

Transcriptomic studies on host-parasite  
interactions in *Schistosoma mansoni*  
intramammalian stages

University of Cambridge

Pembroke College



A thesis submitted for the degree of

*Doctor of Philosophy*

Arporn Wangwiwatsin

The Wellcome Trust Sanger Institute

March 2017



# Declaration

The work presented in this thesis was carried out at the Wellcome Trust Sanger Institute (Hinxton) between October 2012 and March 2017. This dissertation is the result of my own work – contributions from collaborations are clearly referenced. No part of this dissertation has been or is being submitted for any qualification in any other university. This thesis does not exceed the word limit established by the Biology Degree Committee.

# Acknowledgements

First of all I would like to pass a massive gratitude to my supervisor, Matt, for giving me this opportunity, and for the understanding, patience, guidance, and support throughout. Thanks to everyone in the Parasite Genomics team including past and current members. Thanks for all the ideas, the fruitful discussions, and all the fun in team outing. Many people also helped read this thesis: Matt, Gabriel, Patrick, Adam, Kate, Anna, Avril, and Shona. Thank you to my thesis examiners for their contributions to the improvement of this thesis. Thank you Christina and Annabel from the Graduate Office for all their help with all questions I can ever ask. Thank you to Marc who has made every printing request possible.

Many people have been amazingly supportive and helpful throughout my work at the Sanger Institute, sharing their knowledge, helping make my experiment endeavour come true and many have become good friends. Thanks CGaP, Sally Forrest, Christ Kirton, Emma Goss, Theresa Feltwell, Liam Prestwood, Laura Wood, James Hewinson, Gavin Wright, Paul Kellam, Mercedes Pardo, Jyoti Choudhari, Lira Mamanova, Rachael Wash, Sarah Smith, Sally Linsdell, Derek Pickard, Dave Goulding Simon Clare, Cordelia Brandt, Colin Barker, Carol Smee, Dan Mead, Theodoros Roumeliotis (especially for his help with using the IPA), and the Sanger librarians.

Thanks for all support and joy beyond the work at Sanger Institute including Pembroke College, Public engagement teams at the Sanger and Cambridge, the CHaOS Society, Cambridge Thai Society, University Counselling Service, Polly Brown, family, and the husband.

It has been a wonderful experience.

# Summary

The life cycle of the parasitic flatworm *Schistosoma mansoni* is split between snail and human hosts. In humans, it lives in the bloodstream and can survive for over 10 years, during which time, interactions with the host are essential for its survival. Starting the infection, cercariae penetrate human skin and become schistosomules which enter blood vessels. The schistosomules follow blood circulation to the lung and the liver where they develop into adults which migrate to the mesenteric venules for egg-laying. The association with certain host tissues and egg-laying site may involve interactions and adaptations. Moreover, the parasite needs to evade or modulate host immune responses and to optimise its acquisition of host metabolites. Molecular mechanisms involved in these interactions are not completely understood.

To gain deeper understanding into the interactions of *S. mansoni* and its mammalian host, I sought to understand what biological processes are required for successful intramammalian infection; specifically, what guides the tissue tropism of the parasite and what roles do the host tissues play in the infection? In the first part, parasite transcriptomes were produced from *S. mansoni* obtained from experimentally infected mice. The dataset includes novel transcriptomic profile of the lung stage, and covers developmental and egg-laying stages. In the next two parts, co-culture experiments were set up using mechanically transformed schistosomules and cells derived from human tissues. Transcriptomic profiles of the co-cultured schistosomules were investigated in the second part of this thesis; and transcriptomic profiles of the co-cultured human cells were investigated in the final part. The outputs of this thesis provide new insights into the infection biology, provide a large data resource for the research community, and propose avenues for further investigation and characterisation of interaction mechanisms.

