females
DEGs
4000
3500
3000
2500
2000
1500
1000
500
0
38
173
2194
3616
21
28
35
38
Days post infection

B)

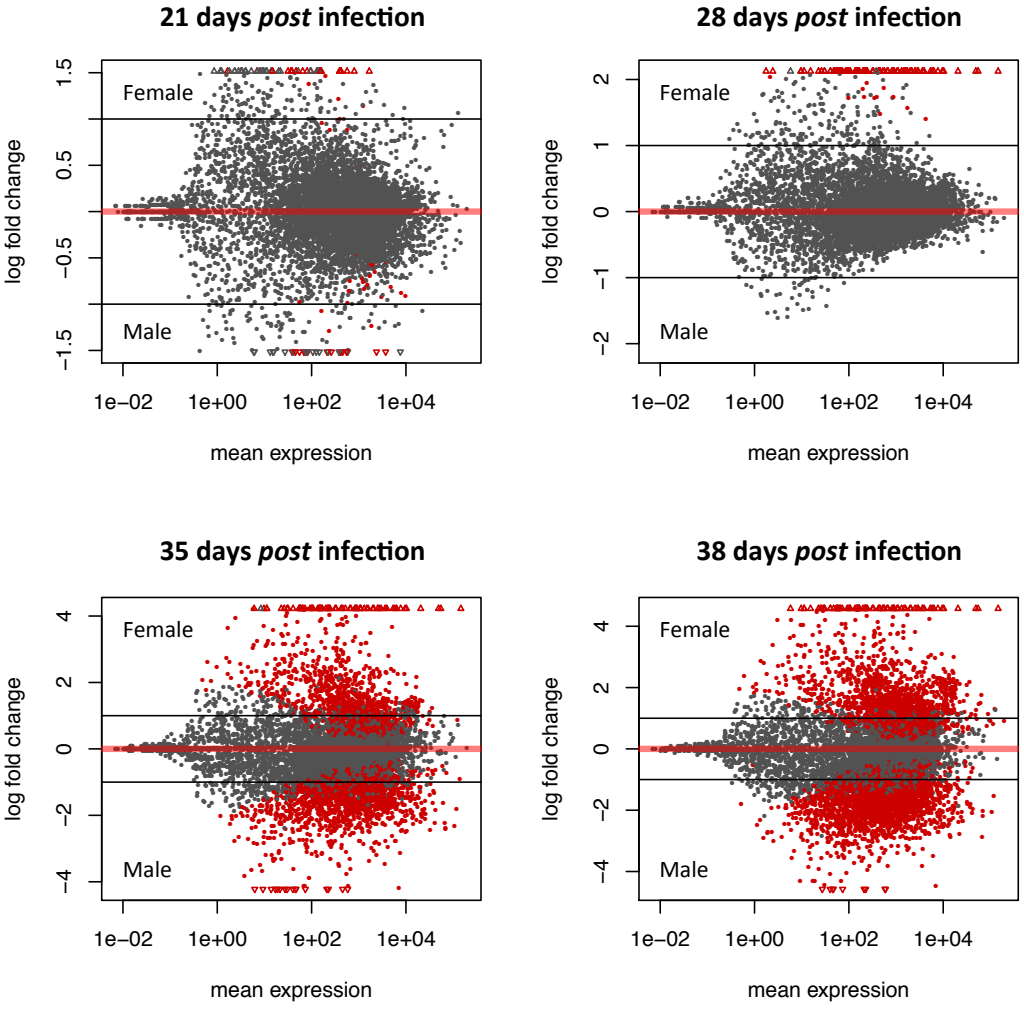

21 days post infection
28 days post infection
35 days post infection
38 days post infection
Female
Male
log fold change
mean expression
1e−02
1e+00
1e+02
1e+04

Figure 3.3: The male and female transcriptome becomes increasingly different, especially from 28 to 35 days *post* infection. A) Number of differentially expressed genes between male and female worms from mixed sex infections at 21 to 38 days *post* infection. Genes were considered to be differentially expressed if the adjusted p-value of < 0.01 and fold-change > 2. B) Log ratio vs mean (MA) plots of pair-wise comparisons of gene expression in male and female worms from MS infections from 21 to 38 d.p.i. Genes with an adjusted p-value of < 0.01 are highlighted in red, the others are grey. The black lines represent the fold change cut-off (log-fold change = ± 1). Triangles indicate genes outside the boundaries of the plots; like other genes, they are red if they are statistically significant, grey if they are not.

In order to identify genes differentially expressed between male and female worms from MS infections at 21, 28, 35 and 38 d.p.i. DESeq2 was used. The threshold for differential expression was an adjusted p-value of < 0.01 and a fold-change in expression > 2 throughout this Chapter. Pairwise comparisons of males and females at each time point showed that there was an increase from 38 DEGs at 21 d.p.i. to over 3600 DEGs at 38 d.p.i. (of 10828 genes in total). The biggest absolute increase occurred from 28 to 35 d.p.i., approximately 1270% (Figure 3.3), as males and females diverge transcriptomically. These changes are responsible for the development of sex-specific features and the increasing proportion of reproductive tissues such as ovaries and vitellaria in female worms.

At 21 d.p.i., male and female worms from MS infections were very similar at the transcriptome level. Of the 38 DEGs (see Figure 3.3a), 22 were up-regulated in females and 16 in males (see Appendix A.2 and A.3 for full lists of DEGs). In female worms several genes involved in cell division and cell fate determination

could be identified (Table 3.1). First, the gene coding for MutS, a protein implicated in mismatch DNA repair, was up-regulated (Table 3.1). Second, female worms express higher levels of the gene coding for Delta-like protein (Table 3.1), which is a ligand of Notch receptors involved in cell fate determination (Zhou *et al.*, 2010). Third, a gene coding for a *single minded*-like protein was up-regulated in female worms (Table 3.1). In *Drosophila*, *single minded* is required for the proper development of midline cells (Nambu *et al.*,1991). At 21 d.p.i. both male and female worms are in the process of developing their reproductive organs (Biolchini *et al.*, 2006) as well as their gut (Basch, 1991) and these gene may be involved in regulating these processes.

Gene ID	Product	Fold Change	Adjusted p-value
Smp_213480	MutS protein	6.54	9.36E-06
Smp_178780	Single minded	4.78	0.00160
Smp_141250	Jagged 1	2.59	0.00874

Table 3.1: **Cell division related genes up-regulated in females from mixed sex infections at 21 days *post* infection.** Fold changes are relative to expression in male worms from mixed sex infections at 21 d.p.i. (see Appendix A.3 for full list). The *p*-value has been adjusted for multiple hypothesis testing.

In males, two nervous system related genes were found to be up-regulated at 21 d.p.i. (Appendix A.3). One of the two genes encoded a sodium chloride dependent neurotransmitter (Smp_160360, 3.54-fold change, adjusted p-value = 3.36E-07). The other encoded a neuropeptide receptor (Smp_133550, 2.0-fold change, adjusted p-value = 0.0067). Collins *et al*. (2010) showed that neuropeptide Y is involved in the development and maintenance of the germline in planarians and identified peptide homologues in *S. mansoni*. This evidence suggests the differential use of neuropeptide receptors in male and female *S. mansoni* in the

mammalian stage of their life cycle. Smp_133550 remains significantly up-regulated in males at 38 d.p.i. (7.1-fold change, adjusted p-value = 2.79E-21) (Appendix A.4). Together, the differentially expressed genes suggest that the development of male and female worms including that of the nervous system has begun to diverge by 21 d.p.i., priming the otherwise transcriptomically very similar males and females for their sex-specific development.

At 38 d.p.i. approximately 33.4% (3616) of all genes were found to be differentially transcribed (Figure 3.3), revealing the extraordinary degree to which mature females from MS infections differ from their male counterparts. 1377 of the 3616 DEGs were up-regulated in females (Figure 3.3). Among the female DEGs were several encoding well known fertility-associated proteins including female specific protein 800 (fs800) and other trematode egg shell synthesis domain containing proteins, two tyrosinases and an extracellular superoxide dismutase all of which were at least 80-fold up-regulated compared to male worms (adjusted p-value < 4.00E-22) (Appendix A.5). Another gene up-regulated in females encodes a homologue of the human TNF receptor associated protein 1 (TRAP1; Smp_155740) (Appendix A.5).

At 38 d.p.i., MS male worms express 2239 genes at significantly higher levels than females (Figure 3.3). A selection of 29 genes pre-dominantly expressed in males was identified in the literature (Fitzpatrick *et al.*, 2005; Piao *et al.*, 2011) and tested for male-biased expression. All 29 genes were found to be differentially expressed (Table 3.2), including genes coding for collagen, calpain and musculature related proteins.

Gene ID	Product	Fold change	Adjusted p-value
Smp_006860	PDZ & LIM protein	4.66	9.42E-23
Smp_022340	PDZ & LIM protein	5.83	6.54E-22
Smp_166920	PDZ & LIM protein	7.50	8.16E-22
Smp_164590	Fibrillar collagen α	4.65	8.03E-12
Smp_123830	Collagen α (xi) chain	4.14	6.81E-12
Smp_196840	Collagen α 1(II) chain	2.59	7.50E-09
Smp_170340	Collagen α 1(IV) chain	5.68	2.01E-07
Smp_156770	Tegumental antigen Sm15	7.26	5.12E-14
Smp_179810	Tegumental antigen Sm8	5.06	8.86E-15
Smp_214180	Calpain	3.48	7.16E-18
Smp_214190	Calpain	3.03	2.08E-21
Smp_020550	Ldl receptor	3.85	1.15E-06
Smp_161930	Actin	4.56	1.49E-18
Smp_059170	Troponin	7.43	1.07E-22
Smp_174520	Dynein light chain	4.08	0.0031
Smp_176300	Dynein heavy chain	2.34	1.13E-06
Smp_212280	Dynein heavy chain	3.11	1.17E-13
Smp_140500	Paramyosin	2.13	0.0003
Smp_021920	Paramyosin	6.85	9.62E-08

Table 3.2: Differential expression analysis confirmed male-biased expression patterns of known male-specific genes from the literature (Fitzpatrick *et al.*, 2005; Piao *et al.*, 2011)**.** This was a comparison of male and female worms from mixed sex infections at 38 days *post* infection. Fold changes are provided relative to female expression; all genes in this table were up-regulated in male worms. The *p*-value has been adjusted for multiple hypothesis testing.

As in MS males at 21 d.p.i., the neuropeptide receptor (Smp_133550) was found to be up-regulated in MS males at 38 d.p.i. (fold change = 7.1, adjusted p-value = 2.79E-21). Additionally, another G-protein coupled receptor (Smp_170020, 9.1-fold change, adjusted p-value = 2.54E-30) with homology to neuropeptide receptors was found to be up-regulated in the male worms (Table 3.3). Overall, a high proportion (13 out of 16) of neurotransmitter-gated ion channels (identified by their PFAM domain annotation) were significantly up-regulated in male worms compared to female worms (Table 3.3). These include three

glutamate-gated chloride channel subunits (2, 3 and 4) (Smp_015630, Smp_104890 and Smp_099500), seven out of eight nicotinic acetylcholine receptors and receptor subunits (Smp_031680, Smp_012000, Smp_132070, Smp_180570, Smp_139330, Smp_157790 and Smp_176310), as well as an acetylcholine receptor subunit α (Smp_101990), a neural acetylcholine receptor subunit α 4 (Smp_037960) and a glycine receptor subunit β (Smp_096480) (Table 3.3). Of the remaining three neurotransmitter receptors identified in the *S. mansoni* genome, one is up-regulated in females (Smp_142690), while expression of the remaining two could not be detected at this stage of the life cycle (Smp_175110, Smp_197600) (Table 3.3). This striking enrichment in neurotransmitter receptor mRNA expression in male worms over female worms raises questions on the role of these proteins in the male reproductive process. When comparing expression of these 13 genes in males to female worms from SS infections, no such enrichment is found (Appendix A.6); only two of the genes, Smp_104890 and Smp_101990, are still expressed at significantly higher levels in male worms than in the immature females (Appendix A.6). This suggests that the relative abundance of different tissue types is responsible for the up-regulation of neurotransmitter receptor genes in male worms. The large proportion of ovarian and particularly vitellarian cells in mature females as well as the relatively lower proportion of muscle (Loker & Brant, 2006) may be responsible for the lower relative expression of neurotransmitter genes in whole females. In contrast, immature female are more male-like than mature females at the transcriptome level, due to the lack of developed reproductive organs.

Gene ID	Product	Average read count	Fold change	Adjusted p-value
Smp_015630	Glutamate-gated chloride channel subunit	353.74	4.61	1.23E-06
Smp_104890	Glutamate-gated chloride channel subunit	244.18	6.03	7.84E-24
Smp_099500	Glutamate-gated chloride channel subunit	38.57	2.75	0.0053
Smp_031680	Nicotinic acetylcholine receptor	176.02	5.13	4.92E-08
Smp_012000	Nicotinic acetylcholine receptor	92.58	3.52	8.81E-05
Smp_132070	Nicotinic acetylcholine receptor subunit	3562.48	2.67	5.57E-16
Smp_180570	Nicotinic acetylcholine receptor	132.26	4.04	9.15E-08
Smp_139330	Nicotinic acetylcholine receptor non α	811.23	4.32	1.98E-18
Smp_157790	Nicotinic acetylcholine receptor subunit	121.98	4.31	2.22E-06
Smp_176310	Nicotinic acetylcholine receptor subunit	699.76	3.58	6.37E-07
Smp_101990	Acetylcholine receptor subunit α	23.03	2.53	0.0053
Smp_037960	Neuronal acetylcholine receptor subunit α	1308.18	2.75	1.76E-05
Smp_096480	Glycine receptor subunit β	520.33	3.44	1.93E-11
Smp_142690	Neuronal acetylcholine receptor subunit α	1875.19	0.42	2.93E-21
Smp_175110	Neurotransmitter gated ion channel	0.00	N/A	N/A
Smp_197600	Nicotinic acetylcholine receptor subunit	0.68	1.5	0.3996

Table 3.3: Neurotransmitter-gated ion channels are largely up-regulated in male worms. The differential expression analysis compares male and female worms from mixed sex infections at 38 days *post* infection. Average read counts are a measure of absolute expression across all samples, and fold changes are provided relative to female expression (fold changes > 1 indicate up-regulation in males). If no expression was measured, fold change and p-value are marked as N/A. The *p*-value has been adjusted for multiple hypothesis testing.

