

Human antibody-mediated immunity against infection
and parasite fecundity in *Schistosoma mansoni* and
Schistosoma haematobium infection

Rebecca Claire Oettle

Darwin College
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

July 2020

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

It does not exceed the prescribed word limit for the Biology Degree Committee.

Human antibody-mediated immunity against infection and parasite fecundity in *Schistosoma mansoni* and *Schistosoma haematobium* infection

Through a combination of epidemiological and proteomic methods, the aim of this thesis was to advance understanding of antibody-driven immune responses that partially protect against infection and reduce fecundity in human schistosome infections.

The relative contribution of exposure to infection and IgE antibody responses against schistosomes were explored using a mathematical modelling approach. Age intensity and serology profiles of sub-cohorts were reproduced to good approximation by incorporating explicitly defined exposure and immunity functions, fitted to field data, into an age- and sex-structured *S. mansoni* transmission model. However, the inferior model fit for males from one tribal group highlights the need to capture heterogeneity in antibody responses that occur with varying cumulative exposure of different demographic groups.

The progressive development of protective immunity against *S. mansoni* infection is associated with cross-reactivity between members of the Tegument Allergen-Like (TAL) protein family. This relationship was explored further and evidence for involvement of an additional TAL family member in the development of IgE-mediated anti-infection immunity found. Evidence supporting similar cross-reactive relationships between *S. haematobium* TAL proteins is provided.

In addition to anti-infection immunity, *S. haematobium* is thought subject to an anti-fecundity immune response. It is proposed that IgG₁ antibodies raised to proteins within whole worm extract are responsible. A proteomic approach used to identify targets of the IgG₁ returned 191 proteins, several of which are closely associated with the parasite-host interface. Three were successfully expressed and analysis identified possible associations between target specific IgG₁ and reduced worm fecundity, although this was strongly confounded by host age and village of residence, a proxy for the force of transmission.

These findings all demonstrate associations between the development of protective antibody responses and variations in exposure behaviour, host age or the force of transmission; the combined contribution of which will determine the level of exposure to immunostimulatory antigens and ultimately development of immunity.

Dedication

The work presented in this thesis would not have been possible without the unending support of my wonderful husband Norman, who has provided the carrot and stick in equal measures.

And, of course, to our amazing son, Joseph, who has provided a buzz of energy and put a smile on my face when I needed it the most.

Acknowledgements

Thank you to my supervisor, Dr Shona Wilson, for providing guidance and feedback throughout this project. I would also like to thank the following people for helping with this research project:

Jakub Wawrzyniak for providing an ear and advice whilst troubleshooting my experiments over coffee.

Dr Anna Protasio at the University of Cambridge for her feedback, questions and advice in our combined helminth lab meetings.

Dr Martin Walker at Royal Veterinary College and Prof. Maria-Gloria Basañez at Imperial College London (ICL) for their time, expertise and long meetings discussing the intricacies of mathematical transmission models.

Charlie Whitaker (ICL) for his enthusiasm for MH-MCMC model fitting.

Dr Colin Fitzsimmons for teaching me vital skills in protein production.

Renata Feret at the Department of Biochemistry, University of Cambridge for her help and guidance in running countless 2D electrophoresis gels and analysing the resulting images. Julie Howard-Murkin and Mike Deery at the Department of Biochemistry, University of Cambridge for giving me an overview of mass spectrometry and running my samples.

Dr Russell Morphew, Dr Iain Chalmers, Prof. Karl Hoffman and the post-graduate students for hosting me at the Barrett Centre for Helminth Research, University of Aberystwyth. Specifically, to Dr Iain Chalmers for a crash course in bioinformatic analysis of proteins and Dr Russell Morphew for time in his lab learning practical skills in protein fractionation.

Ray Owens for hosting me at Oxford Protein Production Facility, UK at Harwell and Heather Rada who provided her time and guidance when performing high throughput protein production at OPPF-UK.

Contents

Declaration.....	i
Abstract.....	iii
Dedication.....	v
Acknowledgements.....	vii
Contents.....	ix
List of figures.....	xv
List of tables.....	xvii
List of appendices.....	xix
List of abbreviations.....	xx

1. Introduction

1.1. <i>Schistosoma</i> life cycle.....	1
1.2. Epidemiology.....	3
1.2.1. Measurement of infection intensity.....	3
1.2.2. Age, exposure and experience of infection.....	5
1.2.3. Control of schistosomiasis.....	8
1.3. Mathematical models of schistosome transmission.....	9
1.4. Adaptive immunity in human schistosomiasis.....	10
1.4.1. Humoral immunity.....	11
1.4.2. Antibody diversification.....	11
1.4.3. Antibody responses associated with protection against <i>Schistosoma</i> infection.....	12
1.4.4. Concomitant immunity.....	14
1.4.5. Tegument allergen-like protein family and anti-infection immunity.....	15
1.4.6. Immune responses associated with anti-fecundity immunity.....	16
1.5. Thesis outline and aims.....	17

2. Materials & Methods

2.1. Description of study sites and cohorts.....	20
2.1.1. Booma, Uganda.....	20
2.1.2. Musoli, Uganda.....	21
2.1.3. Kaladangan, Guenidaga and Kalabougou, Mali.....	21
2.2. Plasma sample processing.....	22
2.3. Preparation of parasite material.....	22
2.3.1. <i>Schistosoma mansoni</i> mRNA extraction.....	22
2.3.2. <i>Schistosoma mansoni</i> adult worm cDNA.....	23
2.3.3. <i>Schistosoma haematobium</i> adult worm cDNA.....	23
2.3.4. <i>Schistosoma haematobium</i> soluble worm antigen (SWA) preparation.....	24
2.4. Two-dimensional gel electrophoresis.....	24
2.4.1. First dimension – isoelectric focusing (IEF).....	25
2.4.2. SDS PAGE gel casting.....	25
2.4.3. Second dimension – SDS PAGE.....	26
2.4.4. Silver staining.....	26
2.5. Two-dimensional Western Blot.....	27

2.6. Identification of spots.....	27
2.7. Mass spectrometry.....	28
2.7.1. Spot cutting.....	28
2.7.2. Trypsin digest.....	28
2.7.3. Mass spectrometry.....	29
2.7.4. Data processing and protein identification.....	29
2.8. Anti-fecundity target bioinformatic analysis.....	30
2.8.1. Spot intensity.....	31
2.8.2. Peptide sequence coverage.....	31
2.8.3. Identification of signal peptides and excreted/secreted proteins.....	31
2.8.4. Gene Ontology (GO) term enrichment.....	32
2.8.5. Differential gene expression analysis.....	32
2.9. Bioinformatic analysis of TAL proteins.....	33
2.9.1. Identification of <i>S. haematobium</i> TAL proteins.....	33
2.9.2. Phylogenetic analysis of <i>S. haematobium</i> TAL proteins.....	33
2.10. Cloning and expression of recombinant TAL proteins.....	34
2.10.1. Recombinant SmTALII expression construct.....	34
2.10.2. Calcium chloride transformation of TG2 <i>E. coli</i> with SmTALII-pGEX-KG.....	38
2.10.3. Expression of TAL proteins in <i>E. coli</i> bacterial expression system.....	39
2.10.4. Purification of TAL proteins from <i>E. coli</i>	40
2.11. One-dimensional gel electrophoresis.....	41
2.11.1. Gel electrophoresis.....	41
2.11.2. Coomassie brilliant blue staining.....	42
2.12. One-dimensional Western blot.....	42
2.13. High throughput (HTP) cloning and expression screening of anti-fecundity targets.....	42
2.13.1. PCR primer design.....	43
2.13.2. HTP PCR reactions.....	43
2.13.3. Analysis of PCR products.....	44
2.13.4. AMPure XP magnetic bead purification.....	44
2.13.5. In-Fusion reactions and HTP transformation into cloning grade <i>E. coli</i>	45
2.13.6. Colony picking and HTP culture.....	45
2.13.7. Construct verification.....	46
2.13.8. Small-scale expression screening in <i>E. coli</i>	47
2.13.9. SDS-PAGE analysis of Ni ²⁺ -NTA purified proteins.....	48
2.14. Scaled-up expression of recombinant anti-fecundity target proteins.....	49
2.14.1. Transformation of NiCo21 <i>E. coli</i> with anti-fecundity target protein OPINF plasmid.....	49
2.14.2. Protein expression and purification.....	49
2.15. Enzyme-linked immunosorbent assay (ELISA).....	50
2.15.1. Coating and contamination assay.....	50
2.15.2. Measurement of antigen-specific antibody levels.....	51
2.15.3. Other antigen-specific antibody levels.....	53
2.15.4. Reciprocal inhibition ELISA.....	53
2.16. Statistical analysis.....	53
2.16.1. Transformation of data.....	53
2.16.2. Definition of the geometric mean.....	54

2.16.3. Analysis of reciprocal inhibition ELISA results for TAL protein cross-reactivity	54
2.16.4. Calculation of post-treatment boosts in antibody responses	54
2.16.5. Definition of positive antibody responders	54
2.16.6. Analysis of <i>S. mansoni</i> TALII specific IgE responses	54
2.16.7. Association between <i>S. haematobium</i> TAL specific IgE responses and baseline CAA	55
2.16.8. Piecewise regression analysis to determine ‘children’ and ‘adults’ age groups	56
2.16.9. Association between <i>S. haematobium</i> TAL specific IgE responses and two-year reinfection status	57
2.16.10. Identification of the anti-fecundity case cohort and matched controls	57
2.16.11. Association between IgG _I responses to <i>S. haematobium</i> anti-fecundity target proteins and worm fecundity	58
2.17. Mathematical modelling	59
3. Introducing anti-infection immunity into a <i>Schistosoma mansoni</i> transmission model	
3.1. Introduction	60
3.1.1. General introduction	60
3.2. Overview of the principal schistosome transmission models	61
3.2.1. Basic model structures	61
3.2.2. Modelling exposure and immunity	63
3.2.3. Mathematical models and control	66
3.3. Chapter aims	67
3.4. Model derivation	68
3.5. Model equations	72
3.6. Data	72
3.7. Modelling cercarial exposure	73
3.7.1. Risk of exposure	73
3.7.2. Age-dependent cercarial exposure function ($\rho(a)$)	74
3.7.3. Relative exposure by sex (ρ_M/ρ_F)	76
3.7.4. Contamination rate ($\kappa_{(a)}$)	77
3.8. Modelling the protective effect of TAL-IgE	77
3.8.1. Strength of protective response	78
3.8.2. Cox Proportional-Hazards model	78
3.8.3. Age-dependent immunity function ($f(I)$)	78
3.8.4. Immunity ‘gain’ parameter (γ_{gain})	81
3.8.5. Longevity of the IgE antibody response	81
3.8.6. Immunity ‘loss’ parameter (γ_{loss})	82
3.9. Model parameters	82
3.9.1. Force of infection (Λ)	82
3.9.2. Basic reproductive number (R_0)	83
3.9.3. Population structure ($\pi(a)$)	83
3.9.4. Human life expectancy ($1/\mu_H$)	84
3.9.5. Adult schistosome lifespan ($1/\mu_W$)	84
3.9.6. Egg excretion rate (ϵ)	85
3.9.7. Density dependent fecundity	85
3.10. Model calibration	87
3.10.1. Distribution of faecal egg count	88
3.10.2. Distribution of IgE optical density	89
3.10.3. Bayesian inference	89

3.10.4. Likelihood specification.....	90
3.10.5. Prior distributions.....	91
3.10.6. Initial conditions.....	92
3.10.7. Metropolis-Hastings Monte Carlo Markov Chain (MCMC).....	92
3.10.8. Adaptive MCMC model.....	93
3.10.9. Model convergence.....	93
3.11. Results.....	94
3.11.1. Model fit.....	94
3.11.1.1. Alur.....	94
3.11.1.2. Bagungu.....	95
3.11.2. Estimated parameter values.....	99
3.11.3. Diagnostics of model fit.....	100
3.12. Discussion.....	105
3.12.1. The exposure function fits the observed data to good approximation.....	105
3.12.2. The immunity function closely describes the protective effect of SmTAL1-specific IgE.....	106
3.12.3. Model convergence and sensitivity analysis.....	108
3.12.4. The Bagungu model demonstrated the need for the inclusion of heterogeneity and immunity in the model.....	109
3.12.5. Future adaptations of the model.....	110
3.12.6. Concluding remarks.....	111

4. Anti-infection immunity and the tegument allergen-like proteins in *S. mansoni* and *S. haematobium*

4.1. Introduction.....	113
4.1.1. Introduction to the TAL protein family.....	113
4.1.2. TAL proteins and reinfection.....	114
4.1.3. Chapter aims.....	116
4.2. Results.....	117
4.2.1. Identification of <i>S. haematobium</i> TAL proteins.....	117
4.2.2. Multiple sequence alignment of <i>S. haematobium</i> TAL proteins.....	118
4.2.3. Phylogenetic analysis of ShTAL proteins.....	122
4.2.4. Cloning, expression and purification of TAL proteins.....	123
4.2.5. Coating and contamination assays.....	124
4.2.6. IgE binding to TAL5 is cross-reactive with TALII in <i>S. mansoni</i> and <i>S. haematobium</i>	126
4.2.7. Individuals seropositive for SmTALII-IgE are a subset of those with a detectable SmTAL5, SmTAL3 and SmTAL1-specific IgE response.....	129
4.2.8. Characterisation of the TALII-specific IgE response, a potential new candidate in the TAL protein reinfection story.....	129
4.2.9. ShTALII responders are a partial subset of ShTAL1, 3 and 5 responders.....	132
4.2.10. The seroprevalence profiles of <i>S. haematobium</i> TAL1, 3, 5 and 11 and their association with age and force of transmission.....	133
4.2.11. Are antigen-specific IgE responses to ShTAL1, ShTAL3, ShTAL5 and ShTALII associated with a lower worm burden?.....	134
4.2.12. Are antigen-specific IgE responses to ShTAL1, ShTAL3, ShTAL5 and ShTALII associated with reinfection status at 2 years?.....	136
4.3. Discussion.....	140
4.3.1. ShTAL protein identification and phylogenetic analysis.....	140

4.3.2.	Analysis of cross-reactive relationships.....	141
4.3.3.	Characterisation of <i>S. mansoni</i> and <i>S. haematobium</i> TALII epidemiology.....	142
4.3.4.	Associations with anti-infection immunity.....	143
4.3.5.	Future study requirements.....	144
4.3.6.	Conclusions.....	145
5.	Anti-fecundity immunity in <i>S. haematobium</i>: identification of antigenic targets and serological analysis	
5.1.	Introduction.....	146
5.1.1.	General introduction.....	146
5.1.2.	Reproductive pairing is required for female sexual maturation.....	148
5.1.3.	Reproductive pairing enhances female nutritional intake.....	148
5.1.4.	Excreted and secreted proteins are involved in interactions between the parasite and host.....	149
5.1.5.	<i>Schistosoma</i> gut proteins are expelled during feeding.....	150
5.1.6.	Previous vaccine candidates and worm fecundity.....	150
5.1.7.	Identifying anti-fecundity targets.....	151
5.1.8.	Chapter aims and objectives.....	151
5.2.	Results	
5.2.1.	Separation of <i>S. haematobium</i> soluble worm antigen by two-dimensional gel electrophoresis.....	152
5.2.2.	Identification of immunoreactive proteins by Western blot.....	152
5.2.3.	Protein identification.....	152
5.2.4.	Bioinformatic analysis of mass spectrometry results.....	155
5.2.4.1.	Excreted/secreted proteins and extracellular vesicles.....	155
5.2.4.2.	Gene Ontology (GO) enrichment analysis.....	156
5.2.4.3.	Proteins identified within the gastrodermis and gastric tract.....	160
5.2.4.4.	Gender-associated gene expression.....	160
5.2.5.	Selection of proteins for downstream recombinant expression.....	162
5.2.6.	High throughput cloning and expression screening.....	167
5.2.7.	Coating and contamination assays.....	170
5.2.8.	Serological profile of anti-fecundity target specific IgG ₁ responses.....	171
5.2.9.	Associations between anti-fecundity target IgG ₁ seropositivity and reduced worm fecundity.....	174
5.3.	Discussion.....	175
5.3.1.	Excreted/secreted and surface expressed proteins.....	175
5.3.2.	Identification of tegumental and cytoskeletal proteins corresponds with the delayed development of anti-fecundity immunity.....	176
5.3.3.	GO terms associated with reproduction were significantly enriched within the candidate anti-fecundity target proteins.....	176
5.3.4.	Enzymes involved in energy production may be excreted during feeding or secreted in EVs.....	177
5.3.5.	What is the significance of the three successfully expressed anti-fecundity targets?.....	178
5.3.5.1.	Putative aminopeptidase W07G4.4 (ShLAP).....	178
5.3.5.2.	Taurocyamine kinase (ShTK).....	179

5.3.5.3. Putative UDP-glucose-4-epimerase (ShGALE).....	179
5.3.6. Regression models do not exclude ShLAP, ShTK and ShGALE as potential anti-fecundity target proteins.....	180
5.3.7. A need for improved annotation of the <i>S. haematobium</i> genome.....	181
5.3.8. Potential future experiments.....	182
5.3.9. Regarding mathematical transmission models.....	183
5.3.10. Concluding remarks.....	183
6. Discussion.....	184
7. References.....	189

List of figures

Chapter 1

1.1.	Illustration of the peak shift	6
------	--------------------------------------	---

Chapter 2

2.1.	pJET1.2/blunt plasmid map.....	36
2.2.	pGEX-KG plasmid map	38
2.3.	Map of pOPINF plasmid.....	46
2.4.	Diagram of microtitre plate layout for coating and contamination ELISA.....	51
2.5.	A) The standard error of residuals from piecewise regression models of CAA and B) CAA by host age (Malian cohort).....	56
2.6.	Worm fecundity by host age (Malian cohort).....	58

Chapter 3

3.1.	Flow diagram of the <i>S. mansoni</i> transmission model.....	70
3.2.	Cercarial exposure function model fit for A) Alur and B) Bagungu females and males.....	76
3.3.	KM curve with drc function, fitted to individual-level serology data for A) total cohort and B) Alur and Bagungu tribes.....	80
3.4.	Underlying demography of study population (Booma, Uganda) in 1998. A) Age distribution for population of Booma, 1998; Sex distribution of females and males from the B) Alur and C) Bagungu tribes.....	84
3.5.	Linear regression between eggs per gram of faeces (epg) and circulating anodic antigen (CAA).....	87
3.6.	Interrater sampling variability across repeated stool samples, analysed by different parasitology technicians.....	88
3.7.	Observed and model predicted age profiles of A) cercarial exposure scores; B) infection intensity (eggs/g faeces), and C) SmTALI-IgE optical density (OD ₆₀₀) by sex for the Alur tribe.....	97
3.8.	Observed and predicted age profiles of A) cercarial exposure scores; B) infection intensity (eggs/g faeces), and C) SmTALI-IgE optical density (OD ₆₀₀) by sex for the Bagungu tribe.....	98
3.9.	Posterior distributions for the six estimated model parameters for the Alur tribe (steps 2500–10000). Distributions are presented as A) histograms and B) smoothed density plots.....	101
3.10.	Posterior distributions for the six estimated model parameters for the Bagungu tribe (steps 2500–10000). Distributions are presented as A) histograms and B) smoothed density plots.....	102
3.11.	Trace output for each parameter estimate from the MH-MCMC chain, run over 10,000 iterations for the Alur tribe A) steps 0 to 10000 and B) steps 2500 – 10000 (burn-in removed).....	103

3.12.	Trace output for each parameter estimate from the MH-MCMC chain, run over 10,000 iterations for the Bagungu tribe A) steps 0 to 10000 and B) steps 2500 – 10000 (burn-in removed).....	104
-------	--	-----

Chapter 4

4.1.	Relative lifecycle stage expression of the <i>S. haematobium</i> TAL1, TAL3, TAL5 and TAL11 proteins.....	116
4.2.	Alignment of amino acid sequences of the <i>S. mansoni</i> and likely <i>S. haematobium</i> TAL protein family.....	120
4.3.	Protein sequence analysis of the <i>S. mansoni</i> and <i>S. haematobium</i> TAL family.....	122
4.4.	A) Agarose gel of cloned SmTAL11; B) Agarose gel of analysis of recombinant <i>E. coli</i> clones	123
4.5.	SDS PAGE gel electrophoresis of purified SmTAL and ShTAL proteins.....	124
4.6.	Coating and contamination assays for <i>S. mansoni</i> and <i>S. haematobium</i> TAL1, TAL3, TAL5 and TAL11.....	125
4.7.	Cross-reactivity of TAL-IgE responses in <i>S. mansoni</i> and <i>S. haematobium</i>	128
4.8.	Venn diagram demonstrating the overlapping seroprevalence of <i>S. mansoni</i> TAL proteins	129
4.9.	Seroprevalence for <i>S. mansoni</i> TAL-specific IgE responses.....	130
4.10.	Venn diagrams demonstrating the overlapping seroprevalence of <i>S. haematobium</i> TAL proteins.....	131
4.11.	Pre- and post-treatment seroprevalence of IgE responses to <i>S. haematobium</i> TAL proteins by age and village, as a proxy for force of transmission.....	132

Chapter 5

5.1.	Two-dimensional gel analysis of <i>S. haematobium</i> soluble whole worm antigen preparation.....	153
5.2.	Two-dimensional Western blots identifying IgG ₁ immunogenic protein spots.....	154
5.3.	Number of genes associated with molecular function (green) and biological process (blue) GO term that are over-represented in the 203 candidate anti-fecundity targets compared to the GO terms annotated to whole <i>S. haematobium</i> genome.....	159
5.4.	Volcano plots showing the relative expression of orthologous <i>S. haematobium</i> genes corresponding to anti-fecundity target proteins in <i>Schistosoma</i> males, females and their respective gonads.....	162
5.5.	Agarose gel showing bands for quality control PCR of <i>S. haematobium</i> cDNA.....	167
5.6.	Agarose gel showing PCR product bands for successfully cloned anti-fecundity target proteins.....	167
5.7.	SDS PAGE gel of small-scale protein expression screening of anti-fecundity targets.....	168
5.8.	Protein fractions eluted from Ni-NTA column A.) ShTK B) ShLAP C) ShGALE.....	168
5.9.	Coating and contamination assays for recombinant ShLAP, ShTK and ShGALE.....	171
5.10.	Seroprevalence of antigen specific IgG ₁ response to ShLAP, ShTK and ShGALE with age, pre- and post-treatment.....	172

List of tables

Chapter 2

2.1.	Selection criteria for downstream expression of anti-fecundity targets.....	30
2.2.	Recombinant TAL protein expression, induction conditions.....	40
2.3.	Recombinant antigen coating concentration and condition.....	52

Chapter 3

3.1.	Parameter definitions and values.....	71
3.2.	Maximum likelihood estimation of age-dependent exposure function parameters for Alur and Bagungu ethnic groups.....	75
3.3.	Maximum likelihood estimated parameter values for SmTAL1-IgE immunity function.....	80
3.4.	Linear regression model parameters eggs (log(epg+1))per worm (log(CAA+1)).....	87
3.5.	Summary of initial model parameter values.....	92
3.6.	Estimated parameter values resulting from 10,000 MH-MCMC model iterations.....	100

Chapter 4

4.1.	<i>S. haematobium</i> candidate TAL proteins.....	118
4.2.	Geometric mean odds ratios describing the association between post-treatment IgE responses to SmTAL antigens and intensity of reinfection at 2 years post-treatment.....	131
4.3.	Association between pre-treatment ShTAL-specific IgE responses and baseline CAA, as a measure of worm burden. Results for the reduced linear regression model, adjusted for age, sex, village and ShTAL-specific IgG ₄ are displayed, for total cohort.....	137
4.4.	Associations between univariant post-treatment ShTAL-IgE responses and measurable urinary egg excretion at two-year reinfection.....	139
4.5.	Association between 9-weeks post-treatment ShTAL-specific IgE responses and reinfection status at two years, adjusted for age, sex, village and ShTAL-specific IgG ₄ , non-significant terms removed. Results are displayed for total cohort.....	139

Chapter 5

5.1.	Number of spots identified in case and control blots for each technical replicate.....	153
5.2.	GO terms relating to molecular function that are significantly over-represented within the 191 proteins identified as immunostimulatory to anti-fecundity case serum.....	158
5.3.	Summary of anti-fecundity target and whole genome genes significantly developmentally regulated in <i>S. haematobium</i> female worm, male worm and egg.....	161
5.4.	List of 48 proteins selected for high-throughput expression screening.....	164
5.5.	List of 48 selected proteins with SignalP and SecretomeP annotation.....	166
5.6.	Successfully cloned and expressed anti-fecundity targets.....	169

5.7.	Association between age group, village and pre-treatment <i>S. haematobium</i> anti-fecundity target specific IgG ₁ seropositivity.....	173
5.8.	Association between seropositivity for AF target specific IgG ₁ and fecundity in individuals aged under 11 years.....	174
5.9.	Association between pre-treatment <i>S. haematobium</i> anti-fecundity target specific IgG ₁ responses and fecundity.....	175

List of appendices

- Appendix 1: Buffer compositions
- Appendix 2: Anti-fecundity target cloning primer construct sequences
- Appendix 3: *S. mansoni* age- and sex-structured transmission model code
- Appendix 4: Supplementary information for Chapter 4
- Appendix 5: Supplementary information for Chapter 5
- Appendix 6: Publications

List of abbreviations

AMP	ampicillin
BP	biological process
CAA	circulating anodic antigen
CaM	calmodulin
Cb	carbenicillin selection
CC	cellular component
CCA	circulating cathodic antigen
cDNA	complementary DNA
CHL	chloramphenicol
CSR	class switch recombination
CSRG	Cambridge Schistosomiasis Research Group
CWRU	Case Western Reserve University
DLC	dynein light chain
DPPII	dipeptidyl peptidase II
ELISA	enzyme-linked immunosorbent assay
EST	expressed sequence tag
ES	excretory/secretory
EV	extracellular vesicle
FC	fold change
FCS	fetal calf serum
FDCs	follicular dendritic cells
FDR	false discovery rate
FOI	force of infection
GALE	galactose/glucose-4-epimerase
GAPD	glyceraldehyde-3-phosphate dehydrogenase
GC	germinal centre
GM	geometric mean
GO	gene ontology
GST	glutathione-S-transferase
HIV	human immunodeficiency virus
HPC	high performance computing
HRP	horseradish peroxidase
HTP	high throughput

ICL	Imperial College London
IEF	isoelectric focussing
Ig	immunoglobulin
IL	interleukin
IPG	immobilized pH gradient
IPTG	isopropyl β -D-1-thiogalactopyranoside
LAP	leucine aminopeptidase
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LLPCs	long-lived plasma cells
MCMC	Markov chain Monte Carlo
MDA	mass drug administration
ME	beta-mercaptoethanol
MF	molecular function
MH	Metropolis Hastings
MHC	major histocompatibility complex
MICE	multiple imputation by chained equations
MS	mass spectrometry
MTP	microtitre plate
MWB	mean worm burden
NES	non-endemic sera
NL	non-linear
NTD	neglected tropical disease
OD	optical density
OLR	ordinary linear regression
OPPF-UK	Oxford Protein Production Facility, UK
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pI	isoelectric point
POC	point-of-care
POP	prolyl oligopeptidase
PSM	peptide-spectrum match
PZQ	praziquantel
SWA	soluble worm antigen
SWAP	soluble worm antigen preparation
SWB	stratified worm burden
SDS -PAGE	polyacrylamide gel electrophoresis
SEA	soluble egg antigen
SHM	somatic hypermutation

SLPCs	short-lived plasma cells
SOC	super optimal broth with catabolite repression
TAL	Tegument Allergen-Like
TBE	tris/borate/EDTA
TK	taurocyamine kinase
Th	T helper
UCP-LF	up-converting phosphor-lateral flow circulating anodic antigen
UDP	uridine diphosphate
WHO	World Health Organization

Introduction

Human schistosomiasis is a parasitic disease of considerable public health importance, caused by trematode worms of the genus *Schistosoma*. Schistosomiasis affects at least 290 million people worldwide (World Health Organization, 2020b), causing significant chronic morbidity. There are three species of primary importance to human health: *S. haematobium*, *S. mansoni* and *S. japonicum*, and two further species, *S. intercalatum* and *S. mekongi*, that are responsible for a much smaller proportion of the human disease burden. These blood-dwelling parasites are transmitted by a cycle of open defaecation or urination, and subsequent contact with contaminated water. Schistosomiasis is therefore a disease of poverty and inequality with 85% of cases found in sub-Saharan Africa, where a lack clean water and sanitation systems perpetuate transmission.

As with many helminth infections, schistosomes are highly aggregated within the host (Guyatt *et al.*, 1994). In endemic populations, a small number of individuals harbour many parasites, whilst the majority of the population carry a much lower parasitic burden, although the relative proportions of these groups is area-specific (Guyatt *et al.*, 1994). It is known that transmission of schistosomes is dependent on a close interplay between varying exposure and susceptibility; understanding this relationship is fundamental to predicting infection dynamics. *S. haematobium* and *S. mansoni* are the most prevalent species found in sub-Saharan Africa and will be the focus of the analysis presented in this thesis.

1.1 *Schistosoma* life cycle

Maintenance of the *Schistosoma* life cycle is dependent on two hosts, the definitive human host, within whom sexual reproduction occurs, and the intermediate freshwater snail host. Infected humans excrete eggs into the environment, via faeces or urine, that contaminate freshwater sources due to a lack of adequate sanitation systems. Viable eggs hatch on contact with water, each releasing a free-living miracidia. Miracidia have limited glycogen stores and are unable to feed, so have between five to six hours to locate and penetrate a suitable snail host of the genus *Biomphalaria* (*S. mansoni*) or *Bulinus* (*S. haematobium*) (Maldonado and Acosta-Matienzo, 1948). The parasite is then amplified through several rounds of asexual reproduction in the snail.

The infected snail subsequently sheds infectious larval stages, known as cercariae. Like the miracidia, cercariae are also unable to feed and are dependent upon a limited glycogen reserve for energy, they therefore have approximately 12 hours within which to find a human host (Lawson and Wilson, 1980; Whitfield *et al.*, 2003). When a human comes into contact with the free-living cercariae, through contact with water, the cercariae penetrates the skin by burrowing headfirst through the epidermis. This is facilitated by the secretion of proteases (Curwen *et al.*, 2006). During skin penetration the cercariae loses its tail and sheds the outer glycocalyx membrane that served to provide environmental protection during the free-living stage (Samuelson and Caulfield, 1982), developing into a schistosomulum. The parasite also develops a seven-layered tegumental membrane outer coat (Hockley and McLaren, 1973). Schistosomes can live for several years, undetected, within the bloodstream of their human hosts and employ several mechanisms by which to avoid recognition by the host immune system (McSorley, Hewitson and Maizels, 2013), shedding of the glycocalyx and the adaptive tegumental membrane are just two such mechanisms (Abath and Werkhauser, 1996; Han *et al.*, 2009; Da'dara and Krautz-Peterson, 2014).

The schistosomulum migrates through the host circulatory system, via the lungs, ultimately developing into an adult worm in the portal vein of the liver, over the course of four to six weeks. It is here that adult worms form copulatory pairs and sexually mature. Schistosomes are unique amongst the platyhelminths in that they are dioecious and dimorphic. When adult worms pair, the female lies within the gynaecophoric canal of the larger, dorsoventrally flattened male worm. Pairing initiates gene expression that induces the female to sexually mature and start producing eggs (Popiel, 1986; LoVerde *et al.*, 2004); daily egg production rates vary by species with each *S. mansoni* female producing approximately 100–300 eggs per day (Cheever *et al.*, 1994) and each *S. haematobium* female approximately 20–200 eggs per day (Cheesbrough, 2005).

It is thought that males have a primary role in pairing due to their greater musculature and mobility, although females may influence pairing through other mechanisms such as chemical signalling (Steinauer, 2009; Lu *et al.*, 2019). The exact nature of the pairing relationship is not well understood, but *in vitro* studies show that constant pairing is fundamental to the preservation of sexual maturity of female worms, since females worms that become separated from their male partner are seen to sexually regress (Popiel and Basch, 1984; Popiel, 1986). Evidence from *in vitro* worm pairing studies also indicates that male worms may have an indirect role in the nutrition of females through facilitating the absorption of nutrients via the tegument (Popiel and Basch, 1984; Cornford, 1986; Gupta and Basch, 1987), supplementing the ingestion of erythrocytes.

Once paired, adult worms migrate against the blood flow to the mesenteric veins surrounding the small intestine (*S. mansoni*) or the vesicular and pelvic venous plexuses surrounding the bladder and reproductive organs (*S. haematobium*). Single females are incapable of making this journey due to their smaller form (LoVerde *et al.*, 2004). The significant role that males play in female development and migration is thought to favour the evolution of a predominantly monogamous mating system (Emlen and Oring, 1977), though there is some evidence that partner changes can and do occur (Steinauer, 2009).

Eggs laid by mature female worms subsequently migrate through the organs of the host, either into the intestinal lumen in the case of intestinal schistosomiasis (*S. mansoni*) or into the bladder in the case of urogenital schistosomiasis (*S. haematobium*), where they are excreted in faeces or urine, respectively. Some eggs can, however, become trapped in the host tissue, stimulating a granulomatous immune response. Chronic infection can consequently result in significant organ damage. In intestinal schistosomiasis this can include abdominal pain, diarrhoea and hepatosplenic damage leading to portal hypertension and oesophageal varices in the most severe cases (Cheever and Andrade, 1967); whilst in urogenital schistosomiasis pathology includes damage to the bladder and reproductive organs (von Lichtenberg, 1975), which can lead to squamous cell carcinoma of the bladder (Vennervald and Polman, 2009) and increased susceptibility to human immunodeficiency virus (HIV) transmission through damage to the reproductive mucosal membrane (Kjetland *et al.*, 2006). Chronic schistosomiasis can therefore result in significant morbidity in the human host.

Reproduction and the resulting production of eggs is essential in perpetuating the lifecycle. The average lifespan of an adult schistosome is thought to be between three to ten years, depending on the species (Wilkins *et al.*, 1984; Fulford *et al.*, 1995). It is thought that a single schistosome worm pair therefore has the reproductive potential to produce 600 billion schistosomes over its lifetime (Gryseels *et al.*, 2006).

1.2 Epidemiology

1.2.1 Measurement of infection intensity

Accurate measurement of worm burden is essential in studying the dynamics of *Schistosoma* infection. In contrast to intestinal helminths, where worms are expelled from the host following treatment, the intravascular location of schistosomes makes direct measurement of worm burden challenging. Infection intensity is most commonly assessed indirectly by measurement of the number of eggs excreted in faecal or urine samples for *S. mansoni* and *S. haematobium*, respectively.

The Kato-Katz faecal examination technique (Katz, Chaves and Pellegrino, 1972) is the standard diagnostic measure recommended by the World Health Organization (WHO) to quantify the prevalence and intensity of infection in *S. mansoni* due to its low cost, non-invasive nature and the relatively low level of technical skill required (Gray *et al.*, 2011). The sensitivity and specificity of the Kato-Katz technique varies by the schistosome species, infection intensity and the number of stool samples examined. Day-to-day variation in faecal egg output is common among parasites and, in epidemiological studies, participants commonly have changes in their test results on consecutive days (Booth *et al.*, 2003). It has been shown that preparation of multiple Kato-Katz slides from stool samples collected over consecutive days improves the sensitivity of the Kato-Katz technique (Booth *et al.*, 2003; Berhe *et al.*, 2004; Bärenbold *et al.*, 2017).

Measurement of excreted *S. haematobium* eggs is performed by urine filtration and microscopy, as recommended by the WHO (Peters, Warren and Mahmoud, 1976). As with the examination of faecal samples, day-to-day variation in urinary egg excretion is common (Feldmeier and Poggensee, 1993). Generally, a 10 ml sample is measured, however, the influence of day-to-day variation has been seen to be minimised by filtration of greater volumes (Doehring, Feldmeier and Daffalla, 1983). Furthermore, it has been suggested that egg excretion follows a circadian pattern, peaking at around midday (Peters, Warren and Mahmoud, 1976; Doehring, Feldmeier and Daffalla, 1983).

An alternative measurement of infection intensity utilises immunodiagnosics to detect active infection. Circulating Anodic Antigen (CAA) and Circulating Cathodic Antigen (CCA) are well described proteoglycans derived from the gut of adult schistosomes and excreted by male and female worms that are both juvenile (reproductively immature) and adult mature worms (Deelder *et al.*, 1989; Deelder *et al.*, 1994). CAA has been identified as a marker that is able to identify single worm infections and is therefore proposed as an effective tool to reliably study drug efficacy, worm burden and anti-fecundity effects (Corstjens *et al.*, 2014). Originally detected in host serum by enzyme-linked immunosorbent assay (ELISA) (Deelder *et al.*, 1989), it has been demonstrated in animal studies that CAA correlates with worm burden (Qian and Deelder, 1983; Deelder *et al.*, 1989). CAA and CCA are also measurable in urine and rapid loss following treatment suggests that their concentration represents live worm burden (De Jonge *et al.*, 1989; Knopp *et al.*, 2015; Kildemoes *et al.*, 2017).

In order to make immunodiagnosics more accessible in remote field locations, a point-of-care (POC) rapid test was first developed for CCA (van Dam *et al.*, 2004); however questions were raised regarding its sensitivity in low *S. mansoni* endemic areas (Peralta and Cavalcanti, 2018) and the specificity and sensitivity of the test in detecting *S. haematobium* infection (Obeng *et al.*, 2008; Sanneh *et al.*, 2017). More recently a rapid POC test was developed for CAA

measurement by means of an up-converting phosphor-lateral flow circulating anodic antigen (UCP-LF CAA) assay and it is reported that UCP-LF CAA has superior sensitivity, even in a low-endemic setting (Knopp *et al.*, 2015). Yet, despite considerable advances in this technology in recent years, immunodiagnostic testing is not as cost-effective as traditional parasitological methods in resource limited settings and is less frequently used (World Health Organization, 2020a).

1.2.2 Age, exposure and experience of infection

The age profile of schistosome infection has a characteristic peak in infection prevalence and intensity during early adolescence. It was initially thought that this pattern of infection resulted from age-specific exposure through water contact. However, several studies demonstrated that age-related patterns of infection remain, even after accounting for variation in exposure (Chandiwana and Woolhouse, 1991; Demeure, Rihet, Abel, Ouattara, Bourgois, *et al.*, 1993; Etard, Audibert and Dabo, 1995). Moreover, peak infection intensity is higher and occurs at an earlier age in high-transmission compared to low-transmission areas. This observation is termed the 'peak shift' and was first acknowledged by Woolhouse *et al.* (1991), with reference to prevalence of *S. haematobium* infection in the Zimbabwe highveld. The following year, Fulford and colleagues identified the same peak shift pattern in *S. mansoni* infection intensity data in Kenya (Fulford *et al.*, 1992) (Fig. 1.1A). The peak shift was then illustrated in a later theoretical study, whereby populations were subjected to variations in the transmission rate (Woolhouse, 1998) (Fig. 1.1B). The concept of the peak shift is considered to support a role for gradually acquired immunity to schistosomes, since the development of an effective immune response is believed to be dependent on the accumulation of sufficient experience of infection. Individuals are thought to be exposed to a greater number and diversity of antigens earlier in their lifetime in high-transmission compared to low-transmission areas, so achieve the threshold for protection at a younger age (Woolhouse and Hagan, 1999). The age profile of infection intensity is consequently believed to reflect the combined contribution of age-dependent exposure and partially protective immunity on the establishment of infection; however, the relative contribution of these two elements is widely debated in the literature (Warren, 1973; Pinot De Moira *et al.*, 2010; Anderson *et al.*, 2015). Moreover, additional factors, including puberty, sex, ethnicity and occupation (Butterworth *et al.*, 1984; Fulford *et al.*, 1998) amongst other population and environmental heterogeneities may have a significant role influence on transmission dynamics (Woolhouse *et al.*, 1998). Many of these risk factors are connected to behavioural variations in the use of water sources (Pinot De Moira *et al.*, 2010). Several authors

have consequently stressed the importance of systematic collection of individual level water contact observations (Anderson *et al.*, 2016).

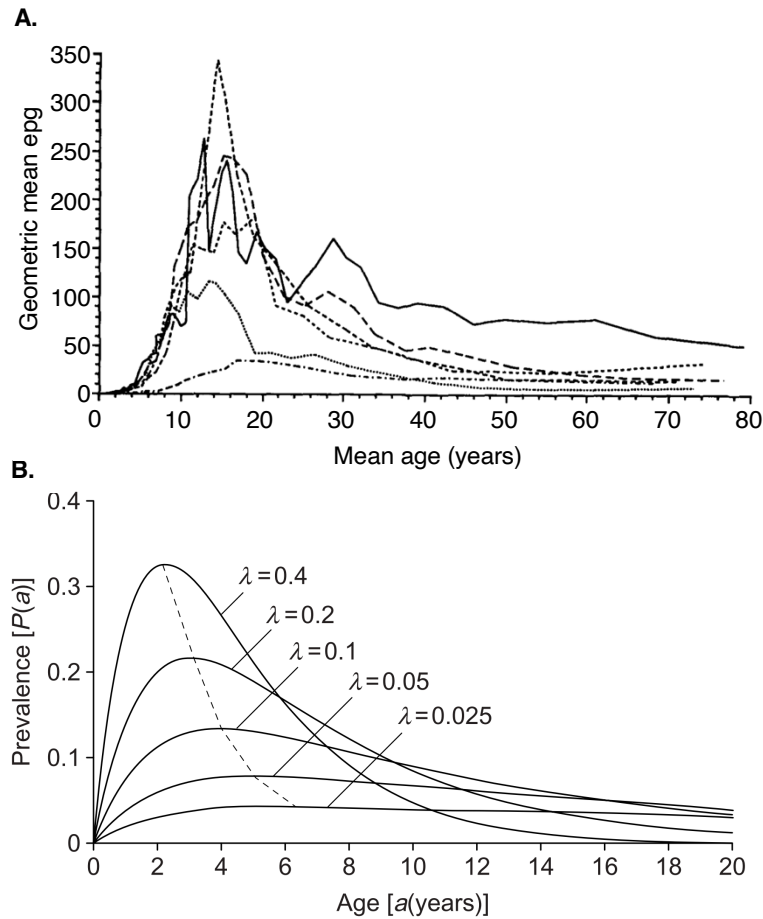


Fig. 1.1 Illustration of the peak shift by A) observations of the shift in peak intensity of infection by Fulford *et al.* (1992) in several villages in the Machakos District, Kenya, and B) the predicted relationship between the prevalence of parasite infection and age in populations subject to different transmission rates from a theoretical study of schistosome infection by Woolhouse (1998), where the dashed line indicates the relationship between peak prevalence and the age of peak prevalence (the peak shift). Copyright permissions granted.

Observations of water contact duration have traditionally been used to estimate risk of human exposure. The probability of exposure to cercariae during any single water contact event is dependent on a variety of biological, ecological and social factors (Fulford *et al.*, 1996). Transmission of schistosomiasis is consequently highly focal and the duration of water contact alone does not accurately describe exposure to infectious cercariae (Fulford, 2000; Pinot De Moira *et al.*, 2007). An alternative approach to estimating cercarial exposure was therefore devised by Fulford (2000). This included analysis of contributory behavioural and ecological

factors which were weighted according to their influence on the risk of exposure. This formula has since been used by others to explore the variations in exposure that can occur on a small geographical scale, for example within a single village (Pinot De Moira *et al.*, 2007).

The infection intensity profile often mirrors the pattern of exposure with age; however, variations in exposure do not always sufficiently explain age-related differences in infection intensity and these differences have also been shown to be associated with the development of partially protective immunity. Experience of infection is determined by both duration of infection and intensity of exposure; yet, this relationship is frequently confounded by age (Butterworth *et al.*, 1988). Populations who are naïve to schistosomiasis can be instrumental in unambiguously distinguishing between the effects of age versus experience of infection *per se*; yet, by definition, these cohorts tend to be rare in endemic regions. One such study was conducted following the construction of a dam the city of Richard Toll in Northern Senegal, in which Sow *et al.* (2011) examined water contact behaviour associated with the ensuing *S. mansoni* outbreak. Whilst foci of *S. haematobium* infection had previously been identified elsewhere in the Senegal River basin, neither *S. mansoni* nor *S. haematobium* infection had been reported in the Richard Toll area (Stelma *et al.*, 1993). The severity of the outbreak was attributed to the lack of acquired immunity in this newly exposed community in addition to the intensity of transmission; yet, despite an assumed lack of immune protection, epidemiological curves reflected the age-related patterns seen in more established endemic situations (Gryseels *et al.*, 1994). The observed frequency and duration of water contact did not, however, explain the high infection intensity of the outbreak (Sow *et al.*, 2011) and whilst the authors found that adjusting for degree of water contact and time of day improved the prediction of infection intensity with age-related exposure, age- and sex-related exposure was not entirely explanatory of infection intensity within this epidemic focus of *S. mansoni* infection. Interestingly, the observed antibody response in this population was found to be largely similar to that seen in chronically infected populations (Van Dam *et al.*, 1996).

Other studies have explored the dynamics of repeated treatment, exposure, reinfection cycles in occupationally exposed cohorts. Satti *et al.* (1996) first observed significantly lower intensities of reinfection amongst canal cleaners with a long history of heavy exposure compared with newly recruited canal cleaners in Sudan, suggesting that greater experience of infection is associated with protection. In a similar study, conducted in car washers in Kenya, Black *et al.* (2010) found that resistance to reinfection can be acquired or enhanced by multiple rounds of treatment following reinfection. Several authors have also suggested that multiple exposure, treatment and reinfection cycles may induce immunological changes which lead to resistance to reinfection (Mutapi *et al.*, 1998; Colley and Evan Secor, 2004). Furthermore, Karanja *et al.* (2002) identified

three types of responders amongst adults with similar levels of occupational exposure, those that were found to either be: resistant to reinfection, develop resistance after multiple treatments or remain susceptible, indicating that history of exposure and cumulative antigen exposure may be associated with the development of a protective immune response against *S. mansoni* infection, independent of age. Additionally, the maintenance of stable endemic equilibrium is dependent on density dependent constraints. In the absence of some form of density regulation, the level of infection would increase uncontrollably. Therefore, whilst exposure to cercariae is essential for transmission, and age-dependent exposure is likely to have a determinant role in shaping the classically peaked age-intensity profile, there is also substantial evidence from epidemiological studies to support a role for immunity in mediating parasite establishment (Hagan *et al.*, 1991; Dunne *et al.*, 1992; Hagan, 1992; Demeure, Rihet, Abel, Ouattara, Bourgois, *et al.*, 1993; Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012).

An in depth understanding of the relationship between age, exposure and acquired immunity is essential to comprehending the drivers of schistosome transmission. This knowledge can subsequently inform policy decisions regarding infection and disease control.

1.2.3 Control of schistosomiasis

Chemotherapy

Schistosomiasis disease management is centred on chemotherapeutic treatment with the anti-helminthic drug, praziquantel. Treatment is administered at a community level through, school-based mass drug administration (MDA) campaigns, since infection intensity tends to be highest in this age group. Special groups of adults who are considered to be at high risk are also treated (World Health Organization, 2006).

Praziquantel kills the adult worms but does not affect the juvenile stages. Moreover, infestation of large bodies of freshwater that are essential to the livelihoods of local communities, such as the Lake Victoria and Lake Albert in East Africa, results in high rates of post-treatment reinfection. MDA is therefore repeated annually. The primary aim of MDA is to reduce the severe morbidity associated with schistosomiasis, by reducing high intensity infections, and to ultimately eliminate schistosomiasis as a public health problem by 2025 (World Health Organization, 2012). A new roadmap for the control of Neglected Tropical Diseases, with targets through to 2030 is due to be released this year (World Health Organization, 2020a). The introduction of routine MDA campaigns in 2003 has reduced the frequency with which high intensity infections are observed at the country-level and, as a consequence, MDA is reducing the incidence of severe schistosomiasis (Deol *et al.*, 2019).

It has been hypothesized that the dioecy of schistosomes may result in non-zero breakpoint of infection when infection intensity is low, whereby single-sex infections do not result in the production of infectious stages. The impact of separate sexes on parasite population dynamics was first recognised by Macdonald (1965). It has therefore been suggested that, under certain epidemiological conditions, widespread treatment of the at-risk population may result in changes to the population biology of schistosomes and may have the potential to contribute towards reduced transmission at the community level (Anderson and May, 1985; French *et al.*, 2010; World Health Organisation, 2017).

Whilst MDA has been shown to be effective in reducing the prevalence of high-intensity infections in several regions, a sustained response is not achieved in all populations (Deol *et al.*, 2019; Kittur *et al.*, 2019). Previously the literature has focused on the concept of breaking schistosome transmission (Rollinson *et al.*, 2013; Anderson *et al.*, 2015). However, it is now recognised that control programmes must consider adaptive strategies in persistent hotspots of schistosome transmission, where MDA has proven less effective at reducing the prevalence and intensity of infection (Wiegand *et al.*, 2017; Kittur *et al.*, 2019). This is likely to require a multifactorial approach to control, including water and sanitation improvement programmes and possibly control of the intermediate host with molluscicides (King, 2017; Li *et al.*, 2019; World Health Organization, 2020a).

1.3 Mathematical models of schistosome transmission

Mathematical models of transmission have a wide range of applications, including the planning and implementation of control programmes, and in many cases, they are able to replicate the fundamental dynamics of infectious diseases. The development of any model is dependent on the balance between parsimony and biological fact. Models are, by their very nature a simplification of reality. Schistosomiasis is highly focal, demonstrating significant spatial heterogeneity, therefore, attempting to characterise all the heterogeneity of *Schistosoma* transmission would result in a mathematically intractable model from which it would be difficult to draw meaningful conclusions. That said, it remains important to capture the principal characteristics of transmission.

A large proportion of the existing schistosome transmission models focus primarily on the relationship between age-dependent exposure and subsequent age-infection profiles, assuming that immunity plays a minor role in the transmission story. There is a significant disconnect between the opinion of theoretical modellers and biologists on this assumption, since substantial evidence from immuno-epidemiological studies exists to demonstrate the

acquisition of partially protective immunity to schistosomiasis. Immunity is therefore widely recognised by field biologists to contribute to variation in host susceptibility; yet immunity remains infrequently estimated and poorly defined in existing models. This has been largely attributed to insufficient data (Anderson *et al.*, 2015).

Several authors have emphasised that the inclusion of acquired immunity in models is necessary to reproduce observed age-intensity profiles (Woolhouse, 1991; Anderson *et al.*, 2015). When immunity is omitted, models may overestimate the positive impact of control achieved by MDA (Gurarie and King, 2005, 2014).

The accuracy of model predictions is dependent on the collective influence of biological assumptions and the data that informs these assumptions. The ideal model will have sufficient empirical basis so that the dependence on data will outweigh the dependence on model assumptions, thus increasing confidence in the validity of the model predictions. The generation of knowledge is therefore essential to informing analysis and intervention strategies and, when studying the dynamics of infectious diseases, interdisciplinary collaboration is essential.

“A closer collaboration between biometricians and parasitologists, and a better acquaintanceship of each with the methods of the other, is one of the most useful things we can work for today”
(Hackett, 1937)

This is particularly important in schistosomiasis where complex social and ecological interactions lend themselves to a systems-based approach (Krauth *et al.*, 2019).

There is an unanswered need for better characterisation of acquired immunity within schistosome transmission models, including the development of mechanistic models that empirically define immunity parameters. This requirement is examined in further detail in Chapter 3, where I draw from the current theory surrounding anti-infection immunity in order to explore the relative contribution of exposure and immunity to schistosome infection with mathematical models.

1.4 Adaptive immunity in human schistosomiasis

Adaptive immunity occurs following exposure to an antigen, either from a pathogen or a vaccination, and encompasses the combined contribution of cell-mediated immunity, facilitated by T lymphocytes (T cells) and humoral immunity, which is mediated by activated B lymphocytes (B cells) and antibodies. Since this thesis focuses on protective immunity driven by antibody responses, this section will focus on the humoral immune response.

1.4.1 Humoral immunity

Naïve B cells are activated after first encountering an antigen. Antigen that is recognised by membrane bound immunoglobulin on the B cell is internalised, processed, and presented to CD4⁺ T helper cells (Th cells) via major histocompatibility complex (MHC) class II molecules (Janeway *et al.*, 2001). Th cells that recognise the MHCII-complexed antigen provide activating signals through co-stimulatory cytokine binding and engagement of CD40 on B cells by its ligand, CD40L, expressed on T cells (Oettgen, 2016). Through these signals the T helper cell stimulates the B cell to proliferate and differentiate into either a memory B cell or a plasma cell. Short-lived plasma cells (SLPCs) comprise the first wave of humoral immune response, which provide initial low-affinity antibodies towards the antigen, whereas memory B cells form memory of the specific antigen so that the immune response can react quickly upon re-exposure to that same antigen. Long-lived plasma cells (LLPCs) are also recognised to be a separate immunological entity, contributing to immune memory (Brynjolfsson *et al.*, 2018).

It is understood that the adaptive immune system has evolved in order to “keep pace with the increasingly sophisticated evasion of pathogens” (Trowsdale and Parham, 2004). In order to recognise and react to the myriad antigens, the adaptive immune system has evolved to generate a broad diversity of antigen-binding specificities, known as the antibody repertoire, generated by the combined process of genetic recombination and mutation.

1.4.2 Antibody diversification

Class switch recombination

Prior to antigen-exposure, naïve B cells exclusively express immunoglobulin (Ig)M and IgD (Janeway *et al.*, 2001). Class switch recombination (CSR) is the process by which the constant region of the immunoglobulin heavy chain proliferating B cells is switched from IgM and IgD to the isotype that can provide greatest protection against the pathogen, IgG, IgE or IgA (Stavnezer, Guikema and Schrader, 2008).

IgG is the most abundant circulating antibody and provides protection primarily against viruses, bacteria and fungi (Janeway *et al.*, 2001). Whereas, IgA is the principal antibody class found in the secretions of mucosal membranes, such as the respiratory tract and digestive system, and serves as the first line of defence against inhaled or ingested pathogens (Woof and Kerr, 2004). Naturally occurring human immunity to helminth infection is however characterised by the expansion of circulating IgE, understood to result from clonal expansion of IgE-producing B lymphocytes (King *et al.*, 1991; Gould and Wu, 2018).

The mechanism of CSR produces antibodies with different effector properties but does not alter immunoglobulin antigen specificity (Stavnezer, Guikema and Schrader, 2008). CSR was previously understood to occur within the germinal centre (GC); however, a recent study proposes that CSR occurs before the GC is formed, prior to somatic hypermutation (SHM) (Flemming, 2019).

Affinity maturation by somatic hypermutation and selection

Affinity maturation is the mechanism through which B cells improve their recognition of antigen by refining the affinity of the B cell receptor (Janeway *et al.*, 2001). Following repeated exposure to an antigen, antibodies of successively greater affinity are produced.

SHM and the selection process for B cells with high-affinity receptors occur in the germinal centres of the peripheral lymphoid tissues (Potter and Capra, 1998). SHM is a fundamental means of generating antibody diversity. During this process, high frequency point mutations accumulate in the immunoglobulin variable gene regions, which can alter the antigen binding affinity (Martin *et al.*, 2015). Cells then migrate from the dark zone of the GC to the light zone, where the selection process that identifies favourable mutations, based on affinity, occurs.

During selection, antigen is presented to B cells by follicular dendritic cells (FDCs). B cells with mutations that increase the binding affinity are positively selected, whilst B cells of lower affinity, or that have acquired stop codons, are subject to apoptosis (Martin *et al.*, 2015). High affinity B cells are enriched accordingly.

1.4.3 Antibody responses associated with protection against *Schistosoma* infection

The current understanding of the mechanisms that confer partial immunity to schistosomes is largely dependent on immuno-epidemiological evidence, since the mouse model does not adequately replicate human resistance to infection.

Early studies by Butterworth *et al.* (1984, 1987, 1988) demonstrated elevated levels of pre-treatment eosinophils, associated with detectable adult worm and egg antigen specific IgM and IgG antibodies, increased susceptibility to reinfection. Similarly, in a study conducted in Sudan, Satti *et al.* (1996) found that specific IgG₂ to egg antigen was higher in both school children and newly recruited canal workers, compared to either adults both without high levels of occupational exposure or chronically infected canal workers. IgG₂ has also been shown to decrease with age and experience of infection (Naus *et al.*, 1998), suggesting that IgG₂ and IgM have a role in blocking the development of protective immunity; it follows that IgG₂ and IgM

are associated with a higher risk of reinfection (Satti *et al.*, 1996). The finding that IgG₂ to soluble egg antigen (SEA) is significantly higher in children and relatively naïve adults, compared to those with greater experience of infection has led to the hypothesis that the early immunogenic stimulus is provided by egg antigens.

Alongside the upregulation of the Th2 response, both plasma IgE and eosinophils increase, positively associated with the Th2 cytokines interleukin (IL)-4 and IL-5, respectively (Walter *et al.*, 2006; Wilson *et al.*, 2013). IL-4 plays a necessary role in the generation and maintenance of *in vivo* IgE responses and has been shown to regulate the polyclonal IgE response both *in vitro* and *in vivo* (Finkelman *et al.*, 1988), whereas IL-5 is a growth and stimulating factor of eosinophils (Rosenberg, Dyer and Foster, 2013). Several immuno-epidemiological studies, examining the three major schistosomes of importance to human health have found that levels of IgE specific to schistosome antigens, particularly antigens derived from the adult worm, are negatively associated with reinfection (Paul Hagan *et al.*, 1991; Dunne *et al.*, 1992; Jiz *et al.*, 2009). Several studies also report that increased schistosome-specific IL-4 and IL-5 are positively associated with protection against reinfection (Medhat *et al.*, 1998; Walter *et al.*, 2006; Bourke *et al.*, 2013; Wilson *et al.*, 2013).

Antigen-specific IgE is low in children compared to adults, gradually increasing with age. On the other hand, antigen specific IgG₄ is highest in children, decreasing through adolescence into adulthood. Correspondingly, the relative amounts of parasite-specific IgE and IgG₄ have been associated with respective protection and susceptibility to schistosome infection in several studies, conducted in a number of different countries (Demeure *et al.*, 1993a; Naus *et al.*, 1998). Earlier studies proposed that the balance between early production of IgG₄ antibodies in children and the gradual increase in IgE antibodies explains the delayed development of protective immunity to schistosomes (Butterworth *et al.*, 1987; Paul Hagan *et al.*, 1991). The inverse relationship between IgG₄ and IgE suggests that these two isotypes antagonise each other during protection, with IgG₄ antibodies blocking IgE effector functions in two possible ways. Firstly, by direct competition with specific IgE for epitopes, neutralising antigen binding sites, thus impeding mast cell and basophil degranulation by preventing IgE cross-linking (Rihet and Demeure, 1991; Demeure *et al.*, 1993b). Secondly, it has been proposed that IgG₄ down regulates effector cell activation after cross-linking the high-affinity IgE receptor, FcεRI, to the inhibitory IgG receptor FcγRIIb (Turner *et al.*, 2005; Strait, Morris and Finkelman, 2006; Aalberse *et al.*, 2009). IgG₄ is therefore likely to be important in disease expression.

1.4.4 Concomitant immunity

The term ‘concomitant immunity’ originates from cancer immunology with reference to the way in which cancer cells harness the host immune response to prevent competition from the development of further tumours. This terminology is now frequently used in parasitology to describe the development of an effective anti-larval immune response, often combined with an existing, mature parasite burden (Brown and Grenfell, 2001). It is suggested that adult parasites, in effect, vaccinate their hosts by releasing cross-reactive larval antigens to prime the host immune system against additional invading larvae, thus limiting parasite burden and preventing overcrowding.

The concept of concomitant immunity to schistosome infection was first described in *S. mansoni*. Smithers and Terry proposed that established live worms facilitate an immune response that prevents establishment of further parasites as a density limiting process (Smithers and Terry, 1967). It was later proposed that host antigens are gradually acquired by *S. mansoni*, either through the direct acquisition of host antigens or by synthesis of antigens that mimic the host (Clegg, Smithers and Terry, 1971). These non-immunogenic, ‘self’ antigens mask the parasite from the host immune system (Smithers, Terry and Hockley, 1969). This is now a well-recognised method by which worms are able to evade the host immune response and explains why adult schistosomes are subsequently able to survive within the vasculature of their host for several years (McSorley, Hewitson and Maizels, 2013). Yet, these masked adult worms are still able to elicit a protective response against invading cercariae, as demonstrated by transfer of adult worms into the portal system of previously uninfected rhesus macaques, which confers protection against subsequent cercarial exposure (Smithers and Terry, 1967). Clegg *et al.* (1971) found that it can take approximately two weeks following infection for the developing schistosomula to acquire host antigens. This delay may provide an opportunity for recognition of parasite proteins by the host immune system.

Furthermore, the acquisition of naturally acquired immunity against infection occurs over several years. This observed delay in acquisition of immunity may be explained by the notion that worm death is necessary for the development of immunity (Mutapi *et al.*, 1998; Woolhouse and Hagan, 1999). As worms die, they release previously cryptic antigens that are then recognised by the host immune system, which subsequently mounts a protective response. Schistosome worms can live for a number of years within their host, so naturally acquired immunity can take some years to subsequently develop. This concept has been referred to as delayed concomitant immunity (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). Dependence on worm death for the development of a protective immune response is supported by evidence from field studies that indicate that antibodies to cryptic tegumental proteins are positively

correlated with protection to schistosome infection (Fitzsimmons, Jones, Stearn, *et al.*, 2012). This hypothesis of delayed concomitant immunity is best described through the responses to members of the *Schistosoma mansoni*-Tegumental Allergen Like (SmTAL) protein family. SmTAL1 (previously Sm22.6), the first of the family to be described, is the major IgE antigen detected on Western blots probed with sera from endemic populations (Dunne *et al.*, 1992).