# Abbreviations

BLAST	Basic Local Alignment Search Tool
C1	Complement component 1
CFH	Complement factor H
CFI	Complement factor I
DAF	Decay-accelerating factor
DMEM	Dulbecco's Modified Eagle's Medium
DMT	Divalent metal transporter
ES	Excretory/secretory products
ESTs	Express Sequence Tags
FPKM	Fragments Per Kilobase of transcript per Million mapped reads
GO	Gene Ontology
GPCR	G-protein coupled receptor
ICAM1	Intercellular adhesion molecule 1
IL-2	Interleukin-2
IPA	Ingenuity Pathways Analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
LDL	Low density lipoprotein
Log <sub>2</sub> FC	Log <sub>2</sub> fold change
LPS	Lipopolysaccharide
MEG	Micro-exon gene
NPC2	Niemann Pick type C2 protein
PCA	Principal component analysis
PDB	Protein Data Bank
RNA-seq	RNA-sequencing
ROS	Reactive oxygen species
SELE	Selectin E
TAL	Tegumental-allergen-like
TGF- $\beta$	Transforming growth factor-beta
TNF- $\alpha$	Tumour-necrosis factor-alpha
VCAM1	Vascular cell adhesion molecule
WTSI	Wellcome Trust Sanger Institute

# Table of Contents

## Chapter 1

<b>Introduction .....</b>	<b>1</b>
1.1 Overview .....	1
1.2 Introduction to schistosomes.....	2
1.2.1 Schistosome life cycle.....	2
1.2.2 Schistosomiasis .....	5
1.3 <i>S. mansoni</i> life in bloodstream .....	7
1.3.1 Host-parasite interactions: the interfaces .....	7
1.3.2 Motility, behaviour, and homeostasis: the nervous system .....	11
1.3.3 Migration route and development .....	12
1.3.4 Metabolic requirements and acquisition .....	17
1.3.5 Immune evasion.....	22
1.3.6 Section summary .....	30
1.4 <i>S. mansoni</i> genome and transcriptome.....	30
1.4.1 Genome and gene annotation .....	30
1.4.2 Transcriptomes .....	33
1.5 This thesis .....	34

## Chapter 2

<b>Materials and methods .....</b>	<b>35</b>
2.1 Parasite materials.....	35
2.1.1 Maintenance of the <i>S. mansoni</i> life cycle .....	35
2.1.2 Snail husbandry.....	36
2.1.3 Mouse maintenance .....	37
2.2 Molecular methods.....	37

2.2.1	RNA extraction from parasites.....	37
2.2.2	RNA extraction from human cells .....	38
2.2.3	Measurements of RNA concentration and purity .....	38
2.2.4	Library production and sequencing.....	39
2.2.5	Overall QC of sequencing outcome.....	40
2.3	Data analysis .....	41
2.3.1	Mapping and quantifying read counts .....	41
2.3.2	Read count and differential expression analysis.....	42
2.3.3	Genes clustering by timecourse expression profile.....	42
2.3.4	GO term enrichment.....	43
2.3.5	Pathway enrichment and pathway network .....	43
2.3.6	Pathway comparison between cell types.....	44
2.3.7	Protein structural prediction .....	45
2.3.8	Protein domains and motif search .....	45
2.3.9	Gene phylogenetic tree.....	45
2.3.10	Artemis and BamView.....	45
2.3.11	Seaview sequence alignment .....	46
2.4	Data presentation .....	46

## Chapter 3

### ***S. mansoni* in experimentally infected mice..... 47**

3.1	Introduction .....	47
3.1.1	Overview.....	47
3.1.2	Host-parasite interactions in intra-mammalian stages.....	47
3.1.3	Progress in transcriptomic and genomic approaches.....	49
3.1.4	Aims and approaches .....	49
3.1.5	Chapter outline.....	50
3.2	Methods.....	50
3.2.1	Mouse infection .....	50
3.2.2	Parasite and tissue collection from infected mice .....	52



3.2.3	RNA extraction and library preparation .....	54
3.2.4	Morphology score .....	54
<b>3.3</b>	<b>Results.....</b>	<b>55</b>
3.3.1	Worm morphology .....	55
3.3.2	RNA quantity and quality.....	58
3.3.3	Sequencing yields.....	58
3.3.4	Sequencing data overall profiles .....	60
3.3.5	Lung stage .....	64
3.3.6	Liver stages.....	76
3.3.7	Adult stages.....	83
<b>3.4</b>	<b>Discussion .....</b>	<b>94</b>
3.4.1	Overview.....	94
3.4.2	Potential effect of collection procedures.....	94
3.4.3	Lung stage signalling.....	95
3.4.4	Lung stage immune evasion .....	96
3.4.5	Complement factor H .....	96
3.4.6	Complement factor I.....	97
3.4.7	Immunomodulation in liver stages .....	97
3.4.8	Implication for intervention.....	98
3.4.9	Expected changes in liver and adult stages .....	98
3.4.10	Liver localisation .....	98
3.4.11	Mesenteric migration.....	99
3.4.12	Micro-exon genes.....	101
3.4.13	Other ways to use this data .....	101
3.4.14	Summary.....	101