To better understand what biological processes the DEGs play a role in in females and males I made use of the KEGG database and performed a pathway enrichment analysis to identify pathways with significantly more DEGs than expected by chance. At 38 d.p.i., the DEGs upregulated in MS females were found to be enriched for 13 pathways involved in DNA replication and protein synthesis (Table 3.4), reflecting the proliferation of ovarian and especially vitellarian cells to drive egg production. In contrast, several lipid-metabolic pathways, developmental signalling pathways and pathways involved in

endocytosis were significantly enriched in males (see Table 3.5). Amongst these, is the TGF-β pathway that has been implicated in the male-female interaction and the stimulation of female fertility (LoVerde *et al.*, 2009) (Table 3.5). Another up-regulated pathway is phosphatidylinositol signalling (smm04070, p-value = 0.020) (Table 3.5). It has been proposed that female fertility is supported by lipid transfer from the male to the female (Popiel & Basch, 1986). Several lipid metabolism-related pathways were indeed significantly enriched in the DEGs upregulated in males at 38 d.p.i., including glycerolipid and glycerophospholipid metabolism (smm00564, p-value = 0.00019; smm00561, p-value = 0.00059) as well as ether lipid metabolism (smm00565, p-value = 0.0082) (Table 3.5). However, this could also simply reflect a higher rate of tegument turnover in male worms, which have a thicker tegument as well as a larger surface area, due to their flattened morphology.

KEGG ID	Pathway	Total Genes	DEGs	Expected	p-value
smm03010	Ribosome	110	82	42.98	7.92E-15
smm03008	Ribosome biogenesis in eukaryotes	62	47	24.23	2.22E-09
smm03050	Proteasome	32	27	12.50	1.34E-07
smm03030	DNA replication	30	22	11.72	1.06E-04
smm03040	Spliceosome	105	59	41.03	1.11E-04
smm03060	Protein export	19	15	7.42	3.84E-04
smm00510	N-Glycan biosynthesis	30	21	11.72	4.13E-04
smm03440	Homologous recombination	18	14	7.03	7.83E-04
smm00970	Aminoacyl-tRNA biosynthesis	34	21	13.29	3.79E-03
smm03013	RNA transport	103	52	40.25	4.50E-03
smm04141	Protein processing in endoplasmic reticulum	89	45	34.78	7.02E-03
smm03410	Base excision repair	20	12	7.82	3.00E-02
smm00240	Pyrimidine metabolism	59	28	23.05	4.37E-02

Table 3.4: Female worms up-regulated KEGG pathways playing a role in protein synthesis and DNA replication. At 38 days *post* infection, genes significantly upregulated in females from mixes sex infections were enriched in 13 KEGG pathways. “Total genes” is the number of genes in the given pathway. “DEGs” provides the number of differentially expressed genes in that pathway. “Expected” provides the number of genes expected by chance to be differentially expressed in each pathway.

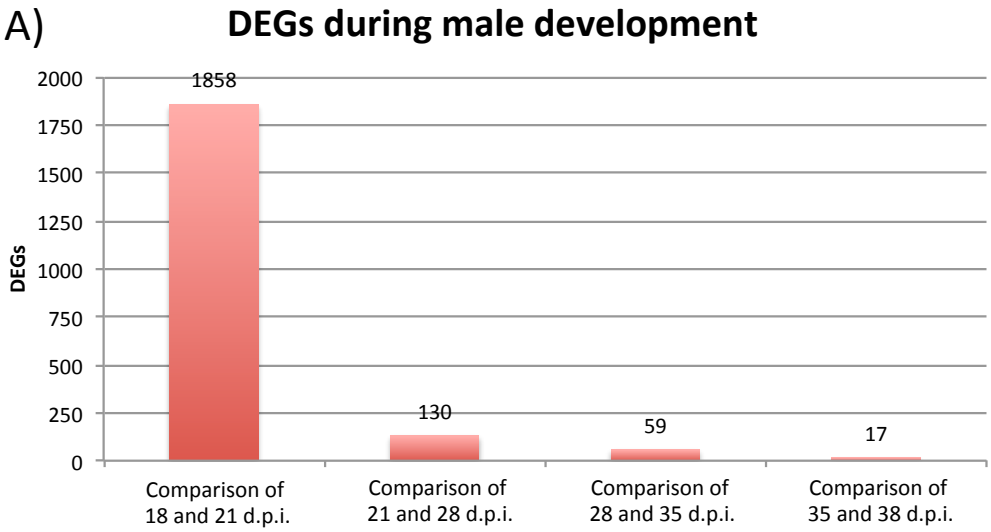
KEGG ID	Pathway	Total Genes	DEGs	Expected	p-value
smm04144	Endocytosis	64	26	10.82	3.15E-06
smm04512	ECM-receptor interaction	9	8	1.52	4.71E-06
smm04068	FoxO signaling pathway	38	18	6.43	8.63E-06
smm04310	Wnt signaling pathway	35	17	5.92	1.02E-05
smm04080	Neuroactive ligand-receptor interaction	15	10	2.54	2.10E-05
smm00564	Glycerophospholipid metabolism	35	15	5.92	1.88E-04
smm00561	Glycerolipid metabolism	17	9	2.87	5.93E-04
smm00500	Starch and sucrose metabolism	14	8	2.37	6.33E-04
smm00380	Tryptophan metabolism	10	6	1.69	2.29E-03
smm00340	Histidine metabolism	5	4	0.85	3.36E-03
smm04142	Lysosome	47	15	7.95	4.96E-03
smm00565	Ether lipid metabolism	9	5	1.52	8.19E-03
smm04340	Hedgehog signaling pathway	9	5	1.52	8.19E-03
smm00514	Other types of O-glycan biosynthesis	4	3	0.68	1.60E-02
smm00010	Glycolysis / Gluconeogenesis	27	9	4.57	1.85E-02
smm04070	Phosphatidylinositol signaling system	32	10	5.41	2.05E-02
smm04630	Jak-STAT signaling pathway	11	5	1.86	2.08E-02
smm04320	Dorso-ventral axis formation	8	4	1.35	2.71E-02
smm04350	TGF-β signaling pathway	16	6	2.71	2.91E-02

Table 3.5 Male worms upregulated KEGG pathways involved in cellular metabolism and various signalling pathways.At 38 days *post* infection, genes upregulated in males from mixed sex infections were enriched in 19 KEGG pathways. "Total genes" is the number of genes in the given pathway. "DEGs" provides the number of differentially expressed genes in that pathway. "Expected" provides the number of genes expected by chance to be differentially expressed in each pathway.

3.2.4 Male development

In male worms, the greatest difference in gene expression was found between 18 and 21 d.p.i. (1858 DEGs in total) (Figure 3.4), suggesting that this was when most of the male-specific developmental changes took place. The number of DEGs dropped sharply after that time, to 130 (21 vs 28), 59 (28 vs 35) and finally 17 DEGs (35 vs 38), (Figure 3.4). Below, some of the results of the differential expression analysis are discussed in detail.

A)


DEGs during male development
DEGs
2000
1750
1500
1250
1000
750
500
250
0
1858
130
59
17
Comparison of 18 and 21 d.p.i.
Comparison of 21 and 28 d.p.i.
Comparison of 28 and 35 d.p.i.
Comparison of 35 and 38 d.p.i.

B)

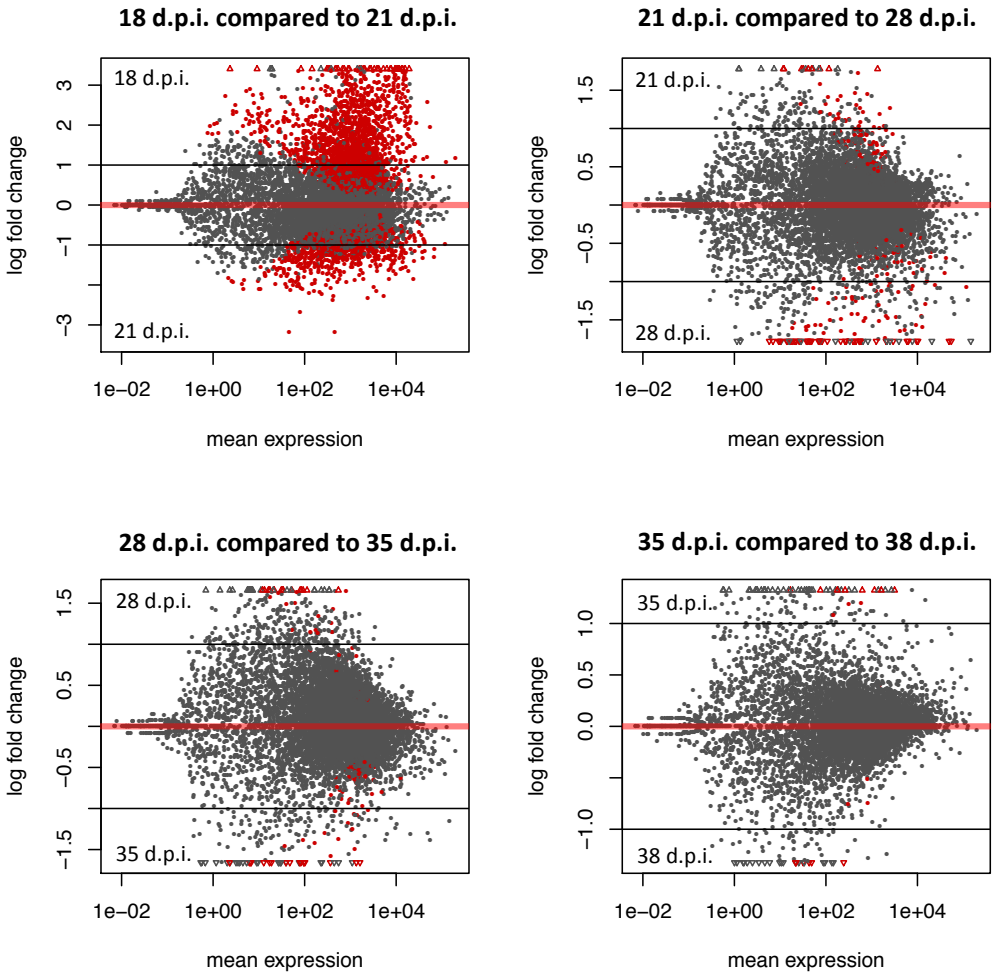

18 d.p.i. compared to 21 d.p.i.
18 d.p.i.
21 d.p.i.
log fold change
mean expression
21 d.p.i. compared to 28 d.p.i.
21 d.p.i.
28 d.p.i.
log fold change
mean expression
28 d.p.i. compared to 35 d.p.i.
28 d.p.i.
35 d.p.i.
log fold change
mean expression
35 d.p.i. compared to 38 d.p.i.
35 d.p.i.
38 d.p.i.
log fold change
mean expression
1e−02
1e+00
1e+02
1e+04

Figure 3.4: The largest number of differentially expressed genes was found between 18 and 21 days *post* infection (d.p.i.) in developing male worms. A) This plot provides the number of differentially expressed genes found by pair-wise comparisons of males from mixed sex infections of different time points. B) Log ratio vs mean (MA) plots showing the mean expression and log-fold change of each pair-wise comparison. Genes marked as red dots have been assigned an adjusted p-value of < 0.01, the other are grey. Black lines represent the fold change cut-off (log-fold change = ±1). Triangles indicate genes outside the boundaries of the plots; like other genes, they are red if they are statistically significant, grey if they are not.

Cytochrome P450-dependent drug metabolism

Praziquantel is currently the main drug to treat schistosomiasis, especially as part of mass treatment programs in endemic areas (Crellen *et al.*, 2016). However, immature worms are known to be less sensitive to praziquantel, making it difficult to clear an infection completely and stop the parasite life cycle in endemic areas (Sabah *et al.*, 1986; Vimieiro *et al.*, 2013). At 18 d.p.i., three out of four genes belonging to the cytochrome P450 dependent drug metabolism pathway were found to be up-regulated in males worms compared to male worms at 21 d.p.i. (Appendix A.7). The relatively high expression of genes involved in cytochrome P450-dependent drug metabolism at 18 d.p.i. may be the reason why immature worms are less susceptible to praziquantel. A Glutathione S transferase (GST) was not only up-regulated in males at 18 d.p.i. compared to males at 21 d.p.i. (Smp_024010, 5.1-fold change, adjusted p-value = 4.50E-05) (Appendix A.8), but also compared to males at 38 d.p.i. (21.9-fold change, adjusted p-value = 1.34E-14) (Appendix A.8). The same gene was also up-regulated in female worms at 18 compared to 38 d.p.i. (18.2-fold change, adjusted p-value = 1.30E-11) (Appendix A.8).

Regulation of cell division in the testes

Compared to 21 d.p.i., the gene coding for the cell cycle regulator Cdc25 (Smp_152200, 6.0-fold change, adjusted p-value = 0.00061) was up-regulated at 28 d.p.i. in male worms, as was follistatin (Smp_123300, 2.3-fold change, adjusted p-value = 0.00027) (Appendix A.9). Follistatin is also associated with the proliferation of spermatocytes and was also found to be down-regulated from 28 to 35 d.p.i. (2.4-fold change, adjusted p-value = 0.00082) (Appendix A.10). Leutner *et al.* (2013) demonstrated that paired male worms express lower levels of follistatin than unpaired males. The down-regulation of the gene encoding follistatin in MS males at 35 d.p.i. compared to those at 28 d.p.i. is probably related to this observation (Appendix A.10). From 21 to 28 d.p.i., a gene encoding the homologue of basonuclin 2 (Smp_138350, 5.4-fold change, adjusted p-value = 1.86E-05) was up-regulated in male worms from mixed sex infections (Appendix A.16).

Aromatic-L-amino acid decarboxylase

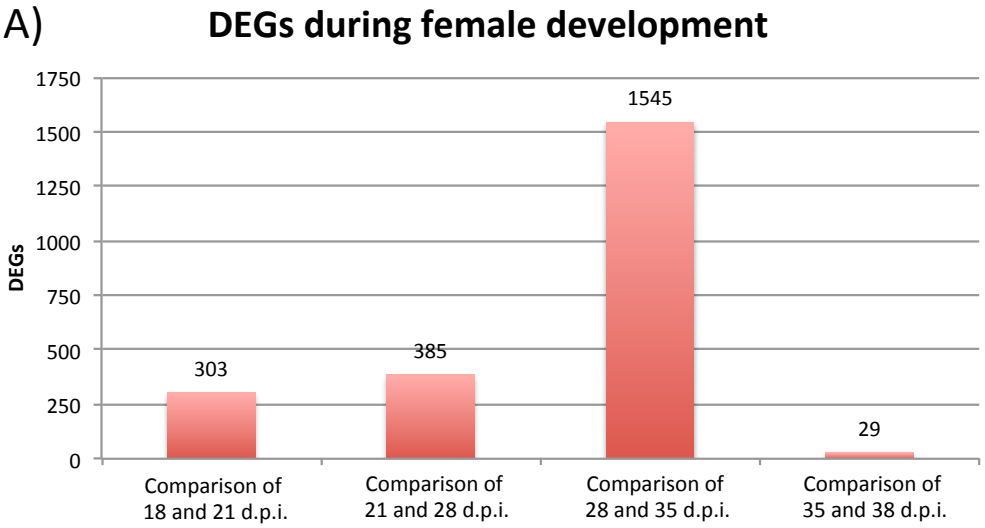
Wang *et al.* (2017) hypothesised that the AADC gene, which they found to be expressed in *S. japonicum* males during mating (see Chapter 1.6.2), was responsible for inducing female sexual maturity. The AADC gene was also found to be up-regulated in male *S. mansoni*. The *S. mansoni* gene (Smp_135230) was significantly up-regulated in male worms at 38 d.p.i. compared to males at 18 d.p.i. (8.50-fold change, adjusted p-value = 2.26E-06) (Appendix A.11). Additionally, it was found to be significantly up-regulated in male worms from MS infections compared to males in SS infections at 35 d.p.i. (8.39-fold change, adjusted p-value = 1.23E-03) and 38 d.p.i. (5.33-fold change, adjusted p-value =

1.99E-03), falling just short of the significance threshold (adjusted p-value < 0.01) at 28 d.p.i. (2.32-fold change, adjusted p-value = 0.0359) (Appendix A.11). The lack of expression in males from SS infections as well as the confirmation of expression in mature MS males (here in *S. mansoni*) adds further weight to the hypothesis of AADC being involved the stimulation of female fertility in schistosomes.