1.4.5 Tegument Allergen-Like protein family and anti-infection immunity

The Tegument Allergen-Like (TAL) protein family is comprised of 13 calcium channel proteins expressed in the schistosome tegument at different levels across all life cycle stages of the parasite. Several members of the TAL protein family have been associated with delayed concomitant immunity. SmTAL1 (previously named Sm22.6) was the first protein in the family to be described and is the main antigen that elicits a detectable IgE response using Western blots probed with sera from endemic populations (Dunne *et al.*, 1992). The level of detectable SmTAL1-specific IgE in the plasma of infected individuals increases with age, corresponding with a significant decrease in reinfection intensity (Webster *et al.*, 1996; Fitzsimmons, Jones, Stearn, *et al.*, 2012).

SmTAL1 is expressed in the 24-hour schistosomule onwards and is predominantly found within the tegument of the adult worm (Fitzsimmons, Jones, Stearn, *et al.*, 2012). As discussed in section 1.1, the tegument is a complex seven-layered membrane that serves to protect the parasite from the host immune response. Tegumental proteins tend to be cryptic, so SmTAL1 is not appreciably exposed to the host by live worms. As a result, an increase in IgE to SmTAL1 is thought to be dependent on antigen exposure following worm death and an increase in SmTAL1-specific IgE is seen following praziquantel treatment, accordingly (Pinot De Moira *et al.*, 2013). SmTAL3 is also expressed in the adult worm and follows a similar specific IgE profile with age. SmTAL3-IgE responders are a subset of SmTAL1-IgE responders (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012; Fitzsimmons, Jones, Stearn, *et al.*, 2012), suggesting that greater exposure to dying worms may be required to stimulate the SmTAL3-IgE response. This may therefore be indicative of an antigen-threshold effect of TAL protein-specific IgE.

It has been suggested that protective immune responses target the invading cercaria (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012); yet, SmTAL1 is expressed from the 24-hour schistosomula onwards and SmTAL3 is expressed in the worm stages only, thus neither protein is appreciably expressed in cercariae. Furthermore, after controlling for age, IgE responses to SmTAL1 and SmTAL3 are not significantly associated with immunity (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). This indicates that these antigens are not likely to be the direct target of

immunity. It has been suggested that when cercaria penetrate the host epidermis and shed their glycocalyx, parasite antigens are exposed to host immune cells, prior to the full development of the tegument (Da'dara and Krautz-Peterson, 2014). There is evidence to suggest that another TAL family member, SmTAL5 is cross-reactive with SmTAL3-specific IgE and, since SmTAL5 is expressed in the early schistosomula, SmTAL5 has been proposed as a potential target of delayed concomitant immunity, although it is likely that other cross-reactive antigens may also contribute to the development of partial immunity (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012).

The association between SmTAL1 and immunity has been repeated across different study populations, however, the hierarchical IgE response to SmTAL1, 3 and 5 has only been shown in a male-only cohort from a village on the shore of Lake Victoria, Uganda. There is therefore a need to demonstrate that these interactions are generalisable to other populations and schistosome species.

1.4.6 Immune responses associated with anti-fecundity immunity

Several vaccine efficacy studies have reported reductions in egg output or worm fecundity, suggesting that immune responses may be able to regulate egg laying activity (Boulanger *et al.*, 1991; Mbanefo *et al.*, 2015; You *et al.*, 2015). In schistosomiasis, worm fecundity is measured by calculating the ratio of excreted eggs to CAA, as a proxy for live worm burden, as described in Chapter 2, section 2.16.10. An age-dependent reduction of measurable *S. haematobium* egg to worm ratios was first observed by Agnew and colleagues in a study conducted in Kenya (Agnew *et al.*, 1996); but, in the same study, the equivalent response was not seen in *S. mansoni*. Possible anti-fecundity immune responses are also seen in some livestock-infecting *Schistosoma* species within the same clade as *S. haematobium* (Bushara *et al.*, 1980), as well as in other trematode species such as *Fasciola hepatica* (Norbury *et al.*, 2018). Evidence from these species that are closely related to *S. haematobium* has made a sizable contribution to the current knowledgebase regarding fecundity of schistosomes, since the murine model is not permissive to *S. haematobium* infection, so laboratory infections are comparatively difficult to maintain. *S. haematobium* is therefore understudied relative to *S. mansoni*, despite *S. haematobium* being the most prevalent schistosome species in Africa.

As with anti-infection immunity, evidence tends to suggest that an anti-fecundity immune response develops over time, with increased experience of infection. Bushara *et al.* (1980) demonstrated suppression of worm fecundity following both experimental and naturally acquired infection of cattle with *S. bovis*. Furthermore, the transfer of worms from 'resistant'

cattle that had exhibited reduced worm fecundity into naïve cattle resulted in large numbers of eggs being excreted from the recipient cattle (Bushara *et al.*, 1983). On the other hand, serum transfer from these 'resistant' cattle resulted in reduced worm fecundity following subsequent experimental infection (Bushara *et al.*, 1983). Furthermore, worm fecundity has been shown to be inversely related to the duration of infection in *S. mattheei*, an ovine schistosome species (Coyne and Smith, 1991). No evidence for immunity-independent density-dependent regulation of fecundity was found by Coyne and Smith (1991).

Glutathione-S-transferase (GST) is the principal antigen that has been associated with a reduction in fecundity following vaccination in several of the major schistosome species, including *S. mansoni*, *S. haematobium*, *S. japonicum* and *S. bovis*. Sm28GST-specific IgA has been shown to neutralise the enzymatic activity and is associated with impaired schistosome fecundity, by limiting both egg production and miracidia hatching capacity (Grzych *et al.*, 1993; Capron *et al.*, 1994). Likewise, vaccination of animals with recombinant *S. bovis* 28kDa GST has demonstrated host species specific effects on worm fecundity. In cattle, administration was shown to reduce worm fecundity, but this effect was not replicated in sheep or goats (Bushara *et al.*, 1993). Moreover, in a study exploring immune responses associated with *S. haematobium* fecundity in Mali, Wilson and colleagues found no association between antigen-specific IgA, IgG₁ and IgE to the *S. haematobium* 28kDa GST and worm fecundity in humans (Wilson *et al.*, 2014). When accounting for age, sex and village, both transmission intensity and IgG₁ specific to antigens derived from the adult worm were significantly associated with notably reduced fecundity (Wilson *et al.*, 2014).

In a mathematical modelling study examining the dynamics of anti-fecundity immunity in *S. haematobium*, Mitchell and colleagues modelled plasma cell populations of IgA and IgG, suggesting that protective immunity involving a reduction in *S. haematobium* fecundity was associated with worm death (Mitchell *et al.*, 2012). Nevertheless, the mechanism by which naturally acquired anti-fecundity immunity may be attained remains to be elucidated.

1.5 Thesis outline and aims

This thesis aims to significantly advance on our understanding of immunity to schistosomiasis. Whilst *S. haematobium* is the most prevalent schistosome species, *S. mansoni* is the best understood of the species important to human health, as a result of its permissibility in the murine model. Therefore, with a focus on *S. mansoni*, the current study uses data from historical studies and leading theory of anti-reinfection immunity to inform parameters in modelling the role of immunity in the transmission potential of schistosomiasis.

As discussed, the transmission of schistosome worms is heavily dependent on the interplay between exposure to infectious cercariae in the environment and the development of protective antibody-mediated immune responses. Existing modelling studies suggest that current MDA programmes may lead to elimination of schistosomiasis by 2030, in line with the WHO targets (World Health Organization, 2012), yet there is an apparent lack of empirically derived exposure and immunity parameters in the existing transmission model literature for schistosomiasis. In Chapter 3, I develop the structure of an existing mathematical model of schistosome transmission to incorporate data-informed exposure and an IgE-driven anti-infection immune response, stimulated by the death of adult worm, with the specific aim of answering the following question:

Can partially protective SmTALI-IgE immunity be incorporated into a *S. mansoni* transmission model that reproduces age-intensity and -serology profiles?

The inclusion of this immune function is supported by several epidemiological studies that have demonstrated an association between worm death, human host serum-IgE and, in turn, reduced rates of reinfection following treatment. Fitzsimmons *et al.* (2012) have, however, proposed that the relationship between SmTALI-IgE and reinfection is in fact more complicated and have implicated additional members of the TAL protein family in the progressive development of the IgE-driven immune response. This finding has not yet been substantiated in other study populations, nor has this relationship been shown in *S. haematobium*. A further objective of the current study is to investigate the relationship between protective antibody responses and their antigenic targets in *S. haematobium*. Chapter 4 therefore combines a data mining approach with wet laboratory experimental work to advance our understanding of the TAL-specific IgE response against infection, specifically asking whether this relationship is generalisable to *S. haematobium*:

Does the anti-infection protective response to *S. haematobium* demonstrate the same dependence on TAL-specific IgE as it does to *S. mansoni*?

In addition to anti-infection immunity, *S. haematobium* and other schistosome species within the same clade (e.g. *S. bovis*) show evidence of an immune response that reduces the number of eggs that female worms excrete. This has been associated with IgG₁ specific to antigens derived from the adult worm but, as yet, the specific immunogenic proteins are unknown. The final aim of this thesis, presented in Chapter 5, is therefore to advance our understanding of the process of anti-fecundity immunity, specifically exploring possible antigenic target proteins:

What are the antigenic targets of IgG₁-mediated anti-fecundity immunity in *S. haematobium*?

2

Materials and methods

2.1 Description of study sites and cohorts

Data and samples from three different archived studies were utilised in this thesis to explore protective immunity against *Schistosoma* species. Data from Booma, Uganda were used in Chapter 3 to examine the contribution of exposure and IgE-driven anti-infection immunity to the patterns of schistosome infection with age, in a mathematical transmission model. Data and samples from studies conducted in Musoli, Uganda and Segou Region, Mali were used in Chapter 4 to explore the IgE response to tegument allergen-like (TAL) proteins in reinfection immunity in *S. mansoni* and *S. haematobium*, respectively. Finally, data and samples from Segou Region, Mali were again used in Chapter 5, this time to advance our knowledge of IgG₁-associated anti-fecundity immunity, through the identification of IgG₁-target *S. haematobium* proteins. The following sections provide a brief outline of the studies, however a detailed description of each cohort and the data collection methods have been published elsewhere, as indicated in the relevant sections.

2.1.1 Booma, Uganda

The village of Booma is located along the eastern shore of Lake Albert, Buliisa District, Uganda. The population is primarily comprised of two tribes (Alur and Bagungu) that live at geographically distinct ends of the village and have differing knowledge, attitudes and behaviour with regard to the use of the lake (Pinot De Moira *et al.*, 2010). A comprehensive longitudinal cohort study was conducted between 1998 and 2000, encompassing prevalence and intensity of infection, serological analysis, observation of water contact behaviour and malacology of the freshwater snail population. A detailed description of the study population and data collection methods can be found in Pinot de Moira *et al.* (2007), including detailed analysis of the microgeographical and tribal variations in exposure to infectious cercariae and calculation of individual-level annual cercarial exposure scores. Measurement of antibody responses has also been previously performed in this study population (Pinot de Moira *et al.*, 2010). The level of detail of the individual-level data available, makes this a unique and valuable study for mathematical modelling analyses. Data from this study are therefore used to parameterise and

fit empirically informed exposure and immunity functions to a schistosomiasis transmission model in Chapter 3. The study sample comprised 248 individuals, ranging in age from 5 to 50 years old (Alur: n = 110; Bagungu: n = 138). Analysis was conducted on a subset of the population for whom parasitology, serology and cercarial exposure scores were available (n = 197; [Alur: n = 85, females: n = 34, males: n = 51] ; [Bagungu: n = 112, females: n = 62, males: n = 50]).

2.1.2 Musoli, Uganda

Musoli is a fishing village situated on the northern shore of Lake Victoria, Uganda. A longitudinal cohort study was conducted between 2005 and 2006, details of which can be found in Fitzsimmons *et al.* (2012). Quantitative parasitology, measuring *S. mansoni* infection intensity, was available at baseline and at nine weeks and two years post-treatment. Serological analysis of the SmTAL proteins has previously been conducted on a subset of 144 males from the original Musoli cohort, which identified an association between SmTAL1, SmTAL3 and SmTAL5 and IgE-mediated immunity to reinfection (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). Data from this study and plasma samples were used in Chapter 4 to advance our understanding of the relationship between the TAL family of proteins, IgE and reinfection. The study cohort comprised a population of 217 male individuals ranging in age from 7 to 60 years old. The population was divided into five age groups of approximately equal size for analysis: aged 7 to 9 (n = 40); 10 to 14 (n = 46); 15 to 24 (n = 38); 25 to 34 (n = 46), and 35 to 60 (n = 47).

2.1.3 Kaladangan, Guenidaga and Kalabougou, Mali

A longitudinal cohort study was conducted in Segou Region, south-central Mali between 2006 and 2007. The study comprised parasitology, serology and measurement of CAA in urine samples in two villages with very high-intensity transmission (Kaladangan and Guenidaga) and one village with moderate to high-intensity transmission (Kalabougou) before treatment and nine weeks after treatment. Egg count data was also available at one-year follow-up timepoint. Full details of the study sites and data collection methods can be found in Wilson *et al.* (2014).

Prior to the publication of the *S. haematobium* genome, Dickinson screened *S. haematobium* expressed sequence tags (ESTs) for sequences orthologous to the *S. mansoni* TAL proteins (Dickinson, 2012). Antibody responses to the TAL protein family were subsequently measured in the Mali plasma samples for IgG₁, IgG₄ and IgE isotypes, as part of the study by Dickinson. Chapter 4 develops the theory surrounding TAL-specific IgE and rates of reinfection, using plasma samples from this study to perform competition assays.

Wilson *et al.* demonstrated an association between soluble worm antigen (SWA)-specific IgG₁ antibody titres and a reduction in worm fecundity scores, suggestive of anti-fecundity immunity (Wilson *et al.*, 2014). The human plasma samples from this study conducted in Mali were also used in Chapter 5 to identify protein targets of this possible IgG₁-driven anti-fecundity immune response.

The study comprised a total of 261 individuals for whom pre- and post- treatment egg counts, urine samples and serology were available. The study cohort was then broken down into those from a moderate intensity transmission village (Kalabougou, n = 158) or those from high intensity transmission villages (Kaladangan and Guenidaga, n = 103). The population was also divided into age groups according to the analysis performed. These either comprised of six age groups to mirror those analysed in the *S. mansoni* (Musoli) cohort: 5 to 6 (n = 52); 7 to 9 (n = 50); 10 to 14 (n = 43); 15 to 24 (n = 39); 25 to 29 (n = 28), and greater than 30 years (n = 49), or a split between children and adults according to infection intensity by CAA or by worm fecundity score (explained in further detail in section 2.16.7 and 2.16.10). Number of individuals within each demographic group: Children (n = 70) and adults (n = 88) in moderate intensity village; Children (n = 32) and adults (n = 71) in high intensity villages for age cut-off 10 years old. Children (n = 73) and adults (n = 85) in moderate intensity village; Children (n = 36) and adults (n = 67) in high intensity villages for age cut-off 11 years old.

2.2 Plasma sample processing

For each of the studies, plasma was isolated from whole blood in the field by centrifugation, and stored at -20°C prior to shipment to Cambridge where they were stored at -80°C. Prior to analysis, plasma samples were treated with 0.3% tributyl phosphate/1% Tween 80 (Sigma, Poole, UK) to inactivate encapsulated viruses by disruption of the lipid membrane (Poulsen and Sørensen, 1993).

2.3 Preparation of parasite material

2.3.1 *Schistosoma mansoni* mRNA extraction

RNA extraction was performed aseptically under a laminar flow extraction hood. A frozen sample of eight six-week old *S. mansoni* adult worms (a gift from Mike Doenhoff, University of Nottingham, UK), stored in 1 ml Trizol® (Invitrogen, Carlsbad, CA, USA) was thawed. The sample was homogenized with three 30 sec pulses with a TissueRuptor® (QIAGEN, Hilden, Germany) and allowed to rest for 5 min to allow complete dissociation of nucleoprotein complexes. 200 µl

chloroform (Sigma, Poole, UK) was added and the sample was shaken vigorously, by hand, for 15 sec and allowed to incubate at room temperature for 2 min. Samples were spun at 12,000 x g for 15 min at 4°C. 400 µl of the clear, colourless aqueous phase was transferred to a sterile, RNase-free Eppendorf Tube. An equal volume of 100% ethanol was added and mixed by shaking vigorously by hand. The sample was then applied to a RNeasy Mini column and RNA extracted, according to the manufacturers protocol (QIAGEN, Hilden, Germany). The quality of the mRNA was measured using NanoDrop™ spectrophotometer.

2.3.2 *Schistosoma mansoni* adult worm cDNA

S. mansoni complementary DNA was synthesized from the mRNA extracted in section 2.3.1, using the iScript™ cDNA Synthesis Kit (#1708890, Bio-Rad Laboratories Ltd., Watford, UK), as per the manufacturer's protocol (available at: <http://www.bio-rad.com/webroot/web/pdf/lsr/literature/4106228.pdf>). 1 µl (28.3 ng) mRNA template was used. The 20 µl reaction mix (4 µl 5x iScript Reaction Mix; 1 µl iScript Reverse Transcriptase; 14 µl Nuclease-free water; 1 µl RNA template) was incubated in a DNA Tetrad 2 Thermal Cycler (MJ Research, Reno, NV, USA), according to the following reaction protocol: priming, 5 min at 25°C; reverse transcription (RT), 20 min at 46°C; RT inactivation, 1 min at 95°C; hold at 4°C.

2.3.3 *Schistosoma haematobium* adult worm cDNA

Complementary DNA from adult *S. haematobium*, Egyptian Strain, NR-48642 was provided by the Schistosomiasis Resource Center for distribution by BEI Resources, NIAID, NIH.

To check the quality of the cDNA, reverse transcription (RT)-PCR was performed to amplify the sequences of two *S. haematobium* reference genes that had been successfully amplified from Sh cDNA in the past, tropomyosin (ShTrop; GenBank L76202.1) and tetraspanin 2 (ShTSP2; Sh_TSP2_EC2). Primer sequences were provided by Dr Gabriel Rinaldi (Wellcome Trust Sanger Institute). The following 20 µl reaction master mix was prepared on ice: 10 µl 2X DreamTaq™ Green PCR Master Mix, (#K1081, Thermo Scientific, Hemel Hempstead, UK); 4 µl nuclease-free water; 2 µl *S. haematobium* cDNA template; 2 µl of each of the following forward and reverse primers (10 µM) were then added to the respective ShTSP and ShTrop reaction mixes: ShTrop forward primer (5'- AAG GAG AAT GCA ATG GAA AGA GCA GT-3'); ShTrop reverse primer (5'- AGC TTT CCT CAC GTT GGG CTG A-3'); ShTSP2 forward primer (5'-TCA CGA CGA ACA TGT TAG CAA-3'); ShTSP2 reverse primer (5'-CCT TTT TAA CGC ATC CCT CTT-3'). The following cycling parameters were used in a DNA Tetrad 2 Thermal Cycler (MJ Research, Reno, NV, USA):

step 1: 94°C for 30 sec; step 2: 94°C for 1 min; step 3: 60°C for 1 min; step 4: 72°C for 1 min; steps 1–4 were repeated 35 times; step 5: 72°C for 10 min; hold at 4°C. The PCR product was analysed on a 1.5% agarose gel run at 100 V for 45 min in TBE (Tris/Borate/EDTA) buffer. The gel was imaged using Azure 600 (Azure Biosystems Inc., Dublin, CA, USA).

2.3.4 *Schistosoma haematobium* soluble worm antigen (SWA) preparation

Lyophilised *Schistosoma haematobium* worm antigen (ShSWA) preparation was obtained from Theodor Bilharz Institute, Egypt (#301015). The preparation was received as lipid-free clear supernatant, ultra-filtered under N₂ pressure, lyophilised in double-distilled water and provided in soluble form (Theodor Bilharz Institute, Egypt). The ShSWA preparation was reconstituted with Chaps/Thiourea rehydration buffer (Appendix 1; A1.3). The rehydration buffer was kindly provided by Renata Feret (Department of Biochemistry, University of Cambridge). The sample pH was checked and adjusted to pH 8.5 with 50 mM NaOH, as required. The protein content of the sample was then quantified using Quick Start™ Bradford Kit 1 (#500-0201, Bio-Rad, Watford, UK) before dividing the preparation into 100µg/20µl aliquots, which were stored at -80°C.

2.4 Two-dimensional gel electrophoresis

Two-dimensional electrophoresis separates proteins on the basis of charge and size. In the first dimension, isoelectric focussing (IEF) separates proteins by isoelectric point (pI value), using an immobilised pH gradient strip, to which an electric potential is applied. Separation in the second dimension, by SDS-PAGE, is performed at 90° to the first dimension.

2.4.1 First dimension – isoelectric focusing (IEF)

A 100 µg sample of ShSWA (5 µg/µl) was incubated with 20 µl chaps/urea sample buffer (Appendix 1; A1.3) for 15 min at 22°C. DeStreak™ Rehydration Solution (GE Healthcare, Chalfont St Giles, UK) and 0.5% immobilized pH gradient (IPG) buffer (pH 3–10, GE Healthcare, Chalfont St Giles, UK) was then added to the sample to a total volume of 250 µl. IPG buffer facilitates uniform conductivity along the length of the gel strip, improving protein resolution and the sensitivity of spot detection. The sample was added to a ceramic strip holder which had been thoroughly cleaned with IPGphor™ Strip Holder Cleaning Solution (GE Healthcare, Chalfont St Giles, UK) prior to use, to minimise protein contamination. The sample was then overlaid with a non-linear immobilised pH gradient gel, pH 3–10 (GE Healthcare, Chalfont St Giles, UK) and

covered with mineral oil. Isoelectric focussing included an overnight gel rehydration phase and was conducted as per the following protocol in a Protean il2 Iso-Electric Focusing Cell (Bio-Rad Laboratories Ltd., Watford, UK):

Step 1: rapid 50V for 10 hours

Step 2: rapid 500V for 1 hour

Step 3: rapid 1000V for 1 hour

Step 4: rapid 8000V for 46000 voltage hours

2.4.2 SDS PAGE gel casting

A 13 cm 12% SDS Page gel was prepared, according to the following protocol. Two glass plates were assembled with a 1 mm gel spacer. 30 ml gel solution was prepared (Appendix 1; A1.3). The gel solution was initially made up without TEMED and ammonium persulfate and filtered through a 5 µm filter. Immediately prior to casting, the TEMED and 10% ammonium persulfate was added to the solution. The flask was gently swirled to mix, taking care not to introduce any bubbles to the solution. The gel was poured into the assembled glass cassette to 10 mm below the top. The gel was then overlaid with 80% ethanol and sealed with parafilm, to minimise exposure to oxygen and create a flat gel surface. The gel was left for a minimum of two hours to allow for polymerisation. The ethanol overlay was then replaced with 1X Tris/Glycine/SDS (TGS) running buffer (#1610772, Bio-Rad, Watford, UK) and sealed with parafilm until ready for use.

2.4.3 Second dimension – SDS PAGE

Following IEF, the IPG gel strip was removed from the strip holder and excess mineral oil removed. The gel strip was equilibrated in reducing buffer (Appendix 1, A1.3) for 15 min at 22°C, with agitation, then rinsed with running buffer (1X TGS, Bio-Rad, Watford, UK). A second equilibration step was completed in alkylation buffer (Appendix 1, A1.3), incubating for a further 15 min at 22°C, with agitation. The gel was again rinsed with running buffer. The IPG gel strip was inserted into the gel mould containing the previously cast 12% SDS PAGE gel and overlaid with sealing agarose with bromophenol blue dye (Appendix 1, A1.3). The gel cassette was assembled into Hoefer SE 600 Ruby apparatus and the upper and lower chambers filled with running buffer. The gel was run under the following standard running parameters at 20°C, until the bromophenol blue dye reached the bottom of the gel:

Step 1: 20 mA for 20 min

Step 2: 40 mA for 3 hr

2.4.4 Silver staining

Two-dimensional gels were stained according to the silver staining protocol, selected for its high sensitivity, as silver staining is able to detect less than 1 ng of protein (Weiss, Weiland and Görg, 2009).

The gel was removed from the cassette and stained according to the following mass spectrometry compatible silver staining protocol. The gel was fixed overnight in colloidal Coomassie fix solution (Appendix 1, A1.3) at 4°C, then washed in Milli-Q water for a minimum of two hours, to regain its shape. The gel was then sensitized in 150 ml 0.02% Na₂S₂O₃ for 1 min. Timing of this step is critical, as a longer incubation time will reduce peptide recovery from the gel. Following sensitization, the gel was rinsed for 20 sec in Milli-Q water. The wash step was repeated for a total of three washes. The gel was incubated in 150 ml cold silver staining solution (Appendix 1, A1.3) for 20 min. Using cold staining solution enhances the staining process. Following staining, the gel was rinsed for 20 sec in Milli-Q water. The wash step was repeated for a total of two washes. The gel was incubated in 150 ml development solution (Appendix 1, A1.3) until the protein bands developed. The reaction was stopped by exchanging the development solution for 150 ml 5% acetic acid. The gel was stored at 4°C in 1% acetic acid, until required for mass spectrometry (MS).

2.5 Two-dimensional Western Blot

Two 2D gels were run in parallel to minimise variation in the experimental conditions between the gels used for the respective anti-fecundity case and control Western blots. Following gel electrophoresis, the gel cassette was wrapped in cling film and placed in the cold room at 4°C for 60 min. The semi-dry transfer system (Biometra FastBlot B33) was pre-cooled to 5°C. All equilibration steps were performed using cold 40% methanol, 1X NuPage® transfer buffer (Thermo Scientific, Hemel Hempstead, UK). The gel was removed from the cassette and equilibrated for 10 min at 4°C, the 0.45 µm nitrocellulose membrane (Amersham, UK) was equilibrated for 2 min at 22°C and several sheets of filter paper were pre-soaked in the equilibration buffer. The transfer stack was then assembled, sandwiching the gel and membrane between the filter paper to ensure tight contact between the gel, membrane and the electrode plates. The transfer was performed at a constant 8 Watts for 25 min. The negatively charged proteins migrate from the gel towards the positively charged electrode and are captured by the

nitrocellulose carrier membrane. The membrane was removed from the stack, briefly rinsed with dH₂O and allowed to air dry. The transfer protocol was repeated for the second gel.

To confirm protein transfer efficacy, each membrane was rehydrated with 1X PBS and rinsed with dH₂O, before staining with REVERT™ 700 Total Protein Stain (Li-Cor, Cambridge, UK, 5 ml/membrane) for 5 min. The membranes were then washed twice for 3 min each with REVERT™ Wash Solution (Li-Cor, Cambridge, UK) and rinsed with PBST (0.1% Tween 20, 1x PBS). The membranes were air-dried and imaged at 1.0 intensity using the 700 nm channel (Odyssey 9120, Li-Cor, Cambridge, UK).

The membranes were blocked overnight at 4°C with blocking buffer (1X PBS, 0.1% Tween 20, 5% Marvel). The blocking buffer was discarded, and the membranes were incubated with pooled plasma from anti-fecundity cases or controls, respectively (section 2.15.10). In each case, pooled plasma was diluted 1:50 with blocking buffer (30 ml/membrane) and membranes were incubated for 2 hours at 22°C. Each membrane was washed three times for 10 min with PBST. The membranes were then incubated with mouse anti-human IgG₁ (#A10630, Thermo Scientific, Hemel Hempstead, UK) at 1:400 concentration, diluted in blocking buffer, for 1 hour at 22°C. The membranes were washed three times, for 10 min each, with PBST and incubated with Alexa-Fluor 790-labelled goat anti-mouse IgG (#A28182, Thermo Scientific, Hemel Hempstead, UK) at 1:15,000 concentration, diluted in blocking buffer, for 1 hour at 22°C. The membranes were washed three times, for 10 min each, with PBST, allowed to air-dry and imaged at 5.0 intensity using the 800 nm channel (Odyssey 9120, Li-Cor, Cambridge, UK).

2.6 Identification of spots

Two-dimensional gel mapping software SpotMap (TotalLab, Newcastle, UK) was used to identify protein spots that were differentially present in the case blots compared to the control blots. To create a master spot map, the Western blots were automatically aligned and the 2D gel was manually aligned. Common spots were used as markers to orient the images. Spots were automatically detected on the case and control blot images using the following settings: smoothing = 1 and background removed. Each case or control Western blot image was then overlaid with the master spot map. Presence and absence of spots on the case and control blots was then recorded and compared. Finally, spots were manually curated to ensure false positive spots, due to staining artefact were removed from the analysis.

2.7 Mass spectrometry

2.7.1 Spot cutting

Spots corresponding to those identified as unique to case blots were punched from a duplicate silver-stained gel using aseptic technique. This was performed under a laminar flow cabinet to minimise sample contamination. Trypsin digest was then performed by Renata Feret in the Department of Biochemistry, University of Cambridge, according to the following protocol.

2.7.2 Trypsin digest

The 2D spots were transferred into a 96-well MTP and were cut into 1–2 mm cubes. 50–100 μ l of silver de-stain solution (Appendix 1, A1.5) was added to cover each sample. The de-staining reaction was incubated for 10 min at 37°C. The liquid was removed using a pipette and the process repeated until the gel pieces were completely clear. The gel pieces were washed several times with HPLC water until colourless. The liquid was removed using a pipette and 100 μ l of 100% acetonitrile (ACN) was added and incubated at 37°C until the gel pieces appeared white and shrunken. The ACN was removed and the gel pieces were incubated for a further 10 min at 37°C to ensure the ACN had fully evaporated prior to the addition of DDT. 50 μ l of digest reduction buffer (Appendix 1, A1.5) was added to cover the samples, which were then incubated for one hour at 56°C. The same volume of digest alkylation buffer (50 μ l, Appendix 1, A1.5) was added to the reaction, which was allowed to proceed for 45 min at room temperature (22°C) in the dark. All of the liquid was then removed using a pipette and 100 μ l of 100 mM NH_4HCO_3 was added for 10 min at 37°C. The liquid was removed using a pipette and 100 μ l of 100 mM NH_4HCO_3 in 50 % ACN was added for 10 min at 37°C. The liquid was removed using a pipette and 100 μ l of 100 % ACN was added to dry the gel pieces. When the gel pieces appeared completely white the liquid was removed using a pipette and the gel pieces left for another 10 min at 37°C to ensure that the ACN had completely evaporated prior to the addition of trypsin. 40 μ l of trypsin mixture (Appendix 1, A1.5) was added to each spot and the digestion reaction was allowed to proceed overnight at 37°C. After digestion, the supernatant was pipetted into a sample vial and loaded onto an autosampler for automated Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

2.7.3 Mass spectrometry

LC-MS/MS experiments were performed by Julie Howard-Murkin in the Department of Biochemistry, University of Cambridge using a Dionex Ultimate 3000 RSLC nanoUPLC (Thermo Fisher Scientific Inc., Waltham, MA, USA) system and an Orbitrap Lumos mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Peptides were loaded onto a pre-column (Thermo Scientific PepMap 100 C18, 5 μ m particle size, 100A pore size, 300 μ m i.d. x 5 mm length) from the Ultimate 3000 auto-sampler with solvent A (Appendix 1, A1.6) for 3 min at a flow rate of 10 μ L/min. The column valve was then switched to allow elution of peptides from the pre-column onto the analytical column. Peptides were separated by C18 reverse-phase chromatography at a flow rate of 300 nL/min on a Thermo Scientific reverse-phase nano Easy-spray column (Thermo Scientific PepMap C18, 2 μ m particle size, 100A pore size, 75 μ m i.d. x 50 cm length). The linear gradient employed was 2–40 % solvent B (Appendix 1, A1.6) in 30 min. The total LC run time was 60 min, including the high organic wash step and column re-equilibration.

The eluted peptides from C18 column LC eluant were sprayed into the mass spectrometer by means of an Easy-Spray source (Thermo Fisher Scientific Inc., Waltham, MA, USA). All m/z values of eluting peptide ions were measured in an Orbitrap mass analyser, set at a resolution of 120,000 and were scanned between m/z 380–1500 Da. Data dependent MS/MS scans (3 sec cycle time) were employed to automatically isolate and fragment precursor ions and generate fragment ions by higher energy collisional-induced dissociation (HCD) (Normalised Collision Energy (NCE): 38 %) in the ion routing multipole. The resolution of the Orbitrap was set to 15000 for the measurement of fragment ions. Singly charged ions, ions with greater than seven charges and ions with unassigned charge states were excluded from being selected for MS/MS and a dynamic exclusion window of 70 sec was employed.

2.7.4 Data processing and protein identification

Post-run, all MS/MS data were converted to .mgf files and the files were then submitted to the Mascot search algorithm (Matrix Science, London UK, version 2.6.0) and searched against the *S. haematobium* protein database (version 1; Bioproject PRJNA78265, 13073 sequences; 5666088 residues), the *S. mansoni* database (version 5.2; Bioproject PRJEA36577, 11774 sequences; 5608274 residues) (www.parasite.wormbase.org/), and a common contaminant sequences containing non-specific proteins such as keratins and trypsin (116 sequences; 38548 residues). Variable modifications of oxidation (M) and deamidation (NQ) were applied as well as a fixed modification of carbamidomethyl (C). The peptide and fragment mass tolerances were set to 20

ppm and 0.1 Da, respectively. A significance threshold value of $p < 0.05$ and a peptide cut-off score of 20 were also applied.

The FASTA sequences of hypothetical *S. haematobium* proteins were submitted to protein BLAST searches against *Schistosoma* species (taxid: 6181). Additionally, the search against the *S. mansoni* database was performed to validate the *S. haematobium* results and provide insight into the identity of hypothetical proteins, since the *S. haematobium* genome used at the time of analysis (version 1.0) was comparatively poorly annotated (Stroehlein *et al.*, 2019).

Peptide ranking provides a measure to evaluate the quality of peptide-spectrum matches (PSMs). A positive protein identification was considered when at least two top-ranking peptides were identified.

2.8 Anti-fecundity target bioinformatic analysis

The anti-fecundity cases (section 2.16.10) had higher IgG₁ responses. Therefore, proteins identified as being highly abundant in SWAP were excluded from downstream analysis due to potential identification bias, as anti-fecundity cases may have seen more dead worms resulting in stronger IgG₁ responses.

Proteins were then selected for downstream expression analysis according to the following search criteria, outlined in Table 2.1

Table 2.1 Selection criteria for downstream expression of anti-fecundity targets

	Protein characteristic	Bioinformatic tool
1	The proteins with the highest MASCOT scores, identified from spots with the highest intensity IgG ₁ signal across both technical replicates	SpotMap software (TotalLab)
2	The protein with the highest peptide sequence coverage identified for each spot	
3	The protein is predicted to be excreted/secreted	SignalP, SecretomeP, Young <i>et al.</i> , 2012; Floudas <i>et al.</i> , 2017
4	The protein is identified within <i>S. mansoni</i> EV proteome	Peer-reviewed literature Nowacki <i>et al.</i> , 2015
5	The corresponding gene is annotated with a GO term associated with a reproductive process	Analysis of GO term enrichment using TopGO in R

6	The corresponding gene is identified as differentially expressed in paired/singe male/female or gonads	Peer-reviewed literature, including: Young <i>et al.</i> , 2012; Lu <i>et al.</i> , 2016, 2017
7	Identified as having a role in reduced worm fecundity	Peer-reviewed literature
8	Identified in proteomic analysis of gut proteins	Peer-reviewed literature, including: Hall <i>et al.</i> , 2011; Figueiredo <i>et al.</i> , 2015
9	Protein has been identified within Schistosoma eggshell	Peer-reviewed literature
10	Protein has been found to be localised to the tegument	Peer-reviewed literature, including: van Balkom <i>et al.</i> , 2005; Braschi, Borges and Wilson, 2006

2.8.1 Spot intensity

Spot intensity was calculated by the spot mapping software (SpotMap; TotalLab, Newcastle, UK). The top five spots with the highest intensity IgG₁ signal across the two 2D Western blot technical replicates were identified and, from these, the proteins with the highest MASCOT scores chosen for downstream expression.

2.8.2 Peptide sequence coverage

Percentage peptide coverage was manually calculated as:

(the number of amino acids spanned by the assigned peptides / the sequence length) X 100

2.8.3 Identification of signal peptides and excreted/secreted proteins

In silico analysis was performed to identify surface-expressed or excreted/secreted (ES) proteins. Data-mining software was used to predict signal peptides, non-classical secretion peptide signatures and transmembrane domains. FASTA sequences were initially submitted to signal peptide prediction using the software SignalP-4.1 Server (available at: CBS, cbs.dtu.dk/services/SignalP/) for prediction of secretory signal peptides (SigP score > 0.45). Protein sequences that did not contain a signal peptide were subjected to further analysis for

prediction of non-classical secretion using the SecretomeP 2.0 Server (available at: <http://www.cbs.dtu.dk/services/SecretomeP/>). Proteins that showed a SecP score greater than 0.5 were considered non-classically secreted proteins. Finally, the TMHMM Server v. 2.0 (available at: <http://www.cbs.dtu.dk/services/TMHMM/>) was used to predict the presence of transmembrane (TM) helices in the protein sequence. All programmes were run with default parameters for eukaryote organisms.

2.8.4 Gene Ontology (GO) term enrichment

GO annotations form a common language, produced by The Gene Ontology Consortium that provides a statement about the function of a gene at the molecular level (The Gene Ontology Consortium, 2019). GO terms represent the shared knowledge of biological functions of genes and proteins across eukaryotic organisms, as supported by scientific literature (Ashburner *et al.*, 2000). Three independent categories comprise: biological process, molecular function and cellular component. Analysis of GO terms was performed using the 'TopGO' package in R (Alexa and Rahnenführer, 2019), with a custom script written by Dr. Anna Protasio, University of Cambridge. The analysis was originally performed on the *S. haematobium* anti-fecundity target gene list, then repeated for their *S. mansoni* homologues to account for poor annotation of the *S. haematobium* genome (Stroehlein *et al.*, 2019).

2.8.5 Differential gene expression analysis

Lu *et al.* (2017) extracted RNA from single and paired worms and performed RNA-seq to profile gene expression in *S. mansoni*. The *S. mansoni* proteins that were identified in the parallel MASCOT search of the LC-MS/MS data (section 2.7.4) were then searched for within the fold-change (FC) and false discovery rate (FDR) data from Lu *et al.* (2017). The differential expression data was used to graphically examine relative transcript expression of the genes that encode the identified proteins in male and female adult schistosomes and their respective reproductive organs. Volcano plots were created using the ggplot2 statistical package in R, where the magnitude of the gene expression ratio is displayed on the x-axis (fold difference (\log_2)) and the significance of the difference in expression between groups ($-\log_{10}$ FDR) on the y-axis. Genes with at least two-fold difference (\log_2 ratios ≥ 1 or ≤ -1) and significance value $\text{FDR} \leq 0.05$ (equivalent to $-\log_{10}$ FDR ≥ 1.3) were considered differentially expressed.

Similarly, transcriptomic data from *S. haematobium* male and female adult worms and eggs, published with analysis of the *S. haematobium* genome (Version 1, (Young *et al.*, 2012)) was used

to examine relative expression of proteins identified as possible targets of anti-fecundity immunity.

2.9 Bioinformatic analysis of TAL proteins

2.9.1 Identification of *S. haematobium* TAL proteins

Nucleotide and amino acid sequences from well-characterised *S. mansoni* TAL proteins (Fitzsimmons, Jones, Stearn, *et al.*, 2012) were used in BLAST searches against the *S. haematobium* genome (Version 1.0, WormBase ParaSite, BioProject PRJNA78265). Protein hits that had at least one EF-hand domain at the N-terminal end, a dynein light chain (DLC) domain at the C-terminal end of the protein and had a significance cut-off of $E < 10^{-5}$ were considered possible members of the TAL protein family. The results of this search were cross-referenced against a transcript search performed by Dr Anna Protasio, University of Cambridge (unpublished data).

2.9.2 Phylogenetic analysis of *S. haematobium* TAL proteins

A phylogenetic analysis of all 13 SmTAL proteins, together with *S. haematobium* proteins identified as possible members of the TAL protein family was conducted, using several representatives of the EF-hand calcium ion binding protein calmodulin from *Schistosoma* and related species as an out-group.

Firstly, alignment of multiple TAL sequences was performed using Multiple Sequence Comparison by Log-Expectation (MUSCLE) (Edgar, 2004). Prior to phylogenetic reconstruction, profiling of conserved domains was performed in Gblocks (version 0.91, (Castresana, 2000)).

Gblocks analysis defines a set of conserved blocks, according to a defined set of thresholds that: determine the degree of conservation; identify stretches of contiguous non-conserved regions; ensure highly conserved positions are flanking the remaining blocks; retain blocks of a minimum sequence length, and remove positions with gaps in the sequence (Castresana, 2000). Further detail regarding the Gblocks method can be found in Castresana (Castresana, 2000). Amino acid sequences tend to be conserved to maintain the structure or function of a protein or specific protein domain. Fewer amino acid replacements occur within the conserved region and when substitutions do occur, they are more likely to be amino acids with similar biochemical properties. In the analysis presented here the Gblocks selection parameters were adjusted for the least stringent block selection settings, allowing for smaller block sizes with a gap in less

than 50% of the sequences and less strict flanking positions (minimum number of sequences for a conserved position: 17; minimum number of sequences for a flanking position: 17).

This process is known as alignment masking, where unreliable sections of the alignment are identified and removed before phylogenetic analysis. This practice has been shown to improve phylogenetic alignment (Talavera and Castresana, 2007) and is routinely recommended to improve tree reconstructions (Kück *et al.*, 2010). Alignments were then manually edited in Jalview (version 2.10.3b1). Phylogenetic trees were constructed in MEGA7 (version 7.0.26 (7170509-i386)) using a maximum likelihood approach, with bootstrapping analysis over 500 iterations and a cut-off of 80% of sequences non-missing amino acids.

2.10 Cloning and expression of recombinant TAL proteins

Recombinant proteins were cloned and expressed in an *E. coli* expression system. For TAL proteins (Chapter 4), the DH5 α *E. coli* strain was used in for cloning due to its high transformation efficiency and ability to stably maintain foreign DNA. The TG2 *E. coli* strain was used for expression (Maizels *et al.*, 1991).

2.10.1 Recombinant SmTAL11 expression construct

Expansion of SmTAL11 by Polymerase Chain Reaction

The plasmids for SmTAL1, SmTAL3, SmTAL5 and ShTAL1, ShTAL3, ShTAL5 and ShTAL11 had previously been cloned and expressed by the Cambridge Schistosomiasis Research Group and were available as plasmids or transformed *E. coli* glycerol stocks, stored at -80°C (Dickinson, 2012; Fitzsimmons, Jones, Stearn, *et al.*, 2012), however SmTAL11 had not yet been successfully cloned or expressed in our laboratory.

The following 20 μ l PCR reaction mix was used to amplify the SmTAL11 coding DNA sequence (CDS), with SmTAL2 (Smp_086480) positive control and non-template negative control: 9.9 μ l sterile water; 4 μ l 5X HF buffer; 0.4 μ l dNTPs; 2 μ l forward primer (10 μ M) (SmTAL11: 5'-ATG GAT CCA TTT TTA CAT GCA TTT-3'; SmTAL2: 5'-AAC TTT ACT TCG TGA TGG TGA TGA-3'); 2 μ l reverse primer (10 μ M) (SmTAL11: 5'-TTA ACA CAA TTG ATA TTG ACC AGG TGT-3'; SmTAL2: 5'-ACC ACG ACA AAA TTC TTC AAA AGT-3'); 2 μ l cDNA (28.3 ng/ μ l); 0.2 μ l Phusion® High-Fidelity DNA Polymerase (Thermo Scientific, Hemel Hempstead, UK). The following cycling parameters were used in a DNA Tetrad 2 Thermal Cycler (MJ Research, Reno, NV, USA): step 1: 98°C for 30 sec; step 2: 98°C for 10 sec; step 3: 69°C for 20 sec; step 4: 72°C for 30 sec; steps 1–4 were repeated for a total of 35 cycles; step 5: 72°C for 5 min; hold at 4°C.

The PCR product was analysed on a 1.5% agarose gel run at 100 V for 45 min. The gel was imaged using Azure 600 (Azure Biosystems Inc., Dublin, CA, USA) and the DNA band gel purified using QIAquick® Gel Extraction Kit protocol (QIAGEN, Hilden, Germany).

Ligation into cloning vector

The purified PCR product was then ligated into pJET1.2/blunt linearized cloning vector (Fig. 2.1) using the blunt end cloning protocol (#K1231, Thermo Scientific, Hemel Hempstead, UK). The following 20 µl ligation mix was made up: 10 µl 2X reaction buffer; 1 µl 0.15 pmol ends purified PCR product (34.6 ng/µl); 1 µl (0.05 pmol ends) pJET1.2/blunt Cloning Vector (50 ng/µl); 7 µl nuclease-free water; 1 µl T4 DNA ligase. The ligation mixture was briefly vortexed, spun for 3–5 sec in a mini centrifuge, then incubated at room temperature (22°C) for 5 min.

Transformation of DH5α *E.coli* cloning strain

A 50 µl aliquot of DH5α competent *E. coli* cells (#C2987H, NEB, Hitchin, UK) was transformed with 5 µl of ligation mixture, using the manufacturer's High Efficiency Transformation Protocol (NEB, Hitchin, UK, available at: <https://international.neb.com/protocols/0001/01/01/high-efficiency-transformation-protocol-c2987>). A parallel control transformation was performed with 0.1 ng of supercoiled vector DNA (pUC19 DNA, #SD0061, Thermo Scientific, Hemel Hempstead, UK) to confirm the transformation efficiency. Transformed DH5α cells were spread on an AMP selection agar plate using a sterile, glass hockey stick spreader and resulting colonies were analysed by PCR.

The pJET1.2/blunt cloning vector carries both an AMP resistance gene and a lethal restriction enzyme gene (Fig. 2.1). When DNA is successfully ligated into the cloning site, the lethal gene is disrupted, therefore only bacterial cells that contain recombinant plasmids are able to form colonies. *E. coli* cells that are transformed with re-circularised vector, without an insert, are killed by the resulting lethal restriction enzyme.

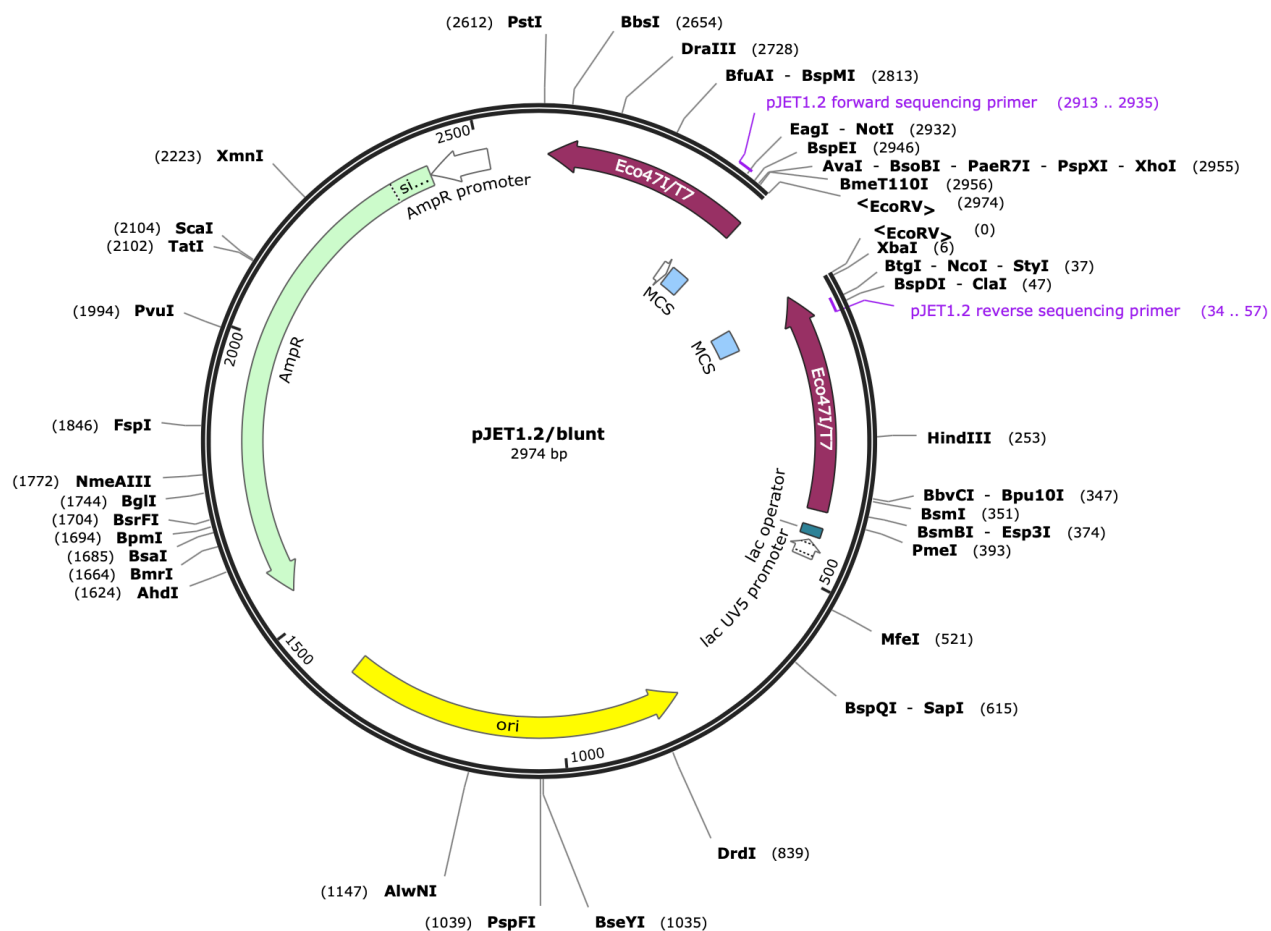


Fig. 2.1 pJET1.2/blunt plasmid map (Thermo Scientific, Hemel Hempstead, UK)

Analysis of recombinant clones

Five colonies were screened for the presence and orientation of the SmTALII DNA insert. The following PCR master mix for five reactions was prepared on ice: 2.4 µl pJET1.2 Forward Sequencing Primer (10 µM) (5'-d(CGA CTC ACT ATA GGG AGA GCG GC)-3'); 2.4 µl pJET1.2 Reverse Sequencing Primer (10 µM) (5'-d(AAG AAC ATC GAT TTT CCA TGG CAG)-3'); 46.2 µl nuclease-free water; 60 µl 2X DreamTaq™ Green PCR Master Mix, (#K108I, Thermo Scientific, Hemel Hempstead, UK). A small amount of each colony to be screened was picked using a sterile pipette tip and resuspended in 20 µl of the PCR master mix. The following cycling parameters were used: 95°C for 3 min; 94°C for 30 sec; 60°C for 30 sec; 72°C for 1 min (1 min/kb). The PCR products were analysed on an 1.5% agarose gel run at 100 V for 45 min.

Isolation of plasmid DNA by mini-prep

2ml mini cultures (2X YT, 100 ug/ml AMP) were inoculated with the same five colonies as those analysed by PCR, in 14 ml snap-cap falcon tubes (#352059, Corning). Cultures were incubated overnight at 37°C with vigorous shaking. Sterile glycerol was added to make 50% glycerol stocks from clones that showed a band of the correct size. Plasmid DNA was isolated and purified from the DH5α using the QIAprep® Spin Miniprep Kit high yield protocol (QIAGEN, Hilden, Germany; available at: <https://www.qiagen.com/gb/resources/resourcedetail?id=ebb38a38-c61c-40a7-910d-9b10d1148022&lang=en>) and eluted into nuclease-free water. Plasmid DNA was then sent for sequencing at the DNA Sequencing Facility (Department of Biochemistry, Cambridge). Plasmids were sequenced with pGEX 5' forward (5'-GGG CTG GCA AGC CAC GTT TGG TG-3') and 3' reverse (5'- CCG GGA GCT GCA TGT GTC AGA GG-3') primers. Results were analysed using SnapGene Viewer (version 4.1.0).

Restriction digest

The following digest mix was prepared: 2 µl 10X FastDigest™ Green Buffer; 1 µl XbaI FastDigest™ restriction enzyme; 1 µl XhoI FastDigest™ restriction enzyme (all ThermoFisher Scientific, Hemel Hempstead, UK); 2 µl pGEX-KG plasmid DNA, or 2 µl SmTAL11-pJET1.2 cloning construct. The reaction was incubated at 37°C for 5 min.

Each digest reaction was loaded directly onto an electrophoresis gel to confirm successful digest by the presence of a band of the appropriate size. The bands were gel purified using QIAquick® Gel Extraction Kit protocol (QIAGEN, Hilden, Germany).

Ligation into expression construct

Recombinant TAL proteins were expressed in *E. coli* as GST fusion proteins. The SmTAL11 DNA fragment was then ligated into the pGEX-KG expression vector (Fig. 2.2), the same vector used for the other TAL proteins of interest.

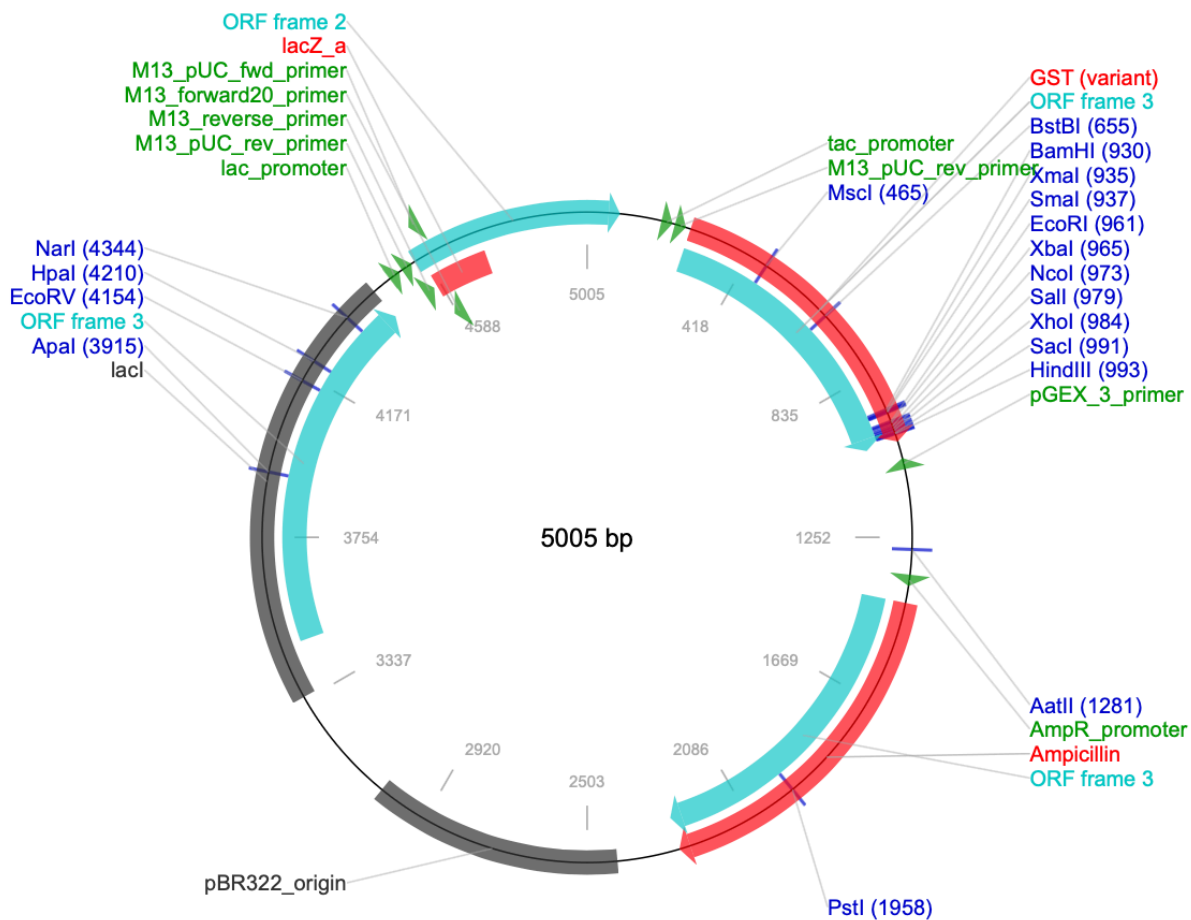


Fig. 2.2 pGEX-KG plasmid map (<https://www.addgene.org/vector-database/2890/>)

2.10.2 Calcium chloride transformation of TG2 *E.coli* with TAL-pGEX-KG plasmid

The same transformation and expression protocols were used for each of the *S. mansoni* and *S. haematobium* TAL proteins.

Sterile 2x yeast extract-tryptone (YT) broth and ampicillin (AMP) negative 1.5% agar plates were prepared. 2x YT was selected since the increased amount of yeast and tryptone extracts in 2x YT compared to LB medium provide additional growth factors and nutrients, enabling the bacteriophage to rapidly multiply without depleting the host cell resources (Green and Sambrook, 2012).

A 1.5% agar/2x YT plate was streaked with TG2 *E. coli* and incubated overnight at 37°C (Multitron Standard, INFORS HT, Basel, Switzerland). Three 10 ml aliquots of sterile 2x YT were

inoculated with three individual *E. coli* colonies. Cultures were incubated overnight at 37°C with shaking at 180 RPM (Multitron Standard, INFORS HT, Basel, Switzerland), to aerate the culture. Each culture was diluted to 100 ml with sterile 2X YT and incubated at 37°C with shaking until an OD₆₀₀ of 0.4–0.5 was reached (Cecil CE 2041 Spectrophotometer 2000 Series). The cultures were then divided between two 50 ml conical Falcon™ tubes and centrifuged at 4700 x g for 10 min. The supernatant was discarded, each cell pellet was resuspended in 25 ml 0.1 M MgCl₂ (Sigma, Poole, UK) and, again, centrifuged at 4700 x g for 10 min. This was repeated with 25 ml ice-cold 0.1 M CaCl₂ (Sigma, Poole, UK). The supernatant was discarded, and the cell pellets resuspended in 2 ml ice-cold 0.1 M CaCl₂. The cell suspension was divided into 250 µl aliquots and incubated on ice for 30 min. The optimal concentration of plasmid was selected by adding either 0 ng, 3 ng, 10 ng, and 30 ng TAL-pGEX-KG plasmid (Fig. 2.2) to four separate 250 µl aliquots, which were then incubated on ice for 45 min, heat shocked at 42°C in a water bath for 2 min, and returned to ice. 800 µl S.O.C. recovery media (#15544034, ThermoFisher Scientific, Hemel Hempstead, UK) was added to obtain maximal transformation efficiency, and the cells were incubated at 37°C for 1 hour with vigorous shaking. 75 µl of each culture was spread on a separate 1.5% agar/2xYT ampicillin selection plate, under aseptic technique. Plates were incubated overnight at 37°C.

Following transformation, five colonies were selected from the 10 ng TAL pGEX *E. coli* plate and cultured overnight at 37°C, 180 RPM (Multitron Standard, INFORS HT, Basel, Switzerland), in 2X YT medium (+ 100 µg/ml AMP) to an OD₆₀₀ of 0.6, indicating that bacterial growth is in mid-log phase. The colonies grown on the 10 ng plasmid plate were selected since they resulted in an even distribution of distinct colonies. Sterile glycerol was added to make 50% glycerol stock cultures that were stored at -70°C, until required.

2.10.3 Expression of proteins in *E.coli* bacterial expression system

The glycerol stock of each TAL transformed *E. coli* was removed from -80°C and put straight on ice. A 1.5% agar/2X YT (+ 100 µg/ml AMP) selection plate was streaked with the transformed *E. coli*, using aseptic technique. The plate was incubated overnight at 37°C. Four sterile 50 ml 2X YT aliquots (+ 100 µg/ml AMP) were inoculated with colonies from the plate. The cultures were grown overnight at 37°C, shaking at 180 RPM. Cultures were expanded to 200 ml (sterile 2X YT, 100 µg/ml AMP) and incubated at 37°C, shaking at 180 RPM until an OD₆₀₀ of 0.6 was reached. The cultures were once again expanded to 600 ml (sterile 2X YT, 100 µg/ml AMP) and monitored until an OD₆₀₀ between 0.6 and 0.8 was reached. The cells were then induced by the

addition of isopropyl β -D-1-thiogalactopyranoside (IPTG; from 1000X stock) and incubated according to the conditions outlined in Table 2.2.

The bacteria were recovered by centrifugation (Beckmann Coulter, JLA 10.5 rotor) 5000 RPM for 10 min at 18°C. The cell pellet was then frozen at -80°C overnight, to aid bacterial lysis. The *E. coli* pellet was resuspended in a total volume of 45 ml of ice-cold PBS/cOmplete™ protease inhibitor cocktail (1 tablet/50 ml; #11697498001, Roche Diagnostics, Mannheim, Germany) before freezing.

Table 2.2 Recombinant TAL protein expression induction conditions

Recombinant protein	IPTG concentration	Temp.	Time
SmTAL1, SmTAL3, SmTAL5, ShTAL3, ShTAL5, ShTAL11	100 μ g/ml	37°C	2 hr
SmTAL11, ShTAL1	10 μ g/ml	37°C	2 hr

2.10.4 Purification of TAL proteins from *E. coli*

The bacterial concentrate was thawed and the NiCo21 pellet resuspended in NPI-10/ cOmplete™ protease inhibitor cocktail (1 tablet/50 ml; #11697498001, Roche Diagnostics, Mannheim, Germany). The cell suspension, in both cases, was passed through a hydraulic cell disrupter (OS Cell Disrupter, Constant Systems Ltd., Daventry, UK) and the lysate was ultracentrifuged at 40,000 x g, 4°C (Avanti JXN-26, JLA 8.1 rotor, Beckmann Coulter, High Wycombe, UK) for 45 min.