## Chapter 4

### ***S. mansoni* in vitro culture with cell lines ..... 103**

4.1	Introduction .....	103
4.1.1	Overview.....	103

4.1.2	<i>S. mansoni</i> and host tissues.....	103
4.1.3	Effects of host environments on <i>S. mansoni</i> .....	104
4.1.4	Aims and approaches .....	105
4.1.5	Chapter outline.....	105
<b>4.2</b>	<b>Methods.....</b>	<b>106</b>
4.2.1	Experimental design .....	106
4.2.2	Preparing human cells: maintaining stock cells .....	108
4.2.3	Preparing schistosomules: media for schistosomules and for the co-culture .....	110
4.2.4	Preparing schistosomules: transforming cercariae into schistosomules	110
4.2.5	Co-culture .....	111
4.2.6	Sample collection .....	112
<b>4.3</b>	<b>Results.....</b>	<b>113</b>
4.3.1	<i>Mycoplasma</i> test.....	113
4.3.2	Worm morphology .....	114
4.3.3	RNA quantity and quality.....	116
4.3.4	Overall profiles of transcriptomes.....	116
4.3.5	Overview of gene expression .....	119
4.3.6	Schistosomules in co-cultured environment: generic responses.....	125
4.3.7	Schistosomule adaptation to HEPG2 environment .....	133
4.3.8	Schistosomule adaptation to HUVEC environment.....	140
<b>4.4</b>	<b>Discussion .....</b>	<b>146</b>
4.4.1	Overview.....	146
4.4.2	<i>In vitro</i> schistosomules and effect of culture methods .....	147
4.4.3	Generic responses .....	149
4.4.4	HEPG2-specific responses.....	150
4.4.5	HUVEC-specific responses.....	151
4.4.6	Summary.....	152

## Chapter 5

### Transcriptomes of cell lines

#### exposed to schistosomes *in vitro* ..... 155

5.1	Introduction .....	155
5.1.1	Overview.....	155
5.1.2	Host responses in <i>S. mansoni</i> infections.....	155
5.1.3	Host tissue responding to <i>S. mansoni</i> .....	156
5.1.4	Aims and approaches .....	156
5.1.5	Outline.....	157
5.2	Methods.....	157
5.3	Results.....	158
5.3.1	RNA quantity and quality.....	158
5.3.2	HUVEC and endothelial cell surface markers .....	159
5.3.3	HEPG2 and liver hepatocytes.....	161
5.3.4	Overall profiles of transcriptomes.....	161
5.3.5	Differential expression between co-cultured vs. worm-free cells .....	163
5.3.6	Pathways between cell types.....	183
5.4	Discussion .....	187
5.4.1	Overview.....	187
5.4.2	On <i>in vitro</i> -adapted cells .....	188
5.4.3	On aged media.....	189
5.4.4	Biological functions affected by parasite co-culture.....	189
5.4.5	On <i>in vivo</i> validation .....	193
5.4.6	Summary.....	193

## Chapter 6

### Discussion and conclusion ..... 195

6.1	Findings and interpretations.....	196
-----	-----------------------------------	-----

6.1.1	Interactions with host defences .....	196
6.1.2	Host reparation .....	198
6.1.3	Responses to environments .....	198
6.2	Limitations .....	201
6.3	Concluding remarks .....	202