3.2.5 Female development

Pairing is thought to commence between three and four weeks *post* infection, allowing female worms to reach sexual maturity around five weeks *post* infection (Biolchini *et al*. 2006). Accordingly, the greatest difference in gene expression in MS female worms was found between 28 and 35 d.p.i. (1545 DEGs) (Figure 3.5). By comparison, there were only 300 DEGs between 18 and 21 d.p.i., 385 between 21 and 28 d.p.i. and 29 between 35 and 38 d.p.i. (Figure 3.5).

A)


DEGs during female development
DEGs
1750
1500
1250
1000
750
500
250
0
303
385
1545
29
Comparison of 18 and 21 d.p.i.
Comparison of 21 and 28 d.p.i.
Comparison of 28 and 35 d.p.i.
Comparison of 35 and 38 d.p.i.

B)

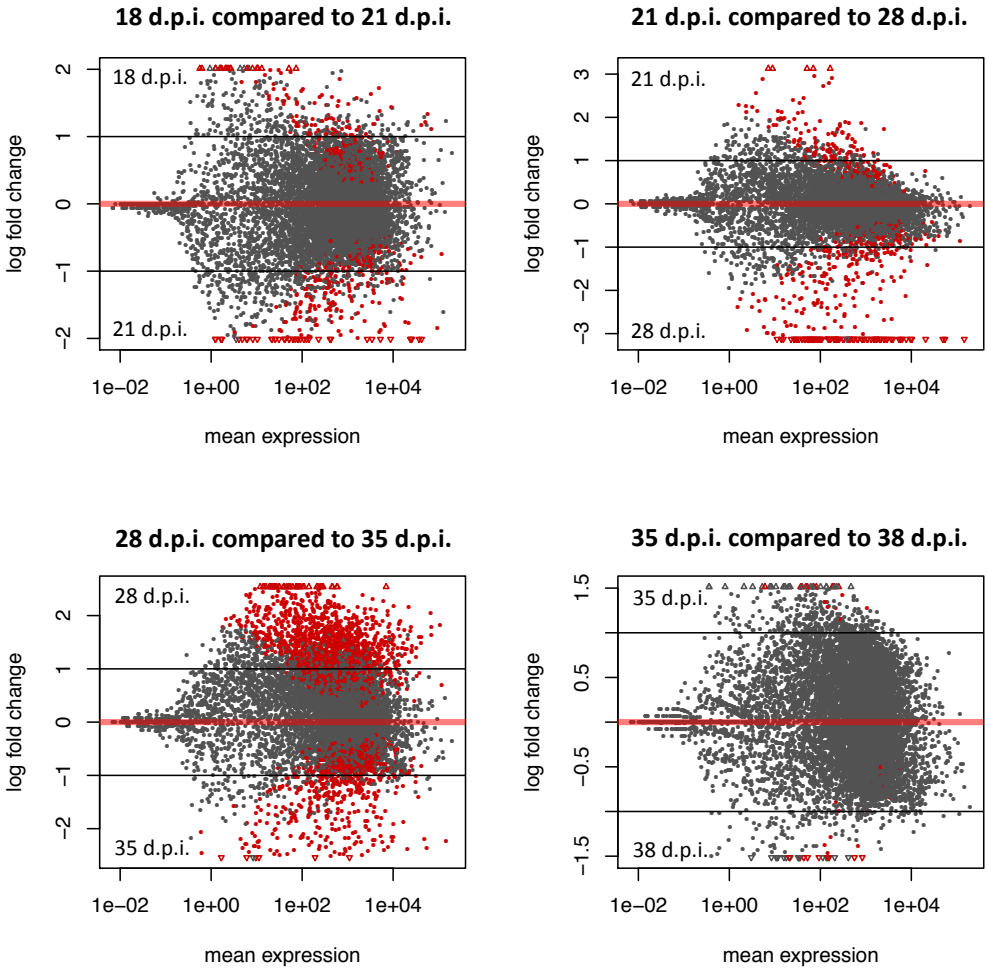

18 d.p.i. compared to 21 d.p.i.
18 d.p.i.
21 d.p.i.
log fold change
mean expression
21 d.p.i. compared to 28 d.p.i.
21 d.p.i.
28 d.p.i.
log fold change
mean expression
28 d.p.i. compared to 35 d.p.i.
28 d.p.i.
35 d.p.i.
log fold change
mean expression
35 d.p.i. compared to 38 d.p.i.
35 d.p.i.
38 d.p.i.
log fold change
mean expression
1e−02
1e+00
1e+02
1e+04

Figure 3.5: The largest number of differentially expressed genes (DEGs) was found between 28 and 35 days *post* (d.p.i.) infection in female worms from mixed sex infections. A) DEGs found by pair-wise comparisons of females from mixed sex infections at different time points. B) Log ratio vs mean (MA) plots show the mean expression and log-fold change of each pair-wise comparison. Genes marked as red dots have been assigned an adjusted p-value of < 0.01, the others are grey. Black lines represent the fold change cut-off (log-fold change = ±1). Triangles indicate genes outside the boundaries of the plots; like other genes, they are red if they are statistically significant, grey if they are not.

Egg production related genes

During the maturation of paired female schistosomes, many fertility related genes are known to become up-regulated (Fitzpatrick & Hoffmann, 2006). This is largely due to the growth and maturation of the reproductive organs including the ovaries and the vitellarium (Fitzpatrick & Hoffmann, 2006). Many of the significantly up-regulated genes were also found in the comparison of male and female worms (Chapter 3.2.3). In this comparison of females from different time points, the most notable up-regulation of fertility-related genes occurred between days 21 and 28, where 266 DEGs were found to become up-regulated (Figure 3.5). The DEGs included all genes encoding trematode egg shell synthesis domains, a tyrosinase (Smp_050270), and the extracellular superoxide dismutase (Smp_095980) (Appendix A.12). Furthermore, two tetraspanins were up-regulated, CD63 antigen (Smp_173150) and a CD63 receptor (Smp_155310) (Appendix A.12). Although CD63 antigen is expressed in mature females, its function remains unclear (Cogswell *et al.*, 2012). At 38 d.p.i., females from MS infections expressed the CD63 receptor and antigen, at 191-fold (adjusted p-value = 5.98E-30) and 39-fold (adjusted p-value = 1.62E-42) higher levels

respectively, than females from SS infections (Appendix A.13). This made them the 21st and 51st of the most differentially expressed genes, the majority of which are thought to be expressed in the vitellarian tissue. Together the DEGs summarised in this paragraph confirmed that female schistosomes from MS infections have not only begun to pair by 28 d.p.i. but also that females have started to develop their reproductive organs at that time. As others recently observed (Biolchini *et al.*, 2006), it appears that pairing takes pace somewhat earlier than reported by Basch (1991). These genes became further up-regulated at 35 d.p.i. and 38 d.p.i. and were consistently expressed at significantly higher levels in females from MS infections rather than SS infections (see Chapter 3.2.8).

Ovary-specific gene expression

While the most differentially expressed genes up-regulated in mature females are thought to be largely expressed in the vitellaria [for example Fitzpatrick *et al.* (2007), Fitzpatrick & Hoffmann (2006), Köster *et al.* (1988)], at 21 d.p.i. (compared to 18 d.p.i.), a least two of the up-regulated genes probably play an important role in the ovaries of mature females. The first encoded a homologue of human sperm associated antigen (Smp_001160, 2.2-fold change, adjusted p-value = 2.08E-05) (Appendix A.14), which plays a role in fertility in mammals (Lin *et al.*, 2001). The other encoded *boule* (Smp_144860, 6.6-fold change, adjusted p-value = 6.34E-05) (Appendix A.14), a conserved transcription factor found in most animals. Amongst others, it has been shown to regulate male and female gametogenesis in the flatworm *Macrostomum ligano* (Kuales *et al.*, 2011). In *S. mansoni*, *boule* was also expressed specifically in the gonads (Chapter 5.2).

WNT signalling

At 18 d.p.i., females express two *wnt* family genes (Smp_196930, 2.5-fold change, adjusted p-value = 0.0027; Smp_167140, 3.1-fold change, adjusted p-value = 0.0067) that became down-regulated at at 21 d.p.i. (Appendix A.15) Additionally, a gene (Smp_062560, 2.2-fold change, adjusted p-value = 0.0067) homologous to a *Secreted frizzled-related protein* 2, a regulator of the Wnt pathway, was up-regulated (Appendix A.15) Frizzled proteins are G-protein coupled receptors that regulate the Wnt signalling pathway (Nusse, 2009). Up-regulation of these three genes suggests, that the Wnt signalling pathway is more active during the earlier parts of their development, where it could regulate the formation of a body axis, cell fate and cell proliferation.

Differential expression of transcription factors

Across this RNA-seq time course, many transcription factors were found to be differentially regulated. In particular a large number of genes coding for proteins regulating transcription became down-regulated at in females worms from MS infections at 28 d.p.i. when compared to 21 d.p.i. (Appendix A.17). Based on their GO term annotations, 37 transcription factor genes were identified: about a third of all the DEGs down-regulated at 28 d.p.i. (Appendix A.17). Such large-scale down-regulation of transcription factors in females from MS infections could suggest that some development processes reach their conclusion at this time. Of the 37 transcription factors, 22 were also expressed at significantly lower levels in females from SS infections at 28 d.p.i. and none were up-regulated (Appendix A.18).

Conversely, as the worms started to pair and females began to initiate the proliferation of their reproductive tissues, especially ovaries and vitellaria, another set of transcription factors was significantly up-regulated (Appendix A.19). Among the 100 most differentially expressed genes at 38 d.p.i. three genes with zinc finger domains were identified (Smp_166560, Smp_087320 and Smp_095350). All three zinc finger genes were found to be up-regulated more than 10-fold (adjusted p-value < 9.00E-22) (Appendix A.19) and it is tempting to speculate that they play a role in regulating transcription of genes involved in female maturation.

3.2.6 Male time series analysis

The previous analysis of pair-wise comparisons made it apparent that differential expression of genes can be found across the time course, with more differential expression found between MS and SS females than males. Next, I wanted to define sets of genes regulated in a pairing dependent manner throughout the course of infection. To do so, I performed a likelihood ratio test to identify genes that were up- or down-regulated over the course of infection in worms from SS or MS infections but not in both infections. Based on this test, 39 genes (Table 3.6) were identified as differentially expressed in males from MS and SS infections across the time course (adjusted p-value < 0.01) in the otherwise very similar male worms (Figure 3.6). These included several genes coding for proteins involved in signalling and signal transduction processes, such as the fibroblast growth factor (FGF) receptor activating protein (Smp_035730, adjusted p-value = 0.0053) and the epidermal growth factor (EGF) receptor (Smp_152680, adjusted p-value = 0.0069), which were found to be expressed at

higher levels in paired males (Table 3.6). These genes are shown to be expressed in the testes in Chapter 5 of this thesis. The FGF and the EGF signalling pathways regulate cell proliferation and differentiation (Eswarakumar *et al.*, 2005; Herbst, 2004). This suggests that paired males have a higher rate of cell proliferation in their testes, which could result in a greater number of spermatocytes being produced.

Gene ID	Product	MS / SS	Fold change 38 d.p.i.	Average read count	Adjusted p-value
Smp_005880	Phosphoenolpyruvate carboxykinase	MS	1.60	3571.83	2.86E-03
Smp_012000	Nicotinic acetylcholine receptor	SS	2.77	129.19	1.10E-05
Smp_012380	Voltage-gated potassium channel subunit β 2	SS	1.59	1174.94	6.87E-03
Smp_013950	Solute carrier family 43	MS	1.66	1119.06	2.03E-06
Smp_015020	Sodium potassium transporting ATPase α subunit	MS	1.71	2041.50	5.04E-03
Smp_019980	Vacuole membrane protein 1	MS	1.19	2745.53	9.17E-04
Smp_020840	Uncharacterised protein	MS	1.74	68.18	4.84E-03
Smp_021960	Uncharacterised protein	MS	1.02	318.04	1.10E-05
Smp_028800	Uncharacterised protein	SS	2.23	39.30	2.49E-04
Smp_035730	FGF receptor activating protein	MS	1.06	934.37	5.33E-03
Smp_044800	Ankyrin repeat and E domain containing protein	MS	1.15	327.46	8.81E-04
Smp_063560	Transmembrane protein 120b	MS	1.22	1080.30	5.33E-03
Smp_114430	Uncharacterised protein	MS	4.63	53.16	5.53E-03
Smp_125210	Aquaporin 9 (Small solute channel 1)	MS	1.49	119.40	6.47E-03
Smp_126290	Uncharacterised protein	SS	2.21	283.56	1.71E-03
Smp_128010	Uncharacterised protein	MS	4.10	146.98	1.10E-05
Smp_134990	Uncharacterised protein	MS	3.01	112.52	2.55E-08
Smp_135020	Putative oxalate:formate antiporter	MS	1.80	249.99	4.84E-03
Smp_135230	Tyrosine DeCarboxylase	MS	3.15	440.13	1.27E-05
Smp_147070	Sodium-coupled neutral amino acid	MS	2.34	1640.36	4.29E-03
Smp_151490	Vertebrate nebulin	MS	1.87	3474.33	6.31E-03
Smp_152680	Epidermal growth factor receptor	MS	1.75	159.89	6.90E-03
Smp_152730	Histone-lysine N-methyltransferase	MS	1.83	39.41	8.32E-03
Smp_158480	AMP binding enzyme	MS	16.87	1332.34	2.12E-09
Smp_158720	SCY1-like protein kinase	SS	1.53	484.09	1.10E-05
Smp_160040	Uncharacterised protein	MS	3.44	15.31	3.10E-03
Smp_167160	Dynamin binding protein	MS	1.51	2472.35	6.90E-03
Smp_168670	cGMP dependent protein kinase 1	MS	1.16	381.69	5.53E-03
Smp_168940	Uncharacterised protein	MS	4.09	293.32	2.72E-03
Smp_169150	Short transient receptor potential channel 3	SS	2.00	144.97	9.86E-03

Smp_172960	Uncharacterised protein	MS	1.36	2005.73	1.24E-03
Smp_173290	Uncharacterised protein	MS	2.22	23.87	1.10E-05
Smp_173300	Uncharacterised protein	MS	2.35	16.12	1.14E-05
Smp_174870	Uncharacterised protein	MS	2.00	487.85	6.75E-08
Smp_176400	Uncharacterised protein	MS	6.56	29.18	7.16E-03
Smp_190210	Uncharacterised protein	SS	5.51	5.27	9.96E-04
Smp_194190	Endothelin-converting enzyme 2	MS	2.47	109.63	4.51E-03
Smp_212750	Uncharacterised protein	MS	2.43	75.40	2.64E-04
Smp_212760	Putative kinesin	SS	2.03	909.99	7.79E-04

Table 3.6: Significant genes found in the time series analysis of male worms. The "MS/SS" column reports whether expression was up-regulated in males from mixed sex or single sex infections at 38 days *post* infection (d.p.i.). The fold change is provided for 38 d.p.i. The average read count is a measure of absolute gene expression. The *p*-value has been adjusted for multiple hypothesis testing.