For isolation of GST-tagged TAL proteins, a 5 ml glutathione sepharose column was poured (Glutathione Sepharose® 4B beads, 17-0756-01, GE Healthcare, Chicago, IL, USA) and equilibrated by passing 5–10 column volumes (CVs) of ice-cold PBS at 0.5 ml/min. The bacterial supernatant was then passed through a 0.45 μ m filter and loaded onto the column at 0.5 ml/min. The column was washed with 10 CVs of ice-cold PBS at 0.5 ml/min. The GST-tagged TAL protein (fusion protein) was eluted from the column with 10 mM reduced glutathione (Sigma, Poole, UK) in 50mM Tris/HCl pH 8.0, collecting 20x 1 ml fractions.

A 15 μ l aliquot of each fraction was retained for SDS PAGE. Peak fractions were identified by measuring absorbance at 280 nm using NanoDrop™ spectrophotometer. Assuming an OD of 1.0 = 1 mg/ml, TAL-GST fusion protein was pooled into aliquots equivalent to 5 mg in 2 ml and frozen at -80°C. A 7.5 μ l aliquot was retained for 1D SDS PAGE.

Cleavage of GST tag from TAL protein by thrombin digest

A 2 ml aliquot of TAL-GST fusion protein (5 mg in 2 ml) was thawed and 0.6 µl Tween 20 (Sigma, Poole, UK) added to give a final concentration of 0.03 %. The fusion protein was added to a 500-unit vial of thrombin (GE Healthcare, Chalfont St Giles, UK, #27-0846-01) and incubated for 2 hours at room temperature on a horizontal shaker, to cleave the GST tag. A 15 µl aliquot was removed for 1D SDS-PAGE.

Purification of TAL from thrombin digest

Following cleavage, free GST was removed by incubating each thrombin digest with 500 µl Q-Sepharose anion exchange beads (GE Healthcare, Chicago, IL, USA), equilibrated with 10 mM reduced glutathione in 50 mM Tris/HCl pH 8.0 for 45 min at 4°C. After each bead incubation step, the sample was recovered by centrifugation at 1000 RPM for 30 sec and a 15 µl aliquot removed for 1D SDS PAGE. The sample was buffer-exchanged into PBS using a PD-10 Desalting column (#17085101, GE Healthcare, Chalfont St Giles, UK), following the manufacturer's spin protocol (ref), then incubated with 500 µl glutathione-sepharose beads (GE Healthcare, Chicago, IL, USA), equilibrated with PBS, for 30 min at 4°C, to remove any remaining free GST. Finally, contaminating thrombin was removed from the sample by adding 250 µl aminobenzamidine-agarose beads (Sigma, Poole, UK), equilibrated with PBS, for 30 min at 4°C. The final protein content of the sample was measured using the NanoDrop™ spectrophotometer, read at A280 nm.

2.11 One-dimensional gel electrophoresis

2.11.1 Gel electrophoresis

One-dimensional gel electrophoresis separates proteins on the basis of size. Proteins are initially reduced to their primary structure and denatured using a reducing agent. SDS then applies a uniform negative charge to each protein, ensuring that the speed at which proteins migrate through the gel is on the basis on size alone.

A sample of each protein (individual protein concentrations are indicated in the relevant figures in Chapter 4 and Chapter 5 results) was reduced to its primary structure by adding 1:4 concentration of NuPAGE® lithium dodecyl sulfate (LDS) Sample Buffer (4X) (Thermo Scientific, Hemel Hempstead, UK) with β-mercaptoethanol (ME) (Appendix 1, A1.2) and heating to 80°C

for 10 min in a heat block. The denatured protein samples were loaded onto a NuPAGE® Novex® 4–12% Bis-Tris pre-cast gel (Thermo Scientific, Hemel Hempstead, UK). The gel cassette was assembled into Bio-Rad mini chamber with 1X NuPAGE® MES [2-(N-morpholino) ethane sulfonic acid] SDS Running Buffer (Thermo Scientific, Hemel Hempstead, UK) and run at 100 Volts for 1 hour.

2.11.2 Coomassie brilliant blue staining

One-dimensional gels were stained according to the Coomassie brilliant blue staining protocol, as the higher protein concentration of single protein bands meant that the increased sensitivity of silver staining was not required. The gel was removed from the cassette, trimmed and rinsed with dH₂O. The gel was then stained with Coomassie brilliant blue stain (Appendix 1, A1.2) for 15 min at 22°C, with gentle agitation on a horizontal shaker. The staining solution was removed, and the gel briefly washed with de-stain solution (Appendix 1, A1.2), before being left to de-stain overnight at 22°C, with gentle agitation.

2.12 One-dimensional Western Blot

A one-dimensional gel was run, as per 2.11.1. The gel was removed from the cast and equilibrated in transfer buffer (Appendix 1, A1.4) and transferred at 30 Volts for one hour onto 0.45 µm nitrocellulose membrane (Amersham BioSciences, Little Chalfont, UK). The membrane was blocked overnight at 4°C with 5% marvel in ELISA wash buffer (Appendix 1, A1.9). Anti-HIS monoclonal AF488 antibody (Thermo Scientific, Hemel Hempstead, UK) was incubated at 1:500 concentration in ELISA incubation buffer (Appendix 1, A1.9) for 1 hour at room temperature (20°C). The membrane was washed and imaged on Azure Biosystems RGB, auto-capture.

2.13 High-throughput (HTP) cloning and expression screening of anti-fecundity targets

HTP cloning and expression screening was completed at the Oxford Protein Production Facility UK (OPPF-UK), according to the following OP PF Standard Protocols (available at: <https://www.oppf.rc-harwell.ac.uk/OPPF/protocols/>).

2.13.1 PCR Primer design

Primers were designed for In-Fusion[™] cloning by Ray Owens at OPPF-UK, using the automated primer design tool (available at: <https://www.oppf.rc-harwell.ac.uk/Opiner/>). The following sequences were added to the PCR primers to enable the In-Fusion[™] cloning step, complementary to the pOPINF vector (Fig. 2.3) for recombination (Berrow, Alderton and Owens, 2009):

Forward cloning tag: AAGTTCTGTTTCAGGGCCCG

Reverse cloning tag: ATGGTCTAGAAAGCTTTA

Full primer sequences can be found in Appendix 2.

2.13.2 HTP PCR reactions

Anti-fecundity immunity targets were cloned in a similar manner to that described for SmTALII in section 2.10.1, with differing reaction conditions and cycling parameters outlined below. 100 µM master stocks were made up with nuclease free water, then further diluted with nuclease free water to make a working stock of 10 µM.

KOD Xtreme[™] Hot Start DNA Polymerase (Novagen 71975-3) master mix was prepared on ice, according to the following reaction mix (for one 50 µl reaction, master mix volume 44 µl) and scaled to the number of reactions required: 25 µl 2X KOD Hot Start Buffer; 10 µl dNTP mix (2 mM); 1 µl KOD Hot Start DNA Polymerase (1 U/µl); 6 µl sterile water; 2 µl template DNA. 44 µl of master mix was added to 3 µl forward primer and 3 µl reverse primer for each reaction. Plates were then sealed with a foil seal and loaded into a Ventri[™] PCR machine (ABI). The following cycling parameters were used: step 1: 94°C for 2 min; step 2: 98°C for 10 sec; step 3: 60°C for 30 sec; step 4: 68°C for 1 min 40 sec; repeat steps 2–4 29 times; step 5: 68°C for 2 min; hold at 4°C.

Cloning was attempted with alternative conditions for the sequences that were not successfully cloned using KOD Xtreme[™] Hot Start DNA Polymerase with the above conditions. Alternative conditions included using KOD Xtreme[™] Hot Start DNA Polymerase, with an annealing temperature of 55°C, or alternative DNA polymerases (Phusion[®] High-Fidelity DNA Polymerase (Thermo Scientific, Hemel Hempstead, UK) or DreamTaq[™] Green PCR Master Mix, (#K1081, Thermo Scientific, Hemel Hempstead, UK)), using the following cycling parameters:

Phusion[®] reaction mix: 9.9 µl sterile water; 4 µl 5X HF buffer; 0.4 µl dNTPs; 2 µl forward primer (10 µM); 2 µl reverse primer (10 µM); 2 µl cDNA (28.3 ng/µl); 0.2 µl Phusion[®]. Cycling conditions: step 1: 98°C for 30 sec; step 2: 98°C for 10 sec; step 3: 69°C for 20 sec; step 4: 72°C for 30 sec; steps 1–4 were repeated for a total of 35 cycles; step 5: 72°C for 5 min; hold at 4°C.

DreamTaq™ reaction mix: 10 µl 2X DreamTaq™; 4 µl nuclease-free water; 2 µl cDNA template; 2 µl forward primer (10 µM); 2 µl reverse primer (10 µM). Cycling conditions: step 1: 94°C for 30 sec; step 2: 94°C for 1 min; step 3: 55°C for 1 min; step 4: 72°C for 1 min; steps 1–4 were repeated 35 times; step 5: 72°C for 10 min; hold at 4°C.

2.13.3 Analysis of PCR products

Following thermal cycling, 5 µl of each PCR product was added to 2 µl of 5X DNA loading buffer (Thermo Scientific, Hemel Hempstead, UK; #R0611) in a 96-well PCR plate, before the products were loaded onto a 1.6 % TBE agarose gel with SYBR Safe stain (Invitrogen, Carlsbad, CA, USA) a column at a time. 5 µl of Hyperladder™ 1kb was used as a reference. The gel was run at 100 V for approximately 40 min, until the bromophenol blue reached the end of the lane.

2.13.4 AMPure XP magnetic bead purification

Good quality PCR products, as assessed by a single clean band on the PCR gel (section 2.13.3), were purified by AMPure magnetic bead-based purification. The AMPure XP bottle was gently shaken to resuspend the magnetic beads. 90 µl AMPure was added to each PCR product to be purified. The bead/PCR reaction was mixed thoroughly by pipetting up and down 10 times and left to incubate for 3–5 min at room temperature. This step binds PCR products of 100 base pairs and larger to the magnetic beads. The plate was placed onto a SPRIPlate 96R magnet (Beckman Coulter, Brea, CA, USA) for 5 min to separate the beads from solution. Once the solution had cleared, it was aspirated from each well with a pipette, taking care not to disturb the bead pellet. With the plate still on the magnet, 200 µl 70% ethanol was added to each well and, without mixing, left to incubate for 30 sec at room temperature. The ethanol was aspirated and discarded. This was repeated for a total of two washes, taking care not to disturb the pelleted beads. The plate was allowed to air-dry for 10 min to allow complete evaporation of any residual ethanol. The plate was removed from the magnet and 30 µl elution buffer (EB, 10 mM Tris pH 8.0) added to each well of the plate and mixed by pipetting up and down 10 times. The plate was placed back onto the magnet and 30 µl of the supernatant was transferred into a fresh PCR plate. 5 µl purified PCR product was added to 2 µl DNA loading buffer and run on a gel as per section 2.13.3 to confirm that the purification process has worked. The concentration of each purified product was measured using NanoDrop™ 8000 Spectrophotometer. One PCR product was gel purified, following the QIAquick® Gel Extraction Kit protocol (QIAGEN, Hilden, Germany).

2.13.5 In-Fusion reactions and HTP transformation into cloning-grade *E. coli*

The following reaction mix was set up for each purified PCR product, using reagents from ClonExpress II One Step Cloning Kit: 0.5 µl (50 ng) linearised pOPINF vector; 3 µl (10–250 ng, on average) PCR product; 1 µl 5X buffer; 5 µl sterile water; 0.5 µl enzyme, and incubated for 30 min at 37°C. pOPINF was a gift from Ray Owens (Addgene plasmid # 26042 ; <http://n2t.net/addgene:26042> ; RRID: Addgene_26042, Fig. 2.3). Transformation was then performed, as follows, for In-Fusion™ HD EcoDry™ Cloning Plus (Takara, Saint-Germain-en-Laye, France). Stellar™ Competent Cells (#636767, Takara, Saint-Germain-en-Laye, France) were thawed on ice for 15 min. 2 µl of each reaction was added to a separate 20 µl aliquot of competent cells. The cells were incubated on ice for 30 min before being heat-shocked for 30 sec at 42°C in a water bath. The cells were returned to ice for 2 min and 120 µl Power Broth (PB) was added, without additional mixing. The cells were incubated for 1 hour at 37°C. Meanwhile, a LB Agar 24-well plate were prepared (Appendix 1, A1.8). Two dilutions of transformed cells were plated per reaction (25 µl transformed cell culture and a 1:5 dilution of culture in PB). The plate was incubated overnight at 37°C.

2.13.6 Colony picking and HTP culture

Several replicate white clones (4–8) were picked per construct, using 200 µl pipette tips to inoculate 1.2 ml PB in the wells of a deep-well block (BD Biosciences, Wokingham, UK). The plate was sealed with a gas-permeable seal (Abgene Inc., Portsmouth, NH, USA) and incubated overnight at 37°C, shaking vigorously at 600 RPM in a floor standing Vertiga incubator (Glas-Col, LLC, Terre Haute, IN, USA). A glycerol stock of each culture was prepared by adding 100 µl culture to 100 µl filter-sterilised LB/30%v/v glycerol. Stocks were stored at -80°C. The deep-well block was resealed with a solid seal (Abgene Inc., Portsmouth, NH, USA) and cells harvested by centrifugation at 6000 x g for 15 min (Beckman JS5.3 rotor, Beckman Avanti). The supernatant broth was discarded.

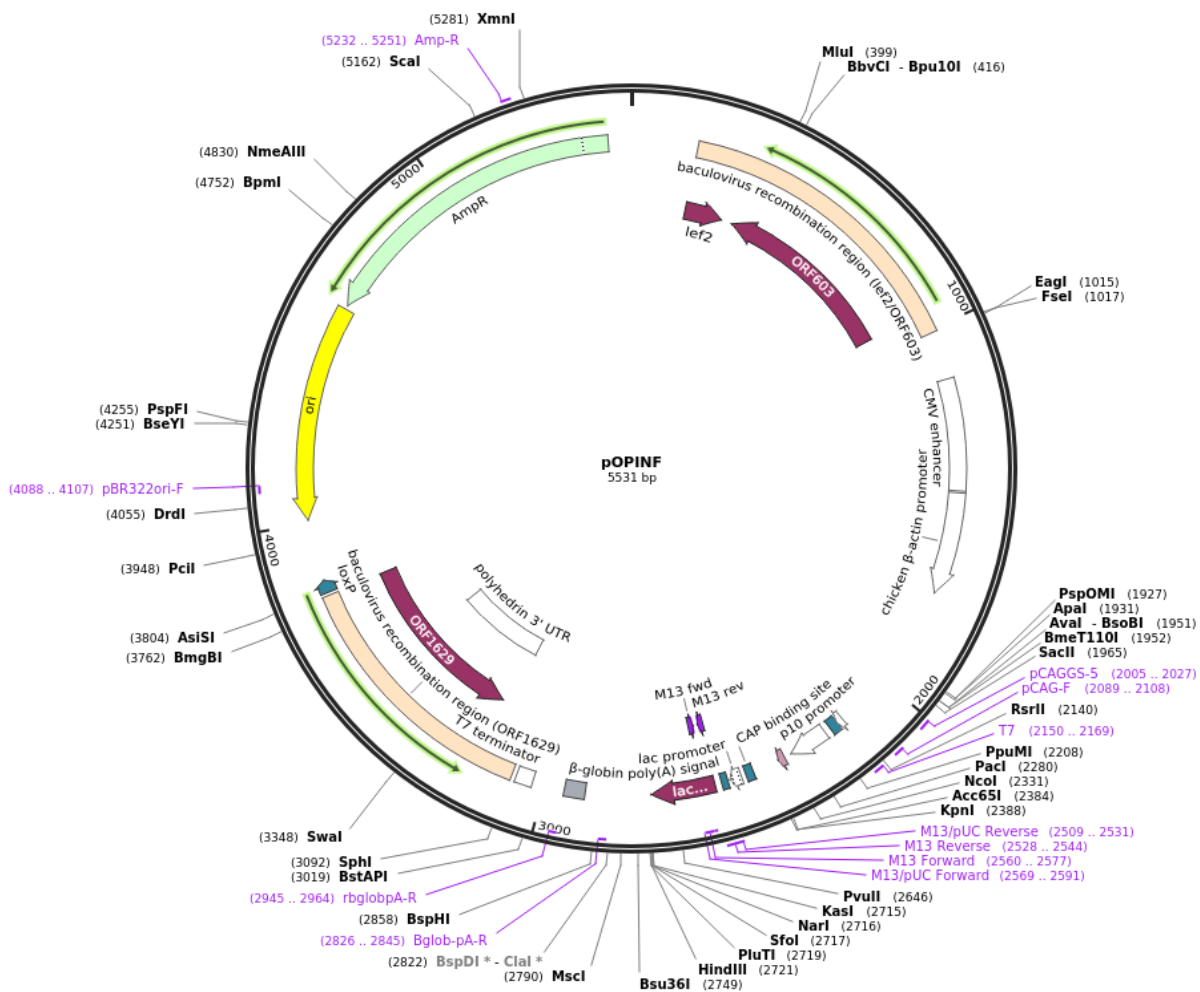


Fig. 2.3. Map of pOPINF plasmid, with restriction site, His3C-lacZ tag site and ampicillin resistance region indicated (Berrow et al., 2009). pOPINF was a gift from Ray Owens (Oxford Protein Production Facility, Harwell, UK) (Addgene plasmid # 26042 ; <http://n2t.net/addgene:26042> ; RRID:Addgene_26042).

2.13.7 Construct verification

Plasmid mini-preps were performed on the Bio-Robot 8000 (QIAGEN, Hilden, Germany), as per the manufacturer's protocol (available at: <https://www.promega.co.uk/-/media/files/resources/protocols/technical-bulletins/0/wizard-sv-96-plasmid-dna-purification-system-protocol.pdf?la=en>), using a Wizard SV 96 DNA purification plate (Promega). Plasmid constructs were then screened by PCR to verify product size. The following master reaction mix was prepared, for one 96-well plate (total reaction volume 25 µl/well): 1275 µl 2X Phusion® High-Fidelity DNA Polymerase (Thermo Scientific, Hemel Hempstead, UK); 954 µl sterile water; 15.3

µl pOPIN forward primer at 100 µM.(5'-GAC CGA AAT TAA TAC GAC TCA CTA TAG GG-3'). The following reactions were then set up on ice: 22 µl master mix; 1.5 µl 10 µM reverse primer (Appendix 2); 1.5 µl appropriate construct plasmid (from section 2.13.5). The plate was sealed with a foil seal and loaded into a Ventri™ PCR machine (Arkansas Biosciences Institute, Little Rock, AR, USA). The following cycling parameters were used: step 1: 98°C for 10 sec; step 2: 98°C for 1 sec; step 3: 60°C for 5 sec; step 4: 72°C for 2 min; repeat steps 2–4 29 times; step 5: 72°C for 2 min; hold at 4°C. A 6.5 µl aliquot of PCR product was then analysed as described in section 2.13.3.

2.13.8 Small-scale expression screening in *E.coli*

For possible anti-fecundity target proteins (Chapter 5), *E. coli* HST08 strain Stellar™ competent cells were used for cloning targets due to its high transformation efficiency. *E. coli* Lemo21 strain was used for protein expression. Lemo21 *E. coli* are derived from BL21(DE3) cells, have a high transformation efficiency and contain the Lemo System™, which allows expression of difficult clones to be tuned to improve protein solubility and folding. This is a useful feature when using high throughput techniques.

The proteins were screened for successful expression in *E. coli*. Lemo21(DE3). Strain cells were thawed on ice for 15 min before heat-shock transformation with 3 µl construct plasmid at 42°C in a water bath for 2 min. The cells were returned to ice and 300 µl PB recovery media added (no selection antibiotic).

Agar plates were prepared as described in section 2.13.5, time including carbenicillin selection (Cb) (50 µg/ml) and chloramphenicol (CHL) (35 µg/ml) and without blue/white screening. Chloramphenicol was added to ensure the Lemo21 plasmid was retained. Antibiotic stocks were diluted 1:1000 from stock solutions in warm agar. 30 µl cells were plated per well and the plate incubated overnight at 37°C, as per section 2.13.5.

The next day, individual colonies were picked using 200 µl pipette tips to inoculate 700 µl PB, supplemented with the appropriate antibiotics [Cb 50 µg/ml; CHL 35 µg/ml]. The plate was sealed with a gas-permeable adhesive seal and incubated overnight at 37°C, shaking vigorously at 600 RPM (Vertiga incubator, Glas-Col, LLC, Terre Haute, IN, USA). 150 µl overnight culture was added to 3 ml Overnight Express™ Instant Terrific Broth (TBONEX, Novagen, Appendix 1, A1.8) [+Cb, +CHL] in a 24-well deep well plate and incubated for a further 3–5 hours at 37°C, with vigorous shaking at 600 RPM (OD₅₉₅ approximately 0.5). The temperature was reduced to 25°C and incubation allowed to continue for a further 20–24 hours.

TBONEX is an autoinduction culture medium for protein production in IPTG-inducible bacterial expression systems. The use of this media enables regulated protein expression without the need to monitor the culture or manually add IPTG during cell growth. This is beneficial in an HTP system since different protein targets may need inducing at different times and for different durations (Grabski, Mehler and Drott, 2005).

The next day, 1 ml of each overnight culture was transferred to a 96-well deep well block and cells were harvested at 6000 x g for 10 min (Beckman JS5.3 rotor, Beckman Avanti). The supernatant media was discarded, and the plate of pelleted cells sealed and frozen for a minimum of 20 min at -80°C, to aid cell lysis. Pellets were allowed to defrost at room temperature for 30 min before being resuspended in 210 µl lysis buffer (NPI-10-Tween, Appendix 1, A1.8), supplemented with 1 mg/ml Lysozyme and 400 Kunitz units/ml DNase Type I (Appendix 1, A1.8). The suspension was allowed to incubate at room temperature for 30 min before the lysate was cleared by centrifuging at 6000 x g for 30 min at 4°C (Beckman JS5.3 rotor, Beckman Avanti). The supernatant was transferred into 20 µl Ni-NTA bead suspension, aliquoted in a flat-bottomed plate (Greiner 655101). Care was taken not to disturb the insoluble pellet, which was saved for potential future analysis. The suspension was mixed for 30 min on a horizontal shaker at room temperature. The 96-well plate was then placed on a 96-well magnet (QIAGEN, Hilden, Germany) for 1 min and the supernatant carefully removed by pipette. The beads were washed twice with 200 µl wash buffer (NPI-20, Appendix 1, A1.8), removing the plate from the magnet to mix on the horizontal shaker for 5 min during each wash. After the second wash step, the plate was placed back on the magnet, the wash buffer discarded and 50 µl elution buffer (NPI-250, Appendix 1, A1.8) added to each well. The plate was mixed on a horizontal shaker for 1 min, then placed back on the magnet for 1 min. The supernatant was then transferred to a fresh 96-well plate for analysis by SDS-PAGE (section 2.13.9).

2.13.9 SDS-PAGE analysis of Ni²⁺-NTA purified proteins

10 µl of each purified protein target was mixed with 10 µl SDS-PAGE loading buffer (Appendix 1, A1.8) in a 96-well plate. The plate was sealed, and the samples boiled for 3 min in a water bath. The samples were loaded into the gel (10 µl/well) with 5 µl low molecular weight (MW) marker (SigmaMarker™ low range, MW 6.5-66 kDa, Sigma, Poole, UK) in well 1 and 5 µl wide MW marker (SigmaMarker™ wide range, MW 6.5-200 kDa, Sigma, Poole, UK) in the final well of the gel. The gel was run at a constant 200 Volts until the blue dye reached the bottom of the gel (approximately 40 min). The gel was stained with Instant Blue (ISB01L, Expedeon, Cambridge, UK) for 45 min before de-staining by replacing the stain with dH₂O until de-stained, replacing the water as necessary.

2.14 Scaled-up expression of recombinant anti-fecundity target proteins

2.14.1 Transformation of NiCo21 *E. coli* with anti-fecundity target protein OPINF plasmid

Poly-histidine tagged (referred to as His-tagged) proteins are often contaminated with significant amounts of endogenous *E. coli* metal binding proteins. NiCo21 *E. coli* (#C2529H, New England Biolabs, Ipswich, MA, USA), derived from BL21(DE3) cells, were therefore selected for the expression of the His-tagged anti-fecundity target proteins, as NiCo21 *E. coli* have been shown to minimise contamination of immobilised metal affinity chromatography (IMAC) fractions (Robichon *et al.*, 2011).

NiCo21 (DE3) Competent *E. coli* cells were transformed with plasmids selected from the 96-well plate produced at OPPF-UK (section 2.13.5), according to the manufacturer's High Efficiency Transformation Protocol (#C2529H, New England Biolabs, Ipswich, MA, USA). Briefly, cells were defrosted on ice for 10 min, then incubated with 3 µl plasmid for 30 min. The cells were then heat-shocked for 10 sec at 42°C and returned to ice for 5 min. 950 µl of S.O.C. (#15544034, Thermo Fisher Scientific, Hemel Hempstead, UK) was added, the cells were incubated at 37°C for 1 hour, with vigorous shaking, and 75 µl of each culture was spread on a separate 1.5% agar/2xYT ampicillin selection plate, under aseptic technique. Plates were incubated overnight at 37°C. Sterile glycerol was added to make 50% glycerol stocks of transformed bacteria.

2.14.2 Protein expression and purification

Bacterial cultures were expanded, induced with 100 µg/ml IPTG at 30°C for 4 hours and the bacteria recovered as described in section 2.10.3. Directly following cell recovery by centrifugation, the NiCo21 pellet was frozen at -80°C. The frozen bacterial concentrate was later thawed and the NiCo21 pellet resuspended in NPI-10/ cOmplete™ protease inhibitor cocktail (1 tablet/50 ml; #11697498001, Roche, Basel, Switzerland). The cells were lysed as described in section 2.10.4.

For isolation of His-tagged anti-fecundity target proteins, proteins were isolated by IMAC on a 1 ml Ni-NTA Superflow Cartridge (#30721, QIAGEN, Hilden, Germany), stored at 4°C. The column was equilibrated with 5 CVs of ice-cold dH₂O, followed by a further 5 CVs of ice-cold NPI-10. The bacterial suspension was passed through a 0.45 µm filter and loaded onto the column at 0.5 ml/min. The column was washed with 10 CVs of ice-cold NPI-20 at 0.5 ml/min, followed by a further two 5 CV washes with ice-cold NPI-75. The flow-through from each step

was collected for downstream analysis. The His-tagged anti-fecundity target protein was eluted from the column with NPI-150 elution buffer, collecting 20x 1 ml fractions.

A 15 µl aliquot of each fraction was retained for SDS PAGE. Peak fractions were identified by measuring absorbance at 280 nm using NanoDrop™ spectrophotometer. A 7.5 µl aliquot was retained for 1D SDS PAGE. His-tagged proteins were stored at 4°C with no further downstream purification required.

2.15 Enzyme-linked immunosorbent assay (ELISA)

2.15.1 Coating titration and contamination assay

A titration assay was performed for each antigen to determine the appropriate coating concentration. This ensured saturated coating of the well surface was achieved, whilst minimising the concentration of sample contaminants. Each antigen was diluted at a two-fold decreasing concentration, starting at 1:10 dilution, in either ELISA coating buffer (for TAL proteins) or NPI-150 (for anti-fecundity target proteins), see Appendix 1, A1.9 for buffer composition. High-binding 96-well half-area MTPs were then coated with 25 µl of diluted antigen, according to Figure 2.4, and either sealed with a foil seal and incubated overnight at 4°C (for wet-coated proteins), or incubated overnight at 37°C in the presence of calcium chloride beads (for dry-coated proteins). The coating conditions for each recombinant protein are outlined in Table 2.3. Plates were washed four times. All wash steps were carried out with ELISA wash buffer, unless otherwise stated, using an automated plate washer (ELx405, BioTek Instruments, Vermont, USA). Next, 25 µl of either anti-rat horseradish peroxidase (HRP)-conjugated antibody (#A9037, Sigma, Poole, UK) (Fig. 2.4 A–B) at 1:2000 dilution in ELISA blocking buffer or blocking buffer only (Fig. 2.4 C–H) was added to each well and the plate incubated for 2 hours at room temperature, shaking on a horizontal shaker. The plates were washed six times and 25 µl of incubation buffer only was added to rows A–B (Fig. 2.4); rabbit anti-bacterial lysate sera (Cambridge Schistosomiasis Research Group) at 1:500 dilution was added to rows C–D (Fig. 2.4), and, for recombinant TAL antigens only, rabbit anti-GST (Sigma-Aldrich, Poole, UK) at 1:1000 dilution to rows E–F (Fig. 2.4), both diluted in incubation buffer. Rat anti-SmTAL serum was available for SmTAL1, SmTAL3 and SmTAL5 only (a gift from Dr Jamal Kalife, Institut Pasteur de Lille to Cambridge Schistosomiasis Research Group) added to rows G–H (Fig. 2.4), diluted in incubation buffer at 1:500. The plates were incubated for one hour at room temperature, shaking, then washed six times. The plates were developed by adding 25 µl of goat anti-rabbit HRP (source) at 1:1000 dilution in incubation buffer to wells in rows C–H (Fig. 2.4) and 25 µl of incubation buffer only to wells in the anti-rat HRP coating inhibition

wells (Fig. 2.4 A–B). The plates were incubated for one hour at room temperature, shaking. 100 µl o-Phenylenediamine dihydrochloride (OPD) substrate solution (Appendix 1, A1.9) was then added to each well and the plates incubated in the dark for 5–20 min. The development reaction was stopped by the addition of 25 µl 2M H₂SO₄ to each well. Absorbance was read at dual wavelength: test wavelength 490nm, reference wavelength 630nm (EL312e, BioTek Instruments, Winooski, VT, USA). The same protocol was followed for His-tagged anti-fecundity target proteins, only measuring coating concentration and bacterial contamination.

	1	2	3	4	5	6	7	8	9	10	11	12
A	1/10	1/20	1/40	1/80	1/160	1/320	1/640	1/1280	1/2560	1/5120	Bl	Bl
B	1/10	1/20	1/40	1/80	1/160	1/320	1/640	1/1280	1/2560	1/5120	Bl	Bl
C	1/10	1/20	1/40	1/80	1/160	1/320	1/640	1/1280	1/2560	1/5120	Bl	Bl
D	1/10	1/20	1/40	1/80	1/160	1/320	1/640	1/1280	1/2560	1/5120	Bl	Bl
E	1/10	1/20	1/40	1/80	1/160	1/320	1/640	1/1280	1/2560	1/5120	Bl	Bl
F	1/10	1/20	1/40	1/80	1/160	1/320	1/640	1/1280	1/2560	1/5120	Bl	Bl
G	1/10	1/20	1/40	1/80	1/160	1/320	1/640	1/1280	1/2560	1/5120	Bl	Bl
H	1/10	1/20	1/40	1/80	1/160	1/320	1/640	1/1280	1/2560	1/5120	Bl	Bl

Fig. 2.4 Diagram of microtitre layout for coating and contamination ELISA. Figure shows serial coating dilution of recombinant protein. Bl: blank wells. Colours represent the different coating or contamination conditions, as follows: anti-rat HRP coating inhibition (blue); rabbit anti-bacterial lysate serum (orange); rabbit anti-GST serum (red); rat anti-SmTALx serum, where 'x' represents either TAL1, TAL3 or TAL5.

2.15.2 Measurement of antigen-specific antibody levels

High-binding 384-well microtiter 600 MTPs (Greiner bio-one Ltd Stonehouse, UK) were washed once with dH₂O using the automated plate washer. Antigen was diluted to an optimum concentration in ELISA coating buffer (Appendix 1, A1.9), as determined by a coating titration assay (section 2.15.1). Plates were then coated with 15 µl per well of diluted antigen and incubated as described in section 2.15.1 for wet- and dry-coated antigens (Table 2.3). The first two columns of each plate were coated with purified human IgG₁ (Human IgG₁ myeloma protein, Sigma, Poole, UK; #15154), IgG₄ (Human IgG₄ myeloma protein, Sigma, Poole, UK; #14639) or IgE (Human IgE myeloma protein, Sigma, Poole, UK; #401152) starting at 30 µg/ml with three-fold dilutions to form a 14-point standard curve. Coating buffer only was added to the final two wells as a blank. The plates were sealed with a foil seal and incubated overnight at 4°C. Standard curves were coated on the same day for wet-coated antigens.

Table 2.3 Recombinant antigen coating concentration and condition

Antigen	Stock concentration (mg/ml)	Coating dilution*	Coating condition
SmTAL1	0.87	1:100	Wet
SmTAL3	0.70	1:100	Wet
SmTAL5	0.66	1:100	Wet
SmTALI1	0.21	1:40	Dry
ShTAL1	0.42	1:80	Dry
ShTAL3	0.48	1:80	Wet
ShTAL5	0.36	1:100	Dry
ShTALI1	0.57	1:100	Wet
ShLAP	0.76	1:160	Dry
ShTK	3.24	1:640	Wet
ShUDP	0.68	1:100	Dry

*recombinant TAL proteins diluted in coating buffer; recombinant anti-fecundity target proteins diluted in NPI 150 buffer.

The plates were washed four times, blocked with 80 µl ELISA blocking buffer per well and incubated for one hour at room temperature. The blocking buffer was then discarded and 15 µl of human sera from the relevant study site (Uganda (Musoli) or Mali), at 1:400 (IgG₁), 1:200 (IgG₄) or 1:20 (IgE) dilution in 10% fetal calf serum (FCS) in incubation buffer, was added per well. A panel of 12 non-infected non-endemic sera (NES) controls was added and the plates were sealed and incubated overnight at 4°C. The addition of FCS minimises non-specific binding. The plates were washed six times and 15 µl of monoclonal mouse anti-human IgG₁ biotin (#555869, BD Pharmingen, Wokingham, UK), monoclonal mouse anti-human IgG₄ biotin (#555882, BD Pharmingen, Wokingham, UK) or monoclonal mouse anti-human IgE biotin (#555858, BD Pharmingen, Wokingham, UK), was added per well at 0.5 µg/ml, diluted in incubation buffer. The plates were incubated at room temperature for 2 hours, protected from light. The plates were washed six times, developed with 15 µl of poly-HRP streptavidin complex (#M2032, Mast Group, Merseyside, UK) per well, at 1:3000 dilution in incubation buffer, and incubated at room temperature for 1 hour, protected from light. The plates were washed six times. 68 µl of OPD substrate solution was added per well and the plates were protected from light and allowed to develop for 5–30 min, until the colorimetric reaction could be seen in the lowest concentrations of the standard curve. The development reaction was then stopped by the addition of 17 µl of 2M sulphuric acid per well. The addition of OPD substrate and acid to each plate was timed to ensure each assay plate had the same development time. Plates were then read at 490/630nm,

as before. To adjust for plate edge effects due to temperature and humidity variations, an in-house VBA Macro was run to adjust optical density (OD) values, accordingly. Antibody concentrations were then extrapolated from the standard curves using the adjusted OD values (GEN5 Software, BioTek Instruments, Vermont, USA).

2.15.3 Other antigen-specific antibody levels

Historical data from the Cambridge Schistosomiasis Research Group were used to explore the patterns of specific antibody responses against *S. mansoni* DLC and EF hand proteins, and *S. haematobium* TAL-IgE for analysis of reinfection immunity targets (Chapter 4).

2.15.4 Reciprocal inhibition ELISA

In order to study cross-absorption between the TAL proteins, plasma from ten individuals who had an IgE antibody response to TAL1, TAL3, TAL5 and TAL11 was used in a competitive inhibition ELISA. The assay was run in parallel for *S. mansoni* and *S. haematobium*.

Following the methods outlined in sections 2.15.1 and 2.15.2, 5–6 µg of each recombinant TAL antigen was coated to 18 individual wells (in triplicate for each plasma pre-absorption condition: TAL1, TAL3, TAL5, TAL11, positive control and negative control) of a high-binding 96-well half-area MTP. A 14-point IgE standard curve (Human IgE myeloma protein, Sigma, Poole, UK; #401152), starting at 30 ng/ml concentration was also coated to each plate. 150 µg/ml of TAL1, TAL3, TAL5 or TAL11 was then added to diluted pooled plasma (1:20 in ELISA incubation buffer) and incubated for 1 hour at room temperature, on an orbital shaker, before being added to the coated plates. The positive control plasma did not contain any competitor TAL protein and the negative control was incubation buffer only. The IgE ELISA then continued as described in section 2.15.2.

2.16 Statistical analysis

2.16.1 Transformation of data

All statistical analysis was carried out in R (version 3.4.3) (R Core Team, 2015). The Shapiro-Wilk test for normality ($\alpha = 0.05$) was applied to all data sets. Skewed data were log-transformed prior to analysis by adding a value of 1 to egg counts and a value of 0.03 to antibody titres to account for zero values.

2.16.2 Definition of the geometric mean

The geometric mean was calculated as the exponent of the mean log-transformed values, minus the value added to account for zero values. This is sometimes referred to as the Williams mean (Williams, 1937; Alexander, 2012).

2.16.3 Analysis of reciprocal inhibition ELISA results for TAL protein cross-reactivity

T-tests were performed to determine the significance of the observed differences in specific IgE binding to solid-phase TAL proteins, following pre-incubation of plasma with competitor TAL proteins. P-values were adjusted using Simes-modified Bonferroni correction for multiple testing (Simes, 1986).

2.16.4 Calculation of post-treatment boosts in antibody responses

The observed post-treatment boosts in specific antibody responses were calculated as:

$$100 * (\exp\{\text{mean}[\ln(\text{post-treatment OD} + 0.03) - \ln(\text{pre-treatment OD} + 0.03)]\} - 1)$$

Paired t-tests were used to test the significance of boosts in TAL specific IgE in Chapter 4 and anti-fecundity target specific IgG₁ in Chapter 5.

2.16.5 Definition of positive antibody responders

To determine individuals with positive antibody responses, antibody levels of 12 uninfected, non-endemic serum (NES) controls were measured. A threshold antibody response was then set as the mean concentration of NES + (3 x SD).

Fisher's exact test was performed to test the significance of the association between the proportion of responders with overlapping specific IgE responses to multiple TAL proteins in the *S. mansoni* (Musoli Uganda) and *S. haematobium* (Segou Region Mali) infected cohorts.

2.16.6 Analysis of *S. mansoni* TAL11 specific IgE responses

The Musoli cohort was divided into five age groups of approximately equal size, defined in section 2.1.2. Kruskal-Wallis test was conducted to examine the differences in SmTAL11 specific pre- and post-treatment IgE antibody responses by age group. *Post hoc* pairwise comparisons were then performed using the Conover-Iman test (Conover and Iman, 1979). The Conover-Iman test is thought to be more powerful than the alternative Dunn's (1964) *post hoc* multiple

comparisons test and is valid when the corresponding Kruskal-Wallis null hypothesis is rejected (Dinno, 2017).

Univariate linear regression models were constructed to explore the relationship between IgE responses and reinfection intensity two years post-treatment. Models were subsequently adjusted to account for age group and nine-week post-treatment egg count, as a marker for clearance of baseline infection following treatment. Having a positive IgE response to multiple members of the TAL protein family (SmTAL1, SmTAL3 and SmTAL5) has previously been associated with reduced reinfection following treatment (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). To explore whether SmTAL1 specific IgE seropositivity also contributes to this relationship, the association between being a quadruple IgE responder (IgE seropositivity to SmTAL1, 3, 5 and 11) and intensity of reinfection at two years was examined in a multivariable model accounting for age group and nine-week egg count. Pairwise interaction terms between the independent variables were included in both multivariable models. Models were sequentially updated with the removal of non-significant terms.

2.16.7 Association between *S. haematobium* TAL specific IgE responses and baseline CAA

As described in section 2.1.3, the three villages that comprise the total Malian cohort can be separated into two groupings demonstrating different scales of infection: moderate to high infection intensity and very high transmission intensity, referred to as moderate intensity and high intensity, respectively, from here on. Village of residence was therefore treated as a proxy for the force of transmission, in the regression models constructed in Chapter 4. The total cohort was divided into five age groups of approximately equal size, defined in section 2.1.3. To describe the epidemiology of ShTAL-specific IgE responders, multivariable models exploring the effect of age group, sex and village of residence on specific IgE seroprevalence to individual TAL proteins were initially constructed.

The association between ShTAL responses and worm burden was then analysed using univariate cross-sectional linear models of pre-treatment ShTAL responses and baseline CAA in the total cohort. Age group, sex, village of residence and TAL-specific IgG₄ for each recombinant TAL protein were then accounted for in multivariable models, with all possible two-way interaction terms built into the multivariable models. All TAL-specific IgG₄ responses were included in the model to explore the possible blocking influence of IgG₄ on the IgE response. Non-significant variables were subsequently removed from the model.

Fitzsimons *et al.* (2012) found that a subset of individuals within the *S. mansoni*-infected Musoli cohort in Uganda had a positive IgE response to several member of the TAL protein family and that being an IgE responder to multiple TAL proteins was significantly associated with reduced reinfection. To explore whether the same, or a similar, relationship stands in an *S. haematobium*-infected cohort, associations between seropositivity for multiple TAL proteins, as an IgE double responder (ShTAL1 and ShTAL3), triple responder (ShTAL1, ShTAL3 and ShTAL5) or quadruple responder (ShTAL1, ShTAL3, ShTAL5 and ShTAL11) were subsequently explored in the multivariable models, with log(CAA+1) as an outcome variable. The same models were subsequently run independently for high and moderate transmission villages.

2.16.7 Piecewise regression analysis to determine ‘children’ and ‘adults’ age groups

Small sample sizes in the regression models for the high and moderate transmission villages dictated that age was reduced to a binary variable according to an age cut-off of 10 years with ‘children’ aged under 10 years and ‘adults’ aged 10 years and older for these models, rather than the five age groups described above. To define ‘children’ and ‘adults’ for the village-level models, piecewise regression analysis was performed to determine the breakpoint of the CAA age profile using data for the total Malian cohort. Breaks at host ages between 7 and 16 years were explored. Breaking the regression at 10 years returned the smallest residual standard error (Fig. 2.5).

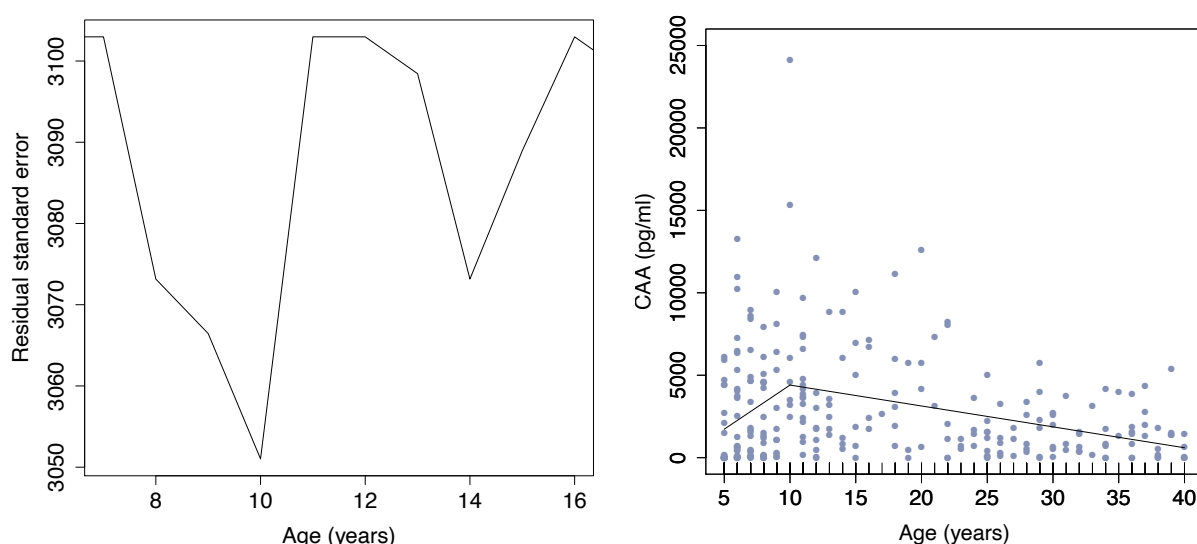


Fig. 2.5. A) The standard error of residuals from piecewise regression models of CAA with breaks at host ages between 7 and 16 years, and B) CAA by host age (Malian cohort). Scatterplot of CAA (picograms of circulating anodic antigen per ml of plasma) versus host age with the fitted regression lines for individuals aged <10 years and ≥10 years.

2.16.9 Association between *S. haematobium* TAL specific IgE responses and two-year reinfection status

Multivariable logistic regression models were constructed to explore the association between post-treatment TAL specific IgE responses and reinfection status two years post-treatment. Reinfection was modelled as a binary response variable to maintain statistical power, since reinfection levels were low. Age, sex, village of residence and TAL-specific IgG₄ responses were accounted for in the models. The models were run for the total cohort, in which age was accounted for by age group, and separately for the moderate and high intensity villages, in which age was accounted for as a binary independent variable, again due to sample size. All TAL-specific IgG₄ responses were included in the model to explore the possible blocking influence of IgG₄ on the IgE response. Non-significant variables were subsequently removed from the model.

2.16.10 Identification of the anti-fecundity case cohort and matched controls

Individual worm fecundity scores were calculated as:

$$[(\text{number of eggs excreted per 10 mL of urine} + 1) / (\text{pg of CAA per mL of plasma} + 1)]$$

Plasma from 10 individuals that had been identified as having reduced egg-to-worm ratios (worm fecundity scores) and plasma from 10 control individuals, matched for age, sex and village of residence was used to detect *S. haematobium* proteins associated with a possible anti-fecundity immune response. Case and control individuals were identified in prior analysis by Dr Shona Wilson (University of Cambridge), according to the following criteria. Anti-fecundity controls were defined as having a worm fecundity score greater than the mean + 2 x standard deviations (SD) of the mean adult fecundity score (individual points above the dashed red line in Figure 2.6), where the adult fecundity score was calculated from individuals aged over 11 years. Anti-fecundity cases, with fecundity scores below the dashed red line in Figure 2.6, were matched for age, sex and village of residence, against the anti-fecundity controls.

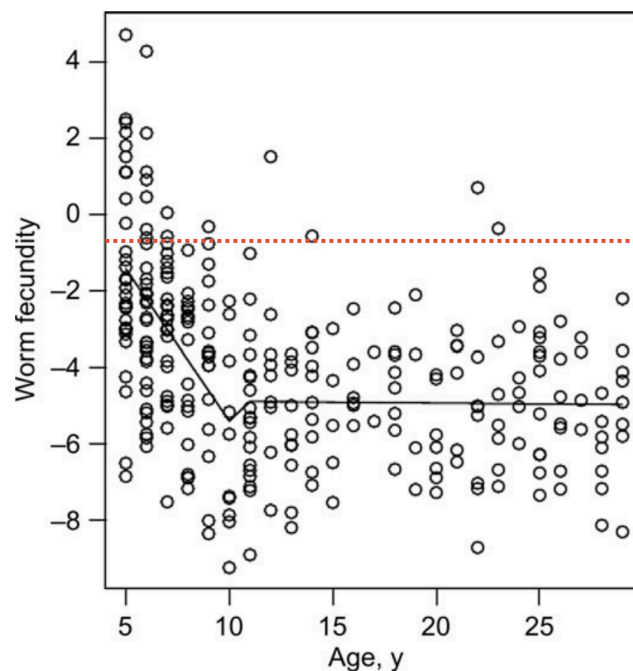


Fig. 2.6. Worm fecundity with host age (Malian cohort). Scatterplot of worm fecundity, versus host age and the fitted regression lines for individuals aged <11 years and ≥11 (taken from Wilson *et al.* (2014), open access). Dashed red line indicates the mean + 2 x SD of the adult fecundity score (individuals aged ≥11): fecundity score cut-off: -0.48.

2.16.11 Association between IgG₁ responses to *S. haematobium* anti-fecundity target proteins and worm fecundity

The effects of age and village, as a proxy for the force of transmission, on anti-fecundity target specific IgG₁ responses were explored using multiple linear regression analysis, with log-transformed IgG₁ titres as the dependent variable and age group, sex and transmission intensity as explanatory variables. Pairwise interaction terms between age and village were included and models were updated in a stepwise manner, sequentially removing non-significant terms. Age groups were also analysed according to the breakpoint at 11 years of age, as determined previously by piecewise regression analysis (Wilson *et al.*, 2014).

Univariable linear regression models were constructed to explore the relationship between IgG₁ seropositivity to recombinant anti-fecundity target proteins.

Data were imputed for *S. mansoni* infection status using multiple imputation using chained equations (MICE) using predictive mean matching over 25 imputations. Host age, village (as a proxy for transmission intensity), sex, $\ln(\text{caa} + 1)$, $\ln(\text{Sh eggs} + 1)$, $\ln(\text{SWA-IgG}_1)$, $\ln(\text{SWA-IgG}_4)$, $\ln(\text{SWA-IgE})$ and $\ln(\text{SWA-IgA})$, $\ln(\text{Sh28GST-IgG}_1)$, $\ln(\text{Sh28GST-IgG}_3)$, $\ln(\text{Sh28GST-IgA})$,

detectable ShTAL1-IgG₁, and IgG₄ and IgE responses. MICE was performed using the ‘mice’ statistical package (van Buuren and Groothuis-Oudshoorn, 2011) in R programming language (R version 3.5.1; (R Core Team, 2015)).

2.17 Mathematical modelling

Mathematical model was constructed and run in R (version 3.5.1; (R Core Team, 2015)). A description of the methods used to define the model parameters, the Bayesian parameter fitting process and the specific statistical packages used can be found in Chapter 3.

3

Modelling anti-infection immunity in an *S. mansoni* transmission model

Preface

The effects of exposure on *Schistosoma* transmission often confound the effects of immunity. This chapter argues that there is a pressing need to model epidemiological patterns of both exposure and immunity, address some of the research requirements raised a review of *The Interdependence between Schistosome Transmission and Protective Immunity* published during my PhD (Oettle and Wilson 2017) (available in Appendix 6).

Acknowledgement of contributions

Professor Maria-Gloria Basañez (Imperial College London) and Dr Martin Walker (Royal Veterinary College, London) provided invaluable expert discussion on parameter definitions and fitting of functions to the data. Dr Martin Walker and Charlie Whitaker (Imperial College London) provided advice and assistance with the model coding, specifically the specific sex structure code and the adaptive MCMC model fitting code, respectively. Cercarial exposure scores and IgE serology data were available from historical analysis by Pinot de Moira *et al.* (Pinot De Moira *et al.*, 2007; Pinot de Moira *et al.*, 2010)

3.1 Introduction

3.1.1 General introduction

Mathematical models that reproduce epidemiological patterns of infection are instrumental to policy-making due to their value in planning and evaluating infection control measures. Models are however designed to provide a representation of real-life scenarios and, by their nature, are unable to incorporate the biological, environmental and behavioural variations that control transmission in their entirety. Instead, judgements must be made regarding which of these variations are most significant or have the greatest potential influence. The accuracy of the

projections made by models is therefore limited by the biological assumptions and data that underpin model calibration. For analysis to be meaningful it is of great importance to select an appropriate model structure and infer model parameters from measured data. Models are therefore, in turn, dependent on the availability and accuracy of observed field data.

As discussed in Chapter 1, observed *Schistosoma* infection profiles demonstrate a characteristically peaked profile with age, with highest infection intensities seen in adolescence. For snail to human transmission to take place, two fundamental processes are required: human contact with the infectious stage of the parasite (cercariae) and establishment of infection, given this contact occurring. It is believed that the lower infection intensity that is generally observed in adults results from the aggregated contribution of these two processes, but there is little agreement in the literature regarding the relative influence of exposure and immunity.

The model presented here will explore the contribution of these two processes in determining the profile of the age-intensity and age-serology curves in a population endemic for schistosomiasis. A comprehensive review of *Schistosoma* models has been performed elsewhere (Anderson *et al.*, 2016); here I discuss salient features of previous models that are relevant to this thesis, with particular reference to exposure and immunity.

3.2 Overview of the principal schistosome transmission models

3.2.1 Basic model structures

One of the earliest epidemiological models of schistosome transmission was the mean worm burden (MWB) model, in which a series of differential equations were used to describe changes in the mean burden of schistosomes within the definitive human host (Macdonald, 1965). The MWB model structure was informed by earlier models that describe the transmission of soil transmitted helminths and remains the basis of a large proportion of schistosome modelling literature today.

Many of the early versions of transmission models that explore schistosome transmission between snails and humans (Hairston, 1965; Macdonald, 1965), are based on simple biological concepts and lack inclusion of the individual heterogeneity that may arise from variations in age, sex, behaviour or susceptibility (Woolhouse *et al.*, 1998; Pinot De Moira *et al.*, 2010). It is believed that this lack of heterogeneity may lead to a failure to sufficiently capture the characteristic overdispersion of parasite infections (Chandiwana and Woolhouse, 1991; Woolhouse, 1991).

Barbour (1978) subsequently introduced individual and spatial heterogeneity into the Macdonald model, to better represent the epidemiological patterns seen in parasitology data, finding that the greater the heterogeneity of water contact, the greater the value of R_0 and the higher transmission. Barbour (1978) concluded that such heterogeneity may therefore significantly affect transmission patterns, a finding that later models tend to agree with (Woolhouse, 1991).

Age is a significant correlate of *Schistosoma* infection intensity and is also associated with both patterns of exposure and the development of protective immunity (Naus *et al.*, 2003). Yet, whilst both Barbour (1978) and, later, Woolhouse (1991) explore heterogeneity in transmission, neither consider the age structure of the host population directly. Anderson & May (1985), on the other hand, modify the Macdonald model to include age structure through partial differential equations, to capture the over-dispersed parasite distribution. Simultaneously, the authors balance the added model complexity resulting from addition of explicit age-structure and assumptions of age-dependent exposure, with removal of the intermediate lifecycle stages. The justification for this is the relatively short survival times of intermediate stages, compared to adult worms within the human host (Maldonado and Acosta-Matienzo, 1948; Lawson and Wilson, 1980). Subsequent MWB models continued to model schistosome infection dynamics with respect mature worms only.

Later, Chan *et al.* (1995) developed a model framework for schistosome transmission based on MWB model of soil transmitted helminths introducing full age structure to the model and different infection and contamination rates. This model was named EpiSchisto® and further iterations of the EpiSchisto® model have since been developed to explore various concepts including acquired immunity (Chan *et al.*, 1996), morbidity (Chan, Guyati and Kingdom, 1996) and control (Chan and Bundy, 1997).

The most recent iteration of the MWB comes from the NTD modelling consortium at Imperial College London (ICL). The ICL model has an age-structured deterministic partial differential framework, with age-intensity profiles generated by age-dependent exposure to cercariae and does not include the potential contribution of acquired immunity.

The Anderson and May (1985) MWB model structure has dominated schistosome modelling literature for several decades and the ICL model is today one of the foremost schistosome transmission models. In the last decade, however, an alternative model structure was independently developed by researchers at the Case Western Reserve University (CWRU). The CWRU model was developed from alternative theory of stratified worm burden (SWB), based on the characteristic overdispersion of schistosomes within the host population (Gurarie, King and Wang, 2010). The SWB model, calibrated with community level *S. haematobium* prevalence

data, explores the transmission dynamics within a network of communities and water contact, or transmission, sites and explicitly models the snail transmission cycle that had been dropped from more recent forms of the MWB model.

As models introduce greater biological complexity, they risk becoming inaccessible to field biologists due to the intricacy of the mathematical theory underpinning the model structure. Furthermore, as the principle of parsimony dictates, when two models are found to reach the same conclusions, the simplest model should be selected (Busemeyer *et al.*, 2015). This enables better model comprehension and mathematical tractability.

It was decided that the model presented here would follow the structure of the EpiSchisto® MWB to enable the explicit parameterisation of individual-level exposure and serology data.

3.2.2 Modelling exposure and immunity

The concept of variation in exposure between different populations has been recognised by researchers for several years (Chandiwana and Woolhouse, 1991; Woolhouse *et al.*, 1998; Pinot De Moira *et al.*, 2010); yet, this variation is frequently studied by examining only the frequency or duration water contact. Whilst it is argued that these observations provide some useful information, there are several additional contributory factors in determining risk of cercarial exposure (Fulford, 2000).

Schistosome transmission is highly focal and microgeographical variations in exposure have subsequently been explored (Pinot De Moira *et al.*, 2007). In so doing, Pinot de Moira *et al.* (2007) and others found that weighted estimates of cercarial exposure were more strongly associated with the prevalence and intensity of reinfection when compared to measurement of duration of water contact alone (Fulford, 2000). Furthermore, in a later study, Pinot de Moira *et al.* (2010) found that differences in exposure explained the observed tribe- and sex-related variations in *S. mansoni* reinfection but not the age-related differences.

The age-dependent exposure patterns included in modifications to the original MWB model by Anderson & May (1985) and others (Chan *et al.*, 1995, 1996; Medley and Bundy, 1996; Chan and Isham, 1998) are largely informed by assumptions over the pattern of exposure. It is often assumed, in the absence of empirical backing, that relative age-dependent exposure gives rise to the age-stratified infection intensity profiles, with exposure modelled as a negative exponential function that reflects the proposed high exposure in children and lower exposure in adults. Furthermore, infection intensity profiles are not considered to be influenced by acquired immunity (Truscott *et al.*, 2017, Anderson *et al.*, 2015).

Acquired immunity to infection with schistosomes is characterised by cumulative experience of exposure, dependence on the degree of antigenic stimulation and provision of partial protection against infection (Woolhouse *et al.*, 1991; Dunne *et al.*, 1992; Naus *et al.*, 2003). Several investigators have cited a shortage of field data, required to inform quantitative estimates for parameter values, and limits to theoretical understanding as reasons for not exploring the contribution of immunity to transmission dynamics in more detail (Allen and Victory, 2003; Anderson *et al.*, 2015). Furthermore, it has been suggested that immuno-epidemiological studies have not, historically, been designed to provide the data required for direct estimation of parameter values (Allen and Victory, 2003; Gurarie, King and Wang, 2010; Restif *et al.*, 2012; Lessler *et al.*, 2015). Thus, when immunity has been included in schistosome transmission models, these immune processes have typically been described phenomenologically, based on assumptions in the absence of data (Anderson and May, 1985; Chan *et al.*, 1996, 2000; Heesterbeek *et al.*, 2015). This remains the case, in spite of increasing immuno-epidemiological evidence. It has, however, been suggested that models lacking either age-dependent parameters or immunity are unable to reproduce age-prevalence or age-intensity curves (Woolhouse, 1991) and may overestimate the positive impact of control.

Adult schistosomes employ several methods of immune evasion (McSorley, Hewitson and Maizels, 2013), host immune responses are therefore thought to be most effective against invading cercarial stages and act to decrease the rate of parasite establishment (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). In the existing modelling literature, immune processes tend to be modelled in the form of concomitant immunity (Griffin, Coutinho and Thomas, 1988; Chan *et al.*, 1996), where current worm burden is related to protection against further infection (Smithers and Terry, 1969). If immunity is related to the current worm burden, as suggested by the theory of concomitant immunity, treatment of initially immune populations would reduce the effect of immunity. Conversely, if immunity takes the form of delayed concomitant immunity, stimulated by the death of adult worms (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012), treatment would increase the acquisition of immunity as a result of increased worm death.

Barbour (1978) first introduced protective immunity into the simple Macdonald model to explain the convex shape of the age intensity curve. In concordance with the findings from Macdonald, Barbour identified that the threshold factor that determines whether or not schistosome transmission can be established is closely related to the proportion of infected snails (Macdonald, 1965; Barbour, 1978); yet, the introduction of human immunity into a more simple model is also able to reflect the data (Barbour, 1978).

Anderson and May (1985) examine a simple phenomenological model of immunity where the immune response is a function of cumulative worm burden and decreases the rate of infection by reducing the force of infection linearly with accumulated experience of exposure and worm burden. In this model, immunological memory is assumed to last for a finite time and is the only constraint on parasite population growth, whilst parasite fecundity and mortality are assumed to be independent of parasite density. Whilst in the EpiSchisto® model, Chan *et al.* (1996) assume that the initial age-intensity profile reflects the effects of acquired immunity, where immunity is again assumed to develop in response to the adult worms, acts against establishment of cercariae but decays exponentially.

Chan and Isham (1998) also model the relative effect of exposure and immunity of the age-infection profile using a phenomenological approach, concluding that varying contact rates account for heterogeneity in infection. Whilst Chan & Isham argue that variability in contact rate generates most of the heterogeneity in the field, the contribution of acquired immunity cannot, as is often the case, be disregarded. Indeed, when the stochastic model (Chan and Isham, 1998) was used in a later paper (Chan *et al.*, 2000), the authors found that, when assuming immunity was lifelong, epidemiological patterns were produced that support the role of acquired immunity.

In a more recent attempt to model the contribution of immunity to the infection intensity profile, Mitchell *et al.* (2012) explored immunity in *S. haematobium* by modelling the protection conferred by IgG plasma cells, looking at both anti-infection and anti-fecundity immunity. However, the kinetics of an IgG response are very different from that of IgE and, as discussed in Chapter 1 anti-infection immunity against *S. mansoni* infection is thought by the immuno-epidemiologists to be IgE dependent.

Although the immunological mechanisms important for resistance to *S. mansoni* infection are increasingly well-characterised, we still have an incomplete mechanistic understanding of the interactions between exposure, immunity and infection. The majority of the modelling studies discussed above base their assumptions of the functional form of contact and immunity on phenomenological conventions used in the literature.

There remains a need to capture the demographic variations in exposure through an empirically informed function in order to deconstruct the relationship between exposure, immunity and infection with schistosomes. In this thesis, I argue that it is essential to consider the role that protective immunity plays in determining transmission dynamics of *S. mansoni*, since underestimation of its effect could lead to significant consequences to public health policy.