## **Chapter 7**

<b>References .....</b>	<b>203</b>
-------------------------	------------

## **Chapter 8**

<b>Appendices .....</b>	<b>235</b>
-------------------------	------------

# List of Figures

Figure 1.1 Life cycle of <i>Schistosoma</i> spp. ....	3
Figure 1.2 The schistosome tegument .....	9
Figure 1.3 Coagulation and complement cascade.....	25
Figure 1.4 Coagulation pathway .....	28
Figure 2.1 Transfer of snail eggs using “The Octopus”.....	37
Figure 2.2 RNA qualities assessed from Agilent Bioanalyzer electropherograms.....	39
Figure 3.1 Experiment layout .....	51
Figure 3.2 Diagram used for morphology scoring.....	55
Figure 3.3 Morphology of <i>in vivo</i> <i>S. mansoni</i> .....	57
Figure 3.4 PCA plots from total reads and down-sized reads.....	59
Figure 3.5 Similarity between 96 clusters.....	61
Figure 3.6 Clusters of genes based on timecourse expression pattern.....	62
Figure 3.7 Clusters of genes based on timecourse expression pattern, with fixed y-axes .....	63
Figure 3.8 Volcano plots of all consecutive pairwise comparison .....	66
Figure 3.9 GO enrichment of genes up-regulated in day-6 compared to day-13 schistosomules .....	68
Figure 3.10 MEGs differentially expressed between day-6 and day-13 schistosomules .....	69
Figure 3.11 Predicted structure of Smp_182770 aligned with structure of human CFH .....	73
Figure 3.12 Homologous relationship of Smp_182770 and alignment of RNA-seq reads to genomic locations.....	75
Figure 3.13 Clusters of genes with high expression in liver stages and GO term enrichment	78
Figure 3.14 GO enrichment of genes up-regulated in day-13 compared to day-6 schistosomules.....	81
Figure 3.15 GO enrichment of genes up-regulated in day-35 compared to day-28 worms.....	86
Figure 3.16 GO enrichment of genes up-regulated in day-28 compared to day-35 worms.....	87
Figure 3.17 Expression profile of GPCRs up-regulated in day-28 compared to day-35 worms .....	88
Figure 3.18 Log <sub>2</sub> FC of genes in cluster 40 in response to pairing.....	93
Figure 4.1 Experiment set up.....	107
Figure 4.2 Gel image of PCR product from <i>Mycoplasma</i> tests .....	114
Figure 4.3 Example images from each time point of each co-culture condition .....	115
Figure 4.4 PCA of transcriptomes from <i>in vitro</i> and <i>in vivo</i> datasets.....	117
Figure 4.5 PCA of transcriptomes from all <i>in vitro</i> parasites and <i>in vivo</i> lung schistosomules .....	118

Figure 4.6 PCA of transcriptomes from all <i>in vitro</i> schistosomules.....	119
Figure 4.7 MA plots of pairwise comparisons between co-cultured and control schistosomules .....	120
Figure 4.8 Differentially expressed genes unique and common to each pairwise comparison .....	121
Figure 4.9 Differentially expressed genes in day-17 schistosomules comparing unique and common changes between cell types .....	122
Figure 4.10 Enriched GO terms in genes up-regulated in day-17 co-cultured schistosomules .....	123
Figure 4.11 Enriched GO terms in genes down-regulated in day-17 co-cultured schistosomules.....	124
Figure 4.12 Generic responses in schistosomules.....	126
Figure 4.13 Homologous relationship of Smp_067800 .....	127
Figure 4.14 Genomic region of Smp_067800 and alignment of RNA-seq reads.....	129
Figure 4.15 Homologous relationship of Smp_052880.....	131
Figure 4.16 Smp_052880 genomic region and amino acid sequence alignment with other BLASTP hits .....	132
Figure 4.17 Volcano plot of pairwise comparison between HEPG2 <i>vs.</i> non-HEPG2 schistosomules.....	134
Figure 4.18 Enriched GO terms in genes up-regulated in HEPG2 compared to non-HEPG2 schistosomules.....	135
Figure 4.19 Enriched GO terms in genes down-regulated in HEPG2 compared to non-HEPG2 schistosomules.....	135
Figure 4.20 Volcano plot of pairwise comparison between HUVEC <i>vs.</i> non-HUVEC schistosomules.....	141
Figure 4.21 Smp_123920 homologous relationship and structural prediction .....	143
Figure 4.22 Expression profiles of Smp_123920 <i>in vivo</i> and <i>in vitro</i> .....	144
Figure 4.23 Smp_016490 domains and expression profiles .....	145
Figure 4.24 Smp_052880 <i>in vivo</i> expression profile .....	150
Figure 5.1 260/230 ratio of RNA extracted from human cells and its relationship with the RNA concentration.....	159
Figure 5.2 Expression levels of endothelial cell surface markers in worm-free HUVEC.....	160
Figure 5.3 PCA from transcriptomes of all cell types, and separate for each cell type.....	162
Figure 5.4 Volcano plots for pairwise comparison in each cell type between co-cultured <i>vs.</i> worm-free conditions .....	164
Figure 5.5 Expression profiles of SELE, VCAM1, and ICAM1 .....	165