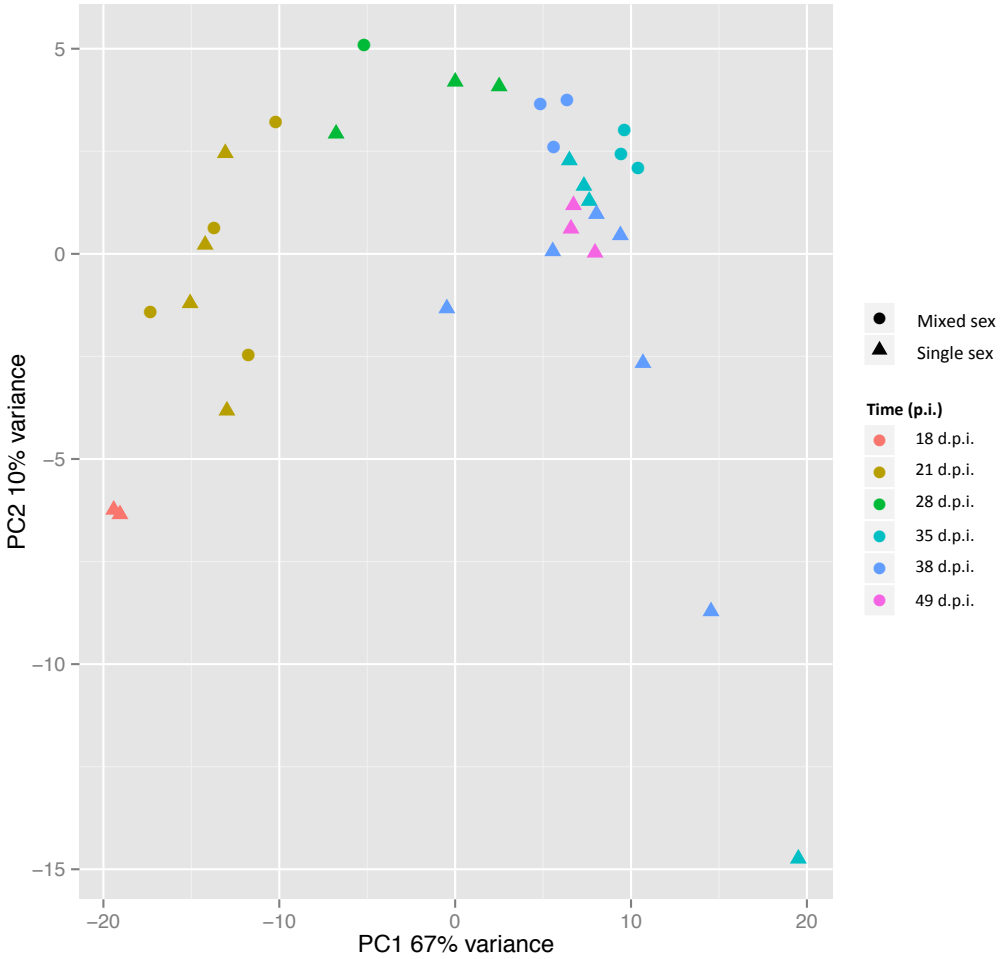


Figure 3.6: Male worms from mixed sex and single sex infections are very similar at the transcriptome levels throughout their development. Principal component analysis plot of male RNA-Seq samples. The first two principal components show 77% of the measured variance. The PC1 axis seems to broadly correlate with worm maturity. Time points indicate the nuber of days *post* infection (d.p.i.). Several outliers can be observed, especially at the late time points.

3.2.7 Female time series analysis

As for the male samples, a likelihood ratio test was used to identify pairing-controlled genes in females and 5173 such genes were discovered (adjusted p-value < 0.01). At 38 d.p.i., 2670 of the genes were up-regulated in MS females and 2503 up-regulated in SS females. The PCA plot (Figure 3.7) shows female worms from MS and SS infection forming separate clusters from 28 d.p.i. onwards. To analyse such a large number of genes, two different clustering methods were used to identify groups of genes with similar expression patterns.

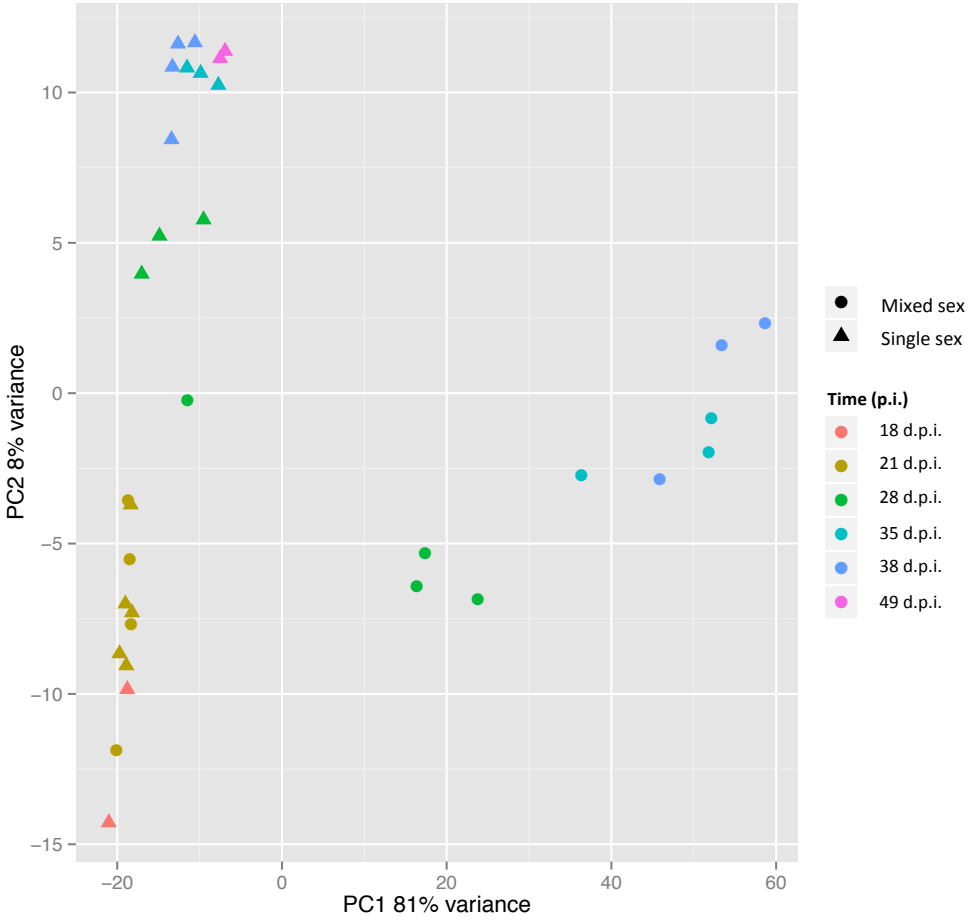


Figure 3.7: Female from mixed sex infections diverge from the single sex samples from 28 d.p.i. In this PCA plot of female RNA-Seq samples, the first two principal components capture 89% of the measured variance. Whereas paired females develop along PC1 over time, with sexual maturation contributing to 81% of the total variation, single females change primarily along PC2 (8%). Time points indicate the nuber of days *post* infection (d.p.i.).

3.2.8 Fertility-related genes

Having identified over 5000 pairing-regulated genes involved in females, I tried two different clustering methods, the Kohonen method (Wehrens & Buydens, 2007) and MBCluster (Si *et al.*, 2014) (Chapter 2.4.11), to split the genes into groups with similar patterns of expression. This was done to ensure the reproducibility of the clustering results and to define a set of putative fertility-related genes. Similar patterns of expression across several samples, especially

in time course experiments, can be evidence for co-expression and suggest roles in the same biological processes (Si *et al.*, 2014). Kohonen is based on a heuristic method designed to produce self-organising maps from data sets using unsupervised learning. Self-organising maps are a two-dimensional representation of multi-dimensional data that aim to place similar data points, e.g. genes close to one another thereby forming clusters of similar genes (Wehrens & Buydens, 2007). In contrast, MbCluster assumes the data to have a negative binomial distribution and uses an Expectation-Maximisation algorithm to estimate model parameters and cluster membership (Si *et al.*, 2014).

For both methods, genes were divided into 16 clusters (Figures 3.8 and 3.9) as higher numbers of clusters usually resulted only in an increasing number of clusters with very small number of genes, even as low as one or two genes (data not shown). In the MBCluster results, one cluster, containing 76 genes, was characterised by the largest overall fold changes and high expression levels in MS females at 28, 35 and 38 d.p.i. (see Figure 3.8). This group of genes included two of the clusters formed by the Kohonen method, which separated the genes further based on their relative expression (see Figure 3.9). Of the two Kohonen clusters, one (Cluster 13) contained genes with particularly high expression in mature females. The 25 genes in that cluster are among the 32 most up-regulated genes in paired females ranging from 540-fold to over 45350-fold up-regulation over the time course (Appendix A.20). Together, these genes account for over 14% of the sequencing reads in the transcriptome of mature females, compared to only 0.0065% of reads in immature females (Appendix A.21). Eleven of the genes were found by InterProScan to code for a trematode eggshell

synthesis domain (Appendix A.20). Also included were genes coding for two tyrosinases (Smp_013540 & Smp_050270) and an extracellular superoxide dismutase (Smp_095980) (Appendix A.20) all of which are known pairing-regulated genes (Fitzpatrick & Hoffmann, 2006).

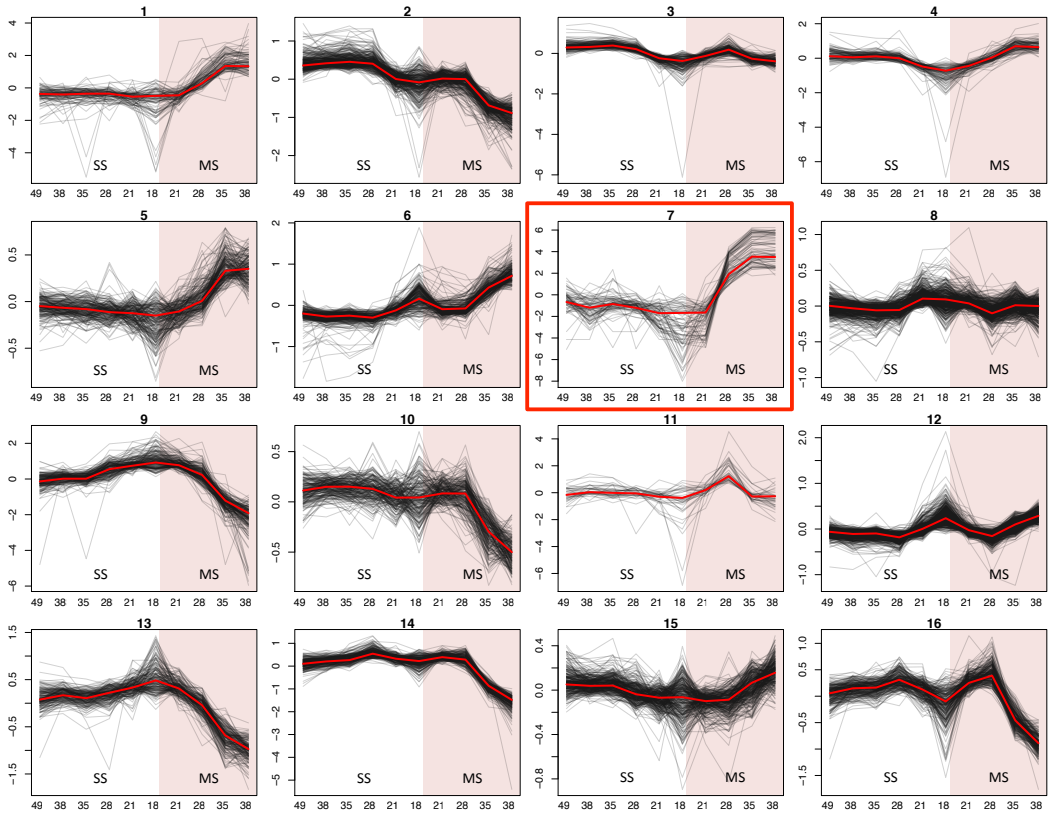


Figure 3.8: Genes were divided into 16 clusters based on expression profiles. The MBCluster plots show gene expression across all female RNA-seq samples. For each gene in a cluster the log2 fold changes (Y-axis) are plotted against the days *post* infection (X-axis). Samples from single sex (SS) infections have a white background; samples from mixed sex (MS) infections have a red background. The red lines represent the cluster average for each time point. The cluster of fertitliy-related genes dicussed above is contained within a red box.