3.2.3 Mathematical models and control

It is understood that treatment of schistosomiasis has two potential effects: the reduction of infection intensity in the treated age groups, as a direct result of treatment-mediated worm death, and the reduction in the other age groups due to the overall reduction in transmission (French *et al.*, 2010; Deol *et al.*, 2019). However, there is some debate as to whether treatment programmes may alter schistosome transmission. While some studies suggest that treatment coverage at a level that does not approach eradication is directly beneficial to those treated, but does not have the wider public health benefit of limiting transmission (Macdonald, 1965; Satayathum *et al.*, 2006), others have observed apparent changes in the force of infection following treatment (French *et al.*, 2010)

Uganda was one of the first countries in which annual preventative chemotherapy was introduced, in 2003. MDA with praziquantel has now been routinely conducted for over 15 years across a number of sub-Saharan African countries in which schistosomiasis is endemic. Praziquantel disrupts the tegumental membrane as it kills adult worms. This enables the host immune system to mount an antigen-specific response to parasite proteins that were previously hidden. Several immunological studies have reported age-dependent resistance to reinfection and identified immunological correlates of resistance (Pinot De Moira *et al.*, 2010; Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012; Colley and Secor, 2014). This epidemiological pattern has often been attributed to the gradual acquisition of immunity, assumed to result from cumulative exposure to antigens released as adult worms die, either coming to the natural end of their lifespan, or through drug-mediated death (Woolhouse and Hagan, 1999). It is thought that a certain threshold of antigen exposure is required in order for the host immune system to generate an effective response. The natural lifespan of adult schistosomes is estimated to range between three to ten years (Fulford *et al.*, 1995); therefore, the time taken to reach the required antigen threshold by natural death of worms may take several years. In high-transmission areas, individuals are exposed to a higher level and greater diversity of antigens earlier in their lifetime, therefore achieving the exposure threshold at a younger age. Since highest infection intensities are seen in children, control programmes have focused their efforts on controlling morbidity in this population through school-based administration of praziquantel. As a result of MDA school-age children may subsequently experience the cumulative exposure required to achieve the antigenic threshold at an earlier age through drug-mediated worm death. This is similar to the antigenic boost that would be delivered by vaccination.

On the other hand, should transmission intensity be reduced by sustained long-term MDA programs, as some studies suggest (Anderson and May, 1985; French *et al.*, 2010), the time taken to accumulate sufficient antigen exposure to mount an effective immune response may be

extended, resulting in an increase in the age at which peak infection intensity is seen. Furthermore, Macdonald (1965) introduced the concept of transmission “break points”, based on the probability of male and female worms meeting and mating within the host in low intensity infections. If control programmes do reduce infection intensity, Macdonald’s break point theory may become more significant.

Two principal questions subsequently arise: i) are MDA control programmes gradually shifting the infection peak, and ii) in which direction; earlier through the earlier development of immunity, or later by reducing transmission? These are important research questions to address, since the answers may have implications for the future direction of control measures (Stothard *et al.*, 2011).

Human populations, environments and biological processes are, by their nature, heterogeneous and it is widely recognised that accurate model predictions require inclusion of heterogeneity (Woolhouse *et al.*, 1998; Chan *et al.*, 2000; Heesterbeek, Anderson, *et al.*, 2015). Specifically, it has been argued that modelling exposure and immunity as distinct processes in dynamic transmission models may improve the models’ predictive capability (Civitello and Rohr, 2014) and enable researchers to distinguish between behavioural and immunological protection.

The aim of the model presented in this chapter is to develop a baseline model framework that includes data-informed exposure and immunity functions to specifically explore the dynamics of exposure and immunity within an *S. mansoni* transmission system. The model is adapted from a previously published schistosome transmission model, EpiSchisto® (Chan *et al.*, 1995). This study is unique in that it utilises comprehensive intensity and serology data, collected at the individual level, in addition to malacological survey data. This enables heterogeneity in exposure and susceptibility to be modelled, including environmental heterogeneity of the water contact sites, by way of the site weighting incorporated into the cercarial exposure score.

3.3 Chapter aims

The aim of this chapter is to explore the relative contribution of exposure and immunity to the age intensity profile when incorporated into a schistosome transmission model. Drawing on insights from previous attempts to model schistosome transmission, this chapter will derive mathematical functions that describe the epidemiological patterns of cercarial exposure and IgE-mediated immunity, exploring the differences determined by sex, age and ethnicity, specifically answering the following question:

Can partially protective SmTALI-IgE immunity be incorporated into a *S. mansoni* transmission model that reproduces age-intensity and -serology profiles?

The specific objectives of this chapter are as follows:

1. To fit an appropriate function that describes tribe and sex variations in the age profile of cercarial exposure scores, using maximum likelihood estimation.
2. To fit a function that describes the protective effect of IgE antibodies raised to tegument allergen-like protein 1 (IgE-SmTAL1) on rates of reinfection, using maximum likelihood estimation.
3. To modify a composite force of infection to include rates of contamination and parasite acquisition with the effects of explicitly parameterised host age-dependent immunological processes that may modulate the force of infection.
4. To utilise MCMC Bayesian methods to identify parameter values to provide specific analysis of age-dependent immunological parameters, including immunological responsiveness (strength of protective response) and loss of immunological memory.

3.4 Model derivation

The model is based on the framework of the EpiSchisto® model developed by Chan and colleagues (Chan *et al.*, 1995). Here, I extend the model to include empirically informed cercarial exposure scores and SmTAL1-specific IgE-based partially protective immunity to *S. mansoni*. This modification is motivated by the apparent lack of strong empirical data informing existing transmission model exposure and immunity parameters to date.

The model consists of a series of partial differential equations, specifying initial and ending conditions as well as the evolution of each compartment over the time series. These describe the dynamics of the schistosome adult worm burden in the definitive human host, by age and sex. The model equations are presented in 3.5. The model was run over a discrete time series of 50 years, to equilibrium, with a timestep of 2 months (60 days). A schematic representation of the transmission model and a list of parameter definitions and values are given in Figure 3.1 and Table 3.1, respectively. The model was constructed in R coding language (R version 3.4.3, (R Core Team, 2015)) and implemented at the individual level, structured by age and sex, and run independently for the populations of two tribes, the Alur and the Bagungu.

The parasite lifecycle in the model focuses on the adult worm stage, divided into live worms (W) and dead worms (D). The transmission and development of the various life cycle stages of *Schistosoma* occur over a range of time scales. The lifespan of the free-living parasite stages is relatively short compared to that of the adult schistosome worms. Therefore, as in Chan *et al.* (1995), the free-living parasite stages are assumed to be at equilibrium.

The force of infection (Λ) describes the rate at which incoming worms establish in the host population. This depends on a combination of host- and parasite-related factors (Eq. 7), described in detail in 3.9.1. Following infection, live worms pass through several (n_W) compartments to reflect a normal survival distribution. The rate at which a worm moves from compartment i to compartment $i + 1$ is $n_W \mu_W$ per year where $1/\mu_W$ is the average worm lifespan in years. For subsequent ages, worm burden is carried from the previous age compartment and lost to the subsequent age compartment (Eq. 1). Worms die as they leave the final (n_W^{th}) compartment, entering a ‘dead worm’ compartment (Eq. 2). This stimulates the generation of IgE antibodies (γ_{gain}) as SmTAL1 protein is released from the damaged worm tegument and presented to the host immune system (Eq. 3). Waning IgE immunity is represented by the immune decay factor, γ_{loss} .

Specifically, the model focuses on the development of IgE-mediated host immunity, that is conferred through natural worm death, describing changes in infection intensity and the level of SmTAL1-specific antibody responses with age. Exposure and immunity to schistosome infective stages varies according to demographic characteristics. The model therefore allows for a distinction between age-, sex- and tribe- specific exposure.

The following assumptions are integral to this model, and are discussed in further detail in the relevant indicated sections:

- Egg counts are drawn from a negative binomial distribution. It is assumed that there is no significant variation in this distribution with age; that is, the over-dispersion parameter remains constant, but the mean varies.
- Replicate slides prepared from repeated egg counts taken over three consecutive days accounts for day-to-day variation in egg excretion.
- The IgE optical density distribution is drawn from a gamma distribution and there is no significant variation in the distribution with age.
- The population is stable, previously untreated and transmission is at equilibrium.
- Egg excretion is proportional to worm burden, no density dependent effect on fecundity is modelled.
- Immunity develops in response to exposure to dead/dying worms and acts to reduce subsequent reinfection by targeting infectious cercarial stage by a form of delayed concomitant immunity.
- Exposure to infectious cercariae varies with age as a function of composite cercarial exposure scores with an underlying gamma distribution.

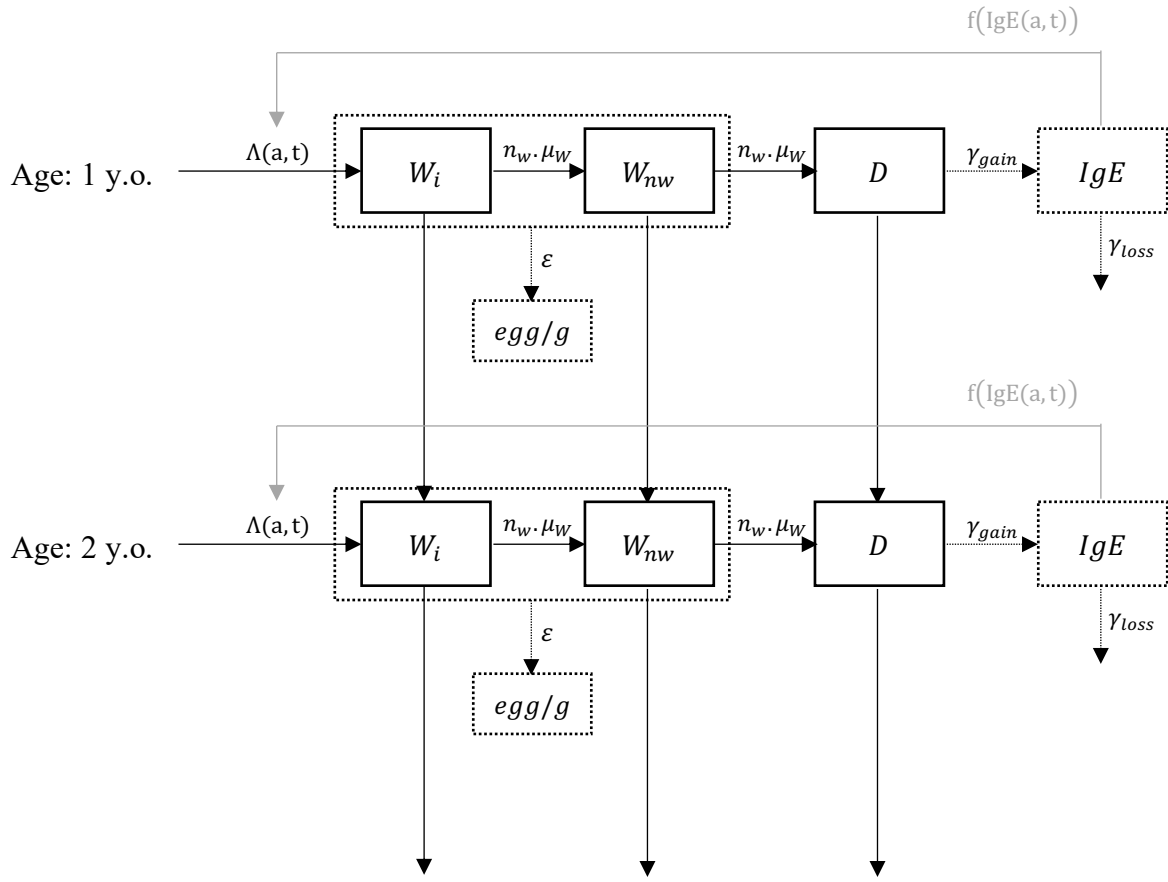


Fig. 3.1 Flow diagram of the *S. mansoni* transmission model. Parameter symbols are defined in Table 3.1. Only two age years are shown for clarity. Solid black lines represent flow between compartments, dotted lines represent model outputs and solid grey line represents the feedback of immunity as a function of IgE antibody response into the force of infection function. The boxed worm compartments W_i to W_{nw} represent the live worm burden.

Table 3.1 Parameter definitions and values

	Symbol	Description	Value & units	Source
<i>Transmission</i>	$\Lambda(a)$	Force of infection	Composite function	(Chan <i>et al.</i> , 1995)
	$*R_0$	Basic reproductive number	Uniform range 0.5 – 8	(Chan <i>et al.</i> , 1995; Woolhouse <i>et al.</i> , 1996; Anderson <i>et al.</i> , 2015)
<i>Parasite</i>	$*\mu_w$	Per capita mortality rate of adult worms	Uniform range 0.06 – 0.29	(A. J. Fulford <i>et al.</i> , 1995)
	$*\varepsilon$	Egg excretion rate per female worm, per day	Uniform range 2 – 20	(Cheever, 1968; Chan <i>et al.</i> , 1995)
<i>Host</i>	$\pi(a)$	Proportion of humans at age a	Function of μ_H	(Turner <i>et al.</i> , 2013)
	μ_H	Per capita mortality rate of human host, Uganda 1998	0.043 yr ⁻¹	Fitted to study data
<i>Exposure</i>	a_{\max}	Maximum recorded age in truncated study population	50 yr	Study data
	ρ_S	Annual cercarial exposure score at age a , in a human of sex s	Estimated (Fig. 3.2, Table 3.2)	Study data; (Pinot De Moira <i>et al.</i> , 2007)
	$*\rho_M/\rho_F$	Relative male-to-female host exposure	Uniform range Alur: 0.5 – 2 Bagungu: 25 – 45	Study data
	κ_S	Rate of environmental contamination with excreta by human of age a and sex s	$\kappa_S = \rho_S$	(Chan <i>et al.</i> , 1995)
	$I(a)$	Protective immunity at age, a , a function of antigen-specific IgE	Estimated (Fig. 3.3, Table 3.3)	Study data
<i>Immunity</i>	$*\gamma_{\text{gain}}$		Uniform range 0 – 1	Uninformed
	$*\gamma_{\text{loss}}$		Uniform range 0.01 – 18	(Vieira, 1988; Lawrence <i>et al.</i> , 2017)
	n_w	Number of worm compartments	8	-
	n_a	Number of age strata	50	Study data
	timestep	Model timestep	60 day	-
<i>Model parameters</i>	duration	Total duration of model run	100 yr	-

* Estimated in the model

3.5 Model equations

The dynamics of the mean worm burden in humans can be described using the following series of partial differential equations:

Worm compartments:

$$\begin{aligned} \frac{\partial W_{s,i}(a,t)}{\partial t} + \frac{\partial W_{s,i}(a,t)}{\partial a} &= \lambda_s(a) - n_W \mu_W W_{s,i}(a,t) & \text{for } i = 1 \\ &= n_W \mu_W W_{s,i-1}(a,t) - n_W \mu_W W_{s,i}(a,t) & \text{for } i = 2 \dots n_W \end{aligned} \quad [1]$$

Dead worm compartment:

$$\frac{\partial D_s(a,t)}{\partial t} + \frac{\partial D_s(a,t)}{\partial a} = n_W \mu_W W_{s(n_W)}(a,t) - W_{s(i)}(a,t) \quad \text{for } i = n_W + 1 \quad [2]$$

For antibody OD levels:

$$\frac{\partial Ab(a,t)}{\partial t} + \frac{\partial Ab(a,t)}{\partial a} = (\gamma_{gain} - \gamma_{loss}) D_s(a,t) \quad [3]$$

3.6 Data

The data used for fitting the model come from a longitudinal reinfection study conducted between May 1998 and July 2000 in the village of Booma on eastern shore of Lake Albert, north-western Uganda. Details of the study can be found in Pinot de Moira *et al.* (2007) and are described briefly in 2.1.1. Two ethnic groups live in geographically distinct areas of the village and have differing rates and patterns of exposure with age. Analysis was restricted to individuals from these two tribal groups. The study sample comprised 248 individuals, ranging in age from 5 to 50 years old (Alur: n = 110; Bagungu: n = 138).

The cercarial exposure function was fitted to data from all study participants for whom an annual cercarial exposure score had previously been calculated (n = 248) (Pinot De Moira *et al.*, 2007). Whereas model calibration was restricted to those individuals for whom baseline, pre-treatment egg counts and serological data were available, in addition to cercarial exposure scores (n = 197). There was no systematic difference in the age (t = 0.471; p = 0.638), cercarial exposure scores (t = 0.269; p = 0.788) or egg excretion (t = 0.0423; p = 0.966) between the sub-cohort used for model calibration and the total cohort data used for estimation of the cercarial exposure

function for either the Alur or Bagungu tribes. Each demographic group was divided into five age bands of approximately equal size. These are defined for each demographic group as follows: Alur females: 7 to 10 years (n = 8), 11 to 14 years (n = 9), 15 to 20 years (n = 8), 21 to 30 years (n = 9), 31 to 50 years (n = 13); Alur males: 7 to 9 years (n = 11), 10 to 14 years (n = 14), 15 to 19 years (n = 14), 20 to 29 years (n = 11), 30 to 50 years (n = 13); Bagungu females: 7 to 9 years (n = 11), 10 to 19 years (n = 15), 20 to 29 years (n = 17), 30 to 38 years (n = 10), 40 to 50 years (n = 16), and Bagungu males: 7 to 10 years (n = 14), 11 to 20 years (n = 15), 21 to 30 years (n = 10), 31 to 35 years (n = 12), 36 to 50 years (n = 17).

3.7 Modelling cercarial exposure

3.7.1 Risk of exposure

The lake is the primary water source for the inhabitants of Booma and there is consequently a high level of water contact across the demographic profile of the village. Detailed water contact observations were performed during this study, a full description of which can be found elsewhere (Pinot De Moira *et al.*, 2007). Briefly, water contact observations were conducted between November 1998 and September 1999 at 19 different sites. Observations took place over an 11-hour period each day. For each water contact event the following information was recorded: name of the individual coming into contact with the lake; the site at which contact was made; the time of day; the activity conducted including the use of any soap; the degree of submersion (number of body parts), and the duration of contact to the nearest minute. Pinot de Moira *et al.* (2007) then used these data to explore risk of cercarial exposure in detail, culminating in individual weighted cercarial exposure scores that reflect the average annual exposure per person. The cercarial exposure scores are calculated as the product of: duration of water contact; time of day; degree of immersion, and site of water contact, for each contact made by each individual and with each variable appropriately weighted (Fulford, 2000; Pinot De Moira *et al.*, 2007). The site weighting of the cercarial exposure score includes analysis of malacological data that was collected in parallel to the human data. The cercarial exposure score therefore encapsulates the intermediate host population dynamics in addition to the definitive host population. Each of the factors within the cercarial exposure score are in turn dependent on age, sex, tribe and geographical location, all of which govern the nature of the activity being performed. Further analysis on the behavioural variations seen in this community is available elsewhere (Pinot de Moira *et al.*, 2010).

As previously described by Pinot de Moira (2007), there is a significant difference in the cercarial exposure scores between individuals from the Alur and Bagungu tribes, despite residing in the

same village. Differences in exposure are also present between males and females particularly in the Bagungu tribe. A mathematical function was fitted to cercarial exposure data *a posteriori*. The different exposure profiles were then introduced into the model to compare the impact of variable exposure on the model output.

3.7.2 Age-dependent cercarial exposure function ($\rho(a)$)

The average annual cercarial exposure scores, derived by Pinot de Moira *et al.* (2007) and described in section 3.7.1, were summarised into a single rate function (Eq. 4).

[4]

$$\rho(a) = (age * c * e^{-(\beta * age)}) + d$$

Cercarial exposure scores are measured on a continuous scale, with no zero values. The distribution of scores is negatively skewed and is assumed to follow a gamma distribution, with mean $E[X] = \kappa\theta$ and variance $Var[X] = \kappa\theta^2$. The likelihood of the parameters ($\theta = \{c, \beta, d\}$), given the data ($x = x_i$) is therefore:

$$p(x|k, \theta) = \prod_{i=1}^N f(x; k, \theta)$$

Where,

$$f(x; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} \quad \text{for } x > 0 \text{ and } k, \theta > 0$$

Initial values for the shape and scale parameters of the gamma distribution were calculated empirically from the sample mean and variance of the cercarial exposure scores:

$$\theta = \frac{Var[X]}{E[X]} \quad \kappa = \frac{E[X]}{\theta}$$

Maximum likelihood estimation was carried out using the Nelder-Meade algorithm from a gamma distribution using ‘optim’ in R (R Core Team, 2015). The resulting parameter estimates

are presented in Table 3.2. The fitted cercarial exposure function for Alur females and males resulted in a curve that peaked in the youngest age group and declined into adulthood. This is similar to the water contact profile that has often been assumed in previous schistosome transmission models. The resulting curves show a similar profile for both females and males of the Alur tribe, although the youngest Alur females exhibited behaviour resulting in slightly higher cercarial exposure scores, this is analysed in detail elsewhere (Pinot De Moira *et al.*, 2007, 2010). In contrast, the fitted cercarial exposure function for males and females of the Bagungu tribe resulted in quite different profiles in relation to each sex and the Alur tribe. The fitted cercarial exposure function resulted in a flatter profile, without a peak in the youngest individuals. Furthermore, Bagungu males had significantly higher cercarial exposure scores compared to Bagungu females. Again detailed analysis of these differences can be found elsewhere (Pinot De Moira *et al.*, 2007, 2010). Observed cercarial exposure scores by age group and the fit of *a posteriori* exposure models are shown in Figure 3.2.

Table 3.2 Maximum likelihood estimation of age-dependent exposure function parameters for Alur and Bagungu ethnic groups.

	Parameter estimate				MLE
	c	β	d	var	
Alur female*	219.33	0.21	182.64	10.66	53.01
Alur male	191.15	0.27	202.73	11.69	226.91
Bagungu female*	- 0.50	0.04	65.30	7.65	55.48
Bagungu male	297.451	0.04	-279.50	16.90	102.56

* The female exposure function was used as a comparator for relative male exposure in the MCMC model and the best fitting parameters for function were fixed accordingly

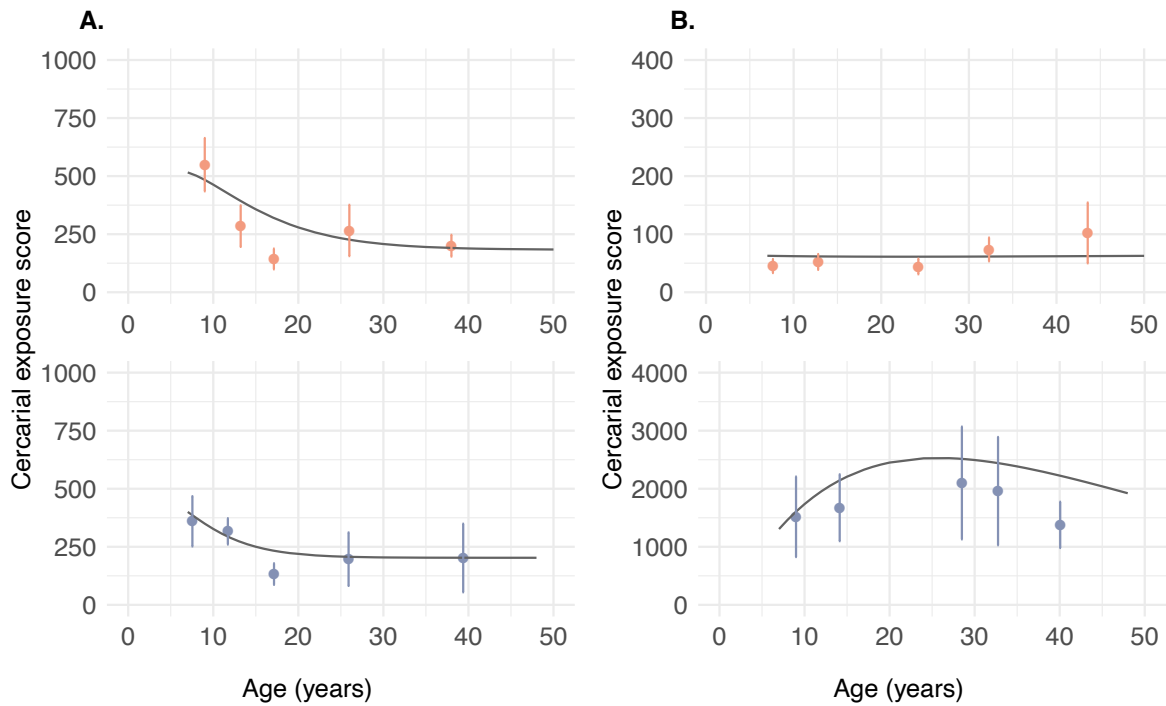


Fig. 3.2 Cercarial exposure function model fit for A) Alur and B) Bagungu females (red) and males (blue). The fit of the exposure function model is represented by the solid line. Points represent the arithmetic mean cercarial exposure by age band, plotted at the mean age of each band. Error bars represent the standard error of the mean (SEM).

3.7.3 Relative exposure by sex (ρ_M/ρ_F)

For simplicity, the same functional form was assumed for females and males from each tribe. This was clearly reasonable for the Alur tribe as the independently fitted functions for each sex demonstrate in Figure 3.2A, whilst the large standard error around the Bagungu male mean suggests that a flat functional form, similar to that fitted for the Bagungu females, may also approximate the data reasonably well. The scale of the exposure scores between males and females of the Bagungu tribe, however, varies significantly (Fig. 3.2B). The relative scale of male exposure scores was therefore estimated in the model, relative to the female exposure function profile. The female exposure function parameters were fixed in the MH-MCMC algorithm for each tribe. ρ_M/ρ_F was then estimated by the model, with a uniform prior distribution ranging from 0.5 to 2 for the Alur and 25 to 45 for the Bagungu, according to the observed difference in the scale of exposure. It was hypothesized that this would demonstrate whether the significantly higher exposure in Bagungu males translates to high infection intensity or serological profiles, or whether infection intensity is over-estimated in the Bagungu males, given the significantly

higher exposure in this subset of the population, which would suggest that other contributory factors are important in mediating infection intensity.

3.7.4 Contamination rate ($\kappa_{(a)}$)

In the absence of data, it is assumed that individuals with the greatest exposure risk are likely to be the highest contaminators, suggesting a proportional relationship between exposure and contamination. Therefore, as in Chan *et al.* (1995), this model assumes that humans contaminate the environment through the excretion of eggs ($\kappa_{(a)}$) at the same rate at which they are exposed to infectious cercariae ($\rho_{(a)}$), therefore:

$$\rho_{(a)} = \kappa_{(a)} \quad [5]$$

3.8 Modelling the protective effect of TAL1-IgE

The control of biological processes by density dependent effects is well recognized in ecology and helminth populations tend to be regulated by multiple density dependent mechanisms (Keymer, 1982; Keymer and Slater, 1987). Studies of helminth infections in both humans and animals have identified a role for the immune system in regulating worm burden in addition to direct density-dependent regulation through competition for resources (Paterson and Viney, 2002; Bleay *et al.*, 2007).

Protective immunity is associated with the IgE response to TAL proteins described in Chapter 1. The protective IgE response is well characterised in *S. mansoni* and IgE to tegument antigen-like protein 1 (TAL1), has been correlated to resistance to reinfection in the three major *Schistosoma* species important to human health and across several geographical locations (Dunne *et al.*, 1992; Santiago *et al.*, 1998; Fitzsimmons *et al.*, 2004). The TAL family of proteins are located in the tegument of the parasite across all life-cycle stages, with varying degrees of expression (Fitzsimmons, Jones, Stearn, *et al.*, 2012). In parameterising the SmTAL1-specific IgE antibody response, the model assumes that IgE antibodies constitute a correlate of risk and that higher titres reduce the probability of reinfection. This assumption is supported by empirical data from Booma, where SmTAL1-IgE levels were significantly negatively associated with reinfection at 12 months (Pinot de Moira *et al.*, 2010).

3.8.1 Strength of protective response

The Kaplan-Meier (KM) estimator is a non-parametric statistic that allows estimation of the survival function (Kaplan & Meier, 1958). In the current study, the survival function was used to describe the relationship between the probability of infection and SmTALI-IgE optical density (OD) at baseline. Analysis was conducted using the ‘survival’ and ‘survminer’ statistical packages in R (R Core Team, 2015). The probability of infection was modelled as a binary outcome, where infection was defined as having a positive Kato-Katz diagnosis (at least one *S. mansoni* egg identified in at least one faecal sample). An object was created using the ‘surv’ function (R Core Team, 2015) to pair infection data with antibody responses, in the same manner that right-censored time data would be analysed in a conventional survival analysis. The ‘survfit’ function (R Core Team, 2015) was then used to create a survival curve from the data. A Cox Proportional-Hazards (CPH) model (Cox, 1972) was used to compare the KM probability curves between males and females and between the two tribes.

3.8.2 Cox Proportional-Hazards model

When the effect of sex was analysed, the CPH model demonstrated that there was not a significant difference in the risk of infection with SmTALI-IgE antibody OD between males and females of the same tribe (Alur Hazard ratio: 0.6429 (95% CI: 0.36–1.14, $p = 0.131$); Bagungu Hazard ratio: 0.7885 (95% CI: 0.49–1.26, $p = 0.318$).

When the CPH model was run using tribe aggregated data with tribe and sex as independent variables and inclusion of an interaction term between tribe and sex, again, no significant difference in infection risk was seen between any of the demographic groups ($p > 0.05$). It was therefore decided to fit a dose response immunity function for the total cohort. The assumption of proportionality was checked graphically for each of the independent variables prior to analysis.

3.8.3 Age-dependent immunity function ($f(I)$)

The infection probabilities and paired antibody titres were extracted from the survival model and a dose-response curve was fitted to the KM curve in order to provide a tractable function that describes the relationship between the IgE antibody response (OD) and probability of infection (successful cercarial invasion).

A dose-response model describes the probability of a specified response from exposure to a specified pathogen in a specified population, as a function of the dose. Dose-response

relationships have been used in a number of studies to examine the relationship between illness outcomes relative to a given dose or amount of infectious organism (Lunn *et al.*, 2019). In this study I adapt this relationship to study probability of infection given an increasing ‘dose’ of antibody, as measured by ELISA.

The generalised four-parameter log-logistic function (Eq. 6) is commonly used to model the relationship typically seen in the association between dose and response (Ritz *et al.*, 2015). This involves the estimation of four parameters ($A1$ – $A4$) in the equation:

[6]

$$f(I) = c + \frac{A2 - A1}{(1 + e^{A3(\log(Ab) - \log(A4))})}$$

Where $A1$ and $A2$ relate to the bottom and top asymptote of the curve, respectively; $A3$ is the slope of the curve, and $A4$ is typically the effective dose (ED) at which the probability is reduced by 50%, also known as ED50. The log-logistic function was fitted to the data extracted from the KM estimator using the dose response curve (drc) statistical package in R (Ritz and Strebig, 2016). The ‘drm’ function utilises the ‘optim’ function to find maximum likelihood estimates of the parameters. Parameters were established for the probability of infection at baseline, given the pre-treatment SmTAL1-specific serum IgE (Table 3.3). The fitted dose response model and KM curves are presented in Figure 3.3.

Table 3.3 Maximum likelihood estimated parameter values for SmTAL1-IgE immunity function.

		Estimate	SE	RSE	AIC	LL
Total	A1	0	NA			
	A2	0.879 ***	0.011	0.042	-659.56	333.78
	A3	0.745 ***	0.020	(186 df)		
	A4	0.046 ***	0.002			
Alur	A1	0	NA			
	A2	0.896***	0.017	0.044	-169.64	88.82
	A3	0.747***	0.035	(48 df)		
	A4	0.056***	0.004			
Bagungu	A1	0	NA			
	A2	0.884***	0.019	0.039	-313.78	160.89
	A3	0.638***	0.010	(86 df)		
	A4	0.022***	0.001			

***p<0.0001; SE: standard error; RSE: residual standard error; AIC: Akaike information criterion; LL: log likelihood; df: degrees of freedom

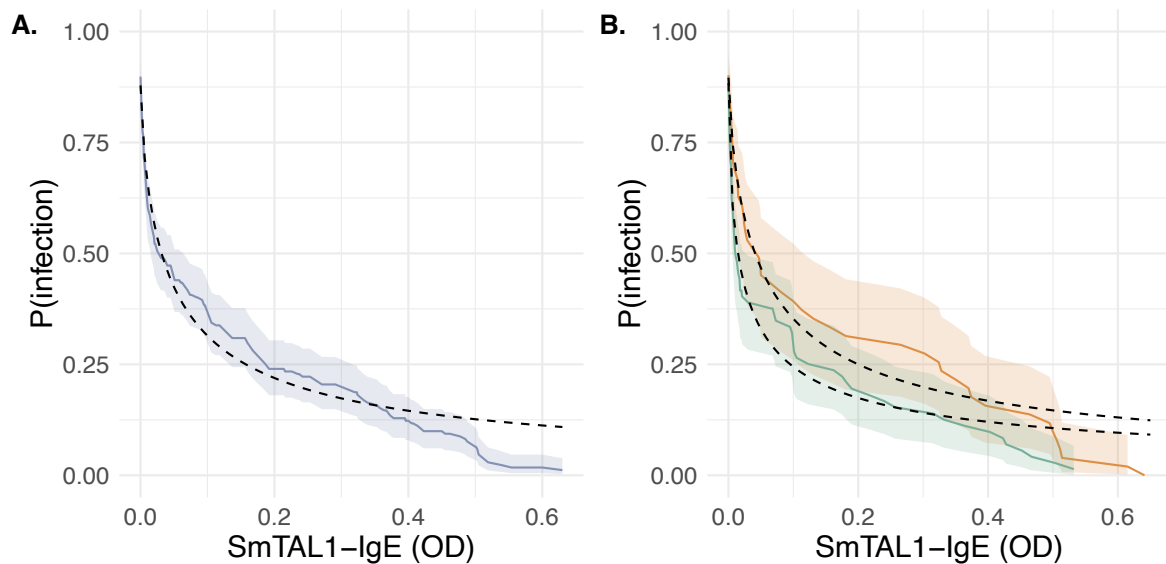


Fig. 3.3. KM curve with drc function, fitted to individual-level serology data for A) total cohort and B) Alur (yellow) and Bagungu (green) tribes. Optical Density (OD) is measured at dual wavelength (490/630 nm). shading indicates 95% confidence interval band. Probability of infection modelled as binary outcome.

3.8.4 Immunity 'gain' parameter (γ_{gain})

The gain of immunity parameter is modelled as a constant rate with a linear relationship between the dead worm compartment and the SmTALI-IgE OD. Due to the uncertainty around this relationship this parameter is fitted within the MCMC algorithm and is assigned an uninformative uniform prior between 0 and 1. It is, however, likely that this is a considerable simplification of the true relationship between exposure to dead worms and the development of immunity.

3.8.5 Longevity of the IgE antibody response

The half-life of circulating serum antibodies varies according to their immunoglobulin class. IgE antibodies are the least abundant in serum and, in the absence of repeated antigenic exposure, the serum half-life of IgE in murine plasma ranges from 2 to 14 hours (Vieira and Rajewsky, 1988; Kubo *et al.*, 2003). Despite displaying the shortest serum half-life of all immunoglobulins, IgE typically remains high in individuals with allergic or parasitic diseases, suggesting that the development of humoral immune memory is required to maintain the long-term effectiveness of antibody responses.

Classical humoral memory provides long-term protection, mediated through both memory B and plasma cells (PCs). Memory B cells are maintained in the absence of antigen, but these cells do not produce antibodies, so are not directly protective. However, re-exposure to antigen stimulates rapid differentiation of memory B cells into antibody-secreting plasma cells (APCs) (Jimenez-Saiz *et al.*, 2017).

It is thought that the development of IgE memory has unique features that differ from conventional B cell memory (Saunders *et al.*, 2019). At least two principal mechanisms through which IgE levels are perpetuated have been suggested: IgG_1^+ or IgE^+ B cells that continually produce short-lived IgE^+ PC (SLPCs) or terminally differentiated long-lived plasma cells (LLPCs) that excrete antigen-specific antibodies for the lifetime of the patient (Moutsoglou and Dreskin, 2016).

It has been suggested that IgE^+ B cells are thought to be poorly recruited to the memory compartment of the immune system (He *et al.*, 2015; Haniuda *et al.*, 2016). Instead, it is proposed that high affinity IgE is produced by sequential class switching from IgM via antigen-experienced IgG^+ B cells to IgE rather than from naïve B cells directly (Aalberse and Platts-Mills, 2004; He *et al.*, 2015; Looney *et al.*, 2016; Jimenez-Saiz *et al.*, 2017). IgE consequently inherits somatic hypermutations from the intermediate IgG phase. This consecutive immunoglobulin class

switch recombination is therefore necessary for affinity maturation of the IgE response (Xiong *et al.*, 2012; He *et al.*, 2017).

Others suggest that LLPCs have an important role in supporting a persistent IgE antibody response (Aalberse and Platts-Mills, 2004; Moutsoglou and Dreskin, 2016). IgE⁺ B cells are short-lived within the germinal centre (Erazo *et al.*, 2007; Yang, Sullivan and Allen, 2012) and have been proposed to favour the differentiation to IgE plasma cells (Young *et al.*, 2012). These antigen-specific plasma cells with long-term immunological memory are maintained preferentially in the bone marrow (Gould and Wu, 2018). Terminally differentiated LLPCs continually secrete antigen-specific antibodies, even after episodes of antigen absence, thus enabling the rapid response by mast cells upon antigen re-exposure. (Gould and Wu, 2018; Saunders *et al.*, 2019)

3.8.6 Immunity 'loss' parameter (γ_{loss})

As described in section 3.8.5, biologically active IgE may be sustained for several months after exposure to the initial stimuli of production through T cell memory and long-lived IgE plasma cells maintaining IgE production (He *et al.*, 2015). However, studies of the longevity of IgE in murine serum suggest that the half-life of IgE is relative short, amounting to days or weeks (Vieira and Rajewsky, 1988). There is no certain numerical estimate of the time taken for decay in antibody response. The immunity decay rate is therefore modelled as a constant rate with an uninformative prior of uniform distribution between 0.01 and 18. When γ_{loss} approaches zero, the duration of immunity is lifelong and when $\gamma_{\text{loss}} = 18$, the IgE response is maintained for approximately 20 days (Lawrence *et al.*, 2017). In this model it assumed that infection is at endemic equilibrium and therefore individuals are continuously exposed.

3.9 Model parameters

3.9.1 Force of infection (Λ):

The force of infection (FOI) is the rate at which susceptible individuals acquire an infectious disease and is modelled as a composite function, comprising several transmission processes (Eq. 7). The transmission-related parameters that comprise the FOI equation include: host age; exposure to the infectious stage of the parasite (cercariae); contamination of fresh water by an individual through open urination or defecation; a density-dependent establishment function, and the reproductive potential of the parasite (basic reproductive number, R_0). The FOI also

incorporates the life span of the parasite and distribution of the host population. Each individual parameter is described further in the following sections.

$$\Lambda(a, t) = \frac{\mu_W R_0 \rho(a) f(I(a, t)) \int \pi(a) \kappa(a) W(a, t) da}{\int \pi(a) \kappa(a) \rho(a) da} \quad [7]$$

3.9.2 Basic reproductive number (R_0)

In epidemiology, the basic reproductive number R_0 determines whether a disease will be transmitted within a population or die out. In order for an infectious agent to spread to endemic equilibrium R_0 must be greater than 1. If R_0 is less than 1, the infection will not persist within the population. For schistosomes, R_0 is defined as the average number of female offspring produced by a female worm (or worm pair) that infected the definitive human host and survived to reproductive maturity in a totally susceptible population (Anderson and May, 1985). For the purpose of this model, R_0 adopts a numerical value. There are relatively few numerical estimates for the value of R_0 in human schistosomes, but the most commonly cited range for R_0 of schistosomes is between 1 and 4 (Anderson and May, 1991). It has, however, been shown that inclusion of heterogeneity in transmission models, through contact rates and/or acquired immunity, can shift the endemic threshold, thus increasing the R_0 (Medley and Bundy, 1996; Woolhouse, Hasibeder and Chandiwana, 1996; Woolhouse *et al.*, 1998). R_0 is therefore estimated within the MH-MCMC model fitting framework with a uniform prior distribution between 0.5 and 8 (Table 3.1), to account for potentially higher values resulting from the inclusion of heterogeneity in the model.

3.9.3 Population structure ($\pi(a)$)

The model considers the acquisition of parasites on a population level. The force of infection therefore incorporates the age structure and average life expectancy of the definitive human host population. The age profile of the population is characterised, to good approximation, by a truncated exponential distribution, described by the function $\pi_{(a)}$ (Eq. 8). It is assumed that the population distribution is stable and does not fluctuate from the distribution at the time of the survey in 1998 (Fig. 3.4).

$$\pi(a) = \frac{\mu_H e^{-\mu_H a}}{1 - e^{(-\mu_H a_m)}} \quad [8]$$

The model will be structured by age and sex for each tribe, due to variation in cercarial exposure patterns between the demographic groups, in particular between Alur and Bagungu populations and between males and females within the Bagungu tribe. The proportional representation of the study population by sex and tribe is shown in Figure 3.4.

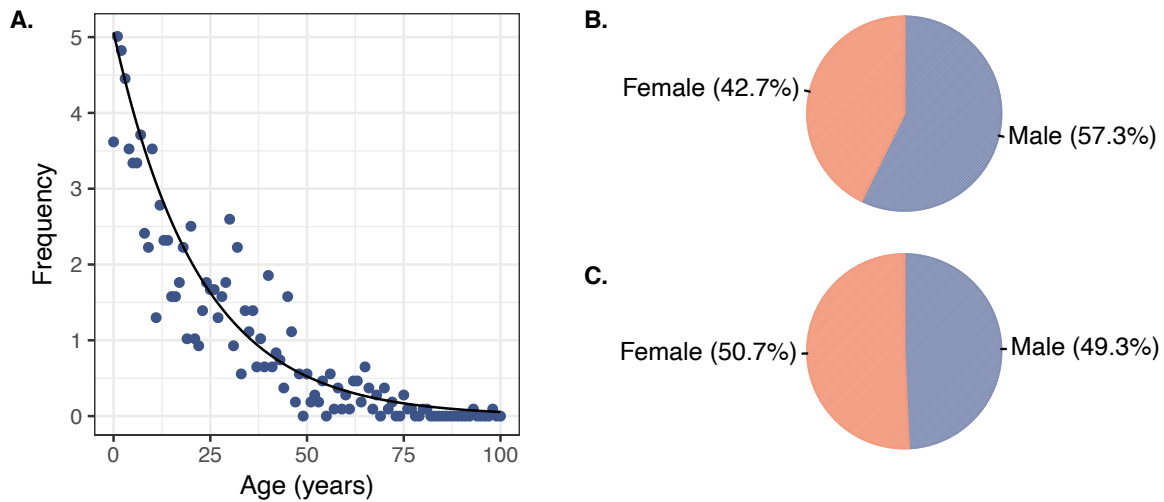


Fig. 3.4 Underlying demography of study population (Booma, Uganda) in 1998. A) Age distribution for population of Booma, 1998. Solid line: fitted probability density function ($\pi(a)$). Sex distribution of females and males from the B) Alur (n = 110) and C) Bagungu (n = 138) tribes.

3.9.4 Human life expectancy ($1/\mu_H$)

An exponential function (Eq. 8) was fitted to the Booma population survey data, resulting in an annual per capita death rate of 0.043 (95% CI: 0.039–0.049, $p < 0.0001$) (Filipe *et al.*, 2005). Schistosomiasis is a disease of chronic morbidity, only resulting in mortality in the most severe cases. Human life expectancy is therefore assumed to be independent of schistosome infection, on the understanding that any increase in mortality as a result of schistosomiasis is likely to be small (Cohen, 1973).

3.9.5 Adult schistosome lifespan ($1/\mu_W$)

The average lifespan of *S. mansoni* has been estimated to be between 5.7–10.5 years (Fulford *et al.*, 1995). In this model the prior was modelled as a uniform distribution between 0.0625 and

0.263, capturing the 95% confidence interval range of worm lifespan ($1/\mu_w$, 3.8 to 16.0), estimated from pre-treatment *S. mansoni* infection intensity and post-treatment reinfection data from two Kenyan villages (Fulford *et al.*, 1995). This range includes the estimates of worm lifespan proposed by others (Feng, Eppert, Milner, *et al.*, 2004). Although it should be noted that there are some reports of schistosomiasis in immigrants to Western Australia in whom *S. mansoni* was diagnosed more than 30 years after leaving an endemic region (Wilkins, 1987; Harris, 1984). This demonstrates the potential longevity of *S. mansoni* within the definitive human host. Worm mortality is also assumed to be independent of density and host age in the model presented here (Fulford *et al.*, 1995).

3.9.6 Egg excretion rate (ϵ)

It is widely recognised that different species of human schistosome have different egg laying characteristics (Jamieson, 2016). The mature female *S. mansoni* worm is estimated to lay between 100 and 300 eggs per day. Autopsy studies suggest that approximately one third of the eggs laid are passed in the faeces (Cheever *et al.*, 1994). Regression analysis indicates that for every 5–8 eggs counted per gram of faeces, one worm pair was recovered by autopsy. Likewise, the EpiSchisto® model previously estimated the daily egg excretion rate per worm pair to be 5.26 eggs per gram per worm per day (Chan *et al.*, 1995). Additionally, Krauth and colleagues found that there is no change in number of eggs per gram of faeces per worm pair with increasing host age (Krauth *et al.*, 2012). The rate of egg excretion, ϵ , is therefore assumed to be constant with age and due to uncertainty around its numerical value will be estimated by the model (Table 3.1), with a uniform prior distribution ranging from 2 to 20.

3.9.7 Density dependent fecundity

Density dependent constraints are essential to maintain population stability. These can take several different forms, including density-dependent establishment or fecundity, either directly or through the development of immunity that modifies these processes (Woolhouse *et al.*, 1991). The contribution of immunity was discussed in detail in section 3.8.

The fecundity of helminths is usually measured by examining the correlation between egg counts and adult worm burden. Direct measurement of worm burden has historically been difficult for schistosome infections as a result of their intervacular location; evidence regarding *S. mansoni* fecundity is therefore divided (Keymer and Slater, 1987). Many field studies now measure circulating anodic antigen (CAA), a by-product of schistosome digestion, in the blood

or urine of infected individuals as a proxy for live worm burden (Corstjens *et al.*, 2014; Ochodo *et al.*, 2015).

Several mathematical models that have examined the relationship between worm burden and egg excretion in *S. mansoni* do not appear to support a role for density dependent fecundity (de Vlas *et al.*, 1992; Chan *et al.*, 1995). Epidemiological evidence tends to concur with these findings (Agnew *et al.*, 1995, 1996). By contrast, processes of density-dependent fecundity are more apparent in other helminth infections (Churcher, Filipe and Basáñez, 2006).

In the absence of strong evidence to the contrary, the relationship between the number of eggs excreted by the human host, as measured by Kato-Katz, and current worm burden is assumed to be linear in the model presented here. This assumption was supported by analysis of egg count and CAA data from the Booma cohort (Fig. 3.5, Table 3.4).

Following Polman *et al.* (2000), a comparison of regression modelling methods was performed in order to take into consideration variation that occurs in the measurement of both egg counts (dependent variable) and CAA (independent variable).

Whilst Polman *et al.* (2000) compared both Deming and ordinary linear regression (OLR) models, here I also examined the results of a Passing Bablok regression model. Both Deming and linear regression models assume normality of the data, whereas the Passing Bablok model does not make any prior assumption on the data distribution. Furthermore, Passing Bablok is not as sensitive to the presence of outliers as either Deming regression and OLR. OLR does not take into consideration variability in the independent variable (CAA measurement). A value of one was added to the data, which were log transformed prior to analysis. Model comparisons were performed using the 'mcr' package in R.

Polman *et al.* (2000) found that variation between regression analysis by Deming regression and ordinary linear regression is low for CAA, concluding that for egg/CAA analysis OLR may be sufficient. Likewise, analysis of the relationship between worm burden and egg excretion in the Booma study population, performed here, indicates that the linear regression model may be sufficient in modelling the egg/CAA relationship (Fig. 3.5). More importantly, the slope of the regression line is not less than one and therefore, although not directly proportional, this analysis does not provide strong evidence in support a reduction in egg excretion with increasing worm burden.

MacDonald (1965) and others (Booth *et al.*, 2003; Berhe *et al.*, 2004) have highlighted the limitations of quantitative parasitology in determining the actual worm burden, relative to the number of excreted eggs. The sensitivity of Kato Katz is understood to be low in low intensity infections, which may result in the underestimation of egg counts in stool samples (Bärenbold

et al., 2017). This may explain the resulting slope greater than one. The sampling and distribution of egg counts will be discussed in more detail in section 3.10.1, relating to the choice of distribution for the likelihood function used in modelling fitting.

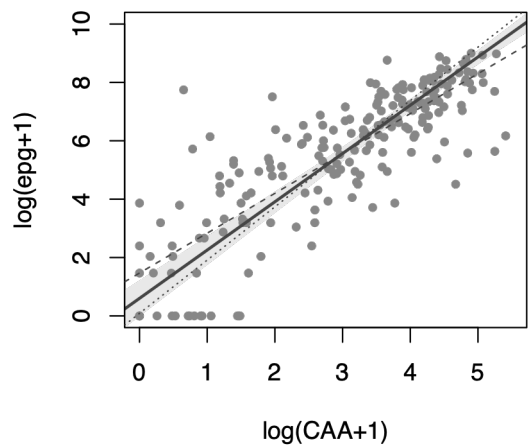


Fig. 3.5 Linear regression between eggs per gram of faeces (epg) and circulating anodic antigen (CAA). The x-axis represents the logarithm of CAA and the y-axis represents the logarithm of eggs per gram faeces. A value of 1 was added to all values prior to transformation to account for zero values. Solid line: Passing-Bablok regression, with shaded 95% confidence interval boundary; dashed line: Linear regression; dotted line: Deming regression. Pearson’s R = 0.825.

Table 3.4. Linear regression model parameters eggs (log(epg+1)) per worm (log(CAA+1))

	Estimate (95% CI)
Intercept	0.601 (-0.008-1.225)
Slope	1.654 (1.496-1.806)

3.10 Model calibration

Parameter estimates were derived by using Metropolis-Hastings Markov Chain Monte Carlo (MH-MCMC) approach to sample from the posterior distribution. The ‘optim’ function in R was used to determine estimated coefficients for this model by maximising the posterior density function. The model was fitted to individual level parasitology (egg counts) and serology (IgE-SmTAL1 OD) data. Only those individuals for whom cercarial exposure, multiple faecal samples and serology data were available were included in the model fitting data (n = 197). Simulations were run for each ethnic group (Alur and Bagungu) using the Robbins-Munro adaptive scaling process to optimise covariance and parameter estimation (Munro & Robbins, 1951). Model parameters that could be directly informed by data were kept fixed. All other parameters were estimated using MH-MCMC (Table 3.1). The R code for the model fitting process can be found in Appendix 3.

3.10.1 Distribution of faecal egg count

Schistosoma worm burden cannot be directly quantified in infected humans due to the intravascular location of adult worms. Parasitological diagnosis of *S. mansoni* infection therefore relies on examination of faecal samples.

Egg counts tend to be over-dispersed with respect to the Poisson distribution. Overdispersion of parasitological data is measured by calculating the variance to mean ratio. In Booma, the variance is considerably higher than the mean, variance-to-mean ratio (VMR) = 1939. This ratio suggests that overdispersion is present within the egg count data and a negative binomial model is therefore appropriate in describing the distribution of parasites within the host (Feng, Eppert, Milner, *et al.*, 2004).

Sensitivity and specificity of the Kato-Katz technique varies by the schistosome species, infection intensity and the number of stool samples examined. Day-to-day variation in faecal egg output is common among parasites and in epidemiological studies participants commonly have changes in their test results on consecutive days. It has been shown that preparation of multiple Kato-Katz slides from stool samples collected over consecutive days improves the sensitivity of the Kato-Katz technique (Booth *et al.*, 2003; Berhe *et al.*, 2004).

Stool samples were collected on three consecutive days and two Kato Katz slides prepared for each sample, to account for variability in observed faecal egg counts (FEC). Analysis found that there was no variation in egg counts between the two slides, analysed by different parasitology technicians (Fig. 3.6). In this study, the sum of egg counts was taken across three samples on three days due the variation in egg output (Van Etten *et al.*, 1997).

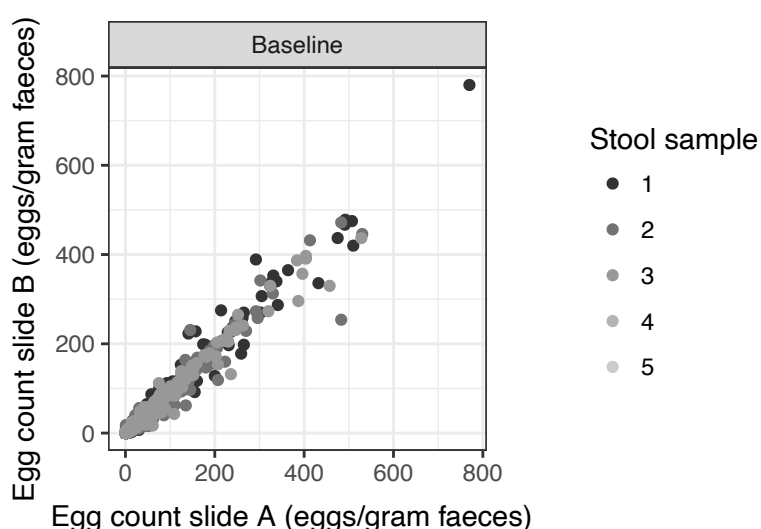


Fig. 3.6. Interrater sampling variability across repeated stool samples, analysed by different parasitology technicians.

3.10.2 Distribution of IgE optical density

Serum IgE is measured by ELISA, resulting in a colorimetric reaction indicating the level of serum specific IgE present for each individual. The optical density (OD), or absorbance, is then measured on a continuous scale, with the lower limit bounded by zero. The upper bound tends to be assay specific and can often be greater than one, though not the case in the data utilised in this chapter. The SmTALI-IgE OD data was therefore assumed to follow a gamma distribution. It was also assumed that there is no variation in the distribution of SmTALI-IgE OD values with age.

3.10.3 Bayesian inference

Epidemiological forecasting is inherently challenging due to unpredictable characteristics and natural variability of the pathogen and their hosts' biology. Further uncertainty arises from the determinants of transmission, such as infectious contact, susceptibility and environmental heterogeneities. Bayesian methods were employed to fit the transmission model to the study data from Booma as adopting a Bayesian approach is understood to generate more robust parameter predictions. In Bayesian inference, the model is being evaluated whilst taking into account information that is already known about different parameters, for example, the likely range of parameter values from published literature or experimental data, referred to as parameter priors.

Bayesian inference aims to estimate the posterior distribution of the parameter θ , via Bayes Theorem:

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)}$$

Where, $p(\theta|Y)$ is the posterior distribution of the parameter of interest θ ; $p(Y|\theta)$ is the likelihood function; and, $p(\theta)$ is the chosen prior distribution of θ . Inference is usually performed ignoring the normalising constant $p(Y)$. The posterior parameter distribution can therefore be written as:

$$p(\theta|Y) \propto p(Y|\theta)p(\theta)$$

That is, the joint posterior distribution is proportional to the product of the likelihood function and prior distribution.

3.10.4 Likelihood specification

The likelihood distribution is the product of the individual parameter likelihoods (Eq. 9), which is equivalent to the sum of log-likelihoods.

[9]

$$p(\mathbf{Y}|\boldsymbol{\theta}) = \prod_{i=1}^n p(y_i|\boldsymbol{\theta})$$

As described in section 3.10.1, egg count data are discrete, measured over an unbounded positive range, where the sample variance is greater than the sample mean. Observations are therefore over-dispersed with respect to a Poisson distribution. The negative binomial model is commonly used to describe data where the Poisson distribution does not explain all variance for skewed data and is therefore selected in this instance.

The negative binomial distribution is an infinitely divisible distribution, that is, if X_i has a negative binomial distribution, then for any positive integer n , there exist independent identically distributed random variables (X_1, \dots, X_n) , whose sum has a negative binomial distribution with mean: $E[X] = n * \mu_i$ and overdispersion: $n * k$. It is therefore assumed that the sum of six repeated egg counts (three samples provided on three consecutive days; two slides prepared from each sample) has the distribution:

$$\sum_{i=1}^n X_i \sim NB(n * \mu_i, n * k)$$

Where the probability density function can be written as:

$$f(X_i; \mu_i, k) = \frac{\Gamma(k+x_i)}{x_i! \Gamma(k)} \left(\frac{k}{k+\mu_i} \right)^k \left(\frac{\mu_i}{k+\mu_i} \right)^{x_i} \quad \text{for } x_i = 0, 1, 2, \dots$$

The likelihood function for N independent and identically distributed observations (x_1, \dots, x_n) is the product of the individual distributions. It therefore follows that the log likelihood function is as follows:

$$\begin{aligned} \ell(\mu_i, k) = & \sum_{i=1}^n \ln(\Gamma(k + x_i)) - \sum_{i=1}^n \ln(x_i!) - n \ln(\Gamma(k)) \\ & + \sum_{i=1}^n nk(\ln(k) - \ln(k + \mu_i)) + x_i(\ln(\mu_i) - \ln(k + \mu_i)) \end{aligned}$$

where μ_i denotes generically the model output, which is a function of the transmission model parameters $(\mu(\theta))$.

Conversely, the SmTALI-IgE OD data was measured on a continuous scale, so the probability of observing the serological data given the model was drawn from a gamma distribution, with the following probability density function:

$$f(x; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} \quad \text{for } x > 0 \text{ and } k, \theta > 0$$

and log-likelihood function:

$$\ell(k, \theta) = (k - 1) \sum_{i=1}^n \ln(x_i) - \frac{1}{\theta} \sum_{i=1}^n x_i - nk \ln(\theta) - n \ln(\Gamma(k))$$

Zero values in the data set were accounted for by the addition of 0.03 to the SmTALI-IgE data.

3.10.5 Prior distributions

Prior probability distributions quantify our existing knowledge about the parameters of interest. The choice of prior can strongly influence the posterior probability. In the model presented in this thesis, priors were treated as low information, or uninformative, with a uniform distribution, since there is considerable uncertainty regarding their true value. This means that all values within the feasible range are equally probable *a priori*. Consequently, the posterior distribution is determined primarily by the data (likelihood function). Table 3.1 defines the prior

distributions used for estimating uncertain parameters in the model. The theory behind the selected ranges is discussed in more detail in each of the relevant sections.

3.10.6 Initial conditions

A summary of the initial values used in the respective Alur and Bagungu model runs is provided in Table 3.5. Initial values for the model estimated parameters were drawn by random number generation from the specified prior distributions (Table 3.1). The exposure (section 3.7.2) and immunity (section 3.8.3) function parameters were estimated *a posteriori* and fixed in the MH-MCMC algorithm and the representative female proportion of the population was directly determined from the baseline population survey (section 3.9.3).

Table 3.5 Summary of initial model parameter values

Parameter	Alur	Bagungu
Model estimated parameters		
R_0	4.5	4.5
μ_w	0.1737717	0.1737717
ε	6	6
γ_{gain}	0.34	0.34
γ_{loss}	4.6	4.6
ρ_M/ρ_F	0.7	38.9
Exposure function parameters (female exposure) (section 3.7.2)		
c	219.3274627	-0.49743697
β	0.2142095	0.04463911
d	182.6394025	65.30170598
Immunity function parameters (universal) (section 3.8.3)		
A_1	0	0
A_2	0.8785873	0.8785873
A_3	0.7453523	0.7453523
A_4	0.0456560	0.0456560
Other		
Population ratio (female:male)	0.427	0.507
nw	8	8

3.10.7 Metropolis–Hastings Monte Carlo Markov Chain (MCMC)

Metropolis–Hastings and other MCMC algorithms are generally used for sampling from multi-dimensional distributions. The MH-MCMC method was selected to estimate several model parameters simultaneously. Simultaneous parameter estimation can move in multiple directions

within the parameter space in any one step, it may therefore be quicker to converge on parameter estimates.

MH-MCMC considers a sequence of random variables $\{X_0, X_1, X_2, \dots\}$, sampled from the distribution $p(X_{t+1}|X_t)$. The Metropolis Algorithm performs a ‘random walk’ through the multi-dimensional parameter space that uses an acceptance-rejection rule to converge to the target distribution (Eq. 10). Each next sample X_{t+1} then depends only on the current state X_t and does not depend on the further history $\{X_0, X_1, \dots, X_{t-1}\}$. This sequence is called a Markov chain.

[10]

$$P(x^i, x^*) = \min \left[1, \frac{\sigma(x^*)}{\sigma(x^i)} \right]$$

MCMC techniques aim to construct chains which draw samples that progressively converge to the target distribution. The initial samples in a chain may follow a very different distribution, especially if the starting point is in a region of low density and may be biased by the initial value. As a result, a burn-in period is typically necessary, where an initial number of samples are discarded. In this model the burn-in period was initially set to 25% (2,500 iterations).

3.10.8 Adaptive MCMC model

The efficiency of convergence of the MCMC algorithm depends on the scaling of the proposal density. A high acceptance rate tends to indicate that the algorithm is staying at the same point and so exploration of the parameter space, or ‘mixing’, is sub-optimal. The Robbins-Munro search method (Garthwaite, Fan and Sisson, 2016) was therefore employed to optimise model fitting. This adaptive method estimates covariance during the MCMC to improve model mixing and aims for a pre-specified long-term acceptance rate, resulting in a covariance matrix that is better informed. The target acceptance ratio in this study was set to 0.234, as suggested by Roberts *et al.* (1997) for a multivariate proposal distribution. The model was set to sample 200 iterations before starting to adapt.