Figure 5.6 Expression of endothelial adhesion molecules between co-cultured <i>vs.</i> worm-free HUVEC .....	166
Figure 5.7 GO enrichment of genes differentially expressed in co-cultured HUVEC compared to worm-free HUVEC .....	168
Figure 5.8 Pathway enrichment of genes differentially expressed in co-cultured HUVEC compared to worm-free HUVEC .....	169
Figure 5.9 Coagulation-related genes up-regulated in co-cultured HUVEC .....	170
Figure 5.10 GO enrichment of genes differentially expressed in co-cultured HEPG2 compared to worm-free HEPG2 .....	174
Figure 5.11 Pathway enrichment of genes differentially expressed in co-cultured HEPG2 compared to worm-free HEPG2.....	175
Figure 5.12 GO enrichment of genes differentially expressed in co-cultured GripTite compared to worm-free GripTite .....	179
Figure 5.13 Pathway enrichment of genes differentially expressed in co-cultured GripTite compared to worm-free GripTite .....	180
Figure 5.14 Immune responses-related genes up-regulated in co-cultured GripTite.....	182
Figure 5.15 Log <sub>2</sub> FC of genes in extracellular matrix organisation pathways.....	184
Figure 5.16 Log <sub>2</sub> FC of genes in coagulation and complement cascade .....	186

# List of Tables

Table 3.1 Average of total read counts .....	59
Table 3.2 Genes in cluster 8 (highly expressed in lung stage).....	64
Table 3.3 Enriched GO terms of genes down-regulated after lung stage .....	65
Table 3.4 Top 20 genes up-regulated in day-6 compared to day-13 schistosomules .....	67
Table 3.5 Genes in cluster 72.....	72
Table 3.6 Top 20 genes up-regulated in day-13 compared to day-6 schistosomules .....	80
Table 3.7 Enriched GO terms (biological process) of genes up-regulated in day-21 compared to day-28 worms.....	83
Table 3.8 Top 20 genes up-regulated in day-28 compared to day-21 worms.....	84
Table 3.9 Enriched GO terms (biological process) of genes up-regulated in day-28 compared to day-21 worms.....	85
Table 3.10 Genes in cluster 40.....	90
Table 4.1 BLASTP output with Smp_067800 amino acid sequence as a query .....	127
Table 4.2 BLASTP output with Smp_052880 amino acid sequence as query .....	130
Table 4.3 Top 20 genes up-regulated in HEPG-2 compared to non-HEPG2 schistosomules	137
Table 4.4 Top 20 genes down-regulated in HEPG-2 compared to non-HEPG2 schistosomules .....	138
Table 4.5 Genes differentially expressed in HUVEC compared to non-HUVEC schistosomules.....	141
Table 5.1 Top differentially expressed genes in HEPG2.....	173



# List of Appendices

Appendix A: Aquarium water, 10X.....	235
Appendix B: Reference genomes used by NPG QC.....	236
Appendix C: Versions of R packages used in data analysis .....	238
Appendix D: Basch media components .....	239
Appendix E: <i>S. mansoni</i> genetrees downloaded from WormBase ParaSite release 9.....	240
Appendix F: Orthologues of Smp_041700 downloaded from WormBase ParaSite release 8 .....	251
Appendix G: List of endothelial cell surface marker from Durr <i>et al.</i> , 2004 .....	252
Appendix H: List of supplementary tables .....	255