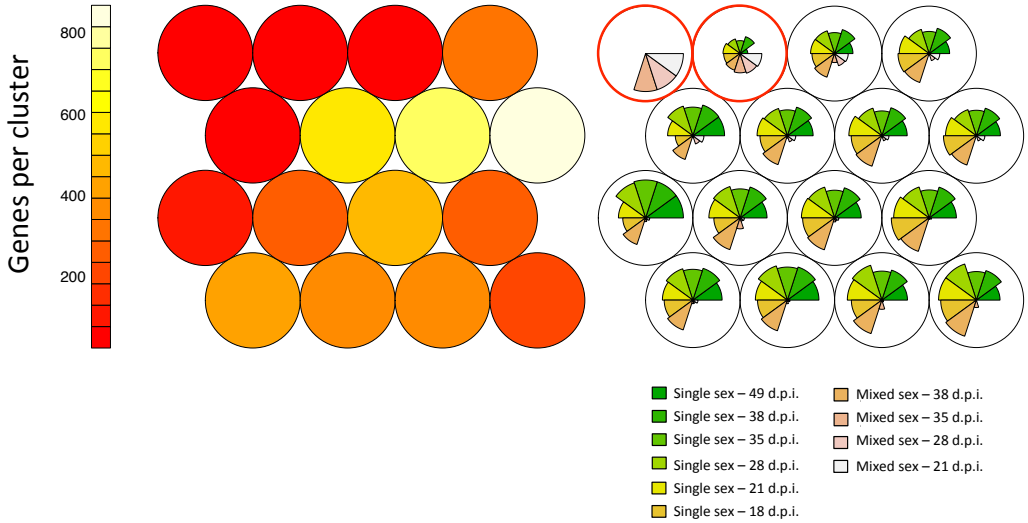


Figure 3.9: Genes were divided into 16 clusters based on their expression using the Kohonen method. Left – heatmap of number of genes per cluster. Right – Average $\log_2$-expression of genes inside each cluster for all time points and for mixed and single sex females. Cluster 13 is on the top left in both maps. The two clusters of fertitliy-related genes dicussed in Chapter 3.2.8, Cluster 13 and 14, are contained in red circles.

To select gene clusters that exhibit significant changes in expression over the time course, genes were filtered based on their average fold changes relative to expression in female worms at 18 d.p.i. The expression levels of seven out of the 16 clusters (1, 4, 7, 9, 11, 13 and 14) identified by the MBCluster method at least doubled or halved over the time course (see Figure 3.10). It showed that the fold changes in gene expression were larger in females from MS than SS infections. A GO term analysis was performed on the sets of genes contained in each of these 7 clusters, identifying a total of 72 significantly enriched unique GO terms.

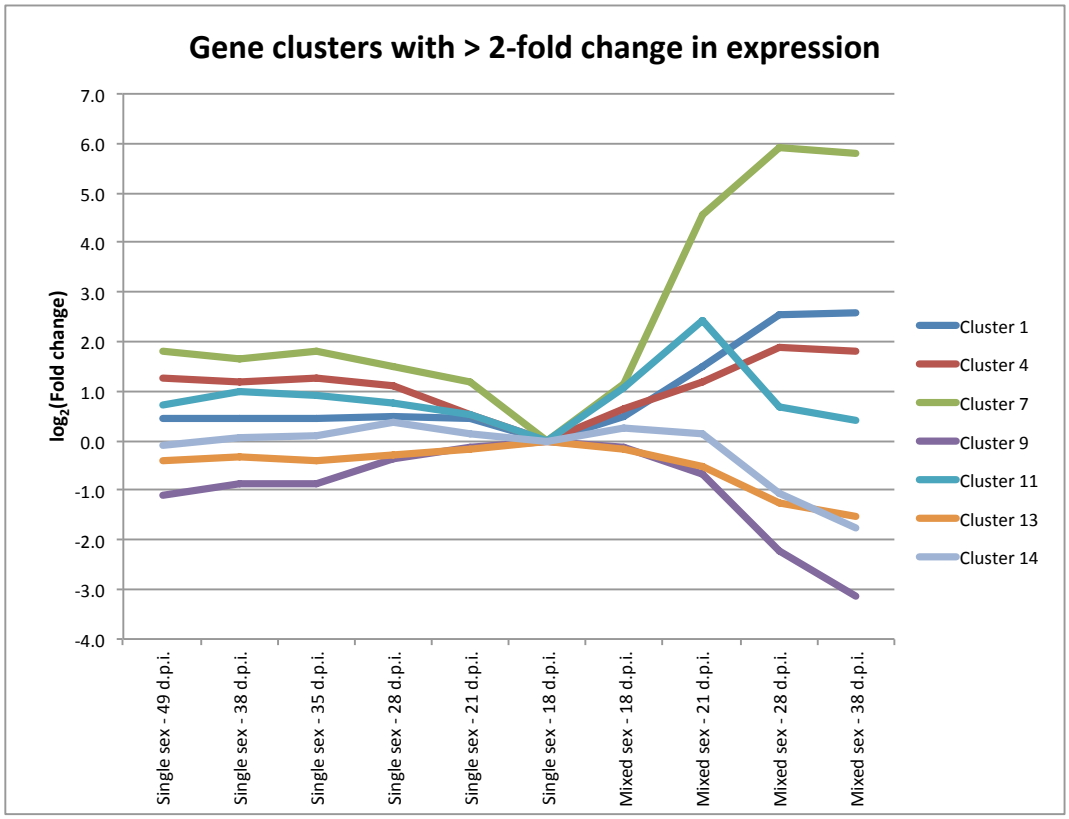


Figure 3.10: The magnitude of changes in gene expression is larger in females from mixed sex than females from single sex infections. Average expression of genes in the MBCluster clusters with greater than 2-fold changes in gene expression across the female samples of the time course. All expression fold changes are measured relative to females from single sex infections at 18 days *post* infection (d.p.i.).

The clusters 9, 13 and 14 were found to be enriched for genes annotated with developmental GO terms such as multicellular organismal development, Wnt signalling and sex determination (Tables 3.7, 3.8 and 3.9). From 18 d.p.i. onwards expression of these genes fell in females from MS as well as SS infections (Figure 3.10). Notably, the down-regulation of these genes was more pronounced in females from MS infections than those from SS infections (Figure 3.10).

GO term	Description	Annotated	Significant	Expected	p-value
GO:0007156	Homophilic cell adhesion	60	22	1.74	2.00E-19
GO:0007275	Multicellular organismal development	286	35	8.3	3.20E-13
GO:0006355	Regulation of transcription	800	57	23.22	9.80E-12
GO:0016055	Wnt signaling pathway	27	7	0.78	0.00026
GO:0007155	Cell adhesion	226	35	6.56	0.00027
GO:0030154	Cell differentiation	78	9	2.26	0.00038
GO:0030238	Male sex determination	6	3	0.17	0.00045

Table 3.7: Significant GO terms for the genes in cluster 9 of the MBCluster analysis. "Annotated" - total number of genes in the genome annotated with a given GO term. "Significant" - the number of genes annotated with that GO term in Cluster 9. "Expected" – the number of genes with that GO term expected by chance. The significance cut-off was p = 0.01. The GO terms are all in the "Biological Process" category.

GO term	Description	Annotated	Significant	Expected	p-value
GO:0006813	Potassium ion transport	77	8	1.99	0.00076
GO:0006941	Striated muscle contraction	145	11	3.74	0.00122
GO:0045596	Negative regulation of cell differentiation	3	2	0.08	0.00195
GO:0030514	Negative regulation of BMP signaling	3	2	0.08	0.00195
GO:0051216	Cartilage development	5	2	0.13	0.00629
GO:0007223	Wnt signaling pathway (Ca modulating)	6	2	0.15	0.00928

Table 3.8: Significant GO terms for the genes in cluster 13 of the MBCluster analysis. "Annotated" - total number of genes in the genome annotated with a given GO term. "Significant" - the number of genes annotated with that GO term in Cluster 13. "Expected" – the number of genes with that GO term expected by chance. The significance cut-off was p = 0.01. The GO terms are all in the "Biological Process" category.

GO term	Description	Annotated	Significant	Expected	p-value
GO:0007186	G-protein coupled receptor signaling	161	30	8.46	2.50E-09
GO:0007155	Cell adhesion	226	44	11.87	4.90E-08
GO:0007156	Homophilic cell adhesion	60	14	3.15	1.70E-06
GO:0006811	Ion transport	389	52	20.44	5.60E-06
GO:0006813	Potassium ion transport	77	14	4.05	3.70E-05
GO:0006468	Protein phosphorylation	337	34	17.71	0.00013
GO:0007160	Cell-matrix adhesion	8	4	0.42	0.00044
GO:0007229	Integrin-mediated signaling pathway	9	4	0.47	0.00076
GO:0006816	Calcium ion transport	49	9	2.57	0.00086
GO:0007165	Signal transduction	547	67	28.74	0.00143
GO:0030245	Cellulose catabolic process	6	3	0.32	0.00255
GO:0007169	Transmembrane receptor protein tyrosine kinase signalling	31	6	1.63	0.00483
GO:0006865	Amino acid transport	22	5	1.16	0.00486
GO:0006814	Sodium ion transport	64	9	3.36	0.0058

Table 3.9: Significant GO terms for the genes in cluster 14 of the MBCluster analysis. "Annotated" - total number of genes in the genome annotated with a given GO term. "Significant" - the number of genes annotated with that GO term in Cluster 14. "Expected" – the number of genes with that GO term expected by chance. The significance cut-off was p = 0.01. The GO terms are all in the "Biological Process" category.

Expression of Clusters 1, 4 and 7 increased over time in both SS and MS females. In all three the effect was greater in females from MS infections than females from SS infections (Figure 3.10). At 38 d.p.i. average expression was around 6- and 3.5-fold higher than at 18 d.p.i., compared with a 1.4- and 2.4-fold increase in SS females for clusters 1 and 4 respectively (Figure 3.10). Cluster 1 contained several genes with a female-biased expression, including gene coding for proteins such as ferritin (Smp_087760) (Appendix A.22), whereas cluster 4 contained several genes encoding saposins, cathepsins and a haemaglobinase, which may be related to digestive processes in mature females (Appendix A.23).

Cluster 7 expression displayed the greatest up-regulation of gene expression amongst all clusters (Figure 3.10). In MS females, expression was on average 60-fold up-regulated at 38 d.p.i. relative to expression at 18 d.p.i. (Figure 3.10). This cluster contained many genes associated with female fertility, particularly egg production, such as genes encoding eggshell synthesis domain containing proteins (Smp_077900, Smp_000290, Smp_191910, Smp_000430 and Smp_077890), tyrosinases (Smp_050270 and Smp_013540), and asuperoxide dismutase (Smp_095980) (Appendix A.24). Cluster 7 was found to be significantly enriched for several GO terms. These included tyrosine metabolic process (GO:0006570, p = 0.00014), phenol-containing compound biosynthetic process (GO:0046189, p = 0.00014) and pigment biosynthetic process (GO:0046148, p = 0.00128) due to the presence of the two tyrosinases in this cluster (Appendix A.25). Both tyrosinases are known to be expressed predominantly in sexually mature females and are thought to catalyse the oxidation of phenolic compounds contributing to egg shell formation (Fitzpatrick *et al.*, 2007). This process can produce free radicals and other oxidising agents (Tada, Kohno, & Niwano, 2014). Superoxide metabolic process (GO:0006801, p = 0.020) (Appendix A.25) was also overrepresented. Superoxide dismutases that are known to be expressed in female worms in a pairing dependant manner (Sun *et al.*, 2014), are also found in this cluster and may serve to protect the female reproductive tissues from oxidative stress. Other up-regulated genes included genes coding for a protein kinase as well as the tetraspanin CD63 receptor (Smp_155310) and a tetraspanin CD63 antigen (Smp_173150) (Appendix A.24).

The role of eggshell synthesis domain containing proteins as well as the tyrosinases in egg production has been demonstrated previously (Fitzpatrick *et al*., 2007). On the other hand, the role of the membrane receptors CD63 antigen (CD63a) and CD63 receptor (CD63R) is not clear and will be the focus of the rest of the chapter. CD63a, but not CD63R, has previously been shown to have a pairing-dependent pattern of expression (Cogswell *et al.*, 2012). As revealed by the time course data, both genes are not only female-specific but also respond to pairing in their expression profiles (Cluster 7 in Figure 3.10).

According to GeneDB (Logan-Klumpler *et al.*, 2012), CD63a is a predicted 24.1 kDa protein with 225 amino acids and CD63R is a predicted 30.6 kDa protein of 276 amino acids. However, tetraspanins often undergo post-translational modifications (Termini & Gillette, 2017), so their actual mass may be substantially higher. Tetraspanins are a group of proteins defined by their four transmembrane regions, which anchor the proteins in the cell membrane (Termini & Gillette, 2017). They generally protrude around 3-5 nm from the cell membrane (Grove, 2014) with two extracellular regions and three short intracellular loops (Termini & Gillette, 2017). Their roles in cell biology involve association with other proteins, including other tetraspanins, and thereby enabling the formation of signalling complexes (Levy & Shoham, 2005).

Next I used qRT-PCR to measure the difference in gene expression between males and females at different d.p.i. for CD63a and CD63R. This was done to confirm the findings of the RNA-Seq analysis. Then, to further investigate the role of CD63a and CD63R in female fertility, *in situ* hybridisations and RNAi

knockdown experiments were performed to determine in which tissues these genes are expressed and whether an effect of these tetraspanins on fertility could be demonstrated.

3.2.9 qRT-PCR analysis of CD63 antigen & CD63 receptor

Pairing-dependent expression of CD63R and CD63a in female worms was confirmed by qRT-PCR (Figure 3.11). Expression in female worms was measured relative to expression in male worms at 21 d.p.i. using PSMD4 (a proteasome 26S subunit) as a reference gene (Liu *et al.*, 2012). CD63R expression in females from MS infections was found to increase dramatically from 21 to 28 d.p.i. and remained high thereafter with expression being over 800-fold higher in females than males at 40 d.p.i. (Figure 3.11).

CD63a expression also increased significantly in females from MS infections from 21 to 28 d.p.i compared to males, and continued to be expressed highly until 40 d.p.i. (Figure 3.12). At 21 d.p.i., females expressed CD63a at 1.25-fold of the male expression, which increased to 56.6-fold at 40 d.p.i., a 45-fold increase in female expression (Figure 3.12).

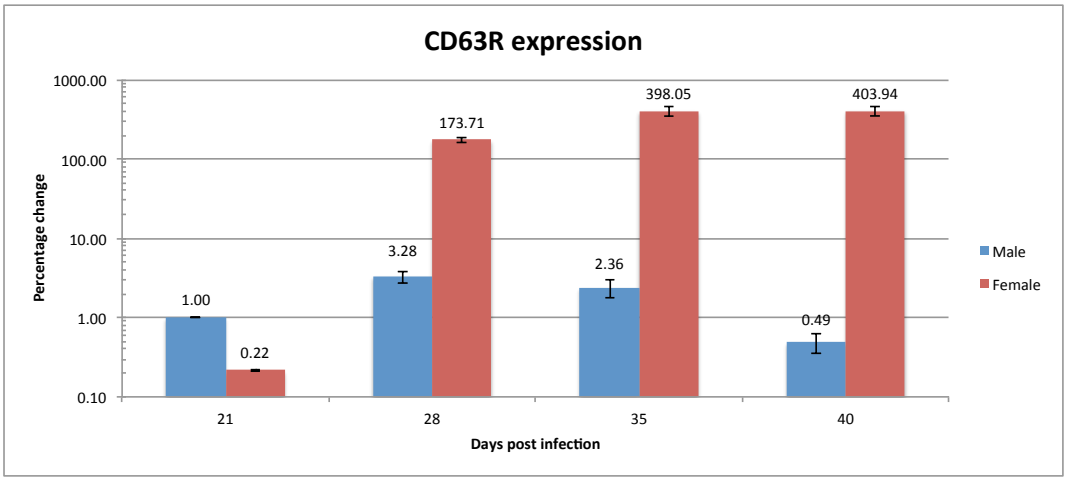


Figure 3.11: Expression of CD63R significantly higher in female worms from 28 d.p.i. onwards. Expression is give as a fold change relative to male expression at 21 d.p.i. PSMD4 was used as internal reference gene for this qRT-PCR analysis. Error bars are the standard error of the mean.