3.10.9 Model convergence

The model was run over 10,000 iterations. The burn-in period was validated by examining the parameter trace plots to confirm whether the model had reached an invariant stage within the burn-in period by visual examination of the chain trajectory.

3.11 Results

3.11.1 Model fit

The fit of each model output was evaluated by visual inspection against the observed data. Characteristics of the age-profiles were examined, including:

- the magnitude of the relative exposure parameter compared to the data;
- the shape of the age-intensity profile;
- the age of peak infection intensity and specific IgE OD, and
- the magnitude of the curve.

3.11.1.1 Alur

The shape of the age-intensity curve generated by the model at equilibrium for plausible fitted parameter values is similar to the observed patterns for the Alur tribe, where peak infection intensity is demonstrated in children and a subsequent decrease in infection intensity occurs through adolescence into adulthood. This appears to mirror the profile of exposure with age. There are however some notable characteristics of the model that differ from the observed data.

The relative magnitude of male exposure that is estimated by the model is slightly overestimated, particularly in the youngest age groups. Conversely, the subsequent peak infection intensity for Alur males is underestimated. The scale of peak infection intensity of the Alur females is however accurately captured by the model. The model estimates that the age of peak infection intensity occurs at an earlier age than that observed by the data for both females and males.

Similar to the age-intensity profile, the age at which specific IgE is predicted to peak by the model is at a younger age than that seen in the observed data. This may result from underestimation of the worm lifespan. Furthermore, the model IgE profile mirrors the post peak decline seen in the infection intensity profile of the model. This is to be expected since the gain in immunity parameter is directly related to dying worms via a constant conversion rate. The model peak shifts from approximately 5 years of age in the infection intensity profile to approximately 7.5 years in the IgE titre profile. A successive shift in the age of peak exposure, infection intensity and IgE response is also observed in the data. This indicates that the model is capturing the delayed development of the IgE response to some extent; however, this effect is not completely captured, since highest mean IgE titres are seen in 15 to 24-year olds in the study data. The inclusion of a protective antibody threshold may better reproduce the delayed peak in the serological profile in future iterations of the model.

The magnitude of the peak OD values is captured reasonably well by the model for the youngest age groups, but the model serology curve then declines beyond the OD values observed in the data, particularly for the Alur males. The decline in IgE OD values predicted by the model is likely to be associated with the immune decay parameter and again may indicate that the constant rate of decay that is assumed in the model is likely to be an oversimplification.

The Alur infection intensity curve clearly mirrors the observed exposure profile. It is therefore difficult to tease apart the relationship between exposure, immunity and infection intensity in this demographic group. The similarity in exposure profiles between Alur females and males also makes it difficult to evaluate the respective contributions of exposure and immunity.

3.11.1.2 Bagungu

The fit of the model to the Bagungu observed data is notably inferior to that of the Alur model. The results of the Bagungu model fit do however provide some interesting points for discussion.

The relative exposure of the Bagungu males is overestimated by the model. The model is fitted to infection intensity and serology data. Since the observed infection intensity of the Bagungu males is significantly lower than their exposure behaviour may suggest, the high prior distribution constraints may limit the model in achieving an accurate fit to the observed infection intensity data.

In contrast to the results for model fitted to the Alur data, the flatter profile of the exposure function for the Bagungu females appears to result in a predicted age intensity curve that peaks in very young children and maintains a plateau with age. The model captures the infection intensity in the oldest age groups but does not reflect the peaked pattern of infection intensity observed in Bagungu females. Moreover, the observed age intensity profile for the Bagungu females appears to demonstrate a peak shift compared to the Alur data, with a lower peak infection intensity occurring at a later host age.

The underestimation of infection intensity in Bagungu females subsequently leads to underestimation of the serological response. The very low IgE OD values predicted by the model for Bagungu females mean that the effect of the partially protective immune response is negligible, and this is reflected in the shape of the model curve, which does not reproduce the decrease in infection intensity in the older age groups.

Contrary to the predicted Bagungu female intensity and immunity profiles, high exposure in the Bagungu males results in very high predicted infection intensity and subsequent IgE OD values creating a sharp peaked profile. It is apparent that exposure does not completely explain the

shape of the age intensity curve and the inclusion of an informed immunity function appears to contribute to the decline in infection intensity observed in adults. The overestimation of infection intensity given the comparatively high exposure in the Bagungu males suggest that additional processes may, however, be involved in mediating infection intensity in this sub-cohort.

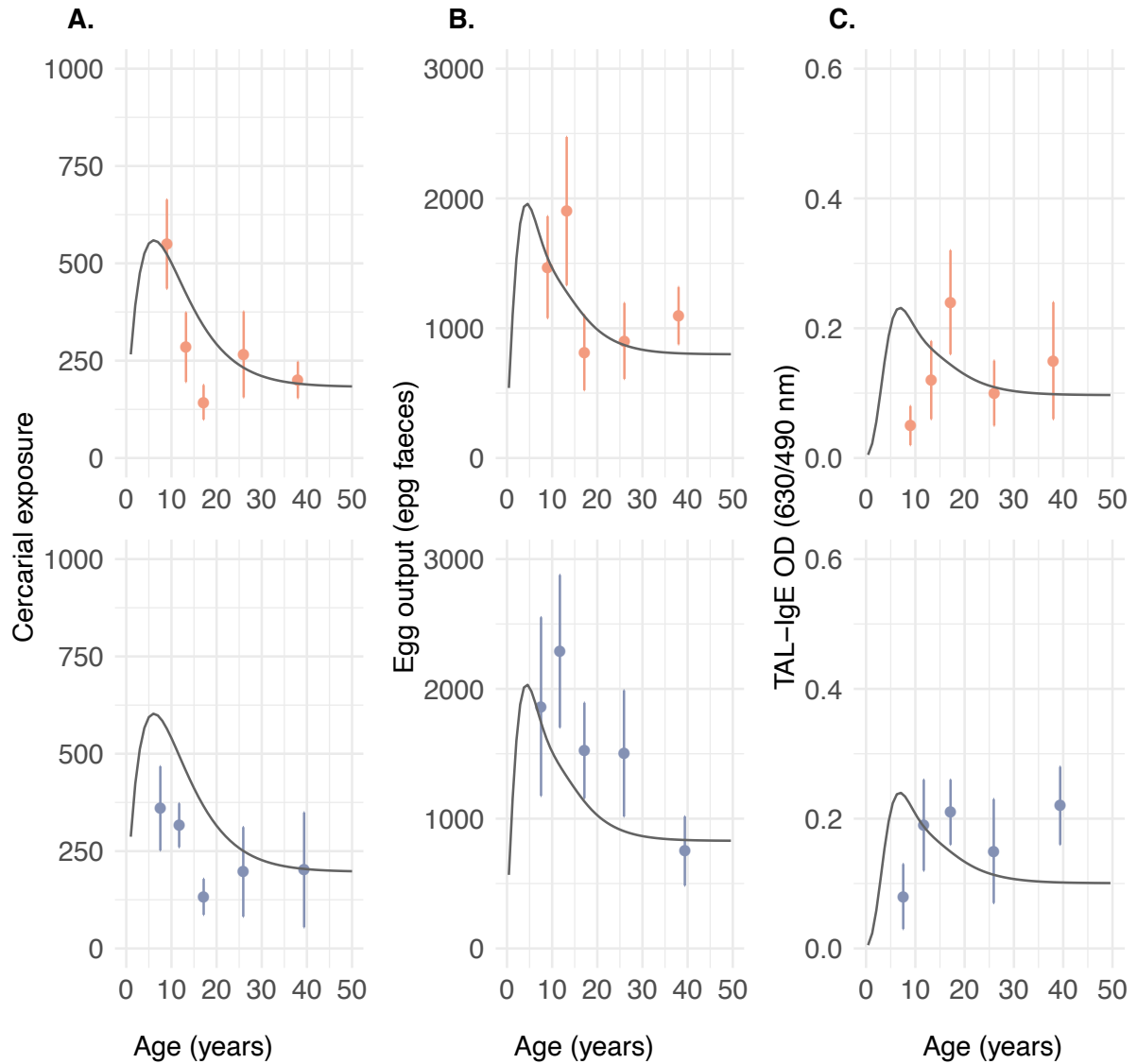


Fig. 3.7. Observed and model predicted age profiles of A) cercarial exposure scores; B) infection intensity (eggs/g faeces), and C) SmTALI-IgE optical density (OD₆₀₀) by sex for the Alur tribe. Mean and standard error of observations are shown (red: females; blue: males) for age groups described in section 3.6. Solid line represents the model, fitted to individual data for the Alur (n = 85).

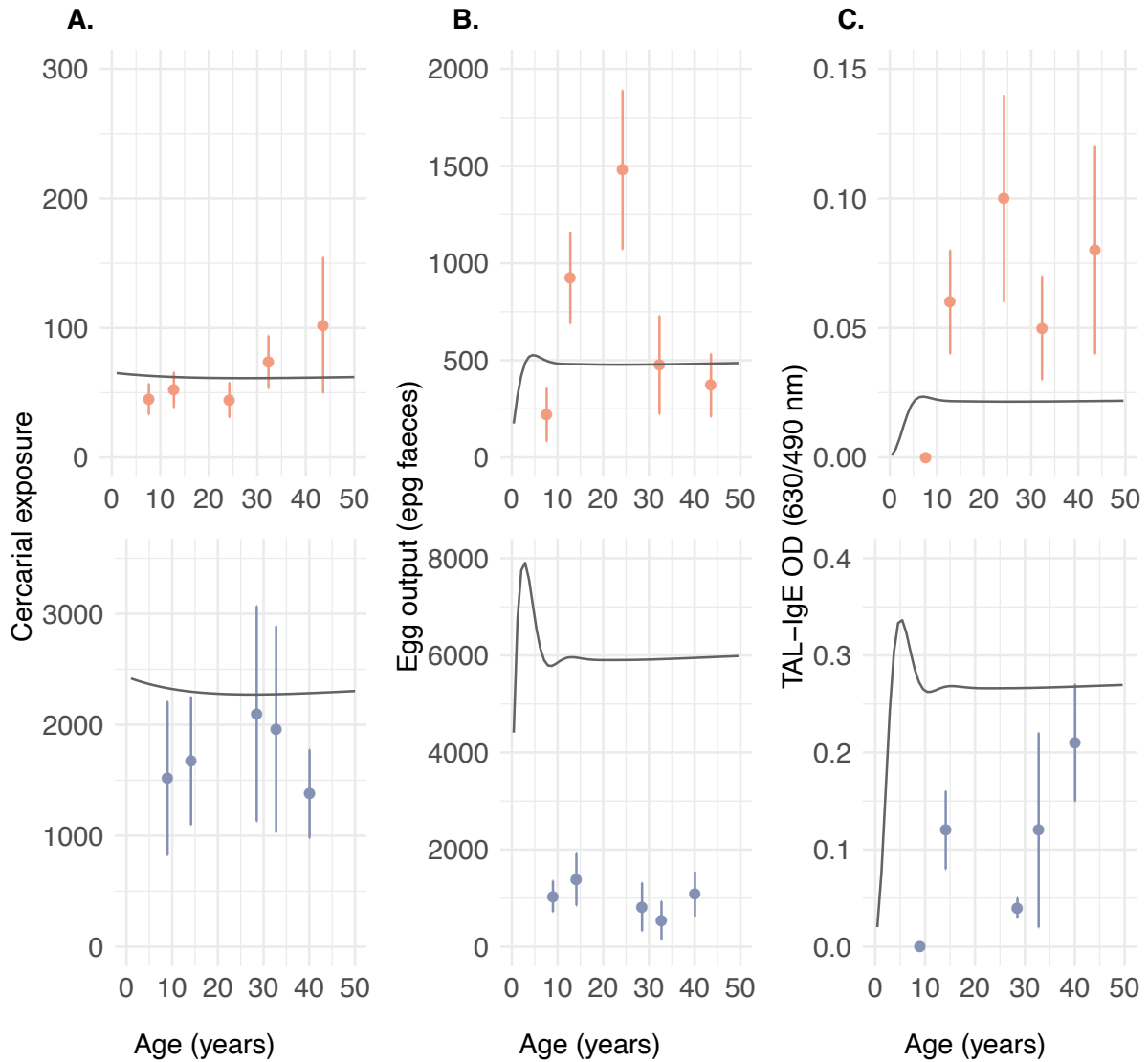


Fig. 3.8 Observed and predicted age profiles of A) cercarial exposure scores; B) infection intensity (eggs/g faeces), and C) SmTALI-IgE optical density (OD₆₀₀) by sex for the Bagungu tribe. Mean and standard error of observations are shown (red: females; blue: males) for age bands described in section 3.6. Solid line represents the model, fitted to individual data for the Bagungu (n = 112).

3.11.2 Estimated parameter values

A summary of the estimated model parameters and model fit is provided in Table 3.6. Estimates of the fitted parameters (R_0 , μ_W , ε) largely lie within the respective published ranges (Table 3.1).

The most commonly cited range of R_0 is between 1 and 4. The median estimate of R_0 for the Alur tribe falls within the upper limit of this range (Anderson and May, 1985). On the other hand, the Bagungu tribe model estimates a higher value for R_0 (Table 3.6). Both estimates are however in line with previous studies that make an assumption of a higher R_0 in simulations with immunity compared with those without immunity (Chan *et al.*, 1996) and those in which population heterogeneity is explicitly modelled (Woolhouse *et al.*, 1998).

The estimated worm mortality rate is similar for each tribe (Table 3.6), relating to a worm lifespan of 4.07 years in the Alur and 3.53 years in the Bagungu tribes. These estimates are towards the lower bounds of the estimated range published in the literature (Fulford *et al.*, 1995).

The median parameter estimates for the egg production rate are again within the published limits for both tribes (Cheever *et al.*, 1994; Chan *et al.*, 1995), despite a broad prior distribution (Table 3.6).

The model estimated immunity parameters (γ_{gain} and γ_{loss}) were the least well described in the literature. The gain of immunity parameter was modelled as a scaling parameter with a constant conversion rate between the dead worm compartment and IgE OD values. Decay of immunity was also modelled as a constant rate of loss of specific IgE antibodies. A higher rate of immune gain was seen in the Alur tribe model compared to the Bagungu (Table 3.6). Serum SmTAL1-IgE was estimated to have a shorter longevity in the Alur model (74.1 days) compared to the Bagungu (80.5 days) model.

Table 3.6. Estimated parameter values resulting from 10,000 MH-MCMC model iterations*

Parameter	Median parameter estimate	
	Alur	Bagungu
R_0	3.980374	5.251514
μ_w	0.2457766	0.283631
ε	7.48516	5.376067
γ_{gain}	0.009051468	0.001937529
γ_{loss}	4.926424	4.534888
ρ_M/ρ_F	1.07815	37.1312
Acceptance ratio	0. 2426	0. 2249
Log posterior	-658.50	-936.75

*Burn-in period removed (Alur: 0–3000; Bagungu 0–2500)

3.11.3 Diagnostics of model fit

On the whole, the histograms demonstrate a generally smooth profile with a normal distribution reflecting the prior distribution that was imposed in the model.

Although visual inspection of the trace plots indicates that the model achieves a degree of convergence for both tribes, it would be expected that the trace would form a tighter horizontal band, similar to that shown for the γ_{gain} parameter in Figure 3.12, rather than the random walk appearance seen for the other parameter estimates. The parameter estimates should therefore be treated with a degree of caution as they may not represent true convergence of the model. The diagnostic plots indicate that the model may not be fully exploring the available parameter space within the 10,000 iterations. Furthermore, the trace plots for the models suggest that there is serial correlation between successive model draws. This could initially be addressed by increasing the effective sample size in order for the chain to explore the sample space more widely.

The target acceptance ratio, described in section 3.10.8, was achieved for both the Alur and Bagungu models (Table 3.6).

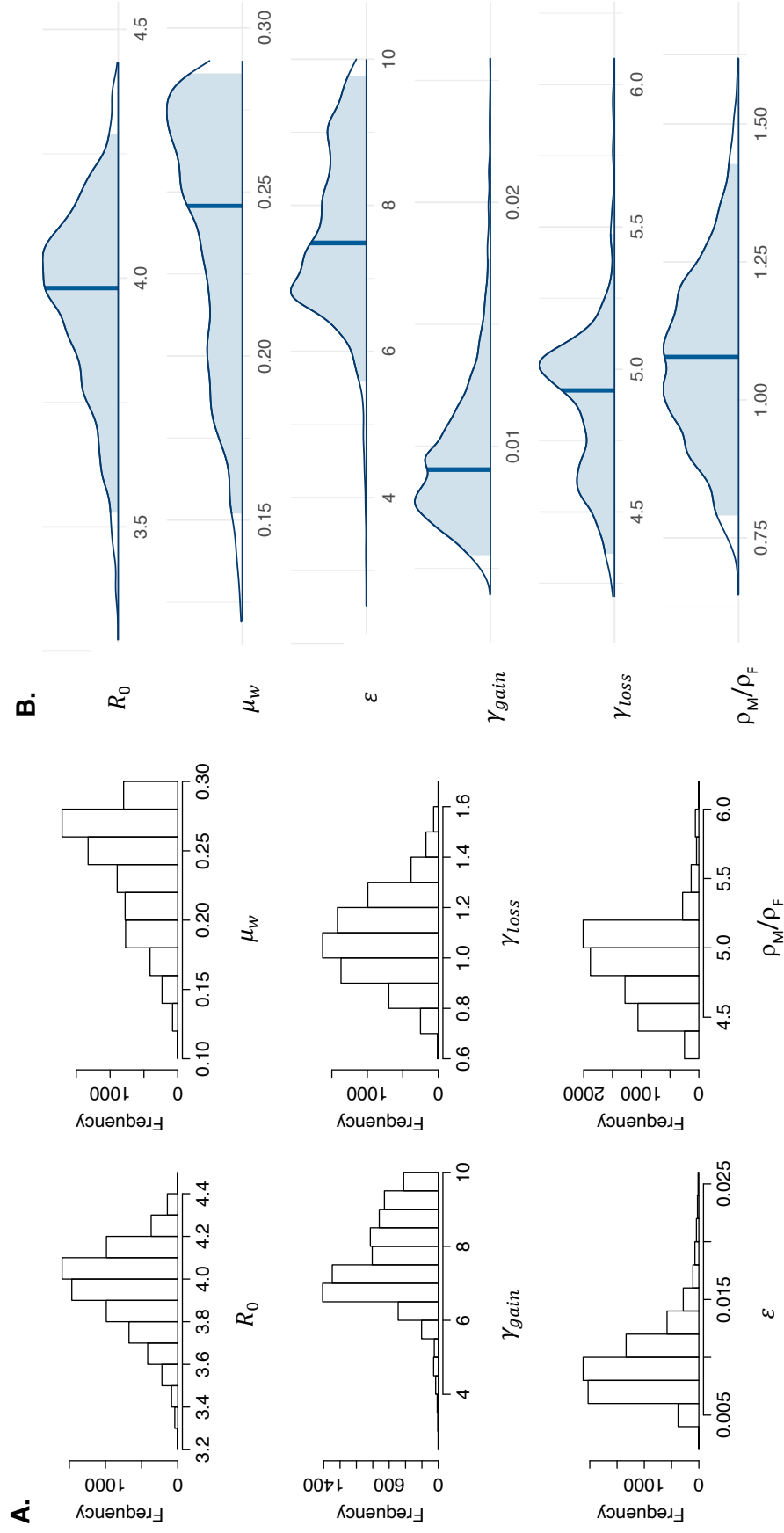


Fig. 3-9 Posterior distributions for the six estimated model parameters for the Alur tribe (steps 2500 – 10000). Distributions are presented as A) histograms and B) smoothed density plots with the median parameter estimate (bold line) and 95% Bayesian credible intervals (shaded region) indicated. The distributions are based on an MH-MCMC chain of 10,000 samples. Burn-in period removed.

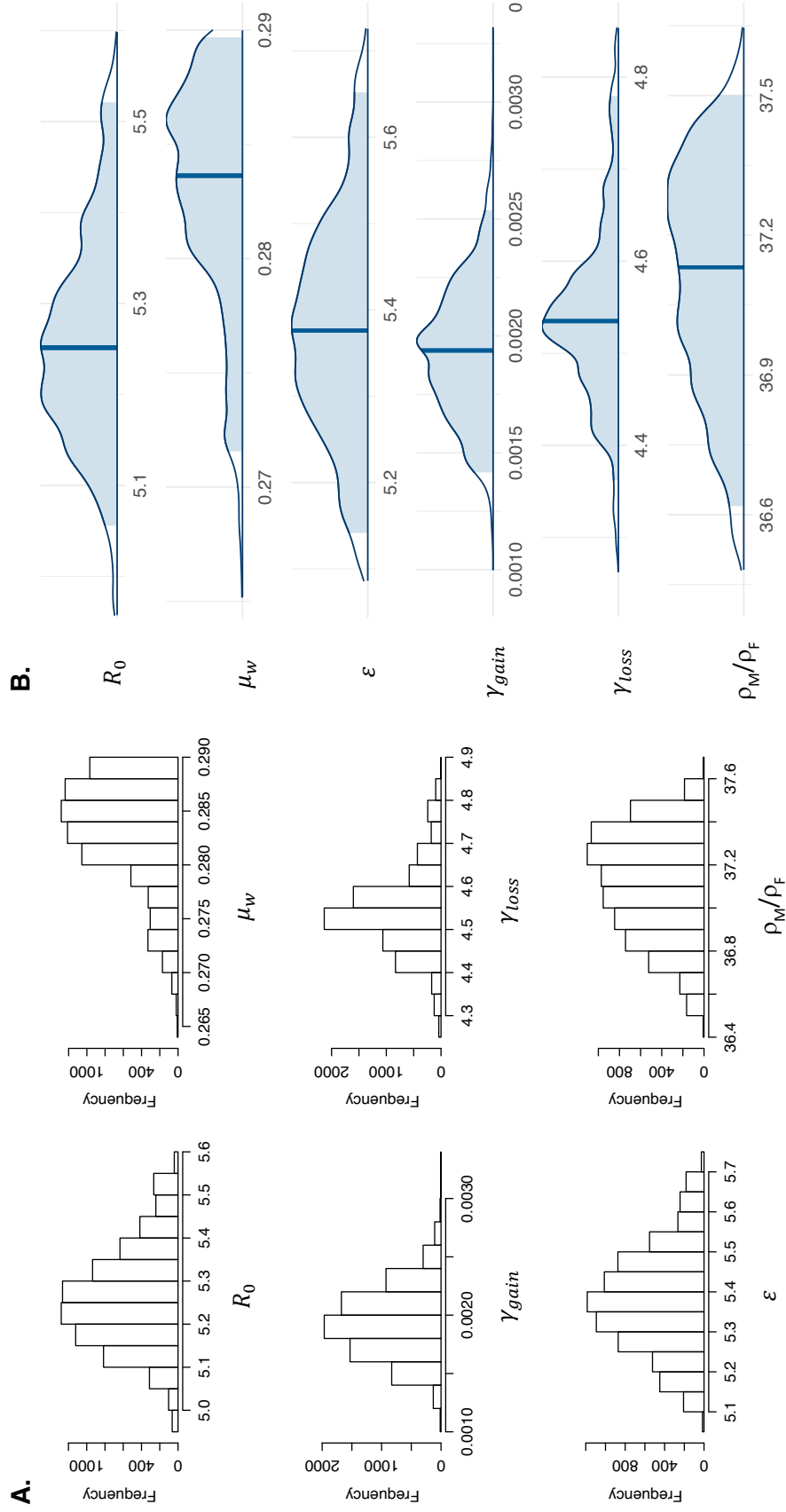


Fig. 3.10 Posterior distributions for the six estimated model parameters for the Bagungu tribe (steps 2500 – 10000). Distributions are presented as A) histograms and B) smoothed density plots with the median parameter estimate (bold line) and 95% Bayesian credible intervals (shaded region) indicated. The distributions are based on an MH-MCMC chain of 10,000 samples. Burn-in period removed.

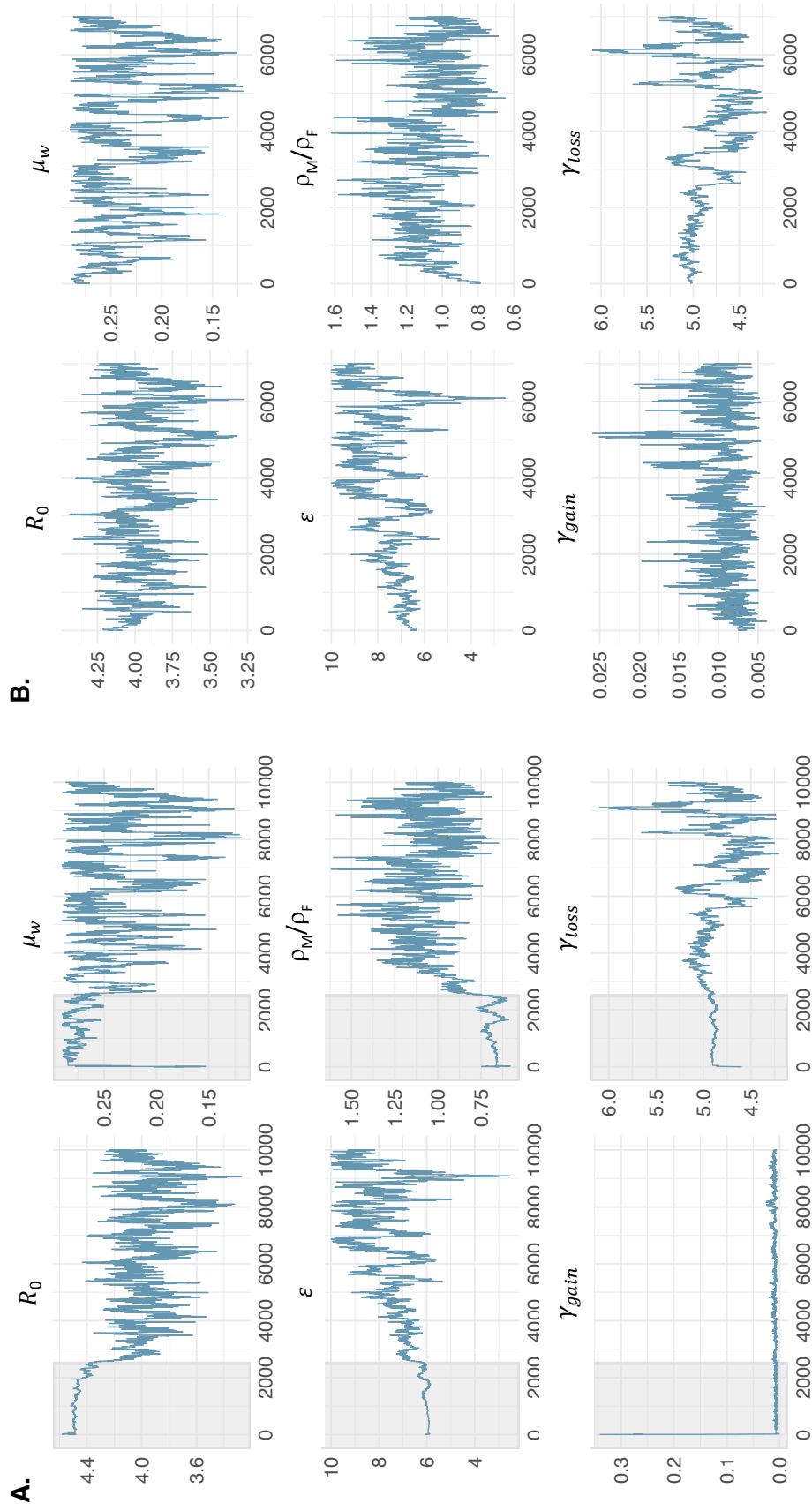


Fig. 3.11 Trace output for each parameter estimate from the MH-MCMC chain, run over 10,000 iterations for the Alur tribe A) steps 0 to 10000 and B) steps 2500 – 10000 (burn-in removed). The shaded region indicates the selected burn-in period 2500.

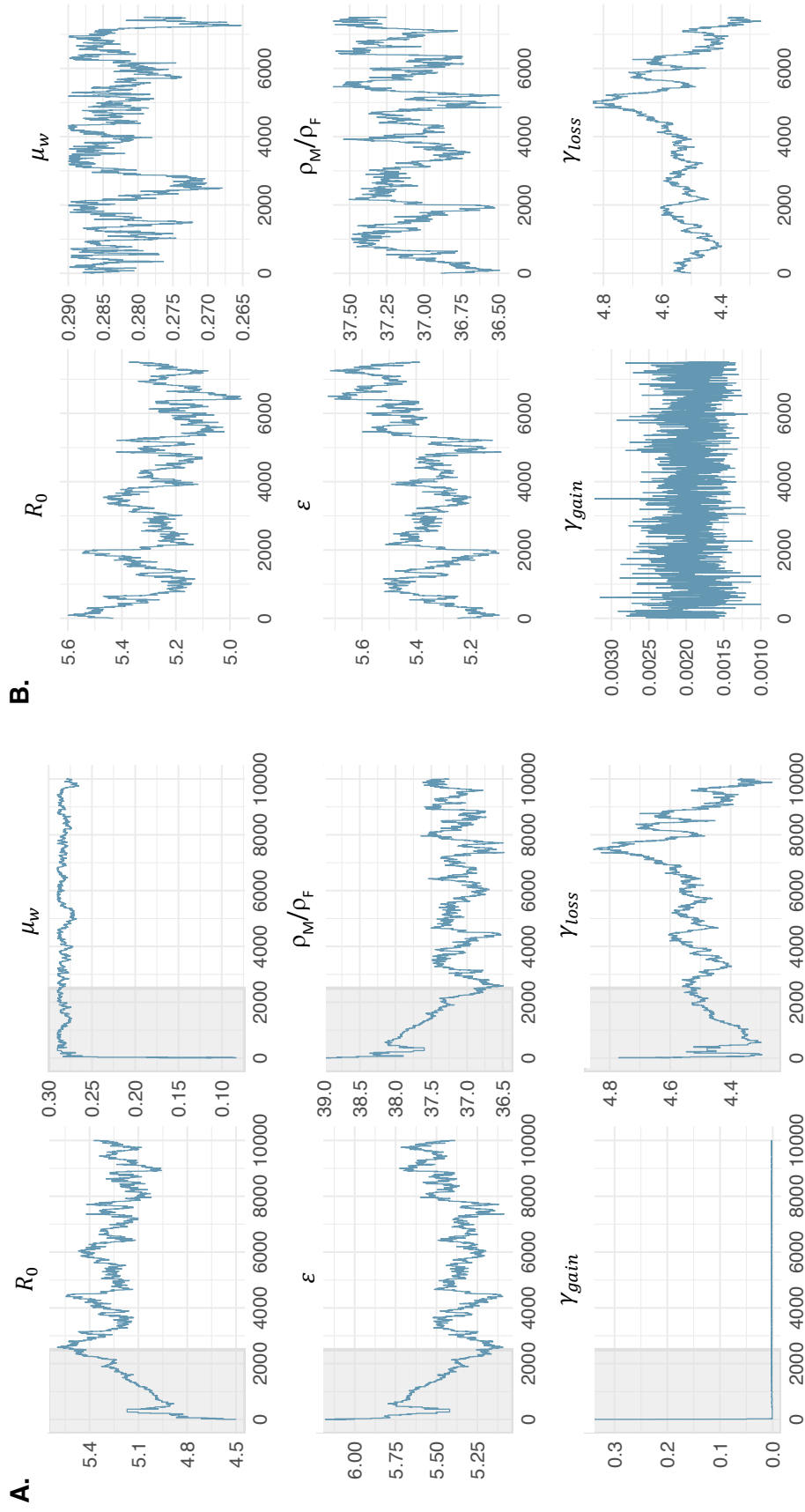


Fig. 3.12 Trace output for each parameter estimate from the MH-MCMC chain, run over 10,000 iterations for the Bagungu tribe A) steps 0 to 10000 and B) steps 2500 – 10000 (burn-in removed). The shaded region indicates the selected burn-in period 2500.

3.12 Discussion

It is imperative to quantify the benefit of large-scale MDA to the wider community, including to those who are untreated via reductions in environmental transmission. In order for these models to predict the requirements for mass-chemotherapeutic control it is imperative to study the aggregated contribution of exposure and immunity. The model presented in this chapter focused on within host dynamics of human schistosomiasis. The aim was to develop empirically informed cercarial exposure and immunity functions that provide a model framework, with parameters estimated by fitting to baseline, pre-treatment data.

3.12.1 The exposure function fits the data to good approximation

Understanding how contact patterns affect the dynamics of macroparasitic infections is one of the main challenges in modelling indirectly transmitted pathogens (Hollingsworth *et al.*, 2015). By including an empirically informed exposure function based on the cercarial exposure scores produced by Pinot de Moira *et al.* (2007), this model incorporates the heterogeneity that is observed in age related patterns of exposure by sex and tribe.

Pinot de Moira *et al.* (2007) concluded that the best fitting model to determine cercarial exposure scores was based on a site weighting dependent on the average number of infected snail intermediate hosts, identified during malacological surveys. In their analysis the authors found that the average number of infected snails was found to be associated with tribal differences in the intensity of reinfection at 12 months (Pinot De Moira *et al.*, 2007). The cercarial exposure scores with a site-weighting based on the number of infected snails were therefore used in the exposure function presented here.

Interestingly, in the earliest versions of the MWB model, the number of infected snails was found to be an important determinant in the level of schistosome endemicity (Macdonald, 1965; Barbour, 1978). The MWB model is commonly criticised for over-simplification of the life cycle and lack of consideration given to the transmission dynamics of the snail intermediate host (Gurarie, King and Wang, 2010). Although snail stages were not explicitly modelled here, rather than modelling variable exposure through water contact duration alone, consideration is given to snail transmission within the exposure function by accommodating site-weighted malacology survey and snail shedding data (Pinot De Moira *et al.*, 2007). The exposure function presented here is therefore distinct from previous models.

For simplicity, the same exposure function was used across all demographic groups with the functional form dictated by specific fitted parameter values that were subsequently fixed within the Bayesian model framework. This was reasonable for the Alur tribe, where the function can

be fitted to the male and female data to good approximation. The relative scale of exposure fitted in the MCMC framework however overestimates exposure in the youngest age groups. The model is fitted to infection intensity and serology data by host age. The overestimated exposure in Alur males may consequently result from the higher observed faecal egg counts in males compared to females. Moreover, the fitted male-specific exposure function (section 3.7.2; Fig. 3.2A) has a slightly flatter profile that accounts for the relatively lower cercarial exposure scores in male children compared to adults. This suggests that using sex-specific exposure function may marginally improve the Alur model fit.

Similarly, there is scope to improve the fitting process for the exposure function in the Bagungu tribe. The exposure patterns by age and sex are more complex in the Bagungu tribe with significant disparities in the exposure behaviour between males and females. The reasons for these differences are explored in detail elsewhere (Pinot de Moira *et al.*, 2010). Whilst a flat exposure profile may describe the observed data, fitting sex-specific functional forms to the data, rather than estimating the sex-specific parameters for the same function may provide a more satisfactory result. Incorporating the male:female relative exposure scaling parameter (ρ_M/ρ_F) into the pre-fitted function may also improve the fit of the exposure function to the data.

The results suggest that the intensity of infection seen in Bagungu males in particular is not explained by cercarial exposure alone, since high exposures do not result in high infection levels, indicating that there is potentially another factor controlling infection intensity.

3.12.2 The immunity function closely describes the protective effect of SmTAL1-specific IgE

In order to elucidate the entwined relationship between exposure and immunity in determining the age-intensity profile, it was important to use the available data to develop an informed immunity function. The immunity function describes the SmTAL1 specific IgE-reinfection probability data to good approximation in this study. Immuno-epidemiological evidence from other field studies however suggests that the relationship between SmTAL protein specific IgE and protective immunity may be more complicated than the relationship described in this chapter, with additional members of the TAL protein family potentially implicated in IgE mediated protective immunity (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). The relationship between IgE and the TAL protein family will be explored further in Chapter 4.

Gain and loss of immunity terms

The immunity gain and decay parameters were the least well described in the literature and, as a consequence, were modelled as uninformative, uniform priors. Considerable identifiability issues were found with respect to estimation of the immune decay parameter in particular. The choice of prior can significantly influence the resulting model fit. Due to uncertainty around the rate of loss of immunity, the uniform prior for γ_{loss} was rather broad. This incorporated the short estimates of serum IgE half-life (Vieira and Rajewsky, 1988) in addition to the potentially extended protective effect resulting from IgE memory (section 3.8.5). Extensive exploration of the parameter space may not have therefore been achieved within the 10,000 model iterations performed here. The immunity loss parameter is also modelled as a constant decay rate, which is likely to be a simplification of the true relationship.

It is likely that the duration of the immune response may significantly influence the age-intensity profile (Chan *et al.*, 1996), so studies that explore this relationship in more detail remain a research priority. This is not a novel assessment and this requirement has been highlighted by several authors previously (Oettle and Wilson, 2017). An unpublished intervention study conducted by the Cambridge Schistosomiasis Research Group in Kenya provides one such example. This study included annual treatment and was conducted in parallel with biannual mollusciciding of the River Kambu, the one source of transmission in the area (Vennervald *et al.*, 2004). Mollusciciding was timed to prevent known seasonal trends in transmission to limit re-infection levels (Kariuki *et al.*, 2013). Preliminary data from this study indicates that, in the absence of continued exposure, serum IgE declines in the two years from baseline. However, histamine release assays of basophil activity indicate that despite a decrease in SWA specific serum IgE, the IgE not only remains biologically active for at least two years in an environment where exposure has been significantly knocked down, but actually increases. This may reflect the importance of the blocking antibody SWA-IgG₄ which was shown to decrease substantially over the timeframe of the study. This would have implications for modelling the immune decay parameter, which would need to consider the difference between measurable serum IgE with continuous exposure, as is often the case in endemic environments, versus measurable serum IgE in a scenario where MDA has reduced transmission to the extent that exposure is intermittent. Others studies have also demonstrated maintenance of long-term IgE memory in the absence of detectable levels of serum IgE (Kubo *et al.*, 2003).

The immunity ‘gain’ parameter is also estimated as a constant conversion rate, fitted to ELISA OD values, since IgE titre field data extrapolated from standard curves was not available from the original Booma study. OD values are, however, assay specific and analysis based on OD

values may consequently result in batch effects. Future development of this function should aim to estimate parameters according to the fit of the model to IgE concentration, whereby ELISA OD values are translated to concentration, or titre, against a standard antibody concentration curve on each microtitre plate. This will improve generalisability to other studies.

Moreover, the relationship between exposure to dead worms and the antigens they release is unlikely to take the form of constant rate of antibody gain, proportional to the number of dead worms seen by the host. In reality, as suggested by the shape of the observed age-serology profile, the rate at which hosts acquire antigen-specific IgE may decrease with cumulative exposure to dead worms. An exponential decrease in the rate of gain, for example, would mean that estimates of the longevity of the IgE response would increase, thus the rate of decay (γ_{loss}) would decrease. It is therefore possible that the current model overestimates the rate of antibody decay. Another theory is that hosts may reach a threshold of exposure to dead worms, whereby any further exposure does not necessarily result in an increase in protection.

As mentioned above, it has been suggested that other antibody isotypes, such as IgG₄ may act to block the IgE response (Hagan *et al.*, 1991; Aalberse *et al.*, 2009). Levels of specific IgG₄ antibodies tend to be higher in children and decline with age. It is thought that the blocking effect of IgG₄, through competition for epitope binding sites may partially explain the observed delay in the development of IgE mediated protective immunity. There is scope to further develop the existing immunity function following validation of this response against additional studies. Incorporating a delayed protective effect of immunity with age, or protective antibody threshold, may better represent the peaked infection profile observed in the data.

3.12.3 Model convergence and sensitivity analysis

The random walk appearance of the trace plots, as opposed to a tight horizontal band implies that there may be some identifiability issues in the model, especially for the Bagungu tribe where the relationship between exposure, immunity and infection intensity is more complicated. This may be addressed through increasing the numbers of iterations, for example to 100,000 iterations, of the MH-MCMC algorithm, to identify whether this is an actual issue with parameter identifiability or whether increased sampling alone could improve convergence. Using more informative priors may also enable the model to converge on appropriate parameter estimates more quickly.

Computing capacity was a limitation in being able to perform these checks in this thesis, due to long model running times. High performance computing (HPC) facilities are available at the University of Cambridge, future analysis will be performed using the University's HPC cluster.

Reliability is a measure of how reproducible the model output is. Running the model on the HPC cluster will also enable a full sensitivity analysis to be performed. Sensitivity analysis, running multiple chains with different initial parameters, will confirm whether the model converges on the same parameter estimates, that is, whether the results are reproducible. Previous schistosome transmission modelling studies have found the model to be particularly sensitive to worm lifespan (French *et al.*, 2015). A low estimate of the worm lifespan (high worm mortality rate) may result in an early peak in infection intensity, similar to that seen in the model presented here. Sensitivity analysis should also be performed by varying the distribution of the worm lifespan, via the number of worm compartments.

3.12.4 The Bagungu model demonstrated the need for the inclusion of heterogeneity and immunity in the model

In contrast to the observed epidemiological profiles of the Alur tribe, the age-intensity profile of the Bagungu tribe does not directly reflect the exposure profile. This has enabled the current model to demonstrate the importance of including an empirically defined immunity function.

Whilst it is clear that there is scope for improvement in the fit of the model, the results presented here do appear to support the contribution of an empirically informed immunity function to the peaked infection intensity profile. Including heterogeneity in exposure, alongside an explicitly parameterised immunity function, has highlighted how higher levels of serum IgE appear to influence the shape of the immunoepidemiological curves, with higher serum IgE resulting in a sharper decline in infection intensity with age and very low levels of serum IgE resulting in an infection intensity curve that reaches a plateau but does not subsequently decline.

It is widely recognized in the helminth literature that ‘wormy’ people are predisposed to heavy infection (Schad and Anderson, 1985): this predisposition is under a combination of immune versus behavioural control. The analysis of these heterogeneous processes is therefore vital in understanding the dynamics of schistosome transmission. The inclusion of both heterogeneity in exposure and immunity (Chan *et al.*, 1996) have been demonstrated to result in increased estimates for R_0 . R_0 is an important parameter in determining the ability of an infection to spread through a population. Greater heterogeneity within the Bagungu tribe compared to the Alur may contribute to the higher estimate for R_0 in the Bagungu.

Furthermore, the inclusion of heterogeneity in the model dictates the need for increased sample sizes. The number of individuals within each demographic group is sizeably reduced when separated in the analysis, thus limiting the power for examining profiles of infection in separate demographic groups, for example by age, sex and ethnicity.

The model assumes a stable population with no migration. However, since this study was conducted there has been a migration of the Bagungu population further inland away from the village and Lake Albert, with a shift in focus towards agricultural employment (D'Udine *et al.*, 2015). Application of this model to present day transmission would therefore need to consider this change to the demographic profile of the population.

3.12.5 Future adaptations of the model

Following a shift in the terminology of schistosomiasis control literature from that of morbidity control and elimination as a public health problem (World Health Organization, 2012), to a focus on breaking transmission, a number of models attempted to predict the requirements and impacts of MDA in reaching proposed elimination targets (Mitchell *et al.*, 2014; Anderson *et al.*, 2015; Gurarie *et al.*, 2015; Truscott *et al.*, 2017). There is scope to further develop the model presented here by incorporating treatment strategies, using district-level data from MDA programmes in Uganda.

Several authors suggest that the presence of immunity may reduce the effectiveness of control (Barbour, 1985; Chan, 1996). Furthermore, when acquired immunity has a strong effect at both the individual and community level, it is thought that withdrawing control interventions may have a substantial impact on the rebound of schistosome infection (Anderson and May, 1985; Chan, 1996). This provides further support for the inclusion of appropriate immunity function in schistosome transmission models. The model here assumes that the total cohort immunity function captures the dynamics of immunity across the different demographic groups. The sample size of the disaggregated data by age did not allow for the immunity function to be fitted independently for each age group. The CSRG has access to several historical datasets including individual level *S. mansoni* parasitology and paired serology samples. It is proposed that these data be used in future studies in order to validate the immunity function presented here. The effector mechanism of immunity is also likely to be more complicated than the function presented in this model, as discussed in section 3.12.2. It is therefore important to advance our existing knowledge of protective immunity against schistosome infection, specifically tailoring this research towards the application of mathematical models. This chapter has made the first steps towards achieving this but highlights the need for optimisation of the immunity function.

Worm pairing, or mating probability, is not considered in this model, since it is assumed that at equilibrium infection intensity will be at such a level that female worms are likely to be paired. The model does not therefore describe the dynamics of a low transmission system. The inclusion of specific worm pairing function may need to be revised in future versions of the model where the inclusion of MDA may reduce infection intensity to a level where mating probability becomes relevant (May, 1977; Bradley and May, 1978). Furthermore the inclusion of

heterogeneity is of particular relevance as sustained MDA programmes have been shown to be successful in reducing the prevalence of schistosomiasis (Deol *et al.*, 2019) highlighting the importance of infection hotspots, where infection intensity remains high.

There are substantial aspects of parameter uncertainty in the model, which are in part addressed through random sampling of relevant parameter ranges by MH-MCMC. However, there is a need for further refinement of the MH-MCMC sampling and sensitivity analysis to ensure the model is converging and producing reliable results.

This model provides a methodological framework for incorporating immunity into mathematical transmission models of *S. mansoni*, specifically. Application of this model to other schistosome species will require species specific validation of the immunity function.

Several authors have compared the ability of different models to reproduce the epidemiological profiles observed in field data. These models are freely applied to both *S. haematobium* and *S. mansoni*, without close consideration being given to species specific modifications (Truscott *et al.*, 2017). For example, the ecology of the snail vector for *S. haematobium* differs from that of *S. mansoni*, with seasonal variations in habitats and human hosts sometimes using multiple different water sources in contrast to *S. mansoni* transmission sites along the African Great Lakes. This demonstrates important differences that must be taken into consideration when modelling the different schistosome species. Moreover, there is evidence to support two distinct immunological mechanisms regulating *S. haematobium* transmission, anti-infection immunity and anti-fecundity immunity. It is questionable that models calibrated for *S. mansoni* transmission dynamics can be applied to *S. haematobium* without significant modification to account for these different processes, and vice versa. The immune processes observed in *S. haematobium* will be examined in more detail in subsequent chapters.

3.12.6 Concluding remarks

The model presented here addresses the need for a model that incorporates both exposure and immunity (Anderson *et al.*, 2016), improving upon the structure of existing models by: 1) using field data to empirically inform the functional forms of exposure and immunity; 2) accounting for uncertainty using MCMC methods for parameter estimation. The model is calibrated to pre-treatment, pre-MDA infection intensities and provides a baseline framework to which treatment can be added in future studies.

The data available for analysis in this chapter have enabled significant improvement to exposure and immunity parameterisation, however there are still limitations in the data. To examine heterogeneity in exposure, immunity and infection and serology profiles, it is necessary to examine the respective

characteristics of different demographic groups, yet it is difficult to also maintain sufficient sample sizes in doing so. The reduced sample sizes that result from division of the total cohort by tribe, sex and age mean that there is uncertainty around how accurately the sample data represent the epidemiology of schistosome infection and subsequent antibody-mediated protective immunity in the entire population. Thus, highlighting the need for future studies to account for heterogeneity in exposure and immunity by ensuring that the total cohort sample size is powered sufficiently to accommodate downstream analysis of the separate demographic sub-cohorts. This is of particular importance when exploring the dynamics of immune processes where seroprevalence is low, such as that observed in some TAL-specific IgE antibody responses.

There is increasing focus on the need to understand why hotspots of high intensity schistosome infection persist, despite long-term MDA activities. It is especially important in these scenarios to incorporate both exposure and immunity into mathematical transmission models in order to better understand the mechanisms that contribute to transmission heterogeneity. The model presented in this chapter provides a significant advance towards understanding the relative contribution of exposure and immunity to *S. mansoni* transmission.

4

The Tegument Allergen Like protein family in *Schistosoma mansoni* and *Schistosoma haematobium*: IgE responses and reinfection

Preface

Current understanding supports a role for Tegument Allergen-Like 1 (TAL1) protein specific IgE as a marker for acquired partial immunity against reinfection in *S. mansoni*, *S. haematobium* and *S. japonicum*. It is proposed that worm death is a prerequisite for the development of specific antibody responses via delayed concomitant immunity and that conferred protection acts against the invading cercariae, limiting the establishment of new infections. Further evidence in *S. mansoni* suggests that additional members of the TAL protein family may contribute to the protective immune response, yet this has not been replicated in other studies, or in the other *Schistosoma* species.

Acknowledgement of contributions

Dr Anna Protasio (University of Cambridge) performed manual curation of the *S. haematobium* genome for potential ShTAL protein sequences to supplement the BLAST searches that were otherwise conducted. Dr Iain Chalmers (University of Aberystwyth) provided useful discussion and advice regarding bioinformatic analysis of ShTAL proteins.

The ELISA data used for the analysis were largely sourced from archived studies conducted by the Cambridge Schistosomiasis Research Group (CSRG) with the exception of SmTALII and ShTAL5. Jakub Wawrzyniak (University of Cambridge) performed the SmTALII ELISA assay using plasma samples from the male cohort from Musoli, Uganda.

4.1 Introduction

4.1.1 Introduction to the TAL protein family

The TAL proteins form a family of calcium binding proteins unique to the platyhelminths (Thomas *et al.*, 2014). Calcium channel signalling is essential to several biological functions in

eukaryotes (Berridge, Lipp and Bootman, 2000) and, as result of their broad involvement across various signalling and regulation pathways, calcium binding proteins have historically garnered interest as potential vaccine targets (Pacífico *et al.*, 2006). Whilst IgE stimulating antigens are not considered to make the most appropriate vaccine candidates, due the association between IgE and anaphylactic responses (Diemert *et al.*, 2012), improving our knowledge of the proposed partial protective immune response is essential for improving our understanding of transmission dynamics and subsequent planning of infection control measures.

The TAL proteins have high levels of sequence and predicted structural similarity, typically consisting of a N-terminal domain containing two EF-hand-like structures and a C-terminal DLC domain (Fitzsimmons, Jones, Stearn, *et al.*, 2012; Thomas *et al.*, 2014). Calcium-binding allergens have been described in many tree pollens and grasses and several environmental allergens are known to show structural homology to TAL proteins (Farnell *et al.*, 2015). EF-hand motifs, in particular, are a feature of many IgE antigens in both plants and animals (Valenta *et al.*, 1998) and EF-hand proteins comprise one of the largest allergenic protein domain families (Moncrief, Kretsinger and Goodman, 1990).

Proteins within the TAL family have been characterised in several *Schistosoma* species (Dunne *et al.*, 1992; Santiago *et al.*, 1998; Fitzsimmons *et al.*, 2004; de la Torre-Escudero *et al.*, 2012; Carson *et al.*, 2018) in addition to other parasitic flatworms, including *Fasciola hepatica*, *Fasciola gigantica*, *Clonorchis sinensis* and *Opisthorchis viverrini* (Kim *et al.*, 2012; Orr *et al.*, 2012; Senawong *et al.*, 2012; Subpipattana, Grams and Vichasri-Grams, 2012). The first members of the TAL family to be independently identified in *S. mansoni* were SmTAL1 (Dunne *et al.*, 1992), SmTAL2 (Francis and Bickle, 1992), and SmTAL3 (Hoffmann and Strand, 1997), previously named Sm22.6, Sm21.7 and Sm20.8, respectively. These proteins have since been characterised in depth (Fitzsimmons, Jones, Stearn, *et al.*, 2012; Thomas *et al.*, 2014), along with recent molecular and biochemical characterisation of the ten additional members of the family that were subsequently identified (SmTAL4-13) (Carson *et al.*, 2018).

4.1.2 TAL proteins and reinfection

Our understanding of the processes governing reinfection are largely informed by evidence from immuno-epidemiological studies, primarily looking at responses in *S. mansoni*, which suggest that TAL proteins, specifically TAL1, are strongly associated with IgE-mediated immunity (Webster *et al.*, 1996; Fitzsimmons, Jones, Stearn, *et al.*, 2012). The level of detectable *S. mansoni* TAL1 (SmTAL1)-specific serum IgE in infected individuals increases with age, in turn corresponding to a decrease in infection intensity (Fitzsimmons, Jones, Stearn, *et al.*, 2012).

There is evidence to suggest that a similar partially protective immune response also occurs in other *Schistosoma* species that infect humans, particularly with respect to TAL1 (Santiago *et al.*, 1998; Fitzsimmons *et al.*, 2004).

Characterisation of *S. haematobium* TAL1 (ShTAL1) has shown that it is closely related to the orthologous *S. mansoni* and *S. japonicum* proteins (Fitzsimmons *et al.*, 2004); however, little is currently known about the remaining members of the ShTAL family. Prior to the publication of the *S. haematobium* genome, Dickinson identified the ShTAL family of proteins from expressed sequence tag (EST) data (Dickinson, 2012); yet, the identity of these proteins has not been verified against the published genome. Furthermore, the relationship between TAL-specific antibodies and the acquisition of protective immunity has not yet been clearly defined in *S. haematobium*.

It has since been proposed that additional members of the TAL protein family may stimulate the protective response via cross-reactive antigen recognition (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). mRNA expression data suggests that each TAL protein is expressed at varying levels across the different life cycle stages. SmTAL1, SmTAL3 and SmTAL11 are largely expressed in the adult worm, whereas SmTAL4 and SmTAL5 are predominantly expressed in the cercaria. Whole-mount immunostaining of cercarial sections demonstrates that SmTAL4 expression appears to be restricted to the tail of the cercaria (Fitzsimmons, Jones, Stearn, *et al.*, 2012), which separates from the cercarial head during skin penetration. SmTAL4 is therefore not indicated as an immune stimulatory protein or a target of the protective response. Conversely, SmTAL5 is expressed throughout the cercaria and early schistosomula and may therefore be exposed to the immune system during invasion and early transformation. Several studies have provided evidence for strong cross-reactivity across the IgE EF-hand family, with a large number of studies demonstrating cross-sensitisation between environmental and food allergens (Tinghino *et al.*, 2002; Verdino *et al.*, 2008; Lopata, Kleine-Tebbe and Kamath, 2016). Evidence suggests that SmTAL5 exhibits IgE cross-reactivity with worm-expressed SmTAL3, with SmTAL3 appearing to be the antigenic source for this IgE response (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). It is this cross-reactive relationship that is proposed to drive delayed concomitant immunity. Since SmTAL11 mRNA is mainly expressed in adult worms (Fitzsimmons, Jones, Stearn, *et al.*, 2012), it is hypothesized that SmTAL11 may also induce specific IgE following worm death, similar to that seen for SmTAL1 and SmTAL3 and may subsequently contribute to the development of delayed concomitant immunity.

Although analysis by Dickinson (2012) suggests that the expression profile of *S. haematobium* differs slightly from that of *S. mansoni*, life stage expression of ShTAL1, 3, 5 and 11 (Fig. 4.1) is

consistent with the hypothesis that antigen specific IgE, cross-reactive to ShTAL5 may also be the driver of partial protective immunity in *S. haematobium*.

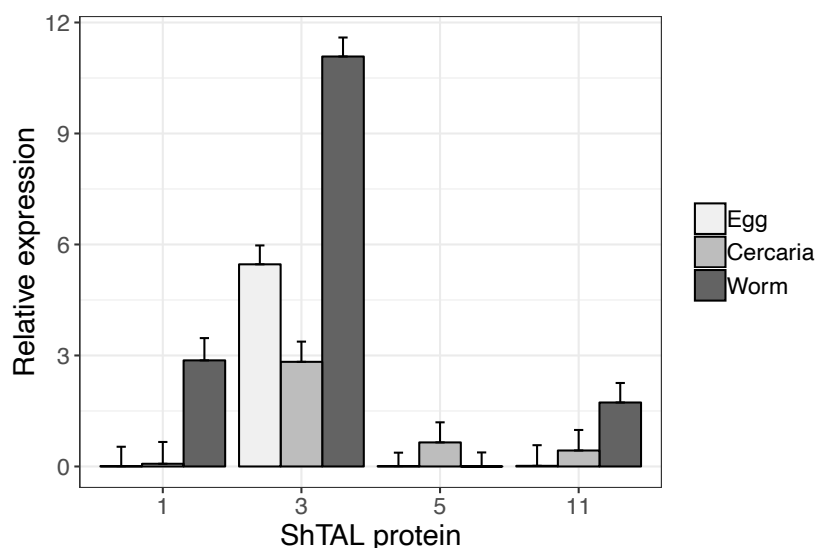


Fig. 4.1 Relative lifecycle stage expression of the *S. haematobium* TAL1, TAL3, TAL5 and TAL11 proteins. Figure produced from existing qPCR data (Dickinson, 2012), where relative expression was calculated using the delta-delta Ct method: Fold gene expression = $2^{-(\Delta\Delta Ct)}$, where $\Delta\Delta Ct = \Delta Ct (\text{Sample}) - \Delta Ct (\text{Control average})$, details of the samples and housekeeping controls can be found in Dickinson (2012).

4.1.3 Chapter aims

The focus of this chapter is to determine whether the ShTAL proteins show a similar mechanism of antibody cross-reactivity to proteins expressed in the invading parasite. Analysis and discussion will therefore centre around those TAL proteins with a previously identified association with protection in *S. mansoni* and those with close phylogenetic affinity, specifically answering the following question:

Does the anti-infection protective response to *S. haematobium* demonstrate the same dependence on TAL-specific IgE as it does to *S. mansoni*?

This will be achieved through the following objectives:

- Identify possible ShTAL proteins from the published *S. haematobium* genome.

- Perform bioinformatic analysis to compare identified ShTAL proteins with recognised SmTAL family members.
- Clone and express recombinant SmTAL11 and explore the role of TAL11 as a potential contributor to the development of delayed concomitant immunity in *S. mansoni* infection.
- Examine the serological profile of recombinant ShTAL1, ShTAL3, ShTAL5 and ShTAL11 and examine the role of these proteins in the development of delayed concomitant immunity in *S. haematobium* infection.

4.2 Results

4.2.1 Identification of *S. haematobium* TAL proteins

Nucleotide and amino acid sequences from the well-characterised *S. mansoni* TAL proteins (Fitzsimmons, Jones, Stearn, *et al.*, 2012) were used in BLAST searches against the annotated *S. haematobium* genome (WormBase parasite BioProject PRJNA78265), as outlined in Chapter 2, section 2.9.1. Prior to annotation of the genome, similar sequence searches were performed against the *S. haematobium* EST database (Dickinson, 2012), but these have not yet been verified against the published genome.

Protein hits that had at least one EF-hand domain at the N-terminal end, a dynein light chain (DLC) domain at the C-terminal end of the protein and had a significance cut-off of E-value $<1e^{-5}$ were considered potential members of the TAL protein family, resulting in 13 proteins with the EF-hand and DLC domains characteristic of the TAL protein family. These were orthologous to the sequences of SmTAL1 – SmTAL5, SmTAL7 – SmTAL13. The BLAST searches also identified an additional, previously uncharacterised, *S. mansoni* protein (Smp_146460; XP_018654160), with an *S. haematobium* orthologue with 85% homology (MS3_08443; A_04044). This potential additional member of the TAL family was also discovered during searches of the *S. haematobium* EST data by Dickinson (2012). A *S. haematobium* sequence orthologous to SmTAL6 was not found using the defined selection criteria. Therefore, those sequences that did not meet the original selection criteria were subject to an inverse BLASTP search back against the *S. mansoni* genome (WormBase parasite BioProject PRJEA36577). This identified a 145 amino acid DLC protein sequence with 77.2% homology to SmTAL6 (Smp_072620.1).

The sequence results from the manual curation of the *S. haematobium* genome concurred with those from the BLAST search, including the identification of a potential new TAL family member, ShTAL14 (MS3_08443).

Table 4.1 *S. haematobium* candidate TAL proteins

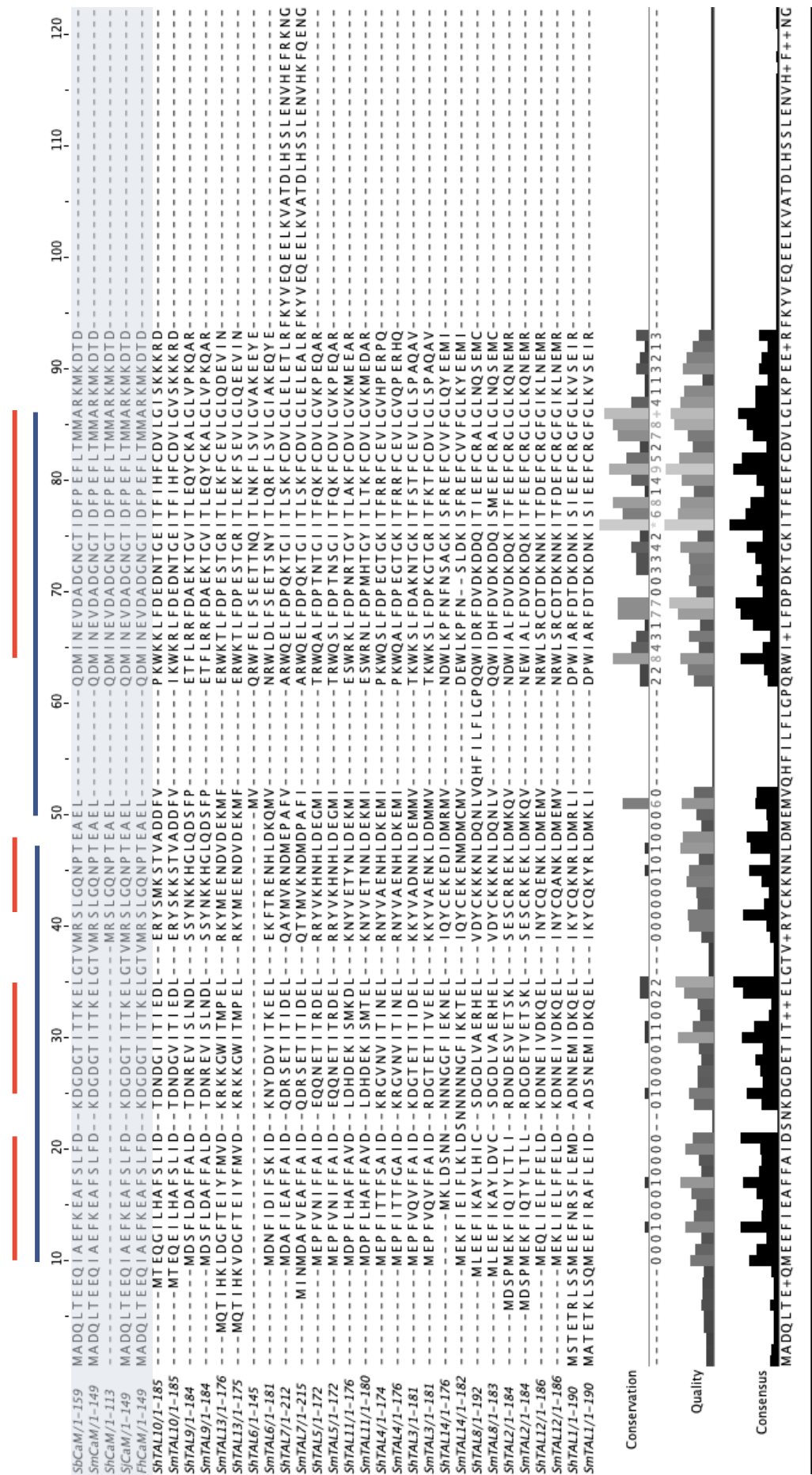
Name	NCBI Accession	Gene	Size (aa)	MW (kDa)	<i>S. mansoni</i> orthologue	E-value	%ID	EF-hand	DLC
ShTAL1	XP_012799745	MS3_08446 (A_08162)	190	22.6	Smp_045200	1.80E-118	87.4	pair	yes
ShTAL2	XP_012797374	MS3_05958 (A_03381)	184	21.7	Smp_086480	2.10E-115	91.3	pair	yes
ShTAL3	XP_012797371	MS3_05952 (A_07883)	181	20.7	Smp_086530	1.40E-111	87.8	pair	yes
ShTAL4	XP_012797370	MS3_05954 (A_03557)	176	20.9	Smp_169190	2.60E-119	95.4	pair	yes
ShTAL5	XP_012797368	MS3_05957 (A_04955)	172	20.6	Smp_195090	6.80E-118	96.5	pair	yes
ShTAL6	XP_012796626	MS3_05176 (A_04145)	145	17.1	Smp_072620	3.2E-78	77.2	no	yes
ShTAL7	XP_012795331	MS3_03822 (A_04865)	212	25	Smp_042140	2.40E-58	90.7	pair	yes
ShTAL8	XP_012797375	MS3_05960 (A_04471)	192	22.7	Smp_086470	2.10E-91	91.7	pair	yes
ShTAL9	XP_012797728	MS3_06334 (A_00804)	184	21.7	Smp_077310	4.00E-111	95.1	pair	yes
ShTAL10	XP_012795756	MS3_04275 (A_05864)	185	21.4	Smp_074460	5.60E-120	87.6	pair	no
ShTAL11	XP_012797369	MS3_05959 (A_05160)	176	20.9	Smp_169200	5.10E-105	82.4	pair	yes
ShTAL12	XP_012799748	MS3_08444 (A_06931)	186	22.1	Smp_045010	7.10E-120	91.4	pair	yes
ShTAL13	XP_012795330	MS3_03823 (B_00228)	175	21	Smp_042150	1.30E-117	93.1	pair	yes
ShTAL14	XP_012799749	MS3_08443 (A_04044)	176	20.8	Smp_146460	1.80E-70	85.1	pair	yes

4.2.2 Multiple sequence alignment of *S. haematobium* TAL proteins

Alignment of the candidate TAL proteins was performed using Multiple Sequence Comparison by Log-Expectation (MUSCLE) (Edgar, 2004). Calmodulin representatives from *S. mansoni*, *S. haematobium* and other closely related species, which contain EF hand domains were used as a reference outgroup. All proteins, with the exception of SmTAL10, ShTAL10 and ShTAL6, had an EF-hand pair domain (InterPro accession: IPR002048) at the N-terminal and DLC domain (InterPro accession: IPR001372) at the C-terminal.

The predicted ShTAL6 was lacking the N-terminal EF-hand domain, as described in section 4.2.1. As reported for SmTAL10 (Fitzsimmons, Jones, Stearn, *et al.*, 2012), the orthologous *S. haematobium* TAL10 sequence, identified by BLAST searches, does not contain a DLC domain at the C-terminal end. Despite the lack of a characteristic DLC domain, SmTAL10 was considered a member of the TAL protein family by Fitzsimmons *et al.* (2012) on the grounds that BLAST searches against the *S. mansoni* genome (taxid: 6183) showed that the SmTAL10 EF-hand region (residues 1 - 72) showed the greatest similarity to the equivalent domain region of several proteins from the TAL protein family (SmTAL3 and SmTAL7), as compared to all other predicted gene products. Similarly, BLAST searches of the ShTAL10 sequence, and specifically the EF-hand domain region (residues 1 - 70, comprising EF-hand 1: 1 - 36 and EF-hand 2: 45 - 70), against the *S. haematobium* genome (taxid: 6185) showed that ShTAL10 is most similar to ShTAL11, ShTAL5, ShTAL4, ShTAL3 ShTAL7 ($<E^{-9}$), as compared to all *S. haematobium* predicted gene products. The multiple sequence alignment for the SmTAL proteins and candidate ShTAL proteins is shown in Figure 4.2 with EF-hand and DLC domains highlighted.

The Gblocks selection, described in Chapter 2, section 2.9.2, using the least stringent block selection settings identified eight conserved blocks, comprising 119 amino acid positions (46% of the original 254 positions). The smallest block comprised five amino acids (Fig. 4.2). All conserved blocks were found to align with either the EF hand pair or DLC domain regions. The resulting blocks were concatenated into a single alignment that was used for phylogenetic analysis.



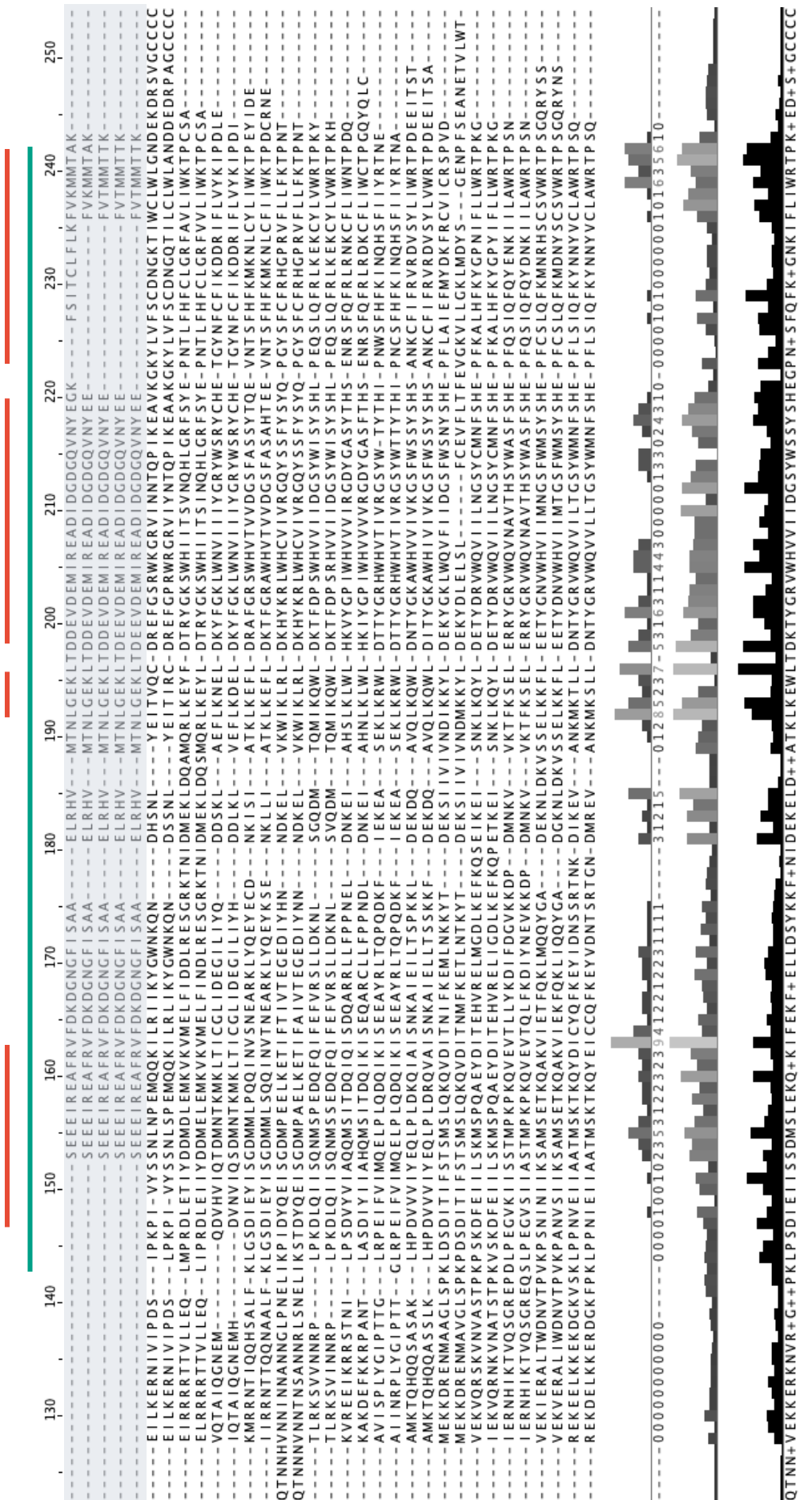


Fig. 4.2 Alignment of amino acid sequences of the *S. mansoni* and likely *S. haematobium* TAL protein family performed using MUSCLE (Edgar, 2004). Alignment of various calmodulin representatives is shown in grey box. The full sequence alignment was analysed using GBLOCKS, the resulting conserved blocks are highlighted in red. The conserved regions have similar or identical amino acid sequences and are generally found within the predicted TAL protein EF-hand pair (marked in blue) and predicted DLG region (marked in green). Accession numbers can be found in Table 4.1

4.2.3 Phylogenetic analysis of ShTAL proteins

To identify whether the ShTAL protein phylogenetic topology reflects that of *S. mansoni*, Maximum Likelihood trees were constructed for the full protein alignment and for the highly conserved regions detected by Gblocks analysis (Fig. 4.3). The tree inferred from the refined alignment improved the Log Likelihood (Gblocks Log Likelihood -4324.81; full alignment Log Likelihood -6736.76). The Gblocks analysis shows two main clades into which the TAL protein family is divided, one clade including *S. mansoni* and *S. haematobium* TAL3, 4, 5, 6, 7, 9, 11 and 13 proteins (Fig. 4.3B, Group 1), whilst the other includes the remaining *S. mansoni* and *S. haematobium* TAL1, 2, 8, 12 and 14 (Fig. 4.3B, Group 2). SmTAL10 and ShTAL10 however, appear to demonstrate more distant phylogenetic affinity.

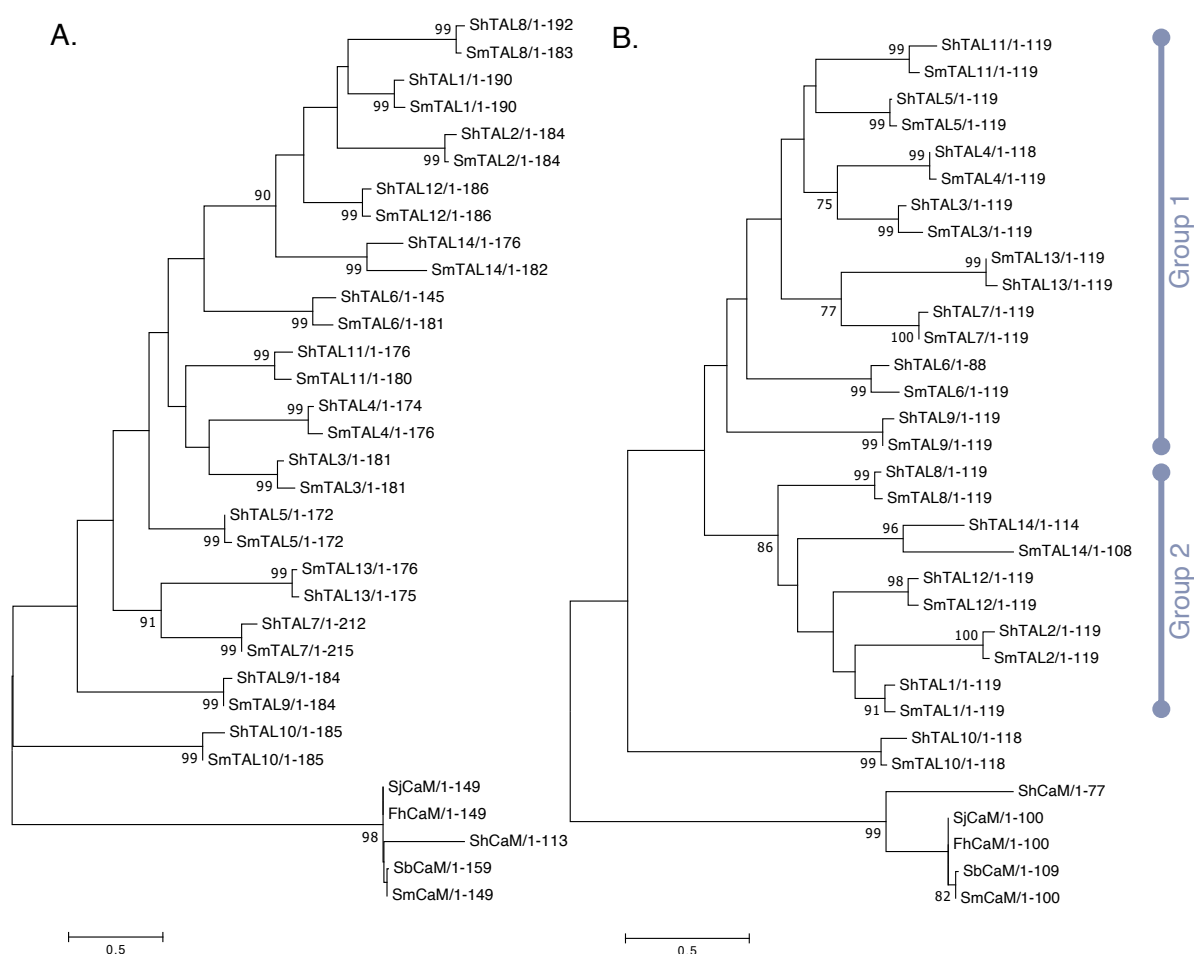


Fig. 4.3 Protein sequence analysis of the *S. mansoni* and *S. haematobium* TAL family. For A) the full protein alignment and for B) the highly conserved regions detected by Gblocks analysis. The evolutionary history was inferred using the Maximum Likelihood method based on Jones-Taylor-Thornton model. Analysis was conducted as described in Chapter 2. Calmodulin (CaM) is a highly conserved calcium binding protein that ubiquitously expressed in eukaryotes. Various representatives from the *Schistosoma*, or closely related, species were included in the analysis as an outgroup. The figure displays the bootstrap consensus tree (1000 replicates), drawn to scale, with branch lengths representing the evolutionary distances used to infer the phylogenetic tree (number of substitutions per site). Bootstrap values with a cut-off of 75% are displayed.

In previous analysis of the *S. mansoni* study population from Musoli, Fitzsimmons *et al.* (2012) found that individuals with a detectable SmTAL5-specific IgE response were a subset of those seropositive for SmTAL1 and SmTAL3 specific IgE. The analysis presented here suggests that SmTAL11 may be more closely related to SmTAL5 than SmTAL3 in phylogeny (Fig. 4.3B, Group 1). This chapter therefore proposes that SmTAL11-specific antibody responses may also contribute to the development of protective immunity. The specific IgE responses to SmTAL1, 3, 5 and 11 were explored in relation to their epidemiological profiles and association with reduced reinfection two years after treatment. IgE responses to the equivalent *S. haematobium* TAL proteins were also examined.

4.2.4 Cloning, expression and purification of TAL proteins

The SmTAL11 DNA sequence was successfully amplified from *S.mansoni* cDNA (Fig. 4.4A). The SmTAL11 PCR product was ligated into the expression vector and transformed into TG2 *E.coli* (as described in section 2.10.1). PCR analysis of the recombinant clones from five selected colonies is presented in Figure 4.4B. The sequence was confirmed to be identical to that published in GenBank (Gene ID: 8347704).

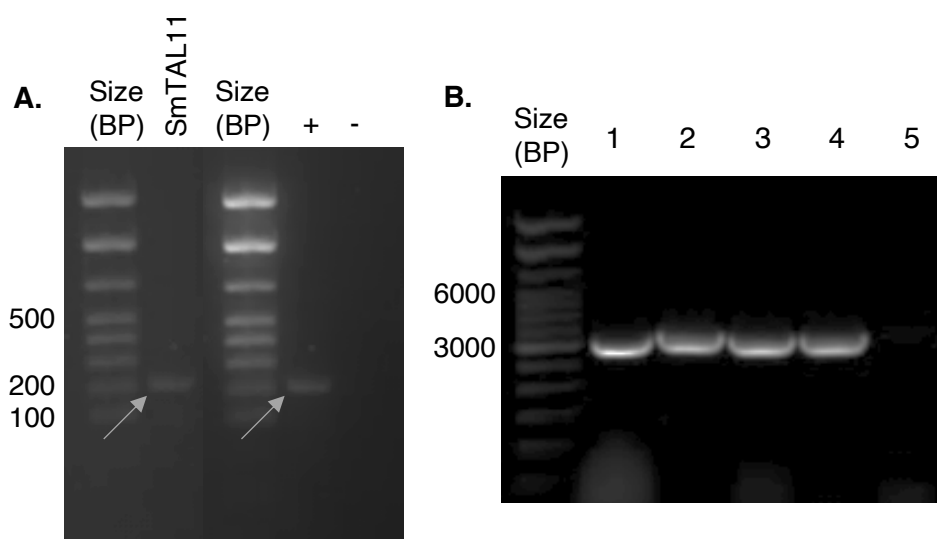


Fig. 4.4 Agarose gel of cloned SmTAL11 A) SmTAL11 PCR product; SmTAL2 positive control (+), and no template negative control (-) (GeneRuler 100 bp DNA ladder, Thermo Scientific, Hemel Hempstead, UK). B) Agarose gel of analysis of recombinant *E. coli* clones. Clones 1 – 4 successfully expressed the SmTAL11 sequence, demonstrating expression construct bands of the correct size (GeneRuler 1kb DNA ladder, Thermo Scientific, Hemel Hempstead, UK).

All *S. mansoni* and *S. haematobium* TAL proteins required for the subsequent analysis were successfully expressed in and purified from TG2 *E. coli* cells, resulting in good protein yield (SmTAL1 [0.87 mg/ml], SmTAL5 [0.66 mg/ml] and SmTAL11 [0.21 mg/ml]; ShTAL1 [0.36 mg/ml], ShTAL 3 [0.48 mg/ml], and ShTAL 11 [0.57 mg/ml]) (Fig. 4.5). The yield achieved for SmTAL11 was comparatively low, but was deemed satisfactory, by coating and contamination assays. SmTAL3 (0.70 mg/ml) and ShTAL5 (0.36 mg/ml) were available from existing recombinant protein stocks made by previous members of the research group.

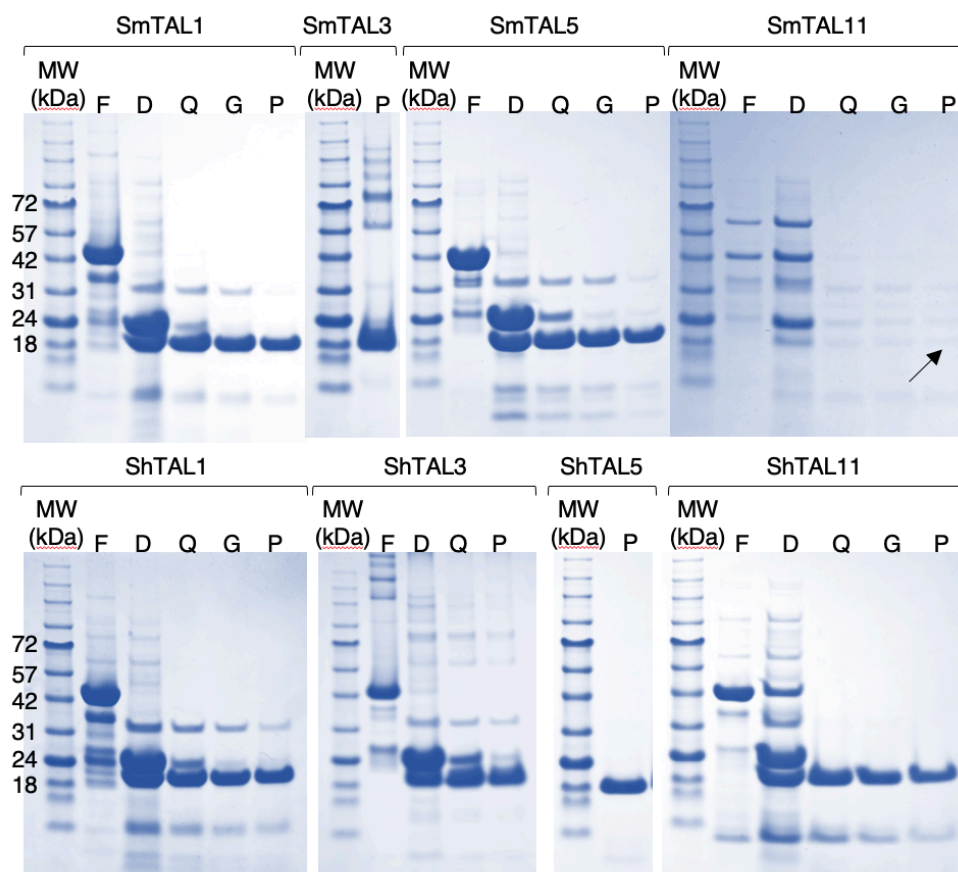


Fig. 4.5 Gel electrophoresis of purified SmTAL and ShTAL proteins. Coomassie blue-stained SDS-PAGE gel of each recombinant *S. mansoni* and *S. haematobium* TAL1, TAL3, TAL5 and TAL11 including fusion protein (F), thrombin digest (D) and subsequent purification steps (Q: Q-sepharose bead purification; G: glutathione sepharose bead purification; P: purified protein following aminobenzamidine agarose bead purification).

4.2.5 Coating and contamination assays

Assays were performed to determine the coating concentration and purity of each recombinant TAL protein, as described in Chapter 2, section 2.15.1. Following coating of the antigen at a gradient of concentrations the microtitre plates (MTPs) were incubated with HRP-conjugated anti-rat antibody, which subsequently binds to any available binding surface of the well that is

not coated by antigen. The antigen coating concentration required for further serological analysis was therefore selected according to the concentration that minimised the anti-rat-HRP signal (blue), whilst also ensuring that signal from contaminant GST (red) and bacterial lysate (orange) was minimised. This ensured that the total available binding surface of the well was coated by antigen. Anti-SmTAL sera from inoculated rats was available for SmTAL1, SmTAL3 and SmTAL5 (a gift from Dr Jamal Kalife, Institut Pasteur de Lille to Cambridge Schistosomiasis Research Group). Therefore, a coating concentration that meets the previously described requirements in addition to maximising anti-TAL serum signal was selected for these proteins. Coating concentrations for downstream serology are indicated in Figure 4.6.

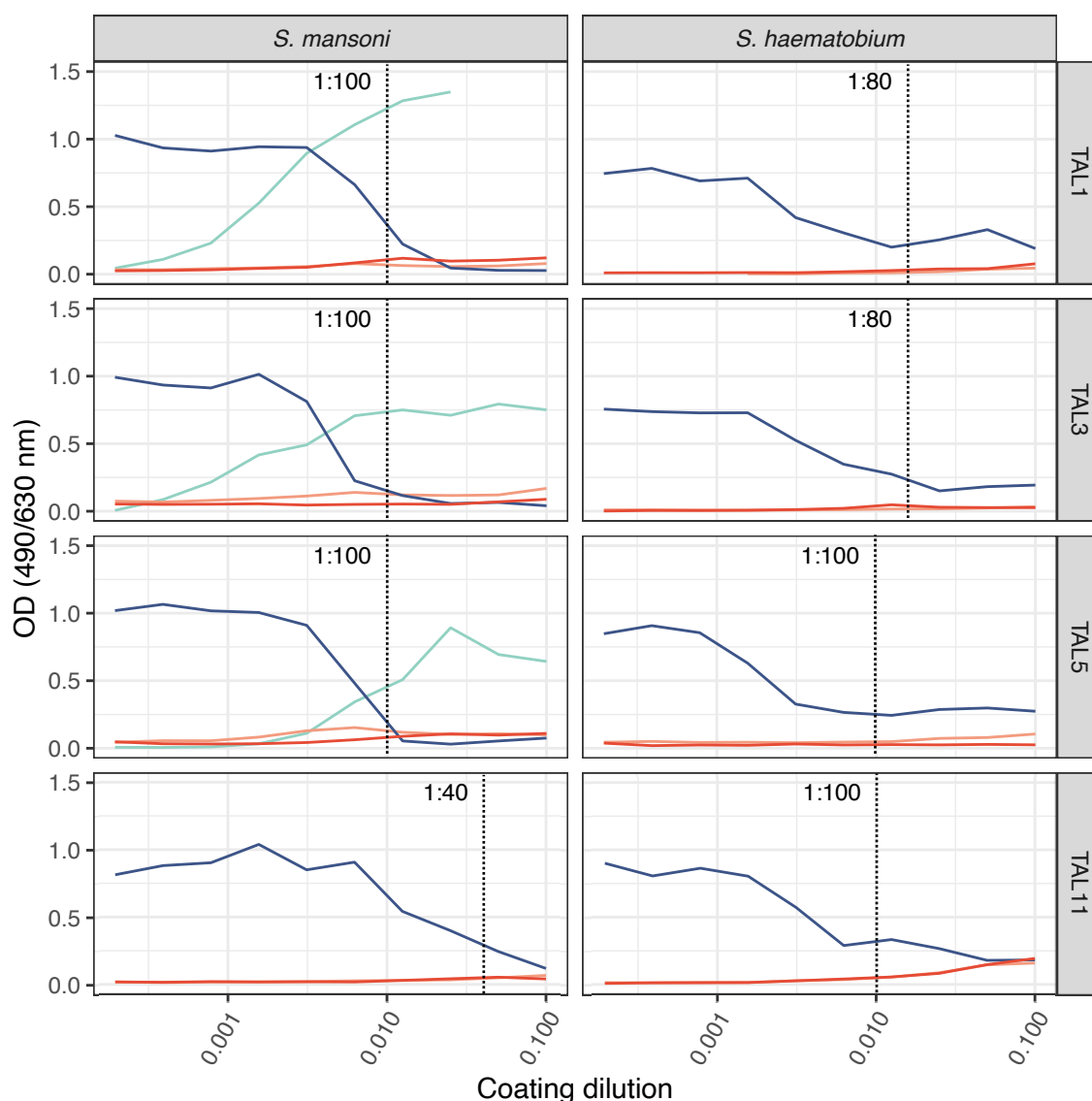


Fig. 4.6 Coating and contamination assays for *S. mansoni* and *S. haematobium* TAL1, TAL3, TAL5 and TAL11. Plots show the binding of anti-rat-HRP to free binding sites where the respective TAL protein has not coated the microtitre well (blue); rat anti-TAL antibodies were available for TAL1, TAL3, and TAL5 following previous experimental work performed by other members of the CSRG (green); anti-GST antibody response indicating remaining GST contamination (red), and the antibody response to bacterial lysate (orange), to indicate whether there is any significant remaining bacterial contamination of the recombinant proteins.

4.2.6 IgE binding to TAL5 is cross-reactive with TAL11 in *S. mansoni* and *S. haematobium*

To ascertain whether IgE to SmTALII demonstrates a similar cross-reactivity profile to that shown by Fitzsimmons *et al.* (2012) between SmTAL3 and SmTAL5, reciprocal inhibition ELISAs were conducted, as described in Chapter 2, section 2.15.4. Assays were repeated with *S. haematobium* antigens, to examine whether the *S. haematobium* TAL proteins also cross react.

The results from the reciprocal inhibition assay, shown in Figure 4.7, demonstrate that binding of plasma IgE to each solid-phase, plate-bound TAL protein was completely inhibited by pre-incubation of the plasma with the respective soluble TAL antigen, as compared to the negative control. This indicates that all available antigen binding sites, or paratopes, had been blocked by 'free' antigen in solution.

Binding of plasma IgE to solid-phase SmTALI does not appear to be cross reactive, since binding was not significantly inhibited by pre-incubating the plasma with SmTAL3 ($t = -0.829$; $p = 0.459$), SmTAL5 ($t = 0.704$; $p = 0.532$) or SmTALII ($t = 1.317$; $p = 0.291$) (Fig. 5.7A). ShTALI binding was not inhibited by pre-incubation with ShTAL3 ($t = -1.288$; $p = 0.268$), ShTAL5 ($t = 0.279$; $p = 0.795$) or ShTALII ($t = 1.989$; $p = 0.157$), again suggesting that ShTALI is not cross-reactive with these members of the TAL protein family (Fig. 5.7E). Binding of plasma IgE to solid-phase SmTAL3 was reduced by almost 50% following pre-incubation of the plasma with both SmTAL5 ($t = 8.497$; $p = 0.001$) and SmTALII ($t = 11.589$; $p = 0.006$) (Fig. 5.7B). Suppression of binding was also seen in the equivalent ShTAL proteins, with suppression of ShTAL3 binding significantly inhibited by pre-incubation with ShTAL5 ($t = 6.667$; $p = 0.021$) and ShTALII ($t = 7.987$; $p = 0.015$), compared to the positive control (Fig. 5.7F). Fitzsimmons *et al.* (2012) showed complete inhibition of IgE binding to SmTAL5 following pre-incubation with TAL3. Here, complete inhibition of binding also occurred following pre-incubation with SmTALII ($t = 1.225$; $p = 0.288$) (Fig. 5.7C). This suggests that SmTALII shares IgE epitopes with SmTAL5. Interestingly, preincubation of plasma with ShTAL3 did not completely suppress IgE binding to ShTAL5 ($t = -31.956$; $p = 0.001$, compared to negative control) in the *S. haematobium* assay, instead resulting in approximately 60% reduction in the IgE signal from the positive control (Fig. 5.7G). Whereas, the level of inhibition following pre-incubation with ShTALII was not significantly different from pre-incubation with either ShTAL5 ($t = -0.794$; $p = 0.480$) or the negative control ($t = -1.970$; $p = 0.188$) across the replicates (Fig. 5.7G). Finally, IgE binding to solid-phase SmTALII was not significantly inhibited by pre-incubating plasma with SmTALI, SmTAL3 or SmTAL5 (Fig. 5.7D); nor was a reduction in ShTALII seen for serum pre-incubated with the equivalent ShTAL proteins (ShTALI ($t = -0.809$; $p = 0.487$), ShTAL3 ($t = 1.684$; $p = 0.175$) or ShTAL5 ($t = 0.598$; $p = 0.592$), all compared to positive control) (Fig. 5.7H).

The results of the cross-reactivity assay therefore suggest that SmTALII and ShTALII may also be cross-reactive components of the TAL-specific IgE-driven partially protective immune response against reinfection.

SmTAL1 EF hand and DLC domain recombinant peptide expression, purification and serological analysis by ELISA were previously performed by other members of the CSRG. Regression analysis of the specific IgE data from these experiments was conducted here to identify potential TAL protein epitopes that may be associated with the partially protective immune response. Univariable models demonstrated a significant association between SmTAL1 EF hand peptide and reinfection at 2 years (GM odds ratio: 0.44; 95% CI: 0.20–0.93; $p < 0.05$), however significance is lost when age and nine-week egg count are accounted for (Appendix 4; Table A4.1). No significant association was found between SmTAL1 DLC peptide and reinfection at 2 years (GM odds ratio: 0.51; 95% CI: 0.22–1.19; $p = 0.12$).

Fitzsimmons *et al.* (2012) found that individuals who were seropositive for SmTAL5 specific IgE were a subset of individuals with a positive TAL-specific IgE response to SmTAL3 and SmTALI. The characteristics of TAL-specific IgE responders were next examined to determine whether this relationship also holds for SmTALII-specific IgE responders.

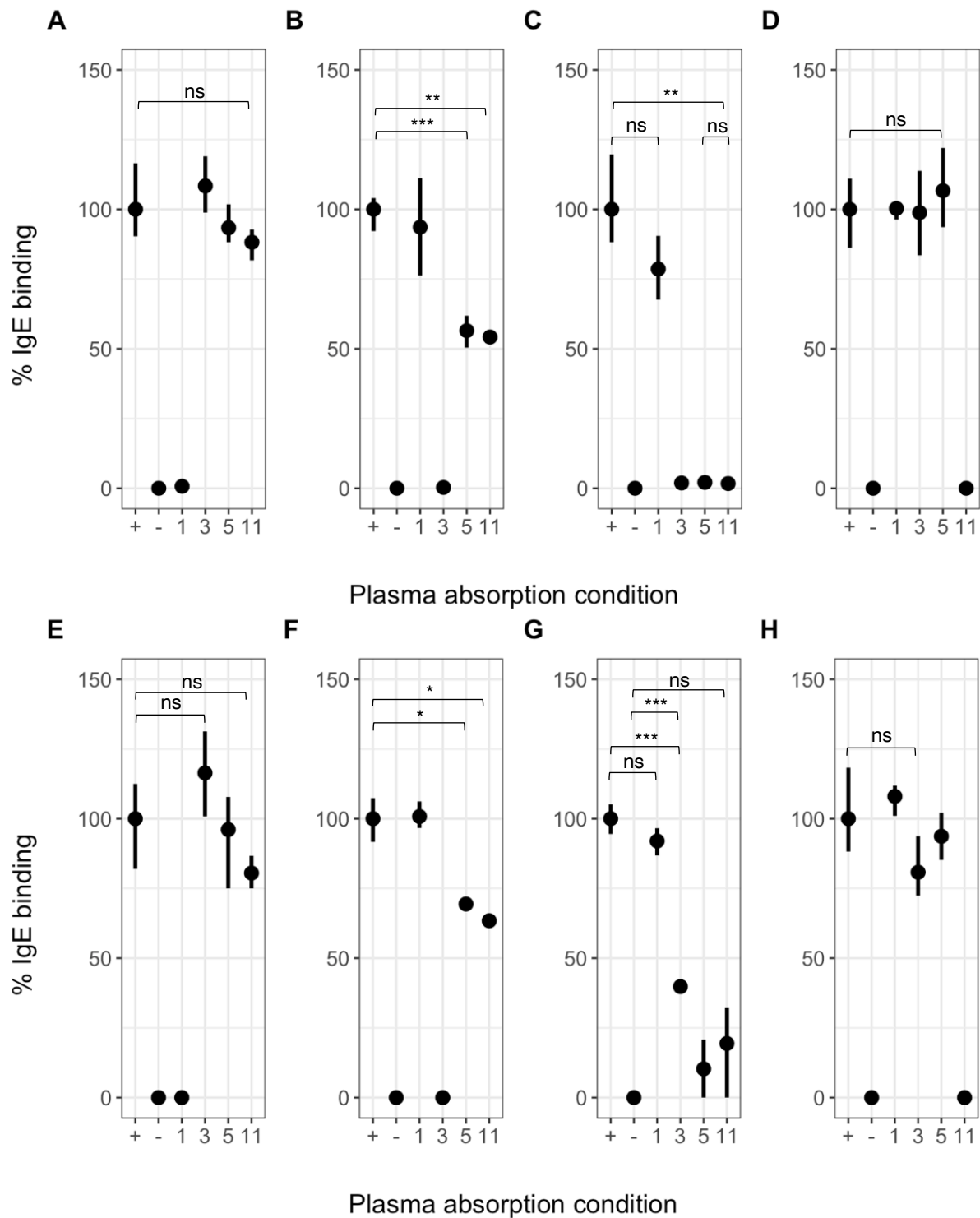


Fig. 4.7. Cross-reactivity of TAL-IgE responses in *S. mansoni* and *S. haematobium*. Reciprocal inhibition ELISA using pooled plasma (n=10) from individuals with demonstrable IgE response to Sm or Sh TAL1, 3, 5, and 11. IgE binding to solid-phase SmTAL1 (A), SmTAL3 (B), SmTAL5 (C), SmTAL11 (D) and ShTAL1 (E), ShTAL3 (F), ShTAL5 (G) and ShTAL11 (H) was inhibited by pre-incubation with antigens Sm or Sh TAL1, 3, 5 or 11, respectively, in solution at 150 ug/ml. Positive control plasma was not pre-incubated with antigen. Negative control contained no plasma. Samples were run in triplicate; points indicate mean percentage IgE concentration relative to positive control. Error bars indicate minimum & maximum value. The significance of binding differences was tested using t-test, with p-values adjusted for multiple testing using the Simes-modified Bonferroni correction (Simes, 1986).

4.2.7 Individuals seropositive for SmTAL11-IgE are a subset of those with a detectable SmTAL5, SmTAL3 and SmTAL1-specific IgE response

Individuals that were found to have a detectable TAL specific IgE response greater than the mean + 3x standard deviations of the non-endemic serum (NES) controls were identified as an IgE 'responder' for a specific TAL. These responders were then classified according to whether they had a positive IgE response for a single TAL protein (single responders), two TAL proteins (double responders), three TAL proteins (triple responders), or whether they had a positive IgE response to all four TAL proteins explored here (quadruple responders). The number of responders within each combination of TAL responses were visualised in a Venn diagram (Fig. 4.8). 134 individuals (61.8%) had a positive response to one or more of SmTAL1, 3, 5 or 11. 79% of IgE responders to SmTAL11 were also found to be responders to SmTAL1, SmTAL3 and SmTAL5.

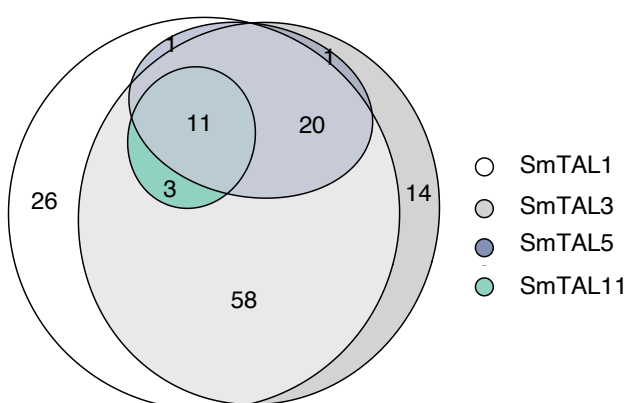


Fig. 4.8. Venn diagram demonstrating the overlapping seroprevalence of *S. mansoni* TAL proteins. Numbers indicate the number of individuals with a detectable antigen specific IgE response, greater than the mean + 3xSD of non-endemic sample controls.

4.2.8 Characterisation of the SmTAL11-specific IgE response, a potential new candidate in the TAL protein reinfection story

It is hypothesised that a reduction in the rate of reinfection is associated with the progressive development of specific IgE antibody responses to SmTAL1 and SmTAL3 that are cross-reactive with SmTAL5, when exposed by the invading parasite. In order to explore this relationship further, the epidemiology of SmTAL11-IgE responses was characterised here for the first time.

Fitzsimmons *et al.* (2012) previously demonstrated that the prevalence of specific IgE responses increases with age for SmTAL1, SmTAL3 and SmTAL5. The seroprevalence of SmTAL11-IgE was generally lower than that observed for SmTAL1, SmTAL3 and SmTAL5, peaking in 15 to 24-year olds. Figure 4.9 shows age group specific antibody responses to SmTAL11 pre- and post-

treatment. Kruskal-Wallis tests, conducted to examine the differences by age in both pre- and post-treatment IgE, found a significant difference in SmTALII-IgE responses, both pre- ($\chi^2=13.74$, $df = 4$, $p = 0.008$) and post-treatment ($\chi^2=17.12$, $df = 4$, $p = 0.002$). *Post-hoc* analysis using the Conover-Iman test (Conover and Iman, 1979) shows that post-treatment SmTALII responses are only significantly different between the oldest and youngest age groups. Increases in specific IgE following treatment showed a strong boost in post-treatment SmTALII-IgE responses ($p < 0.01$), as determined by Wilcoxon matched pair Signed Rank Test.

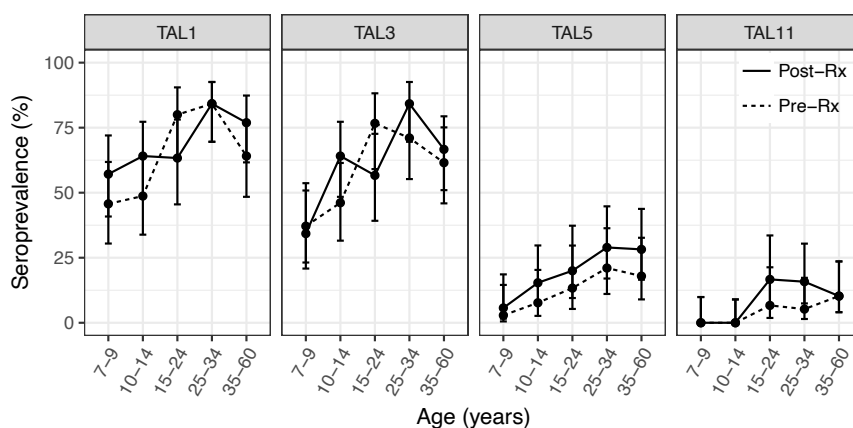


Fig 4.9. Seroprevalence for *S. mansoni* TAL-specific IgE responses at baseline, before treatment (pre-Rx) and 9-weeks after treatment (post-Rx), according to age. Error bars represent the 95% confidence intervals around the seroprevalence for each age group.

Univariate linear regression models, exploring the relationship between IgE response and reinfection intensity showed that seropositivity for post-treatment SmTALII-specific IgE is associated with significantly reduced geometric mean ratio of reinfection intensity at two years (GM ratio: 0.10; 95% CI: 0.03–0.37, $p < 0.001$). A positive nine-week egg count indicates that infection was not cleared, either because of poor treatment efficacy or non-compliance with treatment and will subsequently influence infection status at the reinfection timepoint two years later. When the model was adjusted for age and nine-week post-treatment egg count the effect of post-treatment SmTALII-IgE lost strength but remained significant (GM ratio: 0.18; 95% CI: 0.05–0.65, $p < 0.01$) (Table 4.2). Increasing age was also associated with a strengthening protective effect against reinfection, yet this was only significant in the oldest two age groups (25 years and older) and suggests that factors associated with age, independent of antigen specific IgE responses may also influence reinfection in these individuals. Whilst nine-week egg count had a small increased effect on reinfection, since a positive nine-week egg count indicates that infection was not cleared following treatment, this was not statistically significant.

The influence of being seropositive for combined SmTAL1, SmTAL3, SmTAL5 and SmTAL11 specific IgE on reinfection intensity at two years post-treatment were subsequently examined to explore the relationship between successive TAL protein responses and protective immunity. Similar results were seen as to individual SmTAL protein IgE responses. Results are displayed in Table 4.2. Pairwise interaction terms between the independent variables were not significant and were therefore removed from the models. These results lend support to the hypothesis that SmTAL11 is associated with the development of delayed concomitant immunity. Increasing host age and having a positive IgE response to SmTAL11 was significantly associated with reduced intensity of reinfection at two years post-treatment (Table 4.2, highlighted in red). Being a quadruple TAL responder was also significantly associated with reduced infection intensity at two years (Table 4.2, highlighted in red). This is expected, since SmTAL11 responders are a subset of those triple TAL responders. The responses of ShTAL1, 3, 5 and 11 were subsequently examined to test whether a similar protective response is seen in *S. haematobium*.

Table 4.2 Geometric mean odds ratios describing the association between post-treatment IgE responses to SmTAL antigens and intensity of reinfection at 2 years post-treatment. Models are adjusted for age and 9-week egg count, as a measure of treatment efficacy.

	SmTAL11		SmTAL1, 3, 5, 11	
	GM ratio (95% CI)	p-value	GM ratio (95% CI)	p-value
Age (years) [†]				
10 to 14	0.96 (0.32-2.85)	ns	0.96 (0.32-2.88)	ns
15 to 24	0.42 (0.13-1.41)	ns	0.38 (0.11-1.28)	ns
25 to 34	0.18 (0.06-0.55)	**	0.17 (0.06-0.52)	**
35 to 60	0.11 (0.04-0.31)	***	0.10 (0.03-0.29)	***
9-week egg count [‡]	1.12 (0.87-1.43)	ns	1.14 (0.89-1.46)	ns
SmTAL seropositivity	0.18 (0.05-0.65)	**	0.19 (0.04-0.87)	*

***p<0.001, **p<0.01, *p<0.05, ns non-significant;

[†] values compared to 7- to 9-year-old age group;

[‡] proxy for treatment efficacy

4.2.9 ShTAL11 responders are a partial subset of ShTAL1, 3 and 5 responders

IgE responses in the Malian cohort were initially examined for the presence of individuals seropositive to multiple TAL proteins. As the Venn diagram in Figure 4.10 shows, a similar overlap in TAL protein responders was observed in *S. haematobium* to that seen in *S. mansoni*. Several individuals exposed to *S. haematobium* have positive IgE responses to more than one TAL protein, with 90% of ShTAL3 responders (versus 86% of *S. mansoni* responders) also demonstrating a positive IgE response to ShTAL1. Additionally, 80% of ShTAL5 responders were found to have a positive IgE response to ShTAL1 and ShTAL3, compared to 94% of individuals with responses to the equivalent *S. mansoni* TAL proteins. Finally, overlaying ShTAL11-IgE responders demonstrates that these individuals are also a subset of those study participants seropositive for specific IgE to ShTAL1, ShTAL3 and ShTAL5, with 65% of ShTAL11 responders being quadruple TAL responders. In comparison to *S. mansoni* (Fig. 4.8) there were no complete subsets observed within the pool of *S. haematobium* TAL responders (Fig. 4.10).

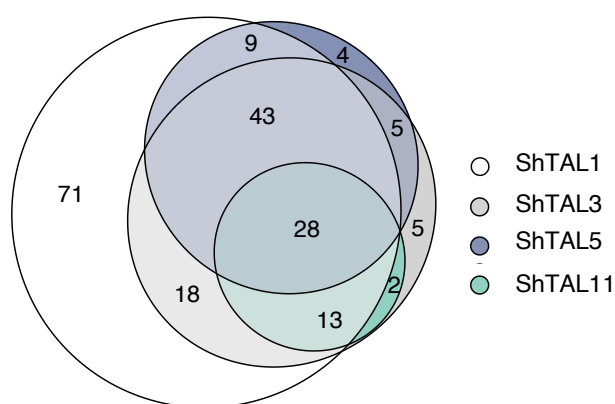


Fig. 4.10. Venn diagram demonstrating the overlapping seroprevalence of *S. haematobium* TAL proteins. Numbers indicate the number of individuals with a detectable antigen specific IgE response, greater than the mean + 3xSD of non-endemic sample controls.

The proportion of individuals that have detectable IgE to a combination of TAL proteins overlap to a lesser extent in *S. haematobium* compared to *S. mansoni*; however, the mechanism controlling reinfection immunity may still be similar between the two species. To better understand the role of TAL proteins in the partially protective immune response in *S. haematobium*, the age profiles of individuals seropositive for ShTAL1, ShTAL3, ShTAL5 and ShTAL11-specific IgE, were examined and compared to those for *S. mansoni* responses from Musoli, Uganda. ShTAL protein ELISA data from the Malian study population were available from a previous study (Dickinson, 2012).

4.2.10 The seroprevalence profiles of *S. haematobium* TAL1, 3, 5 and 11 and their association with age and force of transmission

Multivariable logistic regression models were constructed to explore the influence of age and residing in a high or moderate transmission village, on IgE seropositivity to individual *S. haematobium* TAL proteins. Village of residence was treated as a proxy for the force of transmission. The pairwise interaction between age and village was included, and models were sequentially reduced by the removal of non-significant terms.

Residing in a high transmission village was a significant predictor of having a positive ShTAL1- (GM ratio 3.06; 95% CI: 1.63–6.02; $p = 0.0007$), ShTAL3- (GM ratio : 4.13; 95% CI: 2.40–7.23; $p < 0.0001$), ShTAL5- (GM ratio : 2.51; 95% CI: 1.46–4.35; $p = 0.001$) or ShTAL11-specific (GM ratio 5.24; 95%CI: 2.64–10.94; $p < 0.0001$) IgE response compared to residing in a village in which transmission is moderate (Fig. 4.11). Males were more likely to have detectable IgE responses to ShTAL3 (GM ratio : 1.76; 95% CI: 1.01–3.09; $p = 0.05$), though this was of borderline significance and sex was not a significant predictor of ShTAL1, ShTAL5 or ShTAL11 IgE seropositivity.

The influence of age and transmission intensity on being IgE seropositive to a combination of ShTAL1, ShTAL3, ShTAL5 and ShTAL11 proteins was also examined. Due to the small sample size, age grouping was reduced to a binary variable for these models, with ‘children’, below the age of 10, and ‘adults’, of 10 years old and above. The age cut-off was determined according to piecewise regression analysis, identifying the point of inflection of the CAA age profile for the total cohort (described in section 2.16.7).

When examining responses in the total cohort, adults were more likely to be seropositive for ShTAL1 and ShTAL3 compared to children (GM ratio 2.04; 95% CI: 1.18–3.58; $p = 0.01$), but age was not a significant predictor of detectable IgE to other protein combinations; nor was sex a significant determinant of having a positive IgE response to multiple ShTAL proteins. Village, as a proxy for transmission intensity, was the strongest predictor for having detectable IgE responses to multiple ShTAL proteins ($p < 0.001$).

Post-treatment increases were observed in specific IgE antibody titres to ShTAL1 ($t = -3.257$, $p = 0.001$), ShTAL3 ($t = -4.108$, $p < 0.0001$), and ShTAL5 ($t = -2.0055$, $p = 0.05$) and ShTAL11 ($t = -3.8045$, $p = 0.0002$), in accordance with the patterns seen in *S. mansoni* (Fig. 4.11). The boost in post-treatment ShTAL11-IgE corresponds with predominant expression of ShTAL11 in the adult worm (Fig. 4.1), since worm expressed proteins are thought to be exposed to the host immune system upon worm death. Furthermore, post-treatment antibody boosts were greater in the moderate transmission village compared to the high transmission villages. This could indicate that individuals in the high transmission villages have already been exposed to

antigen at a sufficient level to generate detectable IgE titres. IgE antibody responses were further characterised to elucidate the potentially protective role of TAL11.

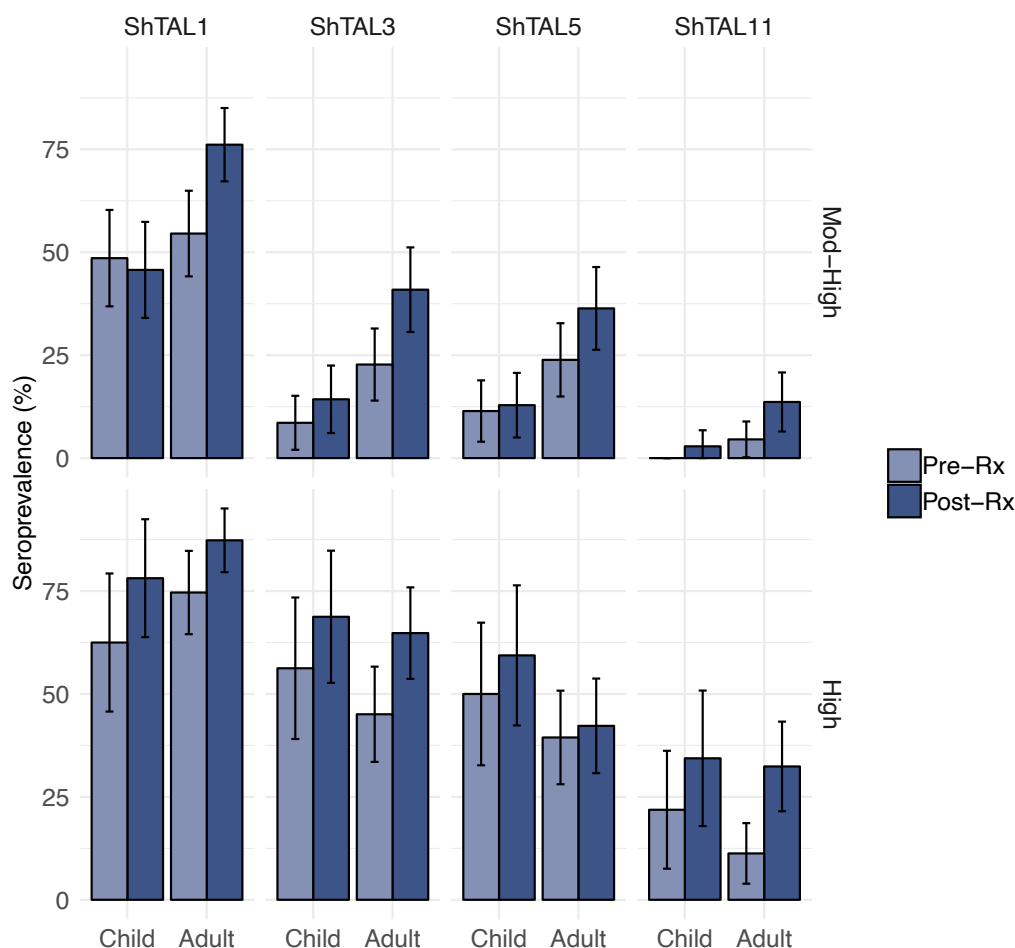


Fig 4.11. Pre- and post-treatment (Rx) seroprevalence of IgE responses to *S. haematobium* TAL proteins by age and village, as a proxy for force of transmission. Error bars represent 95% confidence intervals around the seroprevalence. Children represent individuals less than 10 years old; ‘adults’ represent individuals aged 10 years and older, based on age of peak CAA across whole cohort. Number of individuals within each demographic group: Children (n = 70) and adults (n = 88) in moderate intensity village; Children (n = 32) and adults (n = 71) in high intensity villages.

4.2.11 Are antigen-specific IgE responses to ShTAL1, ShTAL3, ShTAL5 and ShTAL11 associated with a lower worm burden?

As discussed in Chapter 1, section 1.4.6, epidemiological evidence supports the manifestation of processes affecting worm fecundity in *S. haematobium* that do not appear to occur in *S. mansoni* (Agnew *et al.*, 1996). *Schistosoma* parasite load is typically assessed through the measurement

of the number excreted eggs; yet, due to the non-linear relationship between CAA and egg excretion in *S. haematobium*, egg counts are an unreliable measure of worm burden, instead the use of CAA is preferable in assessing *S. haematobium* infection intensity (Wilson *et al.*, 2014). In the present study longitudinal follow-up CAA data was not available, so the association between ShTAL protein responses and worm burden was initially analysed using univariate cross-sectional linear models of pre-treatment ShTAL protein IgE seropositivity and baseline CAA.

Associations between TAL-specific IgE seropositivity and baseline CAA were explored in the total cohort as well as independently for high and moderate transmission villages. In the univariate models for the total cohort, being IgE seropositive to ShTAL1, ShTAL3, ShTAL5 or ShTAL11 was associated with having a significantly higher level of CAA at baseline compared to non-responders in the total cohort. However, when age, village of residence and specific IgG₄ were adjusted for, ShTAL3-IgE (GM ratio: 0.58; 95% CI: 0.29–1.16; $p = 0.12$) and ShTAL5-IgE (GM ratio: 0.79; 95% CI: 0.39–1.60; $p = 0.51$) were negatively associated with baseline infection intensity, though not significantly. Specific IgG₄ responses were however, significantly associated with higher baseline CAA (ShTAL3-IgG₄: GM ratio: 4.56; 95% CI: 2.26–9.17; $p < 0.0001$; ShTAL5-IgG₄: GM ratio: 3.95; 95% CI: 2.04–7.64; $p = 0.0001$). Significant negative interaction terms were found between age and village, though the negative association between both ShTAL3 and ShTAL5 and CAA remained non-significant after accounting for these terms. The influence of sex on baseline infection intensity was not significant in the whole cohort model and was removed during stepwise model reduction. Likewise, when multiple TAL responses were examined for the total cohort (Table 4.3), non-significant negative associations were seen between double (ShTAL1 and 3) and triple responders (ShTAL1, 3 and 5), whereas younger age and TAL-specific IgG₄ appeared to be associated with higher baseline infection intensity, as measured by CAA. Residing in a village with a high force of transmission was also significantly associated with high baseline CAA.

Since village appears to have a significant effect on the development of protective immune responses, further analyses of protection were studied independently for the villages in high and moderate transmission sites, accounting for both age and sex. GM ratios reveal differences in the age of peak infection intensity between the two transmission sites, previously reported elsewhere (Wilson *et al.*, 2014). Age was significantly associated with baseline CAA, peaking in the five to six-year-old age group in the high transmission villages and in 10 to 14-year olds in the moderate intensity village. Furthermore, males are associated with significantly lower baseline infection intensity in high transmission villages and higher, but non-significant, baseline infection intensity in the moderate transmission village compared to females. When positive IgE responses to multiple ShTAL proteins were separated by village, models adjusted

for age, sex and specific IgG₄ found that being seropositive for IgE to a combination of ShTAL proteins (either ShTAL1 and ShTAL3, or ShTAL1, ShTAL3 and ShTAL5) was associated with a lower baseline infection intensity in both high and moderate transmission sites, although these associations were not significant, since statistical power was not maintained following disaggregation of the data by village. Respective antigen specific pre-treatment IgG₄ was however associated with significantly higher CAA at baseline. The results of the reduced models are displayed in Appendix 4 (Table A4.2).

4.2.12 Are antigen-specific IgE responses to ShTAL1, ShTAL3, ShTAL5 and ShTAL11 associated with reinfection status at 2 years?

As discussed in section 4.2.10, a boost in the IgE response to ShTAL1, ShTAL3, ShTAL5 and ShTAL11 was observed following treatment. Although two-year follow-up CAA data was not available to analyse the association between post-treatment ShTAL-IgE seropositivity and intensity of reinfection, logistic regression models were formulated to test whether having a positive post-treatment response to ShTAL proteins was associated with reinfection status at two years.

Reinfection status was modelled as a binary outcome variable and age group, sex and village as independent categorical variables. A positive response was defined as having measurable antigen-specific antibody response greater than the mean plus three standard deviations of an array of non-infected non-endemic control sera (NES). In the total cohort model, having measurable post-treatment IgE to ShTAL5 (GM ratio: 0.38; 95% CI: 0.19–0.76; $p = 0.006$) was associated with significant reduced GM odds ratio of reinfection at two years (Table 4.4). When age, sex, village and positive ShTAL5-IgG₄ response were accounted for in the model ShTAL5-IgE seropositivity remained significantly negatively associated with reinfection at two years (GM ratio: 0.36; 95% CI: 0.13–0.99; $p = 0.05$). ShTAL3-IgE seropositivity also showed a negative association with reinfection status at two years, although this was not significant (GM ratio: 0.55; 95% CI: 0.20–1.44; $p = 0.22$). ShTAL11-IgE was not associated with a reduction in reinfection. Conversely, ShTAL11-IgG₄ was significantly associated with a lack of reinfection at two years. Sex was not a significant predictor of reinfection.

Table 4.3. Association between pre-treatment ShTAL-specific IgE responses (seropositivity) and baseline CAA, as a measure of worm burden. Results for the reduced linear regression model, adjusted for age, sex, village and ShTAL-specific IgG₄ seropositivity are displayed, for total cohort (n = 262).

	ShTALI			ShTALI, 3			ShTALI, 3, 5			ShTALI, 3, 5, II		
	GM ratio (95%CI)	P-value		GM ratio (95%CI)	P-value		GM ratio (95%CI)	P-value		GM ratio (95%CI)	P-value	
Age (years) [†]												
7-9	4.70 (1.86-11.89)	**		4.65 (1.87-11.59)	**		4.64 (1.88-11.48)	***		4.09 (1.65-10.13)	**	
10-14	8.84 (3.29-23.73)	***		8.86 (3.38-23.26)	***		9.65 (3.69-25.21)	***		8.54 (3.27-22.28)	***	
15 - 24	4.97 (1.82-13.56)	**		4.63 (1.71-12.48)	**		5.33 (1.97-14.41)	**		4.24 (1.53-11.74)	**	
25 - 29	3.23 (1.27-8.17)	*		2.68 (1.07-6.73)	*		3.00 (1.20-7.52)	*		2.86 (1.15-7.12)	**	
30+	2.02 (0.67-6.12)	ns		1.90 (0.64-5.67)	ns		2.54 (0.84-7.72)	.		2.40 (0.80-7.22)	ns	
Sex	-	ns		-	ns		-	ns		-	ns	
Village [‡]	5.03 (2.66-9.50)	***		4.09 (2.09-8.00)	***		3.34 (1.69-6.59)	***		2.86 (1.45-5.64)	**	
TALx-IgE [§]	1.06 (0.57-1.96)	ns		0.59 (0.28-1.25)	ns		0.54 (0.22-1.28)	ns		1.02 (0.28-3.73)	ns	
TALI-IgG ₄	2.76 (1.37-5.62)	**		-	ns		-	ns		-	ns	
TAL3-IgG ₄				4.49 (2.22-9.08)	***		2.87 (1.29-6.45)	*		-	ns	
TAL5-IgG ₄							3.03 (1.19-7.71)	*		2.98 (1.23-7.24)	*	
TALII-IgG ₄										2.97 (1.43-6.17)	**	

***p<0.001, **p<0.01, *p<0.05, †p<0.1, ns non-significant; ‡ values are compared to 5- to 6-year-old age group; § values are for high intensity villages compared to moderate infection intensity village; § where TALx represents the indicated combination of TALI, TAL3, TAL5 and TALII responses.