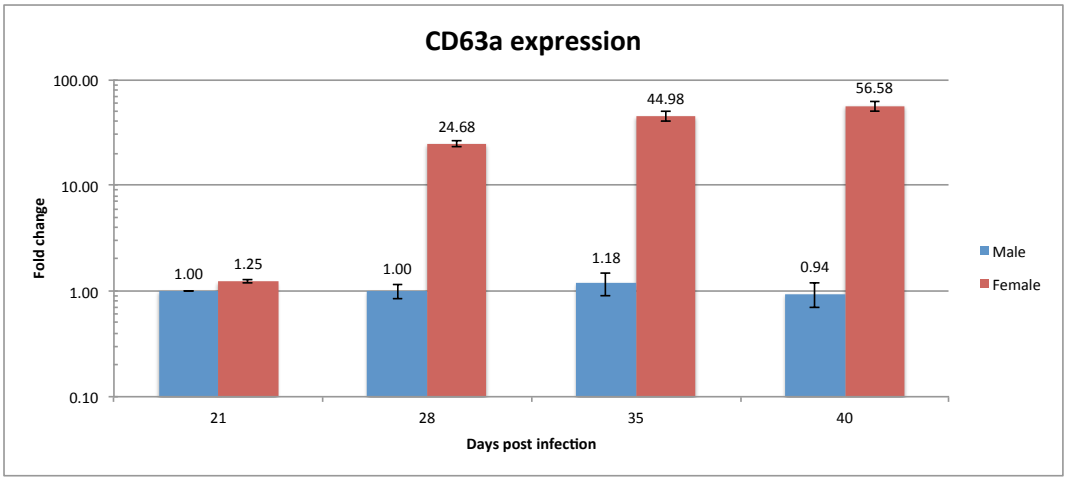


Figure 3.12: Expression of CD63a is higher in female than in male worms. Expression is given as a fold change relative to male expression at 21 d.p.i. PSMD4 was used as internal reference gene for this qRT-PCR analysis. Error bars are the standard error of the mean.

Both RNA-Seq and qPCR data show that there is a strong up-regulation of CD63R and CD63a expression in paired females, whereas expression in males (and, in the case of RNA-Seq data, females from SS infections) remains relatively constant at a low level (Figures 3.11 and 3.12; see Cluster 7 in Figure 3.10).

3.2.10 Whole mount *in situ* hybridisation

The cluster analysis of RNA-Seq data suggested that CD63a and CD63R could be expressed in the vitellaria of MS females and play a role in egg production. *In situ* hybridisation was used to identify the tissues in which expression of the two genes occurs. Tetraspanin-2 (TSP-2) was used as a positive control. TSP-2 is an abundant tegumental protein (Castro-Borges *et al.*, 2011) and its expression at the mRNA level can be demonstrated in the cell bodies of the tegument (see Figure 3.13). These cell bodies lie outside the light brown vitelline lobes but no staining was observed inside the vitelline tissue (see Figure 3.14 for a higher magnification).

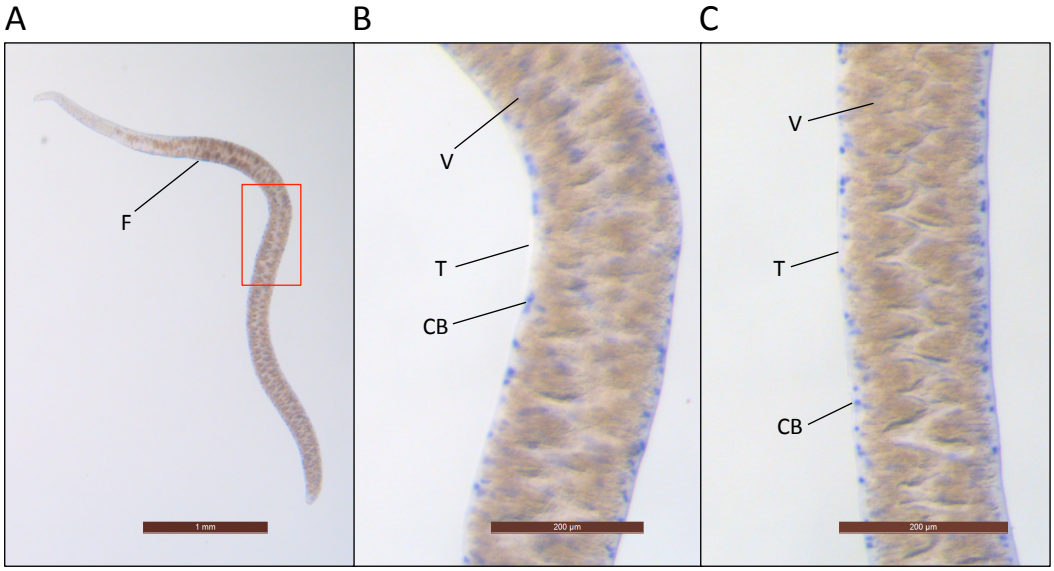


Figure 3.13: Tetraspanin-2 is expressed in the cell bodies of the worm tegument. Alkaline phosphatase staining (Chapter 2.27) of digoxigenin-labelled tetraspanin-2 mRNA in female worms. Expression was not observed in the vitelline tissue. F – Sexually mature female; V – Vitellarium; T – Tegument; CB – Cell body. A) Whole mature female with staining located in the cell bodies of the tegument. The red rectangle indicate where the close-up in B) is derived from. B) Close-up of the female specimen in A). Stained cell bodies can be seen in the tegument. C) Close-up of a different female specimen. Stained cell bodies can be seen in the tegument demonstrating tetraspanin-2 expression there.

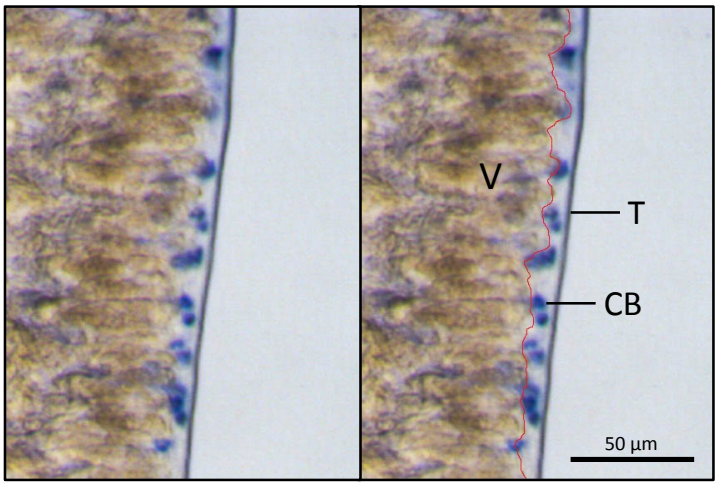


Figure 3.14: A close-up of the stained cell bodies indicates tetraspanin-2 is expressed in the cell bodies of the tegument. The original image is on the left; on the right the border of the vitellarium has been marked in red. V – vitellarium; T – tegument; CB – cell body.

In contrast to TSP-2, CD63a expression was detected inside the vitelline tissue of the 12 female specimens (see Figure 3.15). No expression was detected in the negative control (sense RNA) or in male worms (pictures not shown here).

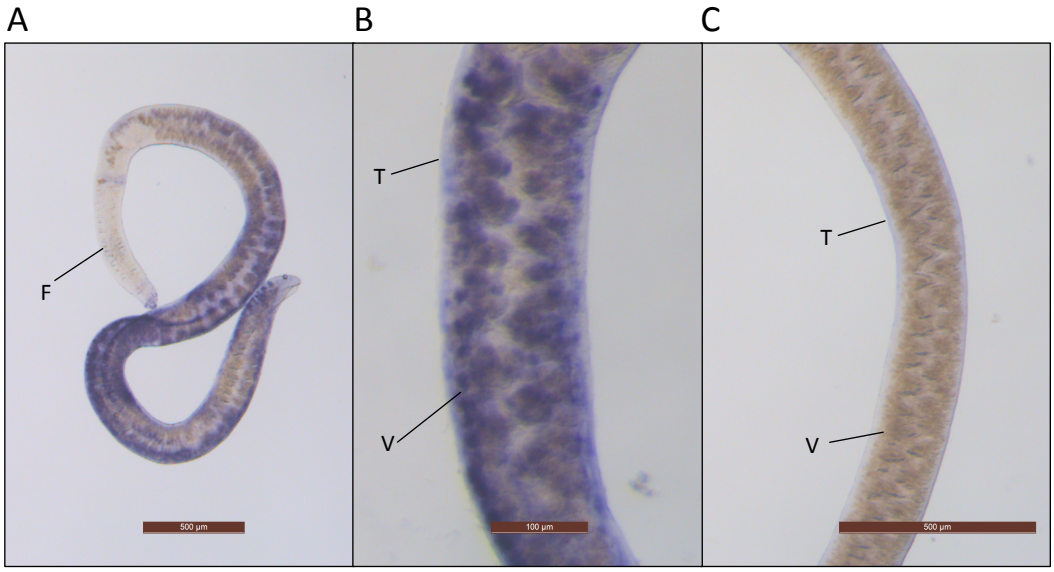


Figure 3.15: CD63 antigen is expressed in the vitelline tissue of female worms. Alkaline phosphatase staining (Chapter 2.27) of digoxigenin-labelled CD63 antigen mRNA in female worms. F – Sexually mature female; V – Vitellarium; T – Tegument. A) Whole female stained for CD63 antigen. B) Close-up female mid-section. C) No signal was found in the negative control.

The pattern of staining observed for the CD63R (see Figure 3.16) was similar to CD63a, showing expression in the vitellaria. Particularly, 14 female specimens displayed CD63R expression in the outer regions of the vitellaria, but little or no staining in the centre of the organ (Figure 3.15). No expression was observed in the female negative control or male worms (Figure 3.15). Erasmus (1975) described the organisation of vitellocytes. The author suggested that there is a gradient of maturity from the least mature (S1) cells at the periphery of the lobes, then S2 and S3 to the most mature (S4) cells in the centre of the female near the vitelline duct.

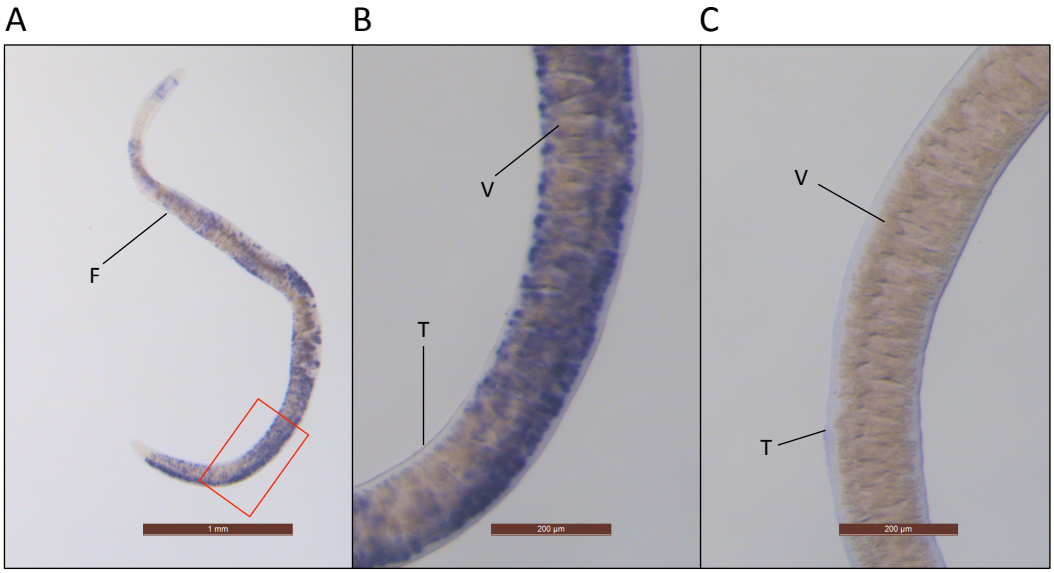


Figure 3.16: CD63 receptor is expressed in the vitelline tissue of female worms. Alkaline phosphatase staining (Chapter 2.27) of digoxigenin-labelled CD63 receptor mRNA in female worms. Staining was found to be most intense in the outer portions of the vitellaria. F – Sexually mature female; V – Vitellarium; T – Tegument. A) Whole female stained for CD63 receptor. A red rectangle indicate the section of the image that the close-up in B) was derived from. B) Close-up of female mid-section.. C) No signal was found in the negative control.

The results of the whole mount *in situ* hybridisation confirm expression in the vitellaria of mature female worms, as suggested by the cluster analysis of RNA-seq data of CD63a (Smp_173150) and CD63R (Smp_155310) (Appendix A.24). However, CD63R expression seemed to take place in the outer regions of the vitellaria (Figure 3.16), possibly due to expression in particular subpopulations of vitelline cells.

3.2.11 RNA interference

In order to determine if CD63a and CD63R play a significant role in the production of eggs, RNA interference (RNAi) was used (Chapter 2.2.2). Soaking in dsRNA was chosen, rather than electroporation, to maintain fertility as much

as possible and the number laid eggs was the measured phenotype. The positive control knockdown, TSP-2, was successful, with qRT-PCR showing a reduction of mRNA levels to 0.5% of those in the negative control (see Figure 3.17).

In contrast, CD63R expression was found to be higher in the treated sample compared to the negative control (see Figure 3.18).

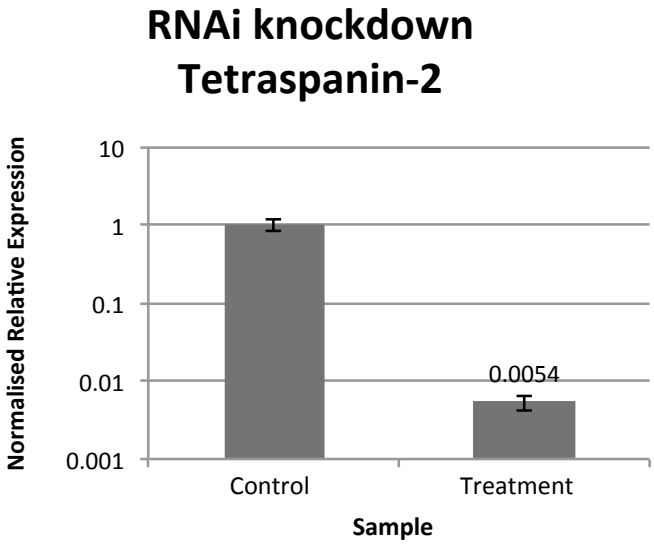


Figure 3.17: Tetraspanin-2 expression was reduced significantly to 0.5% of expression in negative control. Using the 'soaking method', an RNAi knockdown of tetraspanin-2 was performed in mature females. Knocked down tetraspanin-2 expression was measured relative to the negative control samples (soaked in non-specific dsRNA) using qRT-PCR. Two biological replicates were used for treatment and the control group. "Proteasome 26S Subunit, Non-ATPase 4" (PSMD4) was used as internal reference gene. The difference was found to be statistically significant ($p = 0.0083$). Error bars represent the standard error of the mean.