In univariable models, ShTAL3 (GM ratio: 0.41; 95% CI: 0.02–3.08; $p = 0.44$) and ShTAL5 (GM ratio: 0.37; 95% CI: 0.04–2.44; $p = 0.30$) were associated with reduced reinfection at two years in the high transmission villages, however this was not significant. Age was again reduced to a binary variable in the village specific regression models due to the smaller sample size. When age, sex and specific IgG₄ seropositivity were accounted for in multivariable models, none significantly associated with reinfection in the high transmission villages (Appendix 4, Table A4.3). In univariable models for the moderate transmission village, negative associations were found between having measurable post-treatment IgE to ShTAL1 (GM ratio: 0.52; 95% CI: 0.24–1.10; $p = 0.09$), ShTAL3 (GM ratio: 0.32; 95% CI: 0.14–0.73; $p = 0.007$), ShTAL5 (GM ratio: 0.24; 95% CI: 0.10–0.56; $p = 0.001$) and ShTAL11 (GM ratio: 0.58; 95% CI: 0.18–1.87; $p = 0.36$) and reduced odds of reinfection; these associations were significant for ShTAL3- and ShTAL5-IgE.

When age and sex were accounted in the moderate transmission village model, being an adult (10 years of age and older) was highly significantly associated with reduced reinfection in all ShTAL models and the negative association between ShTAL IgE seropositivity and reinfection at two years was lost. All models were sequentially reduced to remove non-significant terms. All pairwise interaction terms were non-significant and were subsequently removed from the model.

Associations between IgE seropositivity to multiple ShTAL proteins and reinfection status at two years post-treatment were subsequently analysed. In univariable models, negative associations were found between IgE seropositivity to ShTAL1 and 3, ShTAL1, 3 and 5 and ShTAL1, 3, 5 and 11 and reinfection status in all cohorts. In the whole cohort model, being IgE seropositive to ShTAL1, 3 and 5 (GM ratio: 0.39; 95% CI: 0.18–0.84; $p = 0.02$) was significantly associated with reduced reinfection. Similarly, in moderate intensity village model, being IgE seropositive to either ShTAL1 and 3 (GM ratio: 0.25; 95% CI: 0.10–0.59; $p = 0.002$) or ShTAL1, 3 and 5 (GM ratio: 0.22; 95% CI: 0.07–0.60; $p = 0.004$) was significantly associated with reduced reinfection. All other associations were non-significant. When age, sex village and specific IgG₄ were accounted for in the models, associations lost significance in the moderate intensity model, but in the total cohort the association between being a ShTAL1, 3 and 5 responder and reduced reinfection at two years remained significant (GM ratio: 0.29; 95% CI: 0.09 – 0.90; $p = 0.03$). Results are displayed in Table 4.5. In the separate models for high and moderate intensity transmission villages a negative association was found between having a positive IgE response to ShTAL1 and 3, or ShTAL1, 3 and 5 but these were not significant. Furthermore, inclusion of responders to multiple ShTAL proteins including ShTAL11 responders, weakened the protective effect; the models do not therefore provide evidence that ShTAL11 responses are protective against reinfection (Appendix 4, Table A4.3).

Table 4.4. Associations between univariant post-treatment ShTAL-IgE responses and measurable urinary egg excretion at two-year reinfection

Specific-IgE	Total cohort		High intensity		Moderate intensity	
	GM ratio (95% CI)	p-value	GM ratio (95% CI)	p-value	GM ratio (95% CI)	p-value
ShTALI	0.79 (0.39-1.54)	ns	1.16 (0.06-9.24)	ns	0.53 (0.24-1.10)	.
ShTAL3	0.69 (0.36-1.32)	ns	0.41 (0.02-3.08)	ns	0.33 (0.14- 0.73)	**
ShTAL5	0.38 (0.19-0.76)	**	0.37 (0.04-2.44)	ns	0.25 (0.10- 0.56)	**
ShTALI1	1.50 (0.65 - 3.80)	ns	2.43 (0.33-49.63)	ns	0.58 (0.18-1.87)	ns

******p<0.01, **.** p<0.1, ns non-significant.

Table 4.5. Association between 9-weeks post-treatment ShTAL-specific IgE responses and reinfection status at two years, adjusted for age, sex, village and ShTAL-specific IgG₄, non-significant terms removed. Results are displayed for total cohort (n = 174).

	ShTALI		ShTALI, 3		ShTALI, 3, 5		ShTALI, 3, 5, 11	
	GM ratio (95% CI)	P-value	GM ratio (95% CI)	P-value	GM ratio (95% CI)	P-value	GM ratio (95% CI)	P-value
Age (years) [†]								
7-9	1.40 (0.31-7.39)	ns	1.50 (0.33-7.97)	ns	1.49 (0.33-7.96)	ns	1.44 (0.32-7.55)	ns
10-14	0.33 (0.07-1.48)	ns	0.39 (0.09-1.66)	ns	0.43 (0.10-1.89)	ns	0.36 (0.08-1.53)	ns
15 - 24	0.03 (0.00-0.13)	***	0.03 (0.01-0.15)	***	0.03 (0.01-0.15)	***	0.03 (0.01-0.13)	***
25 - 29	0.04 (0.01-0.14)	***	0.05 (0.01-0.18)	***	0.06 (0.01-0.22)	***	0.04 (0.01-0.16)	***
30+	0.03 (0.01-0.12)	***	0.03 (0.01-0.14)	***	0.03 (0.01-0.12)	***	0.03 (0.01-0.12)	***
Sex	-	ns	-	ns	-	ns	-	ns
Village [‡]	22.82 (7.38-86.36)	***	28.14 (8.66-113.93)	***	27.05 (8.45-107.53)	***	23.33 (7.57-88.05)	***
TALx-IgE [§]	1.21 (0.45-3.38)	ns	0.51 (0.18-1.40)	ns	0.29 (0.09-0.90)	*	0.95 (0.17-5.28)	ns
TALx-IgG ₄ [§]	-	ns	-	ns	-	ns	-	ns

***p<0.001, *p<0.05, ns non-significant; [†] values are compared to 5- to 6-year-old age group; [‡] values are for high intensity villages compared to moderate infection intensity village; [§] where TALx represents the indicated combination of TALI, TAL3, TAL5 and TALI1 responses.

4.3 Discussion

Partially protective immunity conferred by antigen specific IgE has been extensively studied over several decades (Dunne *et al.*, 1992; Fitzsimmons *et al.*, 2004; Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012). The immune processes seen in humans living in endemic conditions have however proved difficult to reproduce in laboratory animals. This is partly because the size and lifespan of the murine model make it impossible to replicate the prolonged low-level trickle infection that is characteristic of chronic schistosomiasis in humans. Several of the leading hypotheses regarding immunity via exposure to dying adult worms have therefore arisen from immuno-epidemiological field studies (Woolhouse and Hagan, 1999). In this chapter, data from field studies conducted in Uganda and Mali were used to elucidate a possible role for TAL11 in the development of delayed concomitant immunity in *S. mansoni* and characterise the IgE mediated protective response in *S. haematobium*, respectively.

4.3.1 ShTAL protein identification and phylogenetic analysis

Searches of the annotated genome identified orthologous *S. haematobium* proteins for each representative of the SmTAL protein family, mirroring previous findings from EST searches (Dickinson, 2012); in addition to a previously uncharacterised protein with an EF-hand pair and DLC domains analogous to those of the TAL protein family. The identified orthologous proteins were aligned prior to phylogenetic analysis. The quality of a multiple sequence alignment can have a significant influence on the resulting inferred tree. By removing columns that are considered to be less informative to the phylogeny before inferring the tree, phylogenetic analysis of the ShTAL proteins identified two principal clades into which the protein family is divided, consistent with *S. mansoni* studies (Carson *et al.*, 2018).

The Gblocks refined alignment more closely resembles the phylogenetic analysis of the *S. mansoni* TAL proteins performed by Carson *et al.* (2018), which shows two main clades into which the TAL protein family is divided, one clade including SmTAL1, SmTAL2, SmTAL8 and SmTAL12, whilst the other includes the remaining SmTAL3, 4, 5, 6, 7, 9, 10, 11 and 13 proteins (Carson *et al.*, 2018). Contrary to Carson *et al.* (2018), the analysis presented here suggests that SmTAL10 and ShTAL10 have more distant phylogenetic affinity.

There is some debate as to whether alignment masking is beneficial or whether relevant information is lost in the process and Gblocks has received criticism for being overly conservative with criteria for the selection of conserved regions (Tan *et al.*, 2015). This tends to be less problematic in the analysis of orthologous proteins, since sequences tend to be highly

conserved. Still, the least stringent block selection options were used in the analysis presented in this chapter.

The previously uncharacterised protein, named TAL14 in this chapter, was most closely aligned with the sub-clade with TAL1, TAL2, TAL8 and TAL12 and so was not characterised any further in this chapter, but further analysis will be required to confirm its identity as a true TAL protein. Analogous with the phylogeny of *S. mansoni* TAL proteins, the *S. haematobium* TAL phylogeny revealed that the proteins previously associated with IgE protective responses, TAL3 and TAL5, are most closely related to TAL4 and TAL11 proteins, within the same sub-clade. In *S. mansoni*, SmTAL4 expression is limited to the tail of the cercaria, whereas SmTAL5 is expressed throughout the cercaria (Fitzsimmons, Jones, Stearn, *et al.*, 2012). *S. haematobium* mRNA expression data suggest that ShTAL4 and ShTAL5 are also primarily expressed in the cercaria (Dickinson, 2012). ShTAL5 is therefore a plausible target of the IgE-mediated protective response to *S. haematobium* infection. On the other hand, TAL11 and TAL3 are predominantly expressed in the adult worm in both *S. mansoni* and *S. haematobium*. Despite similarities with TAL3, TAL11 has not previously been explored in relation to protective immunity. This chapter therefore provides the first analysis of the association between IgE responses to TAL11 and reinfection.

4.3.2 Analysis of cross-reactive relationships

The IgE response associated with protection is proposed to result from SmTAL3-specific IgE cross-reactive to SmTAL5 that is exposed on the invading cercaria and early schistosomula; hence it is thought that antibodies are not generated to SmTAL5 in sufficient levels to be detected. In the analysis presented here, I find that a similar cross-reactive relationship is seen between SmTAL5 and SmTAL11, since pre-incubation with SmTAL11 results in complete inhibition of IgE binding to SmTAL5. Preincubation with SmTAL5 and SmTAL11 also reduces SmTAL3 binding by approximately 50%. Conversely, pre-incubation with SmTAL3 and SmTAL5 has no effect on SmTAL11 binding. This suggests that SmTAL3 may not be the sole cross-reactive antigen responsible for the development of partially protective immunity. The lack of inhibition of SmTAL11 binding indicates that specific IgE may bind the SmTAL11 epitope with higher affinity compared to the other TAL proteins or that SmTAL11 may have a unique secondary epitope.

Cross-reactivity results for *S. haematobium* show similar cross-reactivity profiles between ShTAL3, ShTAL5 and ShTAL11 to those seen for *S. mansoni* proteins, although a few key differences between the responses are observed. The extent to which ShTAL3 binding is

inhibited by ShTAL5 and ShTAL11 is not as high in *S. haematobium* as compared to the equivalent *S. mansoni* TAL protein inhibition. Furthermore, pre-incubation with ShTAL3 does not appear to completely inhibit ShTAL5 binding. Unlike the equivalent *S. mansoni* protein, ShTAL3 has been shown to have a high level of expression in the egg, in addition to expression in the adult worm. ShTAL3 specific IgG₄ responses may consequently block or regulate the cross reactive IgE response in order to reduce the pathology that would result from high IgE responses to tissue trapped eggs. Further investigation is required to explain these differences. The evidence presented here is suggestive of a similar cross-reactive delayed concomitant immunity driven protective response in *S. haematobium* as that previously seen in *S. mansoni*. The overlapping IgE seroprevalence to multiple TAL proteins suggests that adult worm expressed SmTAL11 and ShTAL11 may also be associated with development of the protective immune response, providing further support to the hypothesis that this is related to experience of infection. This expands upon the relationship between the *S. mansoni* TAL proteins and delayed concomitant immunity proposed by Fitzsimmons *et al.* (2012).

4.3.3 Characterisation of *S. mansoni* and *S. haematobium* TAL11 epidemiology

SmTAL11 seroprevalence is low compared to specific IgE responses to the other SmTAL proteins. The reason for the observed lower IgE seroprevalence is unclear, since SmTAL11 is expressed at a similar level to SmTAL3 in five and seven week old worms, although expression is lower in juvenile worms aged between two and five weeks (Fitzsimmons, Jones, Stearn, *et al.*, 2012).

Protection against reinfection is understood to increase following treatment. This has been attributed to the development of protective IgE antibody response following exposure to cryptic antigens released following worm death (Woolhouse and Hagan, 1999). The strengthening of the specific IgE response to Sm and ShTAL11 post-treatment therefore provides support for involvement of this antigen in the development of protective immunity. Additionally, SmTAL11 IgE seroprevalence was found to be higher in adults compared to children. The association between SmTAL11 and age provides further indication that older individuals have a greater experience of antigens released as worms die, since cumulative exposure to dead worms, and so exposure to immunogenic parasite antigens generally increases with age. However, age does not appear to be significantly associated with increased ShTAL11-IgE levels.

The age-seroprevalence curves for *S. haematobium* show a decrease in the percentage of individuals seropositive for antigen-specific IgE in the oldest age group across all recombinant

ShTAL proteins studied in this chapter. The population in Mali has experienced very high infection intensity. Combined with the experience of other environmental pathogens and parasites, this may represent a degree of immunosenescence or immune exhaustion (Akbar and Henson, 2011). In spite of this, significant associations were demonstrated between older age groups and reduced reinfection with positive IgE responses to several TAL proteins.

4.3.4 Associations with anti-infection immunity

Significant associations between village and TAL responses suggest that the development of successive TAL responses appear to be transmission dependent. This is consistent with the peak shift hypothesis; whereby higher transmission intensity provides exposure to a greater antigenic diversity (Fulford *et al.*, 1992).

As discussed in Chapter 3 section 3.2.2, analyses are dependent on the quality and validity of the available data. *S. haematobium* reinfection studies require longitudinal follow-up with the collection of both egg count and CAA re-infection data to enable accurate determination of the relationship between antigen specific IgE responses and reinfection. Evidence suggests that the relationship between worm burden and egg counts may be under the influence of anti-fecundity processes in *S. haematobium* (Agnew *et al.*, 1996), thus the legitimacy of using egg count as a direct measure of parasite burden in *S. haematobium* is questionable. For this reason, reinfection intensity could not be studied as an outcome variable in linear regression models in this study. Instead, cross sectional pre-treatment antibody responses against baseline CAA found non-significant negative associations between having a response to multiple TAL proteins and infection intensity. This cross-sectional analysis, however, appears to lack the power to identify relationship with TAL-IgE responses. Since IgE responses to ShTAL proteins are boosted following treatment, it was considered that the protective effect of post-treatment IgE responses in relation to reinfection status may be stronger. Regression models were therefore constructed to test this theory.

Logistic regression was used to explore reinfection status at two years post-treatment as a binary outcome with the justification that detectable egg excretion is an indicator of infection, regardless of worm burden. This analysis provides evidence of the protective effect of IgE to ShTAL5, since the association between ShTAL5 specific IgE and reduced reinfection at two years remained significant, despite accounting for strong associations with age and village of residence. Furthermore, having a detectable IgE response to a combination of ShTAL1, ShTAL3 and ShTAL5 was also significantly associated with a lack of reinfection at two years, providing

evidence for the progressive development of a protective IgE, response, similar to that demonstrated in *S. mansoni*.

Furthermore, specific IgG₄ antibody responses were significantly associated with having a detectable egg count at two-year reinfection, which suggests that IgG₄ specific antibodies act to block IgE binding and when these responses were removed from the models the negative association was lost.

Regression models did not lend support to the contribution of ShTALII to partially protective immunity since the association between positive IgE responses to multiple TAL proteins and negative reinfection status at two years was weakened when ShTALII responders were included in the model. However, having a positive SmTALII specific IgE response was strongly associated with reduced reinfection intensity at two years. This association was also seen for those with a positive response to multiple ShTAL proteins (ShTALI, 3, 5 and II), which was expected due to the degree of overlap between ShTAL-IgE responders.

Antibody affinity studies would improve our understanding of this relationship still further; however, there are several challenges with regard to experimental design and the nature of protective IgE antibody responses in serum from schistosomiasis endemic areas. Possible future experiments and the limitations of these are discussed in the following section.

4.3.5 Future study requirements

There are two primary classifications for different epitope topologies, ‘continuous’ or ‘discontinuous’ epitopes, corresponding to whether the antigenic determinant is comprised of a single linear segment or is formed from several segments that are separated in sequence but form a conformational epitope in three-dimensional structure (Atassi and Smith, 1978). Denatured TAL proteins can be detected and visualised by Western Blot using rat antibodies (Fitzsimmons, Jones, Pinot de Moira, *et al.*, 2012); since denaturation removes tertiary structure of the protein, this is suggestive of linear epitope recognition. Furthermore, analysis of ELISAs performed with EF-hand and DLC domain peptides indicate that IgE responses to EF-hand peptides are significantly associated with reduced reinfection intensity at two years. Further investigation of the structure of TAL protein epitopes is required to elucidate the precise nature of the cross-reactive relationship. This may include prediction of linear B cell epitopes, to enable closer examination of epitope-containing peptides from within the EF-hand domain regions, specifically. Should TAL epitopes in fact be discontinuous, requiring native protein

conformation, this would require further research into native protein folding (Dall'antonia *et al.*, 2014).

TAL protein affinity binding kinetics could also be studied using surface plasmon resonance to measure the affinity of a ligand for its receptor, including rate of association and dissociation (on-off rate). On-off rate can also be measured in the presence of competing antigens, to estimate relative avidity, or strength, of the antibody antigen complexes. So, in theory antibody binding dynamics could be directly observed between the different TAL proteins. However, this would require isolation of specific IgE from endemic serum or the production of monoclonal antibodies. Human antibody responses are subject to polyclonal expansion and affinity maturation that cannot be replicated by monoclonal antibodies. Furthermore, humans also have vast genetic polymorphisms that cannot be reflected in animal models. Methods do exist by which IgE can be isolated from serum, but these tend to be challenging due to the comparatively low IgE titres in endemic serum. Instead, negative selection of serum IgE has been attempted, using RF-absorbent or protein G beads to deplete total IgG antibodies from serum samples in order to strengthen the IgE response; yet, experiments performed within our laboratory have failed to remove sufficient total IgG from serum samples, even with multiple repeated treatment. This is likely to be because non-specific serum IgG levels in endemic serum are extremely high as a result of a lifetime of significant immune challenge by numerous other pathogens and parasites. It may also be possible to clone IgE antibodies from memory B cells (Croote *et al.*, 2018), although it is not certain that these antibodies would have the same antigen binding characteristics as those antibodies found in endemic serum.

4.4 Conclusions

In this chapter I have utilised immuno-epidemiological studies to explore the relationship between the TAL protein family and reinfection. This provides the first characterisation of delayed concomitant immunity associated with SmTALII and ShTALI, 3 and 5. However, this approach is not without limitation. It is not possible to account for all variation between comparison groups using statistical methods. There are also limitations of the data, primarily the lack of appropriate measure of worm burden at reinfection for *S. haematobium* but also the sample size of the subset of quadruple *S. mansoni* and *S. haematobium* TALI, TAL3, TAL5 and TALII responders. An alternative study design that took account of CAA and included a greater number of participants, in order to collect appropriate follow-up data, would enable clarification of the association between ShTAL proteins and reinfection.

5

Identification of antigenic proteins associated with anti-fecundity immunity against *Schistosoma haematobium*

Preface

In addition to anti-infection immunity, immunoepidemiological studies indicate that *S. haematobium* is also influenced by an immune response that reduces the number of eggs that female worms excrete. This has previously been associated IgG₁ specific antibodies raised to *S. haematobium* whole worm preparation; yet, the specific immunogenic targets of this response remain unknown.

Acknowledgement of contributions

The anti-fecundity 'case' cohort and the age, sex and village matched controls were identified by Dr Shona Wilson (University of Cambridge), following original analysis of the *S. haematobium* infected cohort from Mali. Trypsin digest, mass spectrometry (LC-MS/MS) and MASCOT searches were performed by Ms Renata Feret and Ms Julie Howard Murkin (Department of Biochemistry, University of Cambridge). Primer sequences for *S. haematobium* housekeeping genes were provided by Dr Gabriel Rinaldi (Wellcome Trust Sanger Institute). Dr Ray Owens (OPPF-UK, Harwell, Oxford, UK) produced primer constructs for high throughput (HTP) cloning. HTP cloning and protein expression screening was performed under the guidance of Heather Rada at OPPF-UK.

5.1 Introduction

5.1.1 General introduction

The first evidence describing reduced worm fecundity was reported in a closely related *Schistosoma* species that infects cattle, *S. bovis* (Bushara *et al.*, 1980). Moreover, serum transfer experiments indicated that the resulting reduction in fecundity is antibody mediated (Bushara *et al.*, 1983).

The potential anti-fecundity immune response was later evidenced in humans in a paper by Agnew *et al.* (1996), who, in a study conducted in Kenya, found that the decrease in excreted egg count transpired earlier than the decrease in worm burden with age in individuals infected with *S. haematobium*. Interestingly, this disconnect did not appear to be mirrored in a separate population infected with *S. mansoni*, analysed in the same study.

The mechanism through which the host immune system targets worm fecundity in *S. haematobium* and species within the same phylogenetic clade was not, however, described in these early studies and there was a subsequent paucity of papers examining anti-fecundity responses. In 2012, a mathematical modelling study exploring immune responses in *S. haematobium* was published. Using a stochastic individual-based model, Mitchell *et al.* (2012) tested the ability of different model structures to reproduce the epidemiological patterns observed in a longitudinal cohort study conducted in Zimbabwe. The authors found that models including a long-lived protective antibody response, linked to reduced worm fecundity and stimulated by worm death, were able to reproduce infection and antibody profiles consistent with the patterns seen in cross-sectional field surveys (Mitchell *et al.*, 2012). This study once again raised the question of anti-fecundity immunity against *S. haematobium*. In their analysis, Mitchell *et al.* (2012) focussed specifically on the relationship with IgA and IgG₁ antibody classes. Support was provided for the involvement of IgG₁, in particular, in later immuno-epidemiological analysis conducted in a Malian cohort (Wilson *et al.*, 2014). The enquiry into *S. haematobium* specific antibody responses and their associations with worm fecundity conducted by Wilson *et al.* found that IgG₁ to soluble whole worm antigen (SWA) was the strongest predictor of reduced worm fecundity (Wilson *et al.*, 2014). As with reinfection immunity, there also appears to be a delay in the development of the anti-fecundity response, as a shift is seen in the age at which the reduction in worm fecundity occurs in high compared to moderate transmission intensity setting (Wilson *et al.*, 2014). Despite this progress, current knowledge is limited as to which proteins are the target of IgG₁-driven protection, and the mechanism by which naturally acquired anti-fecundity immunity may be attained.

Modelling studies by Mitchell *et al.* (2012) postulate a link between worm death and the development of an anti-fecundity immune response, but in order for the immune response to act on worm fecundity, it is logical that the target protein must be accessible whilst the worm is still viable. The dioecy of schistosomes has been a point of intrigue for several decades and a number of studies have explored the relationship between the separate male and female forms (Popiel and Basch, 1984; Basch, 1990); yet, the nature of the specific pathway that drives female sexual maturation still remains unclear. Several theories have been proposed including: transfer of sperm to the female; supplementation of nutritional intake; transfer of neuromediators, and

hormonal or tactile stimulation (Michaels, 1969). Pairing between the male and female worm is discussed further in the following section.

5.1.2 Reproductive pairing is required for female sexual maturation

The maintenance of contact between the male and female worm is a prerequisite for female *S. haematobium* worms to achieve reproductive maturation. Sexual maturation in the female results in differentiation of the gonads, including the ovary and vitellarium, which respectively produce the oocytes and resources for embryogenesis, such as egg-shell proteins (Jones, Jamieson and Justine, 2017). Experiments in which sexually mature female worms are separated from their male pairing (Galanti, Huang and Pearce, 2012) suggest that the female will sexually regress if separated from the male, indicating that some form of continuous signalling occurs during pairing.

Popiel and Basch (1984) demonstrated that the signalling between males and females is localised to regions in which the female is in direct contact with the male, as shown by dissection pairing experiments. This suggests that the developmental response is independent of the central nervous system, with no centralised location for female stimulation. In particular, differentiation of the vitelline gland is dependent on local stimulation and is not disseminated throughout the worm. Furthermore, experiments have demonstrated that female maturation is independent of presence of testes and male reproductive organs (Armstrong, 1965; Michaels, 1969); a possible tactile stimulus (Basch and Basch, 1984) or transfer of a male-secreted hormonal factor (Shaw, Marshall and Erasmus, 1977; Basch and Nicolas, 1989) were therefore hypothesised.

5.1.3 Reproductive pairing enhances female nutritional intake

It has also been suggested that male worms enhance the nutritional intake of glucose and cholesterol by female worms, whilst paired. It was subsequently hypothesised that this male-driven nourishment may contribute to sexual maturation of the female (Cornford, 1986). Although the precise mechanism remains unknown, in pairing experiments performed with various partial sections of male and female worms, Popiel and Basch (1984) demonstrated that adult female worms were able to survive in the absence of a functioning gut, instead obtaining nutrients via uptake through the tegument. Maturation of the vitelline gland and production of mature ova was, however, limited in these females. Nutrients taken via the tegument may therefore be insufficient in attaining the high energy requirements of full reproductive maturity and egg production, and additional ingestion of erythrocytes (red blood cells) may also be necessary (Wang *et al.*, 2015).

5.1.4 Excreted and secreted proteins are involved in interactions between the parasite and host

Schistosomes are known to dynamically interact with their environment, including signalling between the male and female worm, as well as interaction with the host. Excretory/secretory products (ES) have been implicated in these interactions (Hewitson, Grainger and Maizels, 2009) and therefore function closely at the parasite-host interface. This poses an opportunity for recognition of parasite proteins by the host's immune cells.

Classically secreted proteins can be identified by the presence of N-terminal signal peptides, that target the protein to the secretory pathway (Pelley, 2012). Alternative secretion pathways also exist, and proteins may have internal rather than N-terminal signatures that suggest that they are ES. Another atypical mode of protein secretion may involve extracellular vesicles (EVs).

EVs are small membrane-bound vesicles that are classified into two groups, exosomes or microvesicles, depending on their size and biological functionality (Marcilla *et al.*, 2014). EVs play a variety of roles in parasite-parasite and parasite-host interactions and are now understood to be able to execute defined biological functions (Coakley, Maizels and Buck, 2015; Coakley, Buck and Maizels, 2016; Wu *et al.*, 2018) in addition to carrying biologically active molecules to their targets (Margolis and Sadosky, 2019); facilitating communication through mediating the transmission of biological signals and immunomodulatory actions. Platyhelminth EVs, as they are recognised today, were first described in *Echinostoma caproni* and *F. hepatica* (Marcilla *et al.*, 2012), although there is earlier reference made to terms such as membrane-bound vesicles and multivesicular bodies in the literature, including vesicles extruding from the *Schistosoma* tegument (Sturrock, 2001). More recent studies have explored the proteomic composition of the specific EV enriched protein fractions of ES products (Nowacki *et al.*, 2015). These studies found that EV proteins tend to be subset of atypical ES products, which may imply that EVs are a vehicle for non-classical excretion/secretion of proteins (Nowacki *et al.*, 2015). It has been proposed that EVs may constitute the principal means by which proteins are exported in trematodes (Marcilla *et al.*, 2014).

Proteomic analysis of purified EV fraction identified several proteins implicated in schistosome feeding (Sotillo *et al.*, 2016). This concurs with the findings of a later study, which found that enzymes were abundant in exosome-like vesicles in *S. mansoni* (Samoil *et al.*, 2018), including several proteases. A large proportion of the EV proteins recognised by Sotillo *et al.* (2016) were also homologues of vaccine candidates in *S. mansoni* and other trematode species. It was not previously understood how many of these proteins would be exposed to the host immune system

in the live worm; however, EVs may provide one mechanism of exposure to host immune cells. The expulsion of digestive by-products is another.

5.1.5 *Schistosoma* gut proteins are expelled during feeding

Schistosomes have a blind-ended gut, feeding therefore results in regurgitation of waste products back into the host bloodstream (Morris, 1968; Skelly *et al.*, 2014). This process provides another key moment for exposure of parasite proteins to the host immune system, as demonstrated by specific antibody response to CAA, the protein regurgitated as a by-product of digestion that is commonly used as a diagnostic marker of worm burden (Corstjens *et al.*, 2014). Since the ingestion of erythrocytes, and their subsequent digestion, is believed to be necessary for the production of viable eggs by females (Wang *et al.*, 2015), it is likely that any disruption to the course of nutritional intake or absorption would impact fecundity.

Delcroix *et al.* (2007) first explored the host and parasite constituent proteins found in *S. mansoni* regurgitant. MS/MS technology has since improved and subsequent studies have provided a more detailed breakdown of the constituents of schistosome vomitus (Hall *et al.*, 2011; Figueiredo *et al.*, 2015). It has been suggested that these digestive tract proteins are potential candidates for vaccine development (Figueiredo *et al.*, 2015)

5.1.6 Previous vaccine candidates and worm fecundity

One of the WHO leading vaccine candidates, 28 kDa GST had previously shown promise following early vaccination studies that demonstrated an apparent reduction in worm fecundity and egg viability (Riveau *et al.*, 1998). The vaccine potential of GST was subsequently studied in several *Schistosoma* species, with the trial of the recombinant *S. haematobium* vaccine, Bilhvax (rSh28GST) progressing to a phase 3 human study, conducted in Senegal (Riveau *et al.*, 2012). Vaccination with the recombinant Sh28GST vaccine did not, however, prove efficacious despite specific IgG₁, IgG₂ and IgG₄ having been raised in the vaccination arm (Riveau *et al.*, 2018). Furthermore, immunoepidemiological studies looking specifically at worm fecundity did not replicate the association between GST specific antibody responses and a reduction in worm fecundity (Wilson *et al.*, 2014). In recent years the use of protein microarrays to efficiently screen potential vaccine candidates has become widespread (Driguez *et al.*, 2010, 2015; Gaze *et al.*, 2014; Pearson *et al.*, 2015).

5.1.7 Identifying anti-fecundity targets

Since the eggs produced by mature female worms are responsible for the severe schistosomiasis pathology that results from chronic infection, a better understanding of the biological targets of a possible anti-fecundity immune response could have a dual benefit of inform potential future vaccine candidates as well as advancing our knowledge to predict transmission and morbidity in mathematical models.

Mutapi *et al.* (2005) previously defined those schistosome adult worm proteins recognised specifically by IgG in serum samples from study population in Zimbabwe and, in a later study (Mutapi *et al.*, 2008), the authors described the differential responses of antibody isotypes to *S. haematobium* worm antigen, finding that IgG₁ reacted to the greatest number of antigens; a finding that would be expected since IgG₁ is the most abundant isotype in human serum (Janeway *et al.*, 2001). Other studies have defined antibody responses to specific recombinant antigens or to whole worm antigen preparations more generally.

Female worms have been recognised as a major target for vaccination as a result of their production of tissue-damaging eggs (Tebeje *et al.*, 2016). A recent analysis by Lu *et al.* (2019) however, highlights the often neglected importance of males in the reproductive pairing.

5.1.8 Chapter aims and objectives

In the post-genomic era, there has been an increasing focus on large-scale analysis of genes, and their associated protein products, to perform high throughput screening for vaccine candidates; however these have primarily been conducted for *S. mansoni* and *S. japonicum* (Driguez *et al.*, 2010), for which viable animal models exist and the genome is better annotated. Yet evidence for the natural acquisition of an anti-fecundity immune response indicates that this may be specific to *S. haematobium*, amongst the schistosomes important for human health (Agnew *et al.*, 1996).

The unique contribution of this chapter is the analysis of proteins that are specifically associated with reduced worm fecundity scores. There are limited studies that have paired CAA and urine filtration egg count data from which fecundity scores can be calculated. An association has previously been shown between reduced fecundity scores and IgG₁ to SWA (Wilson *et al.*, 2014). The aim of this chapter is to advance the current understanding of specific antigens recognised by IgG₁ isotype antibodies in sera from individuals in whom reduced fecundity scores have been identified. This will be achieved through the following objectives:

- Identify proteins that invoke an IgG₁ response in individuals with reduced fecundity scores compared to age, sex and village matched controls.
- Clone the coding DNA sequence of identified proteins and express and purify recombinant proteins.
- Perform serological analysis of IgG₁ isotype responses to recombinant anti-fecundity target proteins.

5.2 Results

5.2.1 Separation of *S. haematobium* soluble worm antigen by two-dimensional gel electrophoresis

When the soluble proteins of adult *S. haematobium* were subjected to 2D gel electrophoresis, silver nitrate staining of the 2D gel revealed a discrete protein profile. The apparent molecular weight of the polypeptides was observed in the approximate range of 200 to 10 kDa and the polypeptides were distributed according to their respective isoelectric point (pI) values along a non-linear gradient between pH 3 to 10 (Fig. 5.1).

5.2.2 Identification of immunoreactive proteins by Western blot

To identify IgG₁ binding proteins, 2D Western blots of the ShSWA 2D gels were performed in parallel with pooled serum from 10 case individuals with reduced fecundity scores and pooled serum from 10 individuals control individuals, matched for age, sex and village (as a proxy for transmission intensity). A detailed description of the selection process for individuals in the case and control sub-cohorts can be found in Chapter 2, section 2.16.10. Images of the case and control blots are presented in Figures 5.2A and 5.2B, respectively. Following alignment of the gel and immunoblot images using SpotMap software (TotalLab, Newcastle, UK), described in Chapter 2, section 2.6, 32 spots were identified that were uniquely detected by serum from the cases but not by the controls (Fig. 5.1). Table 5.1 provides the numbers of spots identified that were unique to the case blots, unique to the control blots or common to both, for each technical replicate and the combined results. Figure 5.2 C–F illustrates the identification of spots with the spot mapping software, using both the three-dimensional (3D) profile of spot intensity (Fig. 5.2C and 5.2D) and the 2D blot image (Fig. 5.2E and 5.2F) from a sample of the 2D Western blots. Spots that were unique to the cases in both replicates were mapped to the silver stained gel (Fig. 5.1), manually excised and submitted to analysis by mass spectrometry.

Table 5.1 Number of spots identified in case and control blots for each technical replicate

Serum condition	Replicate 1	Replicate 2	Replicate 1 & 2
Case	84	72	32
Control	10	3	0
Common to both	51	40	26

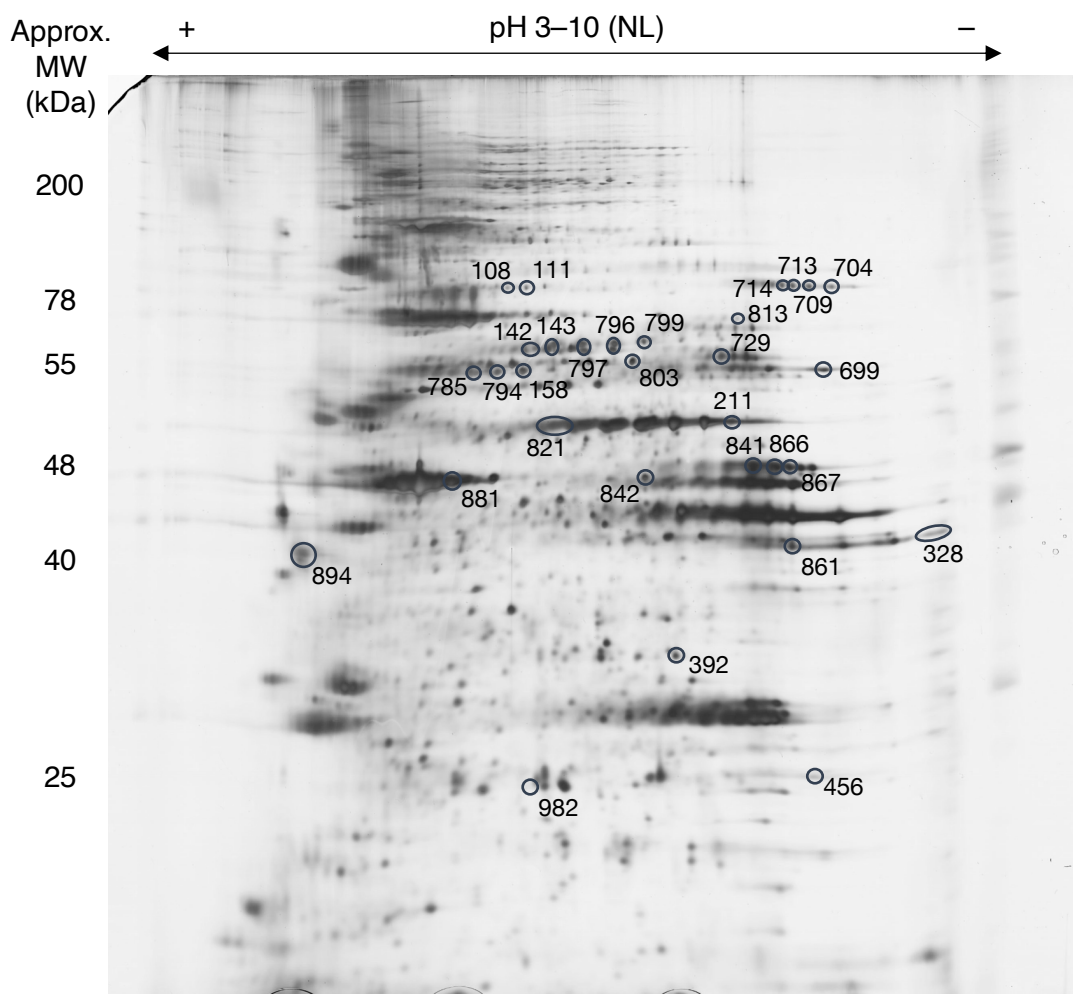


Fig. 5.1 Two-dimensional gel analysis of *S. haematobium* soluble whole worm antigen preparation (Theodor Bilharz Institute, Cairo, Egypt). Proteins were separated by 2D gel electrophoresis with first dimension isoelectric focussing performed along a non-linear (NL) pH gradient of 3–10. The gel was silver stained to visualise the spots. Spots marked on the gel were found to be differentially immunoreactive to pooled case serum compared to control serum. And were cut from the gel and sent for LC-MS/MS analysis. For protein identifications see Appendix 5, Table 5.1

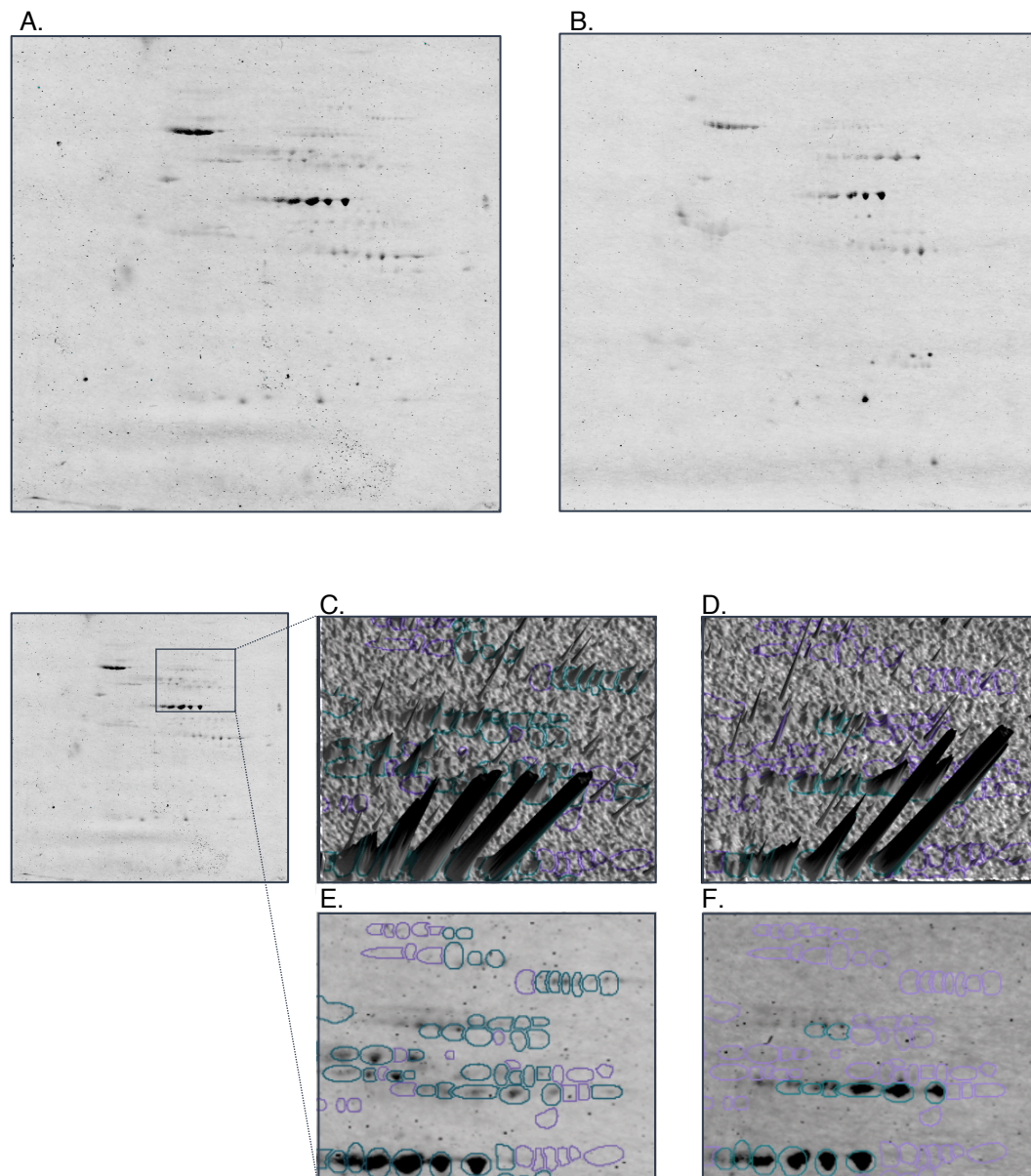


Fig. 5.2. Two-dimensional Western blots identifying IgG1 immunogenic protein spots in A) pooled anti-fecundity case serum (10 individuals) and B) age, sex and village matched control serum (10 individuals). C) and D) show 3D image of spot mapping process in case and control serum blots, respectively. Turquoise: spot present; purple: spot not present; E) and F) show corresponding 2D regions in case and control serum blots, respectively. Inset depicts area from which images of spot mapping process are sampled.

5.2.3 Protein identification

Peptides were identified in the *S. haematobium* database (version 1; Bioproject PRJNA78265, 13073 sequences; 5666088 residues) using the Mascot algorithm (Matrix Science, London UK, version 2.6.0). Search parameters are provided in Chapter 2. This resulted in a total of 459 significant proteins (scores greater than 19 are significant ($p < 0.05$)) and 191 non-redundant proteins (where identical proteins hits were counted as a single record). A full list of identified proteins can be found in Appendix 5, Table A5.1. The 191 non-redundant proteins will from here on be referred to as the *anti-fecundity targets*. A bioinformatic approach was then taken to functionally characterise the anti-fecundity targets.

Since the *S. haematobium* genome is relatively poorly annotated compared to that of *S. mansoni* (Stroehlein *et al.*, 2019), parallel searches were performed against the *S. mansoni* database (version 5.2; Bioproject PRJEA36577, 11774 sequences; 5608274 residues). 171 *S. mansoni* proteins were identified from the MASCOT database searches.

5.2.4 Bioinformatic analysis of mass spectrometry results

5.2.4.1 Excreted/secreted proteins and extracellular vesicles

In order for the specific antibody response to affect live worm fecundity, it logically follows that the target protein would be visible to the host immune system in the live adult worm. There are several means through which proteins are exposed to host antibodies, including expression of the protein on the surface of the parasite, through excretion or secretion (ES), or via release in extracellular vesicles (EVs).

The candidate anti-fecundity targets were initially screened for those that contain either a classical N-terminal or non-classical, internal secretory signature, as described in Chapter 2. Only a small number of proteins were found to include a N-terminal signal peptide (3.1%; 6/191), as determined by SignalP; whereas 27 proteins (14.1%) were identified as likely non-classically secreted protein by searching SecretomeP. These were classified as potential secretory proteins. Nine proteins were predicted to have at least one trans-membrane domain (4.7%; 9/191), two of which were also predicted to have either an N-terminal or internal secretory signal. The remaining majority were not found to contain a signal peptide (85.3%; 163/191).

As discussed in section 5.1.4, proteins released via EVs are believed to mediate a variety of biological and molecular functions understood to include, amongst others, energy generation and host-parasite interactions. Proteins associated with EVs are therefore conceivable candidates for facilitation of the male-female interaction resulting in egg production. *S. mansoni*

protein hits were screened against entries common to previously annotated eukaryotic extravesicular proteins (Nowacki *et al.*, 2015). 18 proteins found to be present in the EV fraction by Nowacki *et al.* (2015) were also immunogenic for IgG₁ in human anti-fecundity case serum in this study (18/171, 10.5%). Eight of these EV proteins have also been shown to be within the top 30 most abundant proteins in SWA (Neves *et al.*, 2015) (Appendix 5, Table 5.2).

It is important to note that *S. mansoni* homologs were used in the analysis of EV associated proteins to inform the functional analysis of the equivalent *S. haematobium* proteins for which data is not available.

5.2.4.2 Gene Ontology (GO) enrichment analysis

To gain a better understanding of the functions of those proteins that were immunogenic to anti-fecundity case serum, GO term enrichment analysis was performed on their respective genes. Although, at the time of analysis, almost half of the *S. haematobium* genome still remained to be annotated (6423/13073 genes; 49.1%), of the 191 proteins identified from IgG₁ immunogenic spots in this study a total of 182 (95.3%) were annotated. 77 level 2 GO terms were enriched with $p < 0.05$, comprising 43 terms relating to biological processes (BP) (Table 5.2); 13 terms relating to cellular components (CC) (Table 5.3), and 21 terms relating to molecular functions (MF) (Table 5.4). Figure 5.3 illustrates the number of genes, from the list of 191 anti-fecundity targets, that are associated with each presented GO term for molecular function (green) and biological process terms (blue). Interestingly, genes associated with reproduction-related GO terms were found to be enriched, including the terms: oocyte microtubule cytoskeleton polarisation; ovulation; oviposition; developmental process involved in reproduction, and embryo development resulting in birth or egg hatching. Genes related to energy production were also found to be enriched, in addition to those involved in digestive tract development, locomotion and body morphogenesis, all of which could feasibly relate to disruption of schistosome pairing, energy requisition and subsequent egg production. The relationship between highly similar GO terms is illustrated in Appendix 5 (Fig. A5.1 and A5.2).

Enriched terms relating to cellular components (CC) are presented in Table 5.3. A quarter of genes of cytosolic origin, corresponding to the anti-fecundity target proteins, were found to be significantly enriched (51/191; 26.7%). It has previously been shown that the most abundant proteins in SWA are most commonly identified and these are often cytosolic (Wilson *et al.*, 2007; Neves *et al.*, 2015). There was also a relatively high representation of structural cell components, such as cytoskeletal parts. Van Balkom *et al.* (2005) found several cytoskeletal proteins to be targets of IgG₁. The list of candidate anti-fecundity target proteins demonstrated

a high representation of tegumental proteins, which may be transiently expressed on the surface, thus eliciting the host immune response. Interestingly genes associated with intracellular and membrane bound organelles were also highly enriched.

Table 5.2 GO terms relating to *biological processes* that are significantly over-represented within the 191 proteins identified as immunostimulatory to anti-fecundity case serum

GO accession	Definition	Ann	Sig	Exp	p.value
GO:0006457	protein folding	55	14	1.79	1.10E-09
GO:0006096	glycolytic process	13	6	0.42	1.50E-06
GO:0006418	tRNA aminoacylation for protein translation	40	11	1.3	3.00E-05
GO:0018996	molting cycle, collagen and cuticulin-based cuticle	9	4	0.29	0.00012
GO:0006108	malate metabolic process	5	3	0.16	0.00032
GO:0055114	oxidation-reduction process	183	15	5.94	0.00049
GO:0008103	oocyte microtubule cytoskeleton polarization	6	3	0.19	0.00062
GO:0005975	carbohydrate metabolic process	103	18	3.35	0.00063
GO:0030835	negative regulation of actin filament depolymerization	6	3	0.19	0.00103
GO:0040011	locomotion	173	13	5.62	0.00122
GO:0010171	body morphogenesis	30	5	0.97	0.00249
GO:0009058	biosynthetic process	940	29	30.53	0.00266
GO:0030728	ovulation	3	2	0.1	0.00308
GO:0050829	defence response to Gram-negative bacterium	3	2	0.1	0.00308
GO:0051098	regulation of binding	3	2	0.1	0.00308
GO:0046716	muscle cell cellular homeostasis	3	2	0.1	0.00308
GO:0006432	phenylalanyl-tRNA aminoacylation	3	2	0.1	0.00308
GO:0019752	carboxylic acid metabolic process	127	27	4.12	0.00436
GO:0006508	proteolysis	259	16	8.41	0.00508
GO:0007264	small GTPase mediated signal transduction	136	11	4.42	0.00559
GO:0040040	thermosensory behaviour	4	2	0.13	0.00602
GO:0042147	retrograde transport, endosome to Golgi	4	2	0.13	0.00602
GO:0072499	photoreceptor cell axon guidance	4	2	0.13	0.00602
GO:0008340	determination of adult lifespan	59	6	1.92	0.01155
GO:0003006	developmental process involved in reproduction	135	11	4.38	0.01324
GO:0009408	response to heat	10	3	0.32	0.01425
GO:0050821	protein stabilization	6	2	0.19	0.01442
GO:0006098	pentose-phosphate shunt	6	2	0.19	0.01442
GO:0006006	glucose metabolic process	6	2	0.19	0.01442
GO:0000910	cytokinesis	18	3	0.58	0.01914
GO:0018991	oviposition	18	3	0.58	0.01914
GO:0070374	positive regulation of ERK1 and ERK2 cascade	7	2	0.23	0.01976
GO:0007611	learning or memory	15	3	0.49	0.02553
GO:0046579	positive regulation of Ras protein signal transduction	8	2	0.26	0.02579
GO:0045448	mitotic cell cycle, embryonic	8	2	0.26	0.02579
GO:0048565	digestive tract development	8	2	0.26	0.02579
GO:0008154	actin polymerization or depolymerization	21	4	0.68	0.03194
GO:0009225	nucleotide-sugar metabolic process	5	2	0.16	0.03228
GO:0043279	response to alkaloid	6	2	0.19	0.03229
GO:0055082	cellular chemical homeostasis	18	2	0.58	0.03238
GO:0045214	sarcomere organization	9	2	0.29	0.03246
GO:0007059	chromosome segregation	52	4	1.69	0.0394
GO:0009792	embryo development ending in birth or egg hatching	101	7	3.28	0.04476

Table 5.3 GO terms relating to *cellular component* that are significantly over-represented within the 191 proteins identified as immunostimulatory to anti-fecundity case serum

GO accession	Definition	Ann	Sig	Exp	p.value
GO:0005737	cytoplasm	880	51	27.5	4.90E-08
GO:0044464	cell part	2253	95	70.41	0.00025
GO:0005875	microtubule associated complex	101	9	3.16	0.00043
GO:0005865	striated muscle thin filament	5	3	0.16	0.00095
GO:0044430	cytoskeletal part	220	23	6.88	0.00107
GO:0043229	intracellular organelle	1513	56	47.28	0.00287
GO:0043227	membrane-bounded organelle	1102	29	34.44	0.00904
GO:0008180	COP9 signalosome	5	2	0.16	0.00909
GO:0005874	microtubule	34	5	1.06	0.01002
GO:0005739	mitochondrion	123	9	3.84	0.01353
GO:0045169	fusome	8	2	0.25	0.02394
GO:0000786	nucleosome	21	3	0.66	0.02616
GO:0005811	lipid droplet	23	3	0.72	0.03331

Table 5.4 GO terms relating to *molecular function* that are significantly over-represented within the 191 proteins identified as immunostimulatory to anti-fecundity case serum

GO accession	Definition	Ann	Sig	Exp	p.value
GO:0051082	unfolded protein binding	28	10	0.89	6.40E-09
GO:0005525	GTP binding	138	17	4.38	1.30E-06
GO:0016616	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	28	8	0.89	6.10E-06
GO:0005524	ATP binding	604	38	19.16	2.00E-05
GO:0004812	aminoacyl-tRNA ligase activity	39	11	1.24	2.10E-05
GO:0004177	aminopeptidase activity	10	4	0.32	0.00018
GO:0070008	serine-type exopeptidase activity	5	3	0.16	3.00E-04
GO:0016615	malate dehydrogenase activity	5	3	0.16	3.00E-04
GO:0008235	metalloexopeptidase activity	13	4	0.41	0.00056
GO:0016868	intramolecular transferase activity, phosphotransferases	7	3	0.22	0.001
GO:0016774	phosphotransferase activity, carboxyl group as acceptor	3	2	0.1	0.00294
GO:0004826	phenylalanine-tRNA ligase activity	3	2	0.1	0.00294
GO:0005093	Rab GDP-dissociation inhibitor activity	3	2	0.1	0.00294
GO:0003924	GTPase activity	62	7	1.97	0.00323
GO:0030170	pyridoxal phosphate binding	23	4	0.73	0.0054
GO:0000287	magnesium ion binding	38	5	1.21	0.00651
GO:0050661	NADP binding	5	2	0.16	0.00939
GO:0046982	protein heterodimerization activity	43	5	1.36	0.01099
GO:0051015	actin filament binding	17	3	0.54	0.01537
GO:0030145	manganese ion binding	7	2	0.22	0.01891
GO:0046933	proton-transporting ATP synthase activity, rotational mechanism	8	2	0.25	0.02469

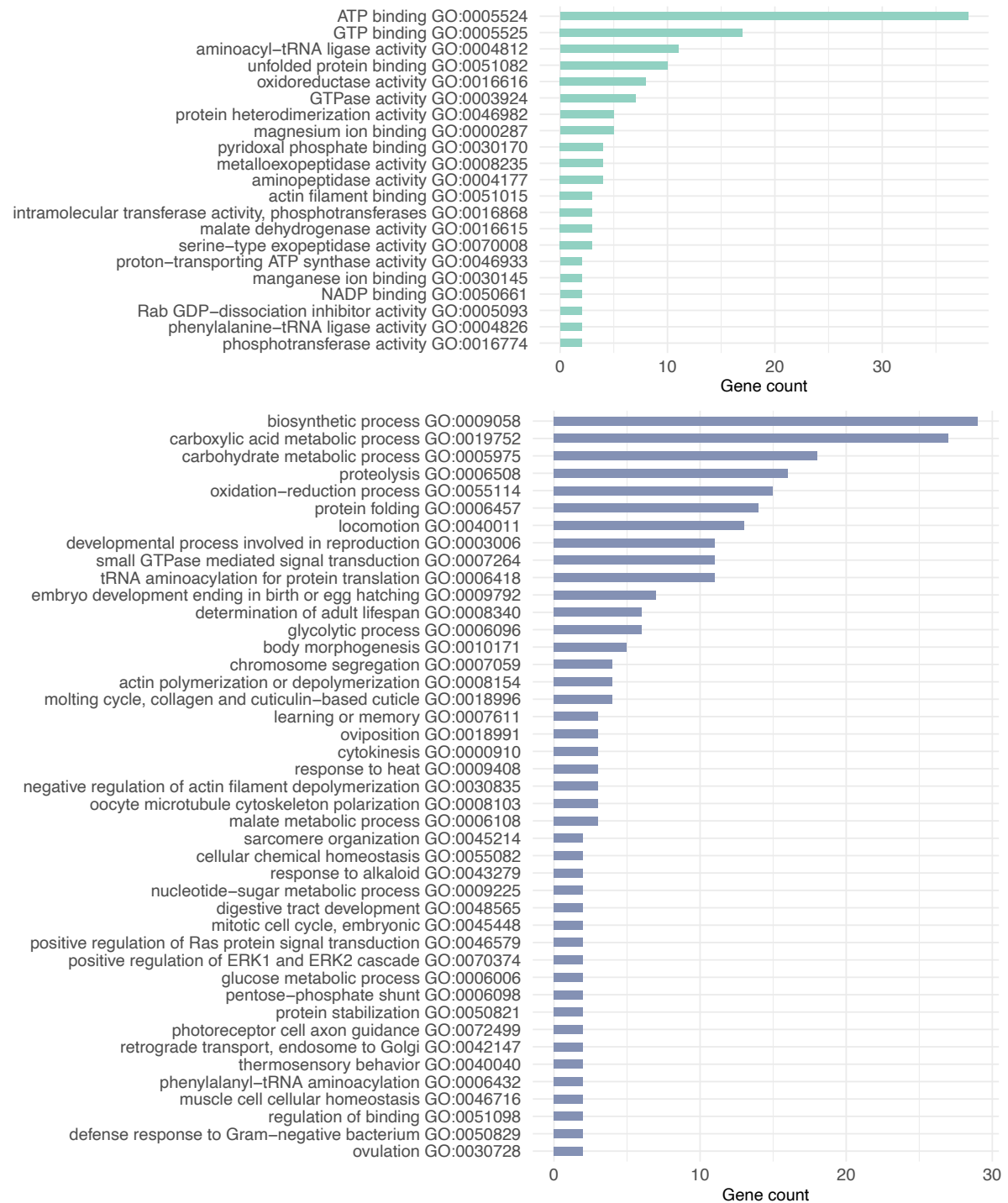


Fig. 5.3 Number of genes associated with molecular function (green) and biological process (blue) GO term that are over-represented in the 191 candidate anti-fecundity targets compared to the GO terms annotated to whole *S. haematobium* genome (v1.0; PRJNA78265 (Young *et al.*, 2012)).

5.2.4.3 Proteins identified within the gastrodermis and gastric tract

The anti-fecundity target proteins were cross-referenced against peer-reviewed literature reporting the proteomic analysis of *Schistosoma* digestive tract proteins (Hall *et al.*, 2011; Figueiredo *et al.*, 2015). 15 of the anti-fecundity target proteins were identified in the gut proteome literature (15/191, 7.9%), representing 21% of the 71 proteins identified by Hall *et al.* (2011). These comprised a variety of cellular functions including proteins from the worm tegument, lysosomal proteins, proteins involved in glycolysis/energy and structural/motor proteins. The 15 in common proteins are presented in Appendix 5, Table A5.3.

Additionally, several peptidases listed in the MEROPS peptidase database (European Bioinformatics Institute (EBI), 2018; Rawlings *et al.*, 2018) were identified as candidate anti-fecundity targets (17/191; 8.9 %). These included lysosomal proteins (Lysosomal protective protein (also known as serine carboxypeptidase A [MEROPS accession: MER0941858]); Dipeptidyl peptidase 2(MS3_05939; MERI053534)); metallopeptidases including a putative amino peptidase with homology to the leucine aminopeptidases (LAP) in *S. mansoni* (85% homology to Smp_030000) (M17 aminopeptidase; MER0929572), thimet oligopeptidase (M3; MER0817167), and a prolyl endopeptidase (MER0991507). A full list of identified peptidases can be found in Appendix 5, Table A5.4.

5.2.4.4 Gender-associated gene expression

Alongside publication of the *S. haematobium* genome, Young *et al.* (2012) published data on protein-coding genes that were inferred to be significantly developmentally regulated in *S. haematobium* adult males, adult females or eggs.

The anti-fecundity target proteins, associated with reduced fecundity were not differentially representative of genes enriched in *S. haematobium* females, males or the egg, since the proportion of genes corresponding to anti-fecundity target proteins that were found to be enriched did not significantly differ from the proportional enrichment for each stage within the whole genome. Proportions were compared using the χ^2 test, results of which are presented in Table 5.5.

Table 5.5 Summary of anti-fecundity target and whole genome genes significantly developmentally regulated in *S. haematobium* female worm, male worm and egg,

Life cycle stage	No. of enriched genes		χ^2	P-value
	Anti-fecundity target (n = 191)	Whole genome* (n = 13,073)		
Adult female [†]	13 (6.8%)	965 (7.4%)	0.15506	0.6937
Adult male [‡]	37 (19.4%)	2010 (15.4%)	1.0372	0.3085
Egg ^{†‡}	29 (15.2%)	1629 (12.5%)	0.45359	0.5006

* Data from Young *et al.* (2012)

[†] versus male

[‡] versus female

Results from a comprehensive study of gender associated gene expression in *S. mansoni* were used to better inform the suggestive expression profiles of the anti-fecundity targets. Lu *et al.* (2016) explored the differential expression of genes in male and female *S. mansoni* worms, in isolated gonads from paired and unpaired worms, and in *S. mansoni* eggs. The anti-fecundity targets were screened for accession numbers that coincided with those presented in the analysis by Lu *et al.* (2016). The presence of differentially expressed genes in males, females and their respective gonads are illustrated in Figure 5.4. Thimet oligopeptidase (MS3_09903) and transketolase (MS3_03775) were both found by Lu *et al.* (2016) to be enriched in paired compared to single females (Fig. 5.4C). Several genes from the anti-fecundity targets were found to be upregulated in the egg compared to the gonads from paired worms (Fig. 5.4F & 5.4G). No significantly differentially expressed genes were found in the single males versus single females or between paired versus single males.

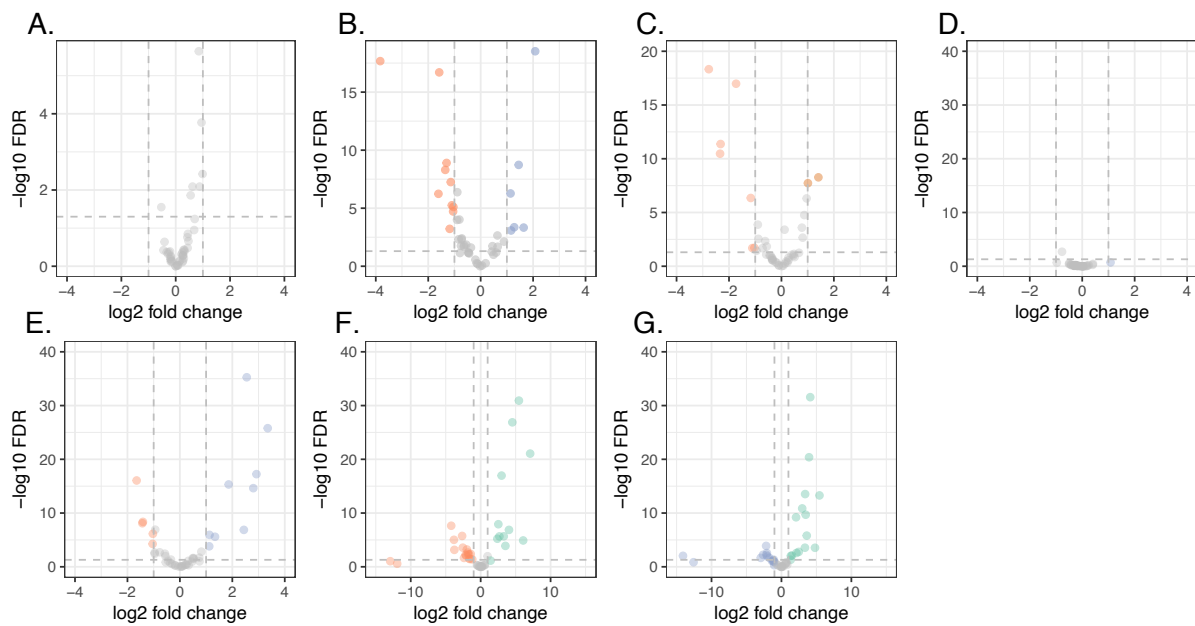


Fig. 5.4 Volcano plots showing the relative expression of orthologous *S. haematobium* genes corresponding to anti-fecundity target proteins in *Schistosoma* males, females and their respective gonads. A) single males vs. single females; B) testes from paired males vs. ovaries from paired females; C) paired vs. single female worms; D) paired vs. single males; E) paired males vs. paired females; F) egg vs. ovaries from paired females; G) egg vs. testes from paired males. Orange: Female worm or gonad; blue: male worm or gonad; green: egg. Differential expression was defined as a log fold change (\log_2FC) greater than or equal to 1 (equivalent to fold change of 2) and FDR (adjusted p-value) < 0.05. Data from Lu *et al.* (2016).

5.2.5 Selection of proteins for downstream recombinant expression

Cloning and expression of all 191 proteins was not within the scope of this thesis. Instead, 48 proteins were initially selected to be taken forward for HTP cloning and expression screening. Details of the selection criteria can be found in Chapter 2, section 2.8, summarised for each protein in Table 5.6. Proteins were initially selected based on those proteins with the top five Mascot scores identified in the spots with the highest spot intensity across the two technical replicates. Proteins for which coverage of unique peptides was highest in each of the differentially immunogenic spots were also selected. Additional proteins of interest were then identified by further manual curation. Several of these proteins were also ES, either through identified peptide signatures (9/48 classically or non-classically secreted peptide signatures; section 5.2.4.1), or in peer-reviewed literature: as ES products (Young *et al.*, 2012; Floudas *et al.*, 2017), EVs (Nowacki *et al.*, 2015) or gut proteins (Hall *et al.*, 2011; Figueiredo *et al.*, 2015). 11/48 proteins had non-classical secretion signatures that met the cut-off for bacterial organisms,

narrowly missing the cut-off for mammals (Table 5.7). Others were selected for enriched GO terms associated with reproductive processes (Tropomyosin (MS3_01364); Serine/threonine-protein phosphatase (MS3_08776)). Proteins that were also identified as differentially upregulated in paired compared to single females worms or isolated gonads were also selected, according to analysis by Lu *et al.* (2016, 2017) (section 5.2.4.4).

Peer-reviewed published evidence was also considered in support of protein selection for downstream analysis. For example, putative aminopeptidase W07G4.4 (A_04283) (ShLAP) has homology to *S. mansoni* leucine aminopeptidase (SmLAP, Smp_030000; 85% homology $E < 0.0001$). McCarthy *et al.* (2004) identified a role for *S. mansoni* and *S. japonicum* LAP in the cleavage of peptides into free amino acids in the gut lumen of schistosomes. This function appears to be conserved in other invertebrate species, LAP has been identified as a digestive tract protein, and is associated with digestion of host blood during feeding (Hatta *et al.*, 2006). Furthermore, LAP null mutants have demonstrated delayed onset or reduced egg-laying in *C. elegans* (Joshua, 2001) and RNAi studies in *S. mansoni* suggest a role in egg hatching, making this protein an interesting candidate for further analysis.

Finally, proteins were removed that were found to be highly abundant in SWAP (Neves *et al.*, 2015) or for which the spot position on the gel was not at the expected molecular weight. The selected proteins covered a range of biological processes feasibly associated with mediation of worm fecundity. A list of the 48 selected proteins are presented in Table 5.6.

Table 5.6. List of 48 proteins selected for high-throughput expression screening

Accession	Product	Size (aa)	MW	Peptides (unique)	Coverage	Selection criteria*
A_07016	Alpha-adducin(MS3_05669)	575	62.5	8(7)		5
A_05991	Annexin A7(MS3_04598)	548	60			3, 4, 8, 10
A_04685	ATP-dependent RNA helicase eIF4A(MS3_01780)	286	33	12(12)	42.7%	3
A_04102	cAMP-dependent protein kinase type II regulatory subunit(MS3_07384)	336	38	6(6)	21.1%	5, 3, 6
A_08296	COP9 signalosome complex subunit 4(MS3_02487)	437	49	11(11)	30.0%	1
A_00035	Cystathionine beta-synthase(MS3_01163)	788	89			6
A_00281	Dihydropyrimidinase-related protein 4(MS3_08954)	463	51	14(12)	38.9%	3
A_04268	Dipeptidyl peptidase 2(MS3_05939)	434	49	4(4)	9.9%	3, 8
A_03435	Elongation factor 2(MS3_08361)	849	95	9(9)	10.1%	9
A_02036	Fragile X mental retardation protein 1-like protein(MS3_07796)	770	83	3(3)	5.1%	3
A_03341	Glutamine synthetase(MS3_07403)	413	46	28(15)	40.4%	2
A_06665	Glutamine--tRNA ligase(MS3_06258)	562	65	24(22)	44.7%	6
A_05311	Glycogen phosphorylase, liver form(MS3_01181)	841	97			9
A_06781	Glycogenin-1(MS3_09086)	282	32	2(2)	10.6%	4
A_00125	hypothetical protein(MS3_04659) - BLAST zinc finger	581	68	9(9)	21.3%	6
A_06542	hypothetical protein(MS3_05150) - BLAST EF hand protein	342	41	11(10)	27.8%	6
D_00394	hypothetical protein(MS3_11411) - BLAST HSP70	372	40.6		11	5
C_00979	hypothetical protein(MS3_11537) - BLAST MF3 Sj eggshell	205	22	2(2)	10.2%	9
A_06413	Kinesin light chain(MS3_07007)	693	79	22(20)	35.2%	6
D_00386	L-lactate dehydrogenase A chain(MS3_10510)	301	33	24(14)	44.2%	2, 3, 9
B_00767	Lethal(2) giant larvae protein 1(MS3_10379)	1358	147	9(8)	8.5%	3
A_01815	Lysosomal protective protein(MS3_04617)	500	57	3(3)	5.8%	3
A_02474	Major egg antigen(MS3_03417)	467	52	7(7)	19.3%	3

B_00584	Major vault protein(MS3_I1386)	16	17	6(6)	46.0%	3
A_07550	NADP-dependent malic enzyme(MS3_01897)	565	64	7(7)	12.9%	
A_05241	Peroxioredoxin-2	219	25	2(2)		3
A_03579	Phosphoglucomutase-1(MS3_09994)	575	63	57(26)	56.0%	1, 2, 3
A_03905	Phosphoglycerate kinase(MS3_04778)	168	18	25(11)	73.2%	2, 9
A_04622	Phosphorylase b kinase gamma catalytic chain, skeletal muscle/heart isoform(MS3_01719)	445	51	3(3)	7.9%	5
A_02645	Plastin-2(MS3_07481)	651	74	9(8)	18.0%	3
A_02310	Prolyl endopeptidase(MS3_01899)	555	63	10(10)	17.0%	10
A_02739	Protein SET(MS3_07372)	254	29	3(3)	11.0%	3
A_04283	Putative aminopeptidase W07G4.4(MS3_08450)	534	58	57(30)	56.0%	2, 3, 7, 8, 10
A_02429	Putative protein disulfide-isomerase ER-60(MS3_10187)	327	37	2(2)	7.0%	
A_01177	Putative UDP-glucose 4-epimerase(MS3_04887)	395	44	2(2)	6.6%	3, 6, 10
A_00749	putative serine--tRNA ligase, cytoplasmic(MS3_06265)	557	64	9(9)	25.9%	6
A_00105	Serine/threonine-protein phosphatase PPI-gamma catalytic subunit(MS3_08776)	292	33.2	4(4)	14.7%	5
A_05946	Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial(MS3_07763)	416	45	13(13)	35.6%	2
A_01496	T-complex protein 1 subunit eta(MS3_01627)	497	54	37(27)	64.0%	1, 2
A_02556	Taurocyamine kinase(MS3_05631)	716	80	68(46)	65.0%	2, 3, 9, 10
A_08102	Thimet oligopeptidase(MS3_09903)	664	76	55(39)	48.0%	2, 6, 8
A_08148	Thioredoxin domain-containing protein C06A6.5(MS3_02305)	230	26	2(2)	10.4%	3
A_04057	Threonyl tRNA synthetase	575	62.5	18(18)	25.0%	6
B_00064	Transaldose	548	60	5(5)	19.0%	6
A_03666	Transketolase(MS3_03775)	286	33	6(6)	17.7%	6, 9
A_06499	Tropomodulin(MS3_06431)	336	38	2(2)	5.6%	6
A_00073	Tropomyosin(MS3_01364)	437	49	25(25)		2, 5
A_00746	Troponin T(MS3_07702)	788	89	5(5)	19.7%	3, 6

*Definitions for selection criteria can be found in Chapter 2, Table 2.1

Table 5.7. List of 48 selected proteins with SignalP and SecretomeP annotation

Accession	Product	SignalP*	SecretomeP†
A_07016	Alpha-adducin(MS3_05669)	-	+/-
A_05991	Annexin A7(MS3_04598)	-	+
A_04685	ATP-dependent RNA helicase eIF4A(MS3_01780)	-	+
A_04102	cAMP-dependent protein kinase type II regulatory subunit(MS3_07384)	-	+/-
A_08296	COP9 signalosome complex subunit 4(MS3_02487)	-	-
A_00035	Cystathionine beta-synthase(MS3_01163)	-	-
A_00281	Dihydropyrimidinase-related protein 4(MS3_08954)	-	+/-
A_04268	Dipeptidyl peptidase 2(MS3_05939)	+	NA
A_03435	Elongation factor 2(MS3_08361)	-	-
A_02036	Fragile X mental retardation protein 1-like protein(MS3_07796)	-	+/-
A_03341	Glutamine synthetase(MS3_07403)	-	+/-
A_06665	Glutamine--tRNA ligase(MS3_06258)	-	-
A_05311	Glycogen phosphorylase, liver form(MS3_01181)	-	-
A_06781	Glycogenin-1(MS3_09086)	-	-
A_00125	Hypothetical protein(MS3_04659) - BLAST zinc finger	-	-
A_06542	Hypothetical protein(MS3_05150) - BLAST EF hand protein	-	+/-
D_00394	Hypothetical protein(MS3_11411) - BLAST HSP70	-	-
C_00979	Hypothetical protein(MS3_11537) - BLAST MF3 Sj eggshell	-	-
A_06413	Kinesin light chain(MS3_07007)	-	-
D_00386	L-lactate dehydrogenase A chain(MS3_10510)	-	-
B_00767	Lethal(2) giant larvae protein 1(MS3_10379)	-	+
A_01815	Lysosomal protective protein(MS3_04617)	+	NA
A_02474	Major egg antigen(MS3_03417)	-	+
B_00584	Major vault protein(MS3_11386)	-	+
A_07550	NADP-dependent malic enzyme(MS3_01897)	-	-
A_05241	Peroxiredoxin-1	-	-
A_03579	Phosphoglucomutase-1(MS3_09994)	-	+
A_03905	Phosphoglycerate kinase(MS3_04778)	-	-
A_04622	Phosphorylase b kinase gamma catalytic chain, skeletal muscle/heart isoform(MS3_01719)	-	-
A_02645	Plastin-2(MS3_07481)	-	+/-
A_02310	Prolyl endopeptidase(MS3_01899)	-	-
A_02739	Protein SET(MS3_07372)	-	-
A_04283	Putative aminopeptidase W07G4.4(MS3_08450)	-	+/-
A_02429	Putative protein disulfide-isomerase ER-60(MS3_10187)	-	+/-
A_01177	Putative UDP-glucose 4-epimerase(MS3_04887)	-	+
A_00749	serine-tRNA ligase	-	-
A_00105	Serine/threonine-protein phosphatase PPI-gamma catalytic subunit(MS3_08776)	-	+/-
A_05946	Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial(MS3_07763)	-	-
A_01496	T-complex protein 1 subunit eta(MS3_01627)	-	-
A_02556	Taurocyamine kinase(MS3_05631)	-	-
A_08102	Thimet oligopeptidase(MS3_09903)	-	-
A_08148	Thioredoxin domain-containing protein C06A6.5(MS3_02305)	-	+/-
A_04057	threonyl tRNA synthetase	-	-
B_00064	Transaldose	-	-
A_03666	Transketolase(MS3_03775)	-	-
A_06499	Tropomodulin(MS3_06431)	-	-
A_00073	Tropomyosin(MS3_01364)	-	-
A_00746	Troponin T(MS3_07702)	-	-

* SignalP score ≥ 0.45 : +; < 0.45 : -† SecretomeP score ≥ 0.6 : +; $0.6 > \text{score} \geq 0.5$: +/-; < 0.5 : -

5.2.6 High throughput cloning and expression screening

The 48 proteins selected for recombinant protein production were subject to high throughput protein cloning and expression protocols, conducted at the Oxford Protein Production Facility (OPPF-UK, Harwell, UK) and described in Chapter 2. Coding DNA sequences were cloned from complementary DNA (cDNA) from adult *S. haematobium*, Egyptian Strain, NR-48642, provided by the NIAID Schistosomiasis Resource Center for distribution through BEI Resources, NIAID, NIH. The quality of the cDNA was confirmed by PCR of two *S. haematobium* housekeeping genes.

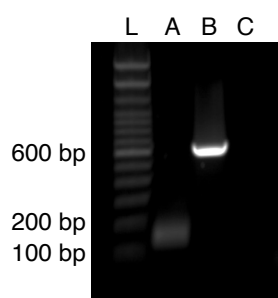


Fig. 5.5. Agarose gel showing bands for quality control PCR of *S. haematobium* cDNA (Egyptian Strain, NR-48642; NIAID Schistosomiasis Resource Center, NIH) with housekeeping genes A) ShTetraspanin 2 (TSP2; 149bp) and B) ShTropomyosin (618bp; GenBank L76202.1). C) negative control with no cDNA. 100 base pair (bp) ladder (L; #15628019, Thermo Fisher Scientific, Waltham, MA,

HTP cloning resulted in visible single bands of appropriate mass for 15 of the 48 proteins (Fig. 5.6). Six of these potential IgG₁-binding proteins were successfully cloned in *E. coli* using the small-scale expression protocol described in Chapter 2, section 2.13.8. The resulting SDS PAGE gel is displayed in Figure 5.7. Subsequently, three recombinant proteins: putative aminopeptidase W07.G4.4 (ShLAP, MS3_08450), taurocyamine kinase (ShTK, MS3_05631) and putative UDP glucose-4-epimerase (ShGALE, MS3_04887) were successfully expressed in *E. coli* during scaled-up production (Fig. 5.8). The one-dimensional Western blot images displayed in Figure 5.8 show bands of an appropriate molecular weight for each of the three expressed proteins: ShLAP at 57.7 kDa; ShTK at 80.4 kDa and ShGALE at 43.9 kDa.

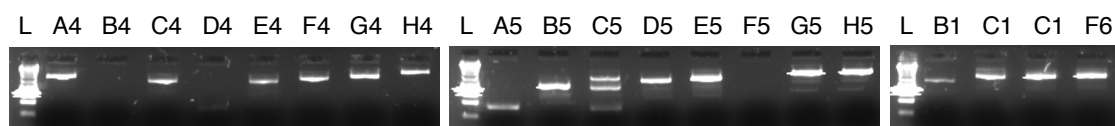


Fig. 5.6. Agarose gel showing PCR product bands for successfully cloned anti-fecundity target proteins. DNA bands were not observed for B4, D4 and F5 and the remaining 28 targets (data not shown). The band observed for A5 was of the incorrect size and so was excluded from downstream expression and C5 resulted in a double band, which was also excluded from further processing. Protein identities are provided in Table 5.8.

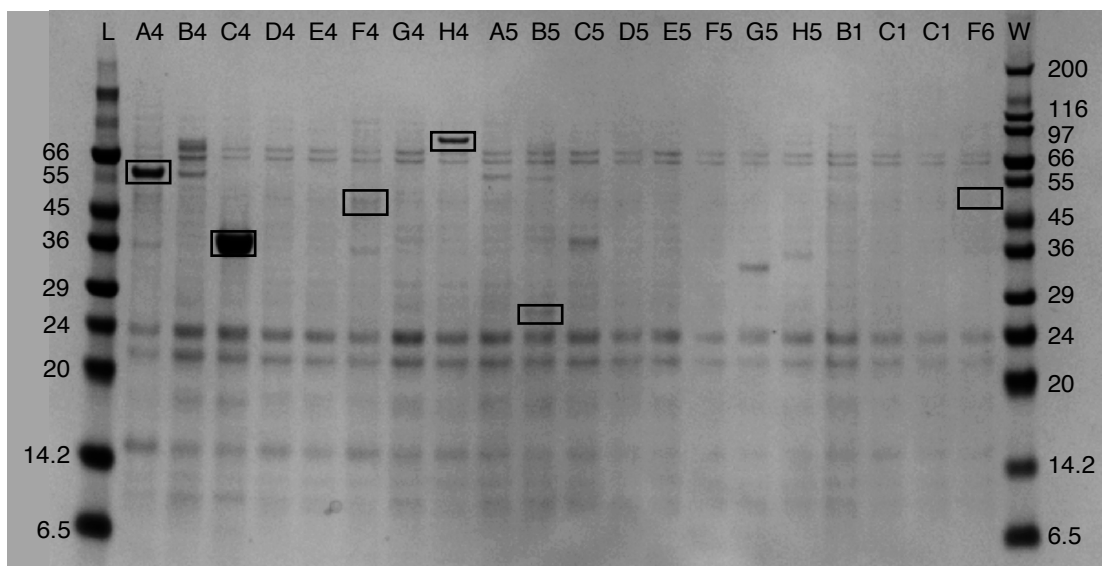


Fig. 5.7. SDS PAGE gel of small-scale protein expression screening of anti-fecundity targets. Boxed bands indicate successfully expressed proteins of the correct molecular weight. Wide range marker (W; Sigma, S8445), low range marker (L; Sigma, M3913) Protein identities are provided in Table 5.8.

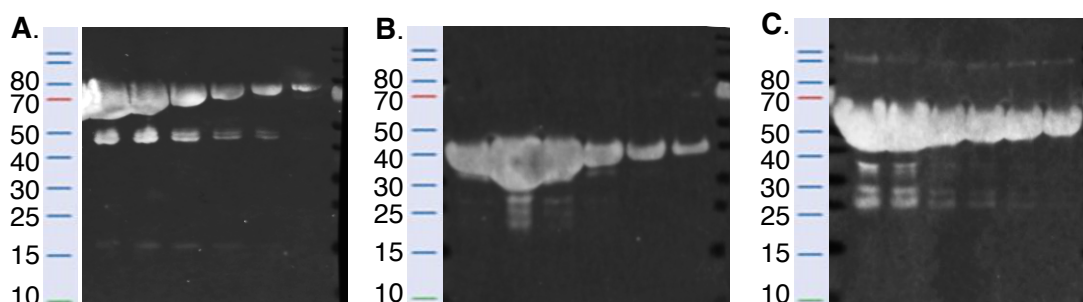


Fig. 5.8. Protein fractions eluted from Ni-NTA column A) Taurocyamine kinase (A_02556; 80.4 kDa), fractions 5–10; B) Putative aminopeptidase (A_04283; 57.7 kDa), fractions; C) UDP 4 glucose epimerase (A_01177; 43.9 kDa) fractions.

Table 5.8 Successfully cloned and expressed anti-fecundity targets

Well ID	Accession	Identity	Cloned (method)	Expression trial	Scaled expression
A4	A_04283	Putative aminopeptidase W07G4.4	+	(X55)	+
B4	A_02429	Putative protein disulfide-isomerase ER-60	-	-	-
C4	A_01177	Putative UDP-glucose 4-epimerase	+	(X55)	+
D4	A_00749	Putative serine--tRNA ligase, cytoplasmic	-	-	-
E4	A_00105	Serine/threonine-protein phosphatase PPI-gamma catalytic subunit	+	(X55)	-
F4	A_05946	Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial	+	(X55)	+
G4	A_01496	T-complex protein 1 subunit eta	+	(X55)	-
H4	A_02556	Taurocyamine kinase	+	(X55)	+
A5	A_08102	Thimet oligopeptidase	-	-	-
B5	A_08148	Thioredoxin domain-containing protein C06A6.5	+	(X55)	+
C5	A_04057	Threonyl tRNA synthetase	-	-	-
D5	B_00064	Transaldose	+	(X55)	-
E5	A_06499	Tropomodulin	+	(X55)	-
F5	A_00746	Troponin T	-	-	-
G5	A_03666	Transketolase	+	(X55)	-
H5	A_05991	Annexin A7	+	(X55)	-
B1	B_00447	ATP-dependent RNA helicase eIF4A	+	(D55)	-
C1	A_08296	COP9 signalosome complex subunit 4	+	(X55)	-
C1	A_08296	COP9 signalosome complex subunit 4	+	(D55)	-
F6	A_04268	Dipeptidyl peptidase 2	+	(X55)	+

X55: xKOD polymerase 55°C annealing temperature; D55: DreamTaq™ polymerase 55°C annealing temperature. Full cloning protocols can be found in Chapter 2, section 2.13.2

ShLAP was identified as a top hit within the MASCOT searches (Score: 2092; emPAI: 76.67) with high peptide sequence coverage (64 matches; 35 unique peptide sequences). Immunolocalisation studies indicate that *S. mansoni* LAP is localised to the gastrodermal cells surrounding the gut lumen and also to the tegument (McCarthy *et al.*, 2004). McCarthy *et al.* (2004) subsequently proposed that SmLAP may have a role in the downstream cleavage of peptides absorbed from the gut, in addition to a possible role in surface membrane remodelling (McCarthy *et al.*, 2004). ShLAP may therefore be important in energy requisition and worm pairing, making it an interesting candidate for downstream analysis.

ShTK was also identified as a top hit within the MASCOT searches (Score: 2993; emPAI: 73.47) with high peptide sequence coverage (80 matches; 54 unique peptide sequences). As a consequence, ShTK was considered a highly significant protein hit. A subsequent search of the literature found that TK is abundantly distributed in the locomotive and reproductive organs of other human-infecting platyhelminths species (Xiao *et al.*, 2013), and is an important enzyme in energy generation and homeostasis (Xiao *et al.*, 2013), further strengthening the argument for its selection for downstream analysis. Homologs have also been identified within the *S. bovis* male worm tegument (Pérez-Sánchez *et al.*, 2008) and in *S. mansoni* egg shell proteome (deWalick *et al.*, 2011).

ShGALE was originally selected as a candidate for downstream expression for its potential exposure at the parasite-host interface, via non-classical secretory pathways and possible tegumental localisation (Balkom *et al.*, 2005; Liu *et al.*, 2012). Protein expression studies in *S. japonicum* indicate that SjGALE expression levels peak on day 23, the point at which worms reach maturity and start producing eggs (Colley *et al.*, 2014). This supports a possible role for SjGALE in the production of energy required for egg production (Liu *et al.*, 2012). Moreover, vaccination studies with a *S. japonicum* homolog (SjGALE) propose an association with reduced worm fecundity (Liu *et al.*, 2012) and a study of differential gene expression in isolated *S. mansoni* gonads indicates that SmGALE is differentially expressed in ovaries of paired females compared to testes of paired male. ELISAs were performed to determine the immunoepidemiological profiles of ShLAP, ShTK and ShGALE using serum samples from the Malian cohort.

5.2.7 Coating and contamination assays

The most appropriate coating concentration and conditions for each of the candidate anti-fecundity targets was determined by ELISA, as described in Chapter 2, section 2.15.1. Following coating of each recombinant protein at a range of dilutions, microtitre plates were incubated

with HRP-conjugated anti-rat antibody, which subsequently binds to any available binding surface of the well that is not coated by the anti-fecundity target protein. Bacterial contamination was detected with rabbit serum primed against *E. coli* lysate.

As with the recombinant TAL proteins, the antigen coating concentration required for further serological analysis was therefore selected according to the concentration that minimised anti-rat-HRP signal (blue), whilst also ensuring that signal from bacterial contaminants (orange) was minimised (Fig. 5.9).

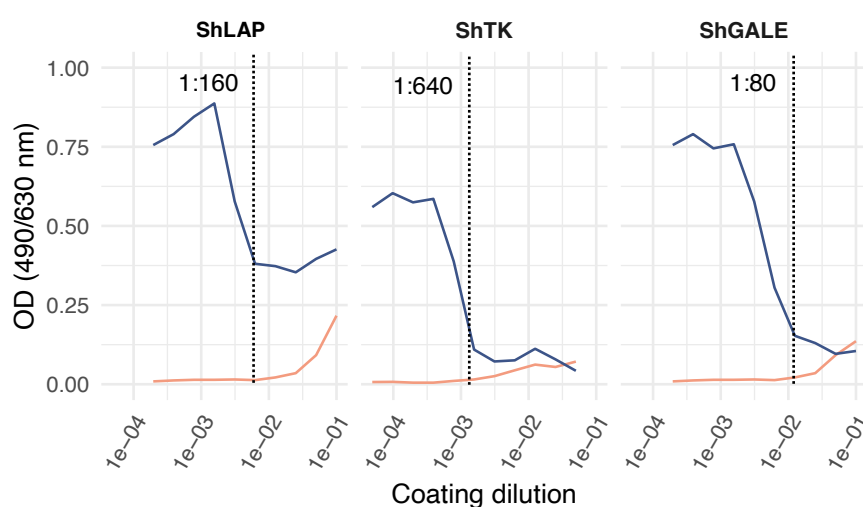


Fig. 5.9 Coating and contamination assays for recombinant *S. haematobium* putative aminopeptidase W07G4.4 (ShLAP; MS3_08450), taurocyamine kinase (ShTK; MS3_05631) and putative UDP-glucose 4-epimerase (ShGALE; MS3_04887). Plots show the binding of anti-rat-HRP to free binding sites where the respective anti-fecundity target protein has not coated the microtitre well (blue) and the antibody response to bacterial lysate (orange), to indicate whether there is any significant remaining bacterial contamination of the recombinant proteins. Dotted lines indicate the selected coating concentration for each recombinant protein.

5.2.8 Serological profile of anti-fecundity target specific IgG₁ responses

Analysis was initially performed on 261 individuals from cohort aged 5 to 40 with baseline and post-treatment serology and baseline urinary egg counts. No significant difference was seen between pre- and post-treatment specific IgG₁ antibody responses to ShLAP ($t = -0.15633$; $p = 0.876$), ShTK ($t = -0.22968$; $p = 0.818$) or ShGALE ($t = -0.12227$; $p = 0.903$), as measured by paired t-test on log transformed titres. This suggests that exposure of the protein to the hosts immune system does not significantly change following treatment, so follows that the proteins are seen by the immune system whilst the parasite is alive.

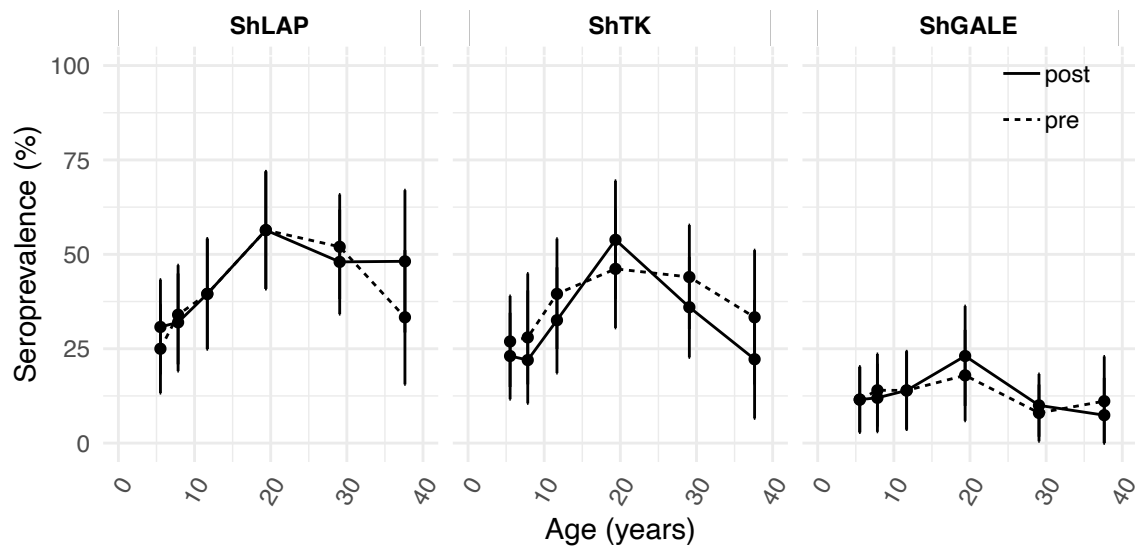


Fig. 5.10. Seroprevalence of antigen specific IgG₁ response to ShLAP, ShTK and ShGALE with age, pre- and post-treatment. Age groups: 5 to 6 years (n = 52); 7 to 9 years (n = 50); 10 to 14 years (n = 43); 15 to 24 years (n = 39); 25 to 34 years (n = 50); 35 to 40 years (n = 27). Points are plotted at the arithmetic for each age band. Error bars show the 95% confidence intervals.

The age profiles of antibody responses displayed in Figure 5.10 show that seroprevalence of ShLAP and ShTK increase with age up to approximately 30 years then, as seen with the ShTAL3, 5 and II-IgE response in Chapter 4, the seroprevalence of antibody responses declines in the oldest age group. As this is suggestive of senescence of the immune system, further regression analysis was performed on a subset of the population aged between five and 29 years of age.

Analysis of the relationship using binomial regression models found that age is a significant predictor of ShLAP seropositivity but not ShTK or ShGALE IgG₁ seropositivity. Individuals are, however, more likely to have a positive IgG₁ response to ShLAP, ShTK or ShGALE if they reside in a village where force of transmission is high compared to moderate (Table 5.9). Sex was not significantly associated with anti-fecundity target IgG₁ seropositivity and was removed from the models. Likewise, pairwise interaction terms between age and village in each of the models were not significant and were subsequently removed.

Wilson *et al.* (2014) demonstrated, in the Malian cohort, that worm fecundity declines with host age until 11 years. No further variation in fecundity is observed in those aged 11 years and over. For standardisation with the analysis presented by Wilson *et al.* (2014), regression models exploring the association between IgG₁ responses to the candidate anti-fecundity target proteins and reduced worm fecundity were analysed in a sub-cohort aged under 11 years.

Table 5.9. Association between age group, village and pre-treatment *S. haematobium* anti-fecundity target specific IgG₁ seropositivity. Results for the reduced binomial regression model, for a reduced cohort aged 5 to 29 years (n = 202).

	ShLAP		ShTK		ShGALE	
	GM ratio (95% CI)	P-value	GM ratio (95% CI)	P-value	GM ratio (95% CI)	P-value
Age (years) [†]						
7 to 9	4.12 (1.24–16.43)	*	0.97 (0.39–2.45)	ns	1.16 (0.34–4.00)	ns
10 to 14	2.43 (0.57–10.96)	ns	1.37 (0.55–3.48)	ns	0.91 (0.25–3.27)	ns
15 to 24	8.25 (2.18–36.79)	**	1.72 (0.68–4.45)	ns	1.16 (0.33–4.11)	ns
25 to 29	4.50 (1.09–20.58)	*	1.58 (0.56–4.46)	ns	0.48 (0.06–2.36)	ns
Village [‡]	12.37 (3.06–59.87)	***	4.30 (2.35–8.03)	***	4.61 (1.95–11.87)	***
Interaction terms						
7 to 9*village	0.09 (0.01–0.57)	*				
10 to 14*village	0.37 (0.05–2.62)	ns				
15 to 24*village	0.13 (0.02–0.89)	*				
25 to 29*village	0.18 (0.02–1.49)	ns				

*** p < 0.0001; * p < 0.05; ns non-significant

[†]Compared to baseline age group 5 to 6 years

[‡]Compared to moderate transmission village

5.2.9 Associations between anti-fecundity target IgG₁ seropositivity and reduced worm fecundity

Univariable linear models were first constructed to test for an association between pre-treatment seropositivity for anti-fecundity target specific IgG₁ and fecundity scores in those aged under 11 years. IgG₁ seropositivity to ShLAP was significantly associated with lower *S. haematobium* fecundity scores (Table 5.10). IgG₁ seropositivity to ShTK and ShGALE were also associated with reduced worm fecundity, although this was not significant. Seroprevalence is low in children under 11 years old (Fig. 5.9), consequently the sample size may not be large enough to detect a significant association for ShTK or ShGALE IgG₁ seropositivity.

Table 5.10 Association between seropositivity for AF target specific IgG₁ and fecundity in individuals aged under 11 years (n = 102)

Antigen	GM ratio (95% CI)	Adjusted R ²	SE
ShLAP	0.18 (0.06–0.53)**	0.082	2.60
ShTK	0.57 (0.17–1.85)	-0.001	2.72
ShGALE	0.68 (0.15–3.08)	-0.007	2.73

** p < 0.01

Co-infection with *S. mansoni* may influence CAA measurement and hence fecundity scores. Models were adjusted for presence of *S. mansoni* infection, as a binary variable, measured by Kato Katz stool microscopy. As discussed in previous analysis of this population, for cultural and religious reasons provision of stool samples by the study population was low (Wilson *et al.*, 2014). In the reduced population of under 11-year olds (n = 102), 27 individuals did not provide a stool sample. The presence or absence of *S. mansoni* infection in these individuals was therefore imputed by multiple imputation using chained equations (MICE) by means of the pattern-mixture model approach in R statistical software, as described in Chapter 2, section 2.16.11. Imputation results are provided in Appendix 5 (Fig. A5.3).

Models were subsequently adjusted to control for the effects of age, sex, village (as a proxy from transmission intensity), the interaction term between age and village, and *S. mansoni* infection intensity using the imputed data. The negative association between antigen specific IgG₁ seropositivity and fecundity lost significance for each antigen when possible confounding variables were accounted for (Table 5.11). Host age and residence in a high transmission village were both significantly associated with reduced worm fecundity. Host sex and the interaction

term between host age and village were not significant in any of the models and were subsequently removed.

Table 5.11. Association between pre-treatment *S. haematobium* anti-fecundity target specific IgG₁ responses and fecundity. Results for the reduced linear regression model, adjusted for *S. mansoni* infection status, host age, village (as a proxy for transmission intensity), for a reduced cohort aged under 11 years (n = 102). Results are shown for the pooled regression models from 25 imputations.

	ShAP		ShTK		ShGALE	
	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value
AF target-IgG ₁	-0.888 (0.599)	ns	0.322 (0.596)	ns	0.206 (0.752)	ns
Age (years)	-0.644 (0.175)	***	-0.683 (0.175)	***	-0.687 (0.175)	***
Village [†]	-1.566 (0.633)	*	-1.905 (0.641)	***	-1.858 (0.639)	**
Sm status	-0.680 (1.205)	ns	-0.520 (1.217)	ns	-0.465 (1.239)	ns

*** p < 0.0001; ** p < 0.01; * p < 0.05; ns non-significant

[†]Compared to moderate transmission village

5.3 Discussion

The production of eggs by mature female schistosomes is essential for maintenance of the parasite lifecycle. It is the eggs laid by female worms that also contribute to pathogenesis that results in chronic schistosomiasis and the subsequent loss of quality adjusted life years (QALYS) amongst more than 240 million individuals affected by schistosomiasis worldwide (King, Dickman and Tisch, 2005).

The candidate anti-fecundity target proteins identified in this chapter represent several biological and molecular functions, including representatives linked to reproduction and energy production. A number of these proteins are also found within the excretory secretory fraction, or may be transiently exposed to the host, for example by their localisation to the worm tegument. Several of these proteins therefore represent promising targets for a protective immune response affecting worm fecundity.

5.3.1 Excreted/secreted and surface expressed proteins

ES and surface expressed proteins are closely associated with the parasite-host interface and therefore represent an opportunity for interaction with host immune cells. 20% of the anti-fecundity target candidates were found to contain secretory or transmembrane domain signatures. More than half of these proteins were predicted to be non-classically secreted. Non-

classical protein secretion via EVs also occurs in schistosomes (Marcilla *et al.*, 2014; Nowacki *et al.*, 2015; Wu *et al.*, 2018) and proteins released into the extracellular environment this way would not necessarily be identified using the traditional motif algorithms. Indeed, it has been suggested that the transport of proteins via EVs may significantly contribute to the secretory proteome of schistosomes (Marcilla *et al.*, 2014). 10% of identified candidate anti-fecundity targets were common to analysis of the EV proteome (Nowacki *et al.*, 2015). EVs have also been found to contain several potential vaccine candidates (Sotillo *et al.*, 2016).

5.3.2 Identification of tegumental and cytoskeletal proteins corresponds with the delayed development of anti-fecundity immunity

For membrane associated proteins to be the target of the anti-fecundity response it would be expected that a boost in the specific antibody response would be seen following treatment-induced worm death, as previously proposed by Mitchell *et al.* (2012). The notable IgG₁ response to cytoskeletal or tegumental proteins may result from anti-fecundity cases having been exposed to more dying worm and having a generally higher IgG₁ response to SWA compared to control individuals. Whilst the cases and controls were matched for age, sex and village, as a proxy for transmission, it was not possible to control for IgG₁ between the two groups due to insufficient individuals within the original cohort to allow this further stratification in the matching process. Additionally, the dynamic nature of the worm tegument and the effectiveness of the worm's ability to disguise itself mean that the protein target may only be transiently exposed, which may also explain the delayed development of anti-fecundity immunity with host age.

5.3.3 GO terms associated with reproduction were significantly enriched within the candidate anti-fecundity target proteins

The analysis of GO terms that were over-represented in genes corresponding to the 191 anti-fecundity target proteins compared with the GO term annotation of the entire *S. haematobium* genome identified several biological processes associated with reproduction. At the time of the analysis the *S. haematobium* proteome was relatively poorly annotated, yet the large majority of anti-fecundity targets identified in this analysis had GO terms assigned. The high representation of reproduction-related GO terms, associated with proteins that are immunogenic to individuals with reduced worm fecundity, is therefore a noteworthy finding.

5.3.4 Enzymes involved in energy production may be excreted during feeding or secreted in EVs

Several enzymes involved in energy production and found within the digestive regurgitant of schistosomes were identified within the candidate anti-fecundity target proteins in this chapter (Hall *et al.*, 2011; Figueiredo *et al.*, 2015). The high energy requirements of egg production, necessitate red blood cell ingestion that is 10 times higher in females compared to males (Lawrence, 1973). In their proteomic analysis of excreted gut proteins, Hall *et al.* (2011) argued that lysosomal excreted peptides, such as cathepsins are of primary importance in digestion. Two lysosomal proteins, lysosomal protein (MS3_04617) and DPPII (MS3_05939), featured in the candidate anti-fecundity targets presented in this chapter. Gut proteins are likely to induce a significant IgG₁ antibody response as they are expelled into the host circulation during schistosome feeding. Furthermore, enzymes are the most abundant protein type found in proteomic analysis of EVs and proteases, in particular, have been noted to be abundant within the EV fraction (Samoil *et al.*, 2018).

Several peptidases were also identified within the 191 proteins differentially immunogenic to case serum, including the putative metallopeptidase ShLAP (MS3_08450). It is likely that more than one protein is the target of an IgG₁-driven immune response and protective responses may progressively develop against related proteins. Peptidases are enzymes that break down peptides into their constituent amino acids. Moreover, peptidases may function within a pathway of peptide degradation requiring the activity of a series of peptidases, rather than being directly associated with digestion of proteins. For example, the *S. mansoni* LAP homologue, SmLAP (Smp_030000), is believed to function intracellularly and is localised to the gastrodermal cells surrounding the gut lumen (McCarthy *et al.*, 2004). SmLAP may therefore have a role in the downstream cleavage of peptides absorbed from the gut lumen (McCarthy *et al.*, 2004).

All three of the proteins successfully expressed and analysed in this chapter are associated with energy generation (McCarthy *et al.*, 2004; Liu *et al.*, 2012; Xiao *et al.*, 2013), although not all the peptidases identified in this study are thought to be secreted by the schistosome. Prolyl endopeptidase (MS3_01899) was identified as a possible anti-fecundity target from the analysis of the MS data presented in this chapter. In characterising an *S. mansoni* homolog, Fajtová *et al.* (2015) demonstrate that *S. mansoni* prolyl oligopeptidase (SmPOP; Smp_213240) is present within the parenchyma and tegument of adult schistosomes, with male worms demonstrating higher signal. Fajtová *et al.* (2015) therefore suggest that SmPOP may exhibit peptidolytic functions at the host–parasite or male–female interface that are specific to the male worm.

Furthermore, SmPOP demonstrates a similar distribution in the parenchyma and tegument as the cysteine peptidase cathepsin B1, which has now been shown to exhibit an association with reductions in worm fecundity in vaccination studies (El Ridi *et al.*, 2014).

It is possible that proteins found in EVs or gut regurgitant are the stimulus for an IgG₁ response in the live worm whilst the target of IgG₁ activity is the protein localised to the tegument. In the female-male pairing relationship, the role of the male is not, limited to nutrition, with bidirectional transfer for molecules between females and males also having been proposed (Lu *et al.*, 2019). Bidirectional signalling may facilitate the males to stimulate reproductive maturity in female worms (Lu *et al.*, 2019). Several of the identified candidate anti-fecundity targets are associated with energy production, whilst also being localised to the worm tegument. There is therefore further scope to explore the interactions between this group of proteins in future analysis.

5.3.5 What is the significance of the three successfully expressed anti-fecundity targets?

5.3.5.1 Putative aminopeptidase W07G4.4 (ShLAP)

Aminopeptidases have a role in a diverse range of biological processes, including feeding, embryogenesis and evasion of the host immune system. Several metallopeptidases were among the full list of candidate anti-fecundity target proteins. The putative aminopeptidase W07G4.4 (ShLAP; MS3_08450) identified in this chapter has significant homology to *S. mansoni* leucine aminopeptidase (SmLAP) and is a member of the M17 metallopeptidase family. Metallopeptidases including LAP have been extensively studied in several *Schistosoma* species and are commonly recognised as a potential vaccine candidate against a number of parasitic helminth infections (Acosta *et al.*, 2008; Pokharel and Rathaur, 2008; Rinaldi *et al.*, 2009; Raina *et al.*, 2011). Vaccination studies in other platyhelminth species have demonstrated mixed results, with reduced worm fecundity observed following vaccination against *F. hepatica* (Acosta *et al.*, 2008), but not *F. gigantica* (Raina *et al.*, 2011). Furthermore, an RNA interference study targeting *S. mansoni* LAP demonstrated significantly inhibited hatching of miracidia from *S. mansoni* eggs, rather than a reduction in fecundity *per se*; nonetheless, a potentially interesting finding in relation to breaking transmission at an alternative point in the lifecycle.

The absence of a post-treatment boost following worm death suggests that the host immune system is exposed to LAP in the live worm. Immunolocalization studies indicate that SmLAP and SjLAP are located on the gastrodermal surface of adult worms and may play a role in cleavage of amino acids during haemoglobin digestion (McCarthy *et al.*, 2004). SmLAP

immunostaining was found to be stronger within the gut of female worms, compared to males (McCarthy *et al.*, 2004).

The analysis presented here demonstrates an increase in IgG₁ response with both age and high compared to moderate force of transmission. Seropositivity of IgG₁ to ShLAP is therefore likely to be linked to cumulative antigen exposure.

5.3.5.2 Taurocyamine kinase (ShTK)

TK is a member of the guanidino kinase family. These enzymes catabolise the reversible Mg²⁺-dependent reaction that maintains the ATP/ADP balance in cells, often in those with high energy processes subject to fluctuation. Accordingly, guanidino kinases have been identified in the motile tail of *Schistosoma* cercariae (Skelly, Stein and Shoemaker, 1993). Differential gene expression experiments in *S. mansoni* also found that SmTK is differentially expressed in the ovaries, compared to the testes, of paired worms (Lu *et al.*, 2016).

As with ShLAP, there was no significant difference between pre- and post-treatment IgG₁ antibody titres in the results presented in this chapter, indicating that treatment does not induce a significant increase in exposure. Furthermore, seroprevalence of IgG₁ to ShTK was not significantly associated with host age. This suggests that a positive ShTK-IgG₁ response is unlikely to develop as a result of worm death.

ShTK is, however, found within the *S. mansoni* secretome (Floudas *et al.*, 2017) and more specifically within the EV fraction, featuring as one of the 35 most abundant EV proteins in Nowacki *et al.* (2015). An *S. bovis* TK homologue has also been identified in the tegument of male *S. bovis* worms (Pérez-Sánchez *et al.*, 2008). Again, this may provide a route through which this protein is exposed to the host immune system and may help explain the increase in specific IgG₁ seroprevalence in high transmission villages. The observed decline in the proportion of IgG₁ responders in the oldest age groups, may be associated with immune senescence, as suggested previously, in Chapter 4.

5.3.5.3 Putative UDP-glucose-4-epimerase (ShGALE)

The observed seroprevalence of ShGALE-IgG₁ responses is comparatively lower than that of ShLAP- and ShTK-IgG₁ responses. ShGALE seroprevalence peaks at approximately 25% seropositivity in early adulthood and, in contrast to the responses seen in ShLAP and ShTK, ShGALE does not demonstrate a significant increase in seroprevalence with age. In this chapter, ShGALE was not found to increase following treatment, again suggesting that exposure is not

significantly boosted following treatment, despite the apparent tegumental location of *S. mansoni* (Sm) and *S. japonicum* (Sj) homologs (Balkom *et al.*, 2005; Liu *et al.*, 2012). It is therefore possible that ShGALE is transiently exposed at a low level during the lifetime of the adult worm. Furthermore, ShGALE was identified as one of the proteins with a non-classically secreted domain signature, providing further opportunity for host exposure.

Interestingly, SjGALE murine vaccination experiments demonstrated a partially protective specific IgG₁ response, resulting in a reduction in worm burden and a substantial reduction in liver egg burden (Liu *et al.*, 2012). So, whilst the association between ShGALE-IgG₁ and fecundity is not significant in the analysis presented here, this may partially be the result of the magnitude of the immune challenge and relatively low sample size of responders in the Malian study. A more significant immunological challenge in a larger population, such as that provided by vaccination may provide stronger support for the role of ShGALE in the anti-fecundity protective immune response.

5.3.6 Regression models do not exclude ShLAP, ShTK and ShGALE as potential anti-fecundity target proteins

The univariate models indicated a significant association between ShLAP and reduced worm fecundity. However, this association is confounded by age and transmission; confirmed by multivariable linear regression models, which show strong significant associations with age, village of residence and the interaction term between age and village. In these models, village is included as a proxy for transmission intensity, since the villages could be clearly separated into those with moderate or high infection intensities.

Force of transmission is closely linked to antigen exposure, since higher infection intensities result in individuals being exposed to greater number and diversity of antigens. This has previously been proposed in relation to cumulative exposure to antigens released as adult worms die but this may also be the case for parasite antigens that are presented to the host immune system whilst the worm is still alive, either via ES products or surface expression.

Protection against reinfection is thought to involve the progressive development of immunity following exposure to several members of a protein family (Fitzsimmons *et al.*, 2012). In the same way, the anti-fecundity immune response is unlikely to depend on a single protein target of IgG₁.

Furthermore, extracellular secreted proteins often have post-translational modifications that are important in determining the tertiary protein structure required for specific antibody epitope binding. These post-translational modifications are not always reproducible in recombinant

protein expression, an issue highlighted by Crosnier *et al.* (2019). It is therefore possible that any association between the anti-fecundity target IgG₁ response and fecundity is weakened by suboptimal epitope structure, thus antibody binding potential, of the recombinant proteins produced in this chapter.

5.3.7 A need for improved annotation of the *S. haematobium* genome

The fragmented state of the *S. haematobium* genome has limitations on functional annotation and cloning. Only 14 of the selected 48 proteins were successfully cloned from *S. haematobium* cDNA, despite a range of PCR conditions and enzymes being tested. This could be due to errors in the primer construct sequences due to poorly characterised exon spanning sequences. Several of the analyses presented in this chapter draw upon previously published gene annotation data. This was performed in order to improve the functional picture of the anti-fecundity target proteins. It should be noted that these results should be reviewed with caution since gene expression does not necessarily equate to translation and protein expression in *S. haematobium*.

Incomplete *S. haematobium* genome sequencing has limited the proteomic methods used in this chapter. Whereas other authors have use an immunomics approach to vaccine candidate discovery, probing a protein microarray with infected sera (Driguez *et al.*, 2010), the *S. haematobium* proteome is not yet characterisation to the quality from which a microarray can be produced. Recently, a ‘next generation’ *S. mansoni* protein microarray consisting of 992 recombinant proteins was produced (de Assis *et al.*, 2016), following earlier proof of principle (Driguez *et al.*, 2015). Few protein microarrays have been developed for eukaryotic multicellular organisms such as schistosomes. The microarray produced by de Assis *et al.* (2016) therefore represents the largest array for a helminth parasite. In spite of this, one of the greatest challenges that construction of protein microarrays for eukaryotic organisms presents is the sheer number of proteins that are encoded for in the genome. The proteins included in the microarray (de Assis *et al.*, 2016) were therefore selected according to the presence of the protein in various layers of the tegument (Braschi, Borges and Wilson, 2006; Castro-Borges, Simpson, *et al.*, 2011; Sotillo *et al.*, 2016), the presence of signal peptides or transmembrane domains (Gaze *et al.*, 2014; Pearson *et al.*, 2015), or seropositivity for proteins from *S. mansoni* soluble worm preparation when screened with sera from individuals living in endemic area, using 2D gel electrophoresis (Ludolf *et al.*, 2014), similar to the methods presented in this chapter.

Since the analysis performed in this thesis, a significantly improved version of the genome has been published, resulting in fewer scaffolds and significantly reduced fragmentation and gaps (Shae. VI: N50 = 0.31 Mb, L50 = 365 scaffolds; Shae. V2 N50 = 4.78 Mb, L50 = 26 scaffolds

(Stroehlein *et al.*, 2019)). Unfortunately, this work was published after completion of the analysis presented here but can be used for future work. Though, this much improved version of the *S. haematobium* genome is still not of the same quality as the published *S. mansoni* genome construct (N50 = 32.16 Mb, L50 = 4 scaffolds), which is significantly less fragmented, enabling better annotation.

Furthermore, Sotillo *et al.* (2019) recently published an in-depth characterisation of the *S. haematobium* proteome. With improved coverage and annotation of encoded proteins, it is therefore not unfeasible that a protein microarray could be constructed for *S. haematobium* in the near future. The results presented in this chapter could inform the selection of proteins chosen for a future protein microarray for *S. haematobium*.

5.3.8 Potential future experiments

Neves *et al.* (2015) found that 20% of the mass of soluble worm antigen preparation is contributed to by only 5 proteins: enolase, aldolase, actin, glyceraldehyde-3-phosphate dehydrogenase (GAPD), fatty acid binding protein, whilst 50% of the mass is representative of 18 proteins. In this chapter, 12 of these 18 proteins have been identified as targets of IgG₁ response in anti-fecundity cases. These proteins were therefore ruled out these as possible candidates, as it is likely that these proteins are differentially immunogenic to anti-fecundity case sera due to a generally higher IgG₁ response in anti-fecundity cases compared to controls. It is possible that the higher IgG₁ response results from greater cumulative exposure to worm death.

Several authors report that different methods of SWA preparation and the specific detergents used can result in different protein profiles (Castro-Borges, Dowle, *et al.*, 2011). Furthermore, carbohydrate residues, or glycans, are well-known stimulants of IgG₁ antibody response.

A series of further experiments could be performed to identify potential additional immunostimulatory proteins by using alternative detergent preparations or by stripping the glycans from proteins prior to analysis. This would distinguish between responses that may result from the limits to the SWA preparation methods versus true anti-fecundity IgG₁-mediated responses. Exploring serum responses in individuals with reduced fecundity to isolated gut proteins, ES products or EV secreted proteins, in particular, may prove particularly informative in light of the results presented in this chapter. This would, however, be dependent on the availability of *S. haematobium* parasites, which, due to the difficulty in maintaining the life cycle compared to *S. mansoni*, few laboratories produce. There is also scope for further optimisation of cloning and more extensive characterisation of the targets already selected, especially in light of publication of the improved genome.

Finally, alternative methods of recombinant protein expression, such as the mammalian expression system presented by Crosnier *et al.* (2019), may improve the specific antibody binding capacity of endemic serum. Together with the recently improved *S. haematobium* genome assembly (Stroehlein *et al.*, 2019) and characterisation of the proteome (Sotillo *et al.*, 2019) this methodology has the potential for enhanced detection of candidate anti-fecundity targets.

5.3.9 Regarding mathematical transmission models

The leading *S. haematobium* transmission model is the stratified worm burden (SWB) model produced by Case Western Reserve University (CWRU) (Gurarie, King and Wang, 2010). Gurarie *et al.* (2015) adapt this model to predict the long-term impact of MDA in Africa. In line with the findings of Wilson *et al.* (2014) and the results presented in this chapter, the SWB model finds that maximal worm fecundity decreases with host age (Gurarie *et al.*, 2015). However, the authors also found that values for the biological parameters that describe worm fecundity are consistent for each age group across villages subject to different risk of infection (Gurarie *et al.*, 2015). To the contrary, the findings of Wilson *et al.* (2014) suggest that the age at which the decline in fecundity occurs is likely to be dependent on the force of transmission. Transmission intensity is also found to be significantly associated with the IgG₁-mediated anti-fecundity immune response in the analysis presented here.

5.3.10 Concluding remarks

Developing our understanding of the mechanisms behind the immune processes that reduce worm fecundity in *S. haematobium* is of great importance in providing future options for disease control. A more in depth understanding of anti-fecundity immunity is vital to improving transmission and morbidity modelling studies. The immune processes that govern *S. haematobium* infection intensity and transmission are still relatively poorly understood; however, this chapter has identified several potential targets of an IgG₁-driven anti-fecundity immune response, providing the foundations for further analysis of these promising candidates.

6

Discussion

The development of antibody-mediated protective immune responses is closely associated with variations in exposure behaviour, host age and the force of transmission; the combined contribution of which determine the level of exposure to immunostimulatory antigens. These differences have been demonstrated across the three experimental chapters that comprise this thesis.

Incorporating partially protective SmTAL1-IgE immunity into a *S. mansoni* transmission model

In Chapter 3, the vastly different exposure profile of males from the Bagungu tribe provided an opportunity to explore the influence of different exposure behaviour on infection intensity and development of serological responses.

The explicit parameterisation of the partially protective SmTAL1-IgE immune responses was successfully incorporated into the *S. mansoni* transmission model presented in Chapter 3. This model was largely able to reproduce the age-intensity and age-serology profiles of the observed field data, particularly for the Alur tribe. Females and males from this tribe both demonstrate the ‘typical’ exposure profile that is often assumed in the modelling literature, where exposure is high in children and subsequently declines into adulthood. It is this ‘typical’ exposure profile that has led many authors to assume that the infection intensity profile is determined by differences in exposure behaviour alone, with the contribution of acquired immunity assumed to be negligible (Anderson *et al.*, 2015). It is only by examining the combined influence of heterogeneity in exposure alongside the development of protective immunity that the intricacies of the relationship between exposure and immunity are highlighted.

In the Bagungu male population, the level of exposure to infectious cercariae is substantially higher than any other demographic group examined in the analysis. Furthermore, the Bagungu male exposure profile is not mirrored in the age-intensity or age-serology profiles for this sub-cohort, suggesting that there may be additional factors that influence the relationship between exposure and infection intensity. The complex relationship between exposure, infection intensity and the development of immunity in the Bagungu male sub-cohort highlights the

difficulties that can be experienced when working with epidemiological data collected from observational studies. Humans are naturally diverse, leading to heterogeneity and stochasticity in observed epidemiological patterns. Extensive multidisciplinary analysis has previously been performed on this population, including social science and anthropology, malacology, parasitology and immunology to elucidate the complex patterns of exposure and immunity, the details of which can be found elsewhere (Pinot De Moira *et al.*, 2010).

A study that examined the prevalence of periportal fibrosis, a long-term sequela of chronic schistosomiasis, in Booma and a neighbouring village (Bugoigo), found that the prevalence of periportal fibrosis in adults was significantly higher in Booma, despite similar parasitological measurements between the two villages (Booth *et al.*, 2004). Moreover, the duration of residency in Booma was a critical risk factor for the development of periportal fibrosis, with end-stage periportal fibrosis associated with long-term exposure to schistosome infection (Booth *et al.*, 2004). This indicates that Booma had a degree of stability amongst its population that was not observed in the neighbouring villages. Additionally, a study of two populations with different occupational exposure profiles by Black *et al.* (2010), conducted in the Kisumu Region of Kenya, on Lake Victoria, found that resistance to reinfection is dependent on immunological status and history of exposure. The protective immune response was found to be augmented by treatment in this study, with individuals with a longer history of exposure acquiring protective immune repossess at a faster rate (Black *et al.*, 2010). Bagungu males are therefore likely to have experienced substantial cumulative exposure to parasite antigen over a number of years and, as a consequence are likely to have developed an effective protective immune response against reinfection. This may explain the similar observed infection intensity profile compared to Bagungu females, despite the differences in exposure between the two demographic groups.

There does not, however, appear to be significant difference in the parameterisation of IgE-mediated immunity function, presented in Chapter 3, across the four demographic groups examined, as would be expected if the SmTALI-IgE antibody response is assumed to be responsible for curtailing the high infection intensities in Bagungu males. SmTALI-IgE is a marker for exposure to dead worms, the first step in the pathway towards the development of partially protective immunity. This serological marker does not, therefore, fully explain the complete IgE-mediated immune response. Again, an expected conclusion, since it is highly unlikely that a single antigen-antibody interaction would provide the 'silver bullet' to protection against reinfection. Instead, evidence suggests that anti-infection immunity is likely to involve additional members of the TAL protein family as well as possible IgG₄ blocking antibody responses. A relationship that is explored further in Chapter 4.

Chapter 3 emphasises the importance of modelling host immune responses considering the potential confounding that occurs between behavioural variation in exposure, age and immunity.

The anti-infection protective response to *S. haematobium* demonstrates a same dependence on TAL-specific IgE as it does to *S. mansoni*

Progressive development of IgE-mediated protective immunity was further explored in Chapter 4. It has been proposed that, in addition to SmTAL1, other members of the TAL protein family may be associated with the development of IgE-mediated immunity to reinfection. This is thought to involve the development of a cross-reactive IgE response between SmTAL3 and SmTAL5, two closely related members of the TAL protein family; yet, prior to the analysis presented in this thesis, this association had not been substantiated in other study locations or in other *Schistosoma* species.

Examination of the phylogenetic relationships between the TAL proteins, identified the potential role for an additional member of the TAL protein family, SmTAL11, in the development of acquired anti-infection immunity due to its close phylogenetic relationship with SmTAL5. The regression analysis presented in Chapter 4 suggests that the anti-infection protective response to *S. haematobium* demonstrates a similar dependence on TAL-specific IgE as it does to *S. mansoni*, supporting a role for IgE seropositivity to ShTAL1, 3 and 5 in protection against detectable egg count two years after treatment; yet the contribution of ShTAL11-IgE to anti-infection immunity was not found to be significant.

Again, this chapter finds an important association between the development of antigen-specific antibody responses and exposure. Significant differences were seen between the observed antibody response in the villages subject to high infection intensities compared to the village with moderate infection intensity, indicating that exposure, linked to the force of transmission, may have