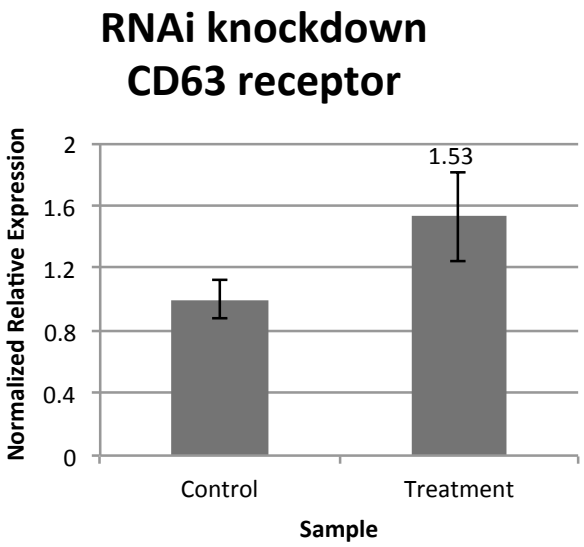


Figure 3.18: An increase in CD63R expression was observed relative to the negative control. RNAi knockdown in mature female worms, soaked in CD63a-specific dsRNA, was measured relative to negative negative control worms (soaked in non-specific dsRNA) using qRT-PCR. Two biological replicates were used for both treatment and control. "Proteasome 26S Subunit, Non-ATPase 4" (PSMD4) was used as internal reference gene. The difference was found to be statistically non-significant ($p = 0.086$). Error bars represent the standard error of the mean.

For CD63a, a moderate knockdown of around 35% was measured (see Figure 3.19). In negative controls, which did not undergo reverse transcription, amplification was only observed over six cycles after amplification in the samples suggesting that the DNase treatment successfully removed most potential gDNA contamination (Appendix A.25). In the no template control, amplifcations was only observed over 10 cycles after amplification in the sample showing that little non-specific amplification was taking place (Appendix A.25). TSP-2 is expressed in the tegumental cell bodies (Tran *et al.*, 2010), which may be more accessible to the dsRNA than the vitellarian tissue deeper inside the worm. This may explain why the TSP-2 knockdown was successful compared to the CD63a and CD63R knockdowns.

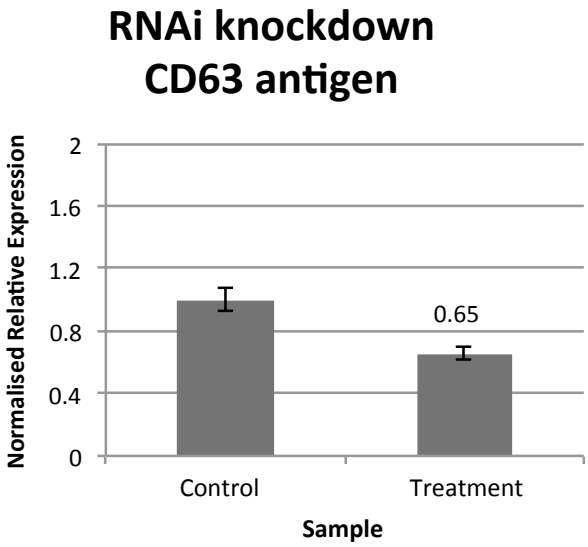


Figure 3.19: A moderate knock of CD63a expression of 35% was observed. RNAi knockdown of CD63R was measured in mature females relative to the negative control (worms soaked in non-specific dsRNA) using qRT-PCR. Two biological replicates were used for both treatment and control. "Proteasome 26S Subunit, Non-ATPase 4" (PSMD4) was used as internal reference gene. The difference was found to be just short of statistical significance ($p = 0.053$). Error bars represent the standard error of the mean.

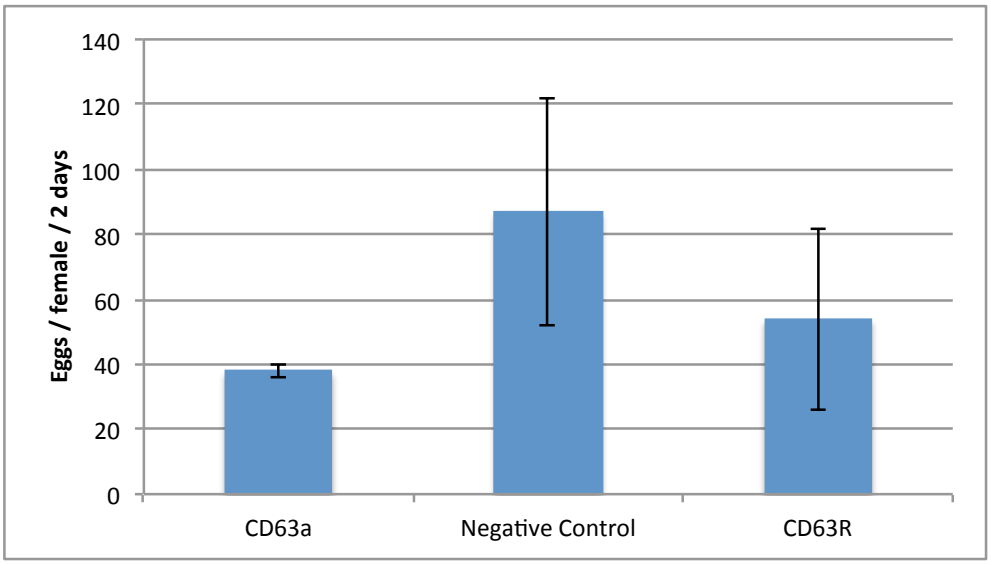


Figure 3.20: Reduction in laid eggs by the CD63a & CD63R knock-down worms was not statistically significant. An ANOVA showed that the treatment groups did not differ sginificatnly from the negative control (p = 0.485). Egg production was measured as a phenotype of RNAi knockdown. Eggs were counted two days after RNAi knockdown began. The number of eggs is given per female. Error bars represent the standard error of the mean of the three biological replicates.

A moderate reduction in eggs was observed in both knockdown groups relative to the negative control group (see Figure 3.20), although no strong reduction in gene expression could be demonstrated and in the case of CD63R a 50% higher expression was measured compared to the negative control. However, because only two biological replicates could be set up for each condition, an analysis of variance (ANOVA) found there to be no statistically significant difference between the groups (p = 0.485).

In the RNA-seq data, high levels of CD63R expression were observed in mature females. However, the qRT-PCR analysis indicated low expression of CD63R in the control and the treated samples. Therefore, I investigated whether there was a change in mRNA expression between freshly perfused worms and those

cultured *in vitro* for seven days, as had been done for the knockdown experiment. From the batch of worms used for the RNAi soaking experiment several worms had been placed in Trizol directly following the perfusion, which were used as *in vivo* control.

The qRT-PCR analysis showed that CD63R expression was 230-fold higher in freshly perfused worms compared to worms after seven days *in vitro*, whereas CD63a expression was only 14-fold higher in the freshly perfused worms compared to worms after *in vitro* culture (see Figure 3.21), probably due to the fact that CD63a is expressed at higher base levels in all worms compared to CD63R. Some down-regulation of these genes *in vitro* had been anticipated even in the negative control because of the regression of female fertility in cultre (Galanti *et al.*, 2012). However, the extent to which expression of CD63a and CD63R declined was unexpected. The down-regulation in mRNA expression with such a large effect size both in the control and RNAi treated worms is likely to interfere with an accurate measurement of the effect of the RNAi treatment and obscure the effect of the RNAi knock down.

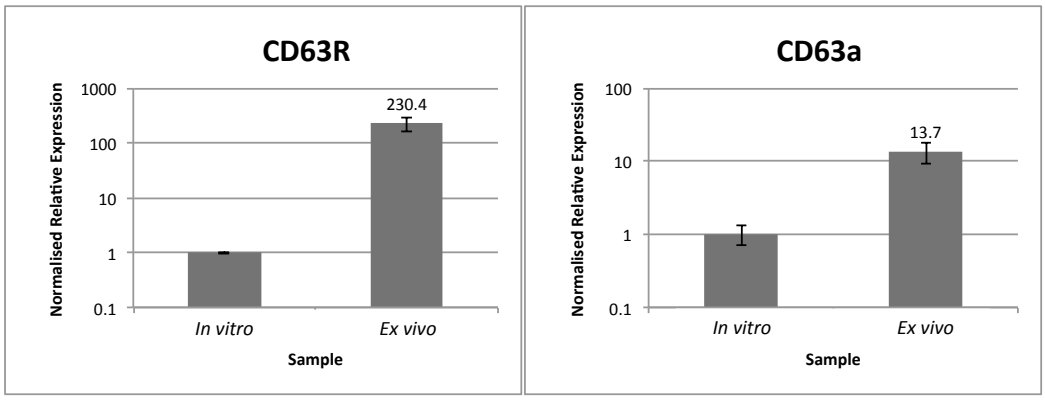


Figure 3.21: Expression of CD63 receptor (left) and CD63 antigen (right) was significantly higher prior to *in vitro* culture. Expression was measured in freshly perfused (*ex vivo*) worms and worms cultured for seven days (*in vitro*). For both qRT-PCR reactions PSMD4 was used as internal reference gene. The differences in expression were statistically significant for both genes (CD63R: $p = 0.016$; CD63a: $p = 0.036$). Error bars represent the standard error of the mean.

The experiment showed that TSP-2 mRNA levels could be depleted by approximately 99.5% by soaking worms in dsRNA (Figure 3.17). However, the soaking did not reduce CD63R mRNA and only resulted in a mild reduction of CD63a mRNA. To facilitate better access of dsRNA molecules to the deeper tissues such as the vitellarium, the treatment was repeated, using electroporation instead of simple soaking as well as a shortening of the experiment from one week to 48 hours. A potential disadvantage of electroporation is the greater level of stress caused to the female and the impact on their egg production this might have, which could obscure any subtle effect the knockdown might have on egg production. The length of the experiment was also shortened to lessen the effect of *in vitro* culturing and reduce the decline in CD63a and CD63R expression in the negative control that was demonstrated previously.

In this RNAi experiment, a very moderate reduction of about 16% in CD63R mRNA levels was observed (Figure 3.22). This reduction was well within the standard error of the normalised expression levels (see Figure 3.22). The results of this experiment did not allow the null hypothesis to be rejected. With more replicates it might be possible to determine, whether the small observed effect seen was reproducible or due to chance. Nonetheless, this result is an improvement on the previous soaking knockdown where an increase in CD63R mRNA levels was observed in the treated sample relative to the negative control.

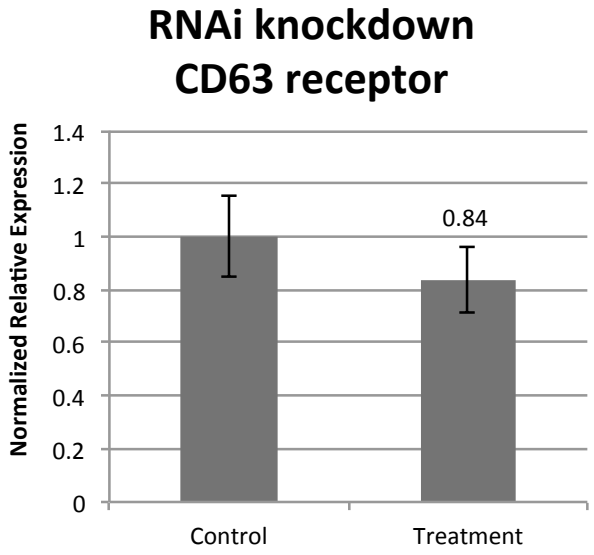


Figure 3.22: No significant knock-down of CD63 receptor expression was observed. RNAi knockdown of CD63 receptor using the 'electroporation method' was measured relative to the negative control sample (electroporated with non-specific dsRNA) using qRT-PCR. PSMD4 was used as internal reference gene. The difference in expression was not found to be statistical significant ($p = 0.822$). Error bars represent the standard error of the mean.

In the CD63a knockdown experiment, a statistically significant ($p = 0.044$) 47% reduction in CD63a mRNA levels was measured with much smaller standard errors measured compared to the CD63R knockdown (see Figure 3.23). In this experiment, the knockdown was found to be more successful than the soaking

experiment [reduction by ~35% (Figure 3.19)]. Taken together, these results suggest that CD63a is targetable by RNAi.

For measuring expression of both CD63R and CD63a, PSMD4 was used as a reference gene. The negatve controls, that had not been reverse transcribed, only showed amplification about seven cycles after the samples suggesting that the DNase treatment successfully removed most gDNA contamination (Appendix A.26). The no-template controls showed no amplification until about ten cycles after the samples suggesting that there was little non-specific amplification (Appendix A.26).

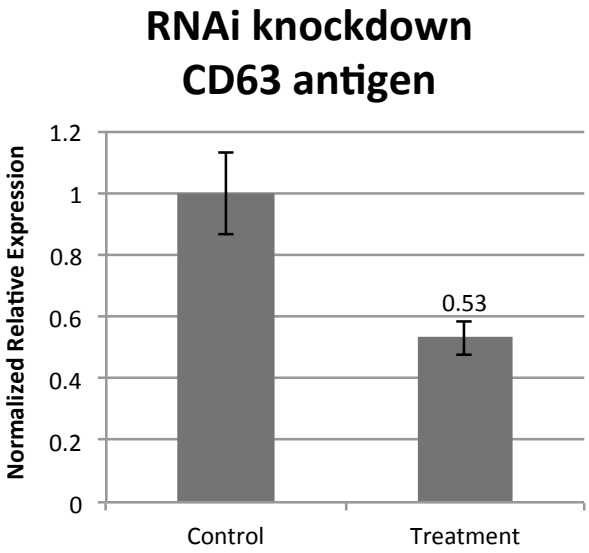


Figure 3.23: A significant knock-down in CD63 antigen expression was observed. RNAi knockdown of CD63 antigen using the 'electroporation methods' was measured relative to the negative control sample (electroporated with non-specific dsRNA) using qRT-PCR. PSMD4 was used as internal reference gene. The difference in gene expression was statistical significant ($p = 0.044$). Error bars represent the standard error of the mean.

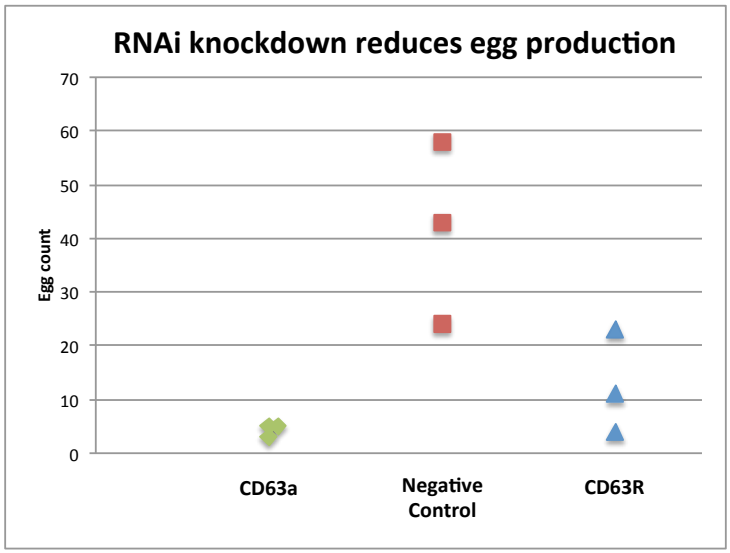


Figure 3.24: The number of laid eggs was reduced significantly in the CD63a and CD63R knockdown worms. The number of eggs per female is plotted for the knockdown of CD63 antigen, the negative control and the CD63 receptor; three replicates were available for each condition. The eggs were laid during the 48 hours following RNAi electroporation. An ANOVA showed that the treatment groups laid significantly fewer eggs compared to the negative control ($p = 0.0156$). The analysis also showed that both, the CD63a and CD63R knockdown resulted in a significant reduction of eggs compared to the negative control (CD63a: $p = 0.00679$; CD63R: $p = 0.0201$).

Again, a reduction in egg laying was observed in both knockdown experiments (see Figure 3.24). A 9.6-fold reduction of egg laying was observed in worms with knocked down CD63a expression and a 3.2-fold reduction in worms with knocked down CD63R expression, relative to the negative control (Figure 3.24). Here, three biological replicates were analysed and the observed reduction in eggs laid was greater than in the soaking experiment (Figure 3.20). As a result, an ANOVA showed that the reduction in eggs laid was found to be statistically significant ($p = 0.0156$) but also that CD63a and CD63R knockdowns significantly affected egg production compared to the negative control (CD63a: $p = 0.00679$; CD63R: $p = 0.0201$). Notably, the number of eggs was overall much lower (approximately 10%) compared to the egg counts in the soaking

knockdown experiment (compare Figures 3.20 and 3.24), which might be due to the effect of electroporation on the females. Although only a moderate reduction of mRNA was observed for CD63a, the significant reduction in egg production is strongly suggestive of a real effect and not chance. This is especially true, as CD63a expression was reduced to a similar extend in the soaking experiment where egg laying was also reduced (albeit not statistically significant) and therefore showing the same trend (compare Figures 3.20 and 3.24).

CD63R expression, however, was reduced to a much smaller extent (Figure 3.23), yet a phenotype was measured at the egg laying level (Figure 3.24). This was also observed in the soaking experiment, where a (non-significant) reduction in egg production was observed (Figure 3.20). At worst, this could suggest that the egg laying phenotype is not a suitable phenotype, as egg production can be sensitive to many other factors, such as stress and injury caused during the perfusion. If this is the case, then a lot of random fluctuations should be seen within replicates. To some extent this can be seen, certainly in the negative control, but also CD63R (Figure 3.24). However, the egg counts for the CD63a were remarkably similar (5, 5 and 3 eggs in each well respectively) and much smaller than even the lowest negative control sample (Figure 3.24). A similar trend was observed in the soaking experiment, giving me further confidence that the observed effect of the CD63a knockdown on egg production was real.

3.3 Discussion

The aim of this chapter was to perform an analysis of gene expression of the developing male and female worms inside the mammalian host. The differential expression analysis reproduced many findings in the published literature, validating the approach. The analysis of male worms showed clearly that the male transcriptome was largely unaffected by pairing, with only a handful of genes consistently differentially expressed between males from SS and MS infections (Leutner *et al.*, 2013). Females on the other hand are known to mature in a pairing dependent manner from the fourth week of infection (Biolchini *et al.*, 2006), with my RNA-Seq data showing the greatest increase in differentially expressed genes between females from SS and MS infections between 21 and 28 d.p.i. Also, in MS infections, egg laying was observed from 35 d.p.i. onwards in addition to egg production related gene expression reaching its peak at 35 and 38 d.p.i. This includes genes encoding eggshell synthesis domain containing proteins, tyrosinases, extracellular superoxide dismutases, as well as genes coding for the tetraspanin CD63 antigen and receptor, showing that paired female worms reached sexual maturity around this time. Another up-regulated gens was TRAP1, a heat shock protein mainly localised to the mitochondria, that protects cells from oxidative stress by inhibiting the respiratory complex II and protecting the cell from apoptotosis (Guzzo *et al.*, 2014). Tyrosinases expressed in the vitellaria are essential for proper egg development in schistosomes (Fitzpatrick *et al.*, 2007). However, tyrosinase activity is generally thought to result in oxidative stress, as it has been shown in other systems such as mosquitoes (Christensen *et al.*, 2005). The extracellular superoxide dismutase

and TRAP1 may help cells in the vitellaria to cope with the oxidative stress resulting from the tyrosinase activity to prevent damage to the tissue and allow a high rate of continuous egg production. Together, the results confirmed many findings of the published literature (Fitzpatrick & Hoffmann, 2006; Leutner *et al.*, 2013) but provided a more comprehensive picture of when these sex specific genes become differentially expressed.

When comparing male and female worms at different time points, but particularly at 38 d.p.i., a significant up-regulation of nicotinic receptors was observed in male worms (Table 3.3). Nicotinic receptors represent a group of drug targets in many parasitic nematodes which are targeted by anthelmintics such as levamisole, pyrantel and oxantel (Williamson *et al.*, 2009). These drugs cause the ligand-gated ion channels to open, resulting in "prolonged muscle contraction and spastic paralysis of the worms" (Williamson *et al.*, 2009). Such receptors have recently also become of interest as a drug targets for *S. mansoni* as described by MacDonald *et al.* (2014). The authors observed that the *S. mansoni* nicotinic acetylcholine receptors do not exert an excitatory effect on the neuromuscular system as in mammals and nematodes, but instead function as inhibitory neuromuscular receptors (MacDonald *et al.*, 2014). MacDonald *et al.* (2014) localised two of these proteins to the worm surface as well as in nerves. Therefore, the up-regulation of neurotransmitter receptor genes in male worms could be the result of the greater surface area as well as greater proportion of muscle tissue in male worms in comparison to mature female worms.

During the analysis of the juvenile male worm, genes related to the cytochrome P450 related drug metabolism were found to be up-regulated. Cytochrome P450-dependent metabolism plays an important role in processing praziquantel in humans (Li *et al.*, 2003) but whether this is true for schistosomes is not clear. This pathway was found to contain several differentially expressed genes encoding glutathione S transferases, one of which (Smp_102070) has previously been shown to bind praziquantel and play an important role in the detoxification of *S. mansoni* (McTigue *et al.,* 1995; Cardoso *et al.*, 2003). Another GST encoding gene (Smp_024010) was also found to be down-regulated after 18 d.p.i.. From day 18 to 38, its expression decreased 21.9 and 18.2-fold in male and female worms respectively, potentially making the worm more susceptible to the effects of praziquentel. McTigue *et al.* (1995) showed that homodimers of GST bind praziquantel at the dimer interface, not in their substrate-binding site and concluded that praziquantel inhibits GST activity, which leads to a build-up of reactive oxygen species and toxic metabolites inside the parasite, in particular cytotoxic free radicals. Cardoso *et al.* (2003), on the other hand, concluded that the binding of praziquantel to GST had a protective effect on the worm by keeping praziquantel away from its real cellular target [for example voltage operated calcium ion channel (Nogi *et al.*, 2009)]. Given the rapid effect of praziquantel on schistosomes (Xiao *et al.*, 1985), a build-up of reactive oxygen species and toxic metabolites does not seems as plausible as the conclusion of Cardoso *et al.* (2003). However, in either case, a greater expression of GST encoding genes such as Smp_102070 and Smp_024010 would confer a protective effect on the schistosomes.

When examining the expression of testes related genes, a number of interesting similarities and differences were observed between males from SS and MS infections. The comparison between male worms from MS infections at 21 and 28 d.p.i. revealed that they appear to use a similar mechanism to prevent the premature proliferation of prospermatogonia as mammals. Basonuclin 2 (Smp_138350), a testes specific gene in mammals and *S. mansoni*, was found to be up-regulated in male worms from SS and MS infections prior to pairing at around 28 d.p.i. Vanhoutteghem *et al.* (2014) reported that "basonuclin 2 is required for proper mitotic arrest, prevention of premature meiotic initiation and meiotic progression" of germ cells. In knock out mice, prospermatogonia multiply excessively leading to accumulation of spermatocytes and ultimately apoptosis during meiotic prophase (Vanhoutteghem *et al.*, 2014). Adult male worms are known to produce sperm in the absence of a female partner (Armstrong, 1965); therefore, the role of this gene in schistosomes may be to allow male worms to mature sufficiently before investing energy into sperm production. However, differential regulation of genes potentially involved in sperm production were also observed. Neves *et al.* (2005) demonstrated morphological differences between the testes of males from SS and MS infections, despite both producing sperm. In the male time series analysis, genes encoding the FGF receptor activating protein (Smp_035730, adjusted p-value = 0.0053) and the epidermal growth factor receptor (Smp_152680, adjusted p-value = 0.0069) were found to be up-regulated in paired males. These genes were shown to be expressed in the testes at significantly higher levels than in other organs (see Chapter 5). Both FGF and the EGF signalling pathways regulate cell proliferation and differentiation (Eswarakumar *et al.*, 2005; Herbst, 2004).

This may assist to regulate sperm production in male worms and allow unpaired male worms to conserve energy.

In female worms from SS and MS infections, a much larger number of differentially expressed genes were observed than in male worms. When examining females between 21 and 28 d.p.i. a large proportion of the significantly regulated genes were found to encode transcription factors. This included approximately a third of the genes up-regulated at day 21 (37/119), suggesting that some developmental processes finished then, but also that the regulation of the transcriptome is changing significantly around the time pairing is thought to occur. Interestingly, a comparison of females from SS and MS infections at 28 d.p.i. showed that down-regulation of these genes encoding transcription factors occurs independently of pairing at this stage of the worm development. Female worms became developmentally arrested around this time in the absence of male worms (Figure 3.2), which stimulate further maturation. The down-regulation of these genes coding for transcription factors probably marks the beginning of this developmental arrest. However, at least six genes encoding transcription factors were found to be up-regulated at 28 d.p.i., four of which were found to be consistently associated with sexual maturity in female worms, with expression being pairing-dependent. Expression of these four genes was detected at very low levels in most samples, but a significant increase in expression in females from MS infections was measured from 28 d.p.i. onwards. Although their precise function remains unknown, their expression patterns over the RNA-Seq time course caused them to be clustered with genes expressed

predominantly in the vitellaria of mature female worms. This is further evidence of a changing regulation of transcription in paired females.

Using two clustering algorithms, MBCluster and Kohonen, a set of 76 female specific genes putatively involved in egg production was identified, a subset of which has previously been shown to play important roles in female reproduction (for example Fitzpatrick *et al.*, 2007; Fitzpatrick & Hoffmann, 2006). Due to their functions as cell surface receptors and their potential involvement in signalling, the CD63 antigen and tetraspanin CD63 receptor stood out. Further investigation with qRT-PCR confirmed the RNA-Seq expression patterns. Both genes were expressed in paired females, beginning around 28 d.p.i. To get more insight into the role of these genes and add evidence for their role in egg production whole mount *situ hybridisation* (WISH) and RNAi knockdown experiments were performed for both genes. The WISH experiments showed that both genes were expressed in the vitellarian tissue of the female worms, although the CD63 receptor was expressed in the outer portions of the vitellarium, whereas the CD63 antigen was found to be expressed throughtout the vitellarian tissue.

The RNAi knockdown experiment proved challenging. In the first round of knockdowns, using a soaking technique, the positive control, TSP-2, worked well with a reduction of mRNA to 0.5% of the original expression. However, knockdown efficiency for CD63a and CD63R was far less satisfactory. CD63a expression had been reduced by 35% whereas expression of CD63R was measured to be 53% higher after RNAi treatment compared to the control samples. This unexpected result may have been caused by the effects of the *in*

vitro culture. Worms were found to have significantly lower expression of CD63a (~14x) and CD63R (~230x) after *in vitro* cultre, compared to freshly perfused worms (Figure 3.21). It has previously been shown (Galanti *et al.*, 2012) that female worms lose maturity and regress outside their host even in the presence of male worms; however, the extent of this effect was greater than expected. In a repeat of this experiment, I opted to reduce the length of the *in vitro* culture after RNAi treatment to 48h, instead of 7 days. Furthermore I chose to use electroporation rather than soaking in the second experiment to maximise probe penetration, at the expense of worm fitness. This experiment proved to be more successful with a measured reduction of CD63R mRNA of 16% (instead of a measured increase) and a reduction of CD63a mRNA of 47% (instead of 35%). Finally, the number of eggs laid by the RNAi treated worms was significantly reduced for both CD63R and CD63a by 3.2 and 9.6-fold respectively, suggesting that both genes play a role in the production of eggs. The CD63 antigen is known in humans to form complexes with integrins and act as part of signalling complexes that regulate cell proliferation and differentiation (Tugues *et al.*, 2013). For example CD63a is known to interact with syntenin-1 – an adaptor protein regulating vesicle formation and TGF-β mediated SMAD activation (Hwangbo *et al.*, 2016) – and other adaptor proteins of the clathrin-dependent endocytosis pathway (Charrin *et al.*, 2014) as well as the "tissue inhibitor of metalloproteinase-1" and integrins (Jung *et al.*, 2006). The interactions with syntenin-1 and the adaptor proteins occur at the same location on the CD63a protein and are thought to be mutually exclusive (Berditchevski & Odintsova, 2007). By interacting with syntenin-1, CD63a might play a role in the regulation of TGF-β signalling as well as the formation of vesicles, potentially in the

vitellocytes of the maturing female. Mature (S4) vitellocytes are known to contain a large number of vesicles containing eggshell protein precursors (Cogswell *et al.*, 2012). The CD63 receptor, on the other hand, may play a role in endocytosis via clathrin-dependant as well as independent pathways (Berditchevski & Odintsova, 2007) and may aid in the formation of the vitelline cells' secretory vesicles that will later go on to form the eggshells. Another mechanism by which CD63 antigen may regulate vitelline proliferation, namely apoptosis, will be examined in more detail in Chapter 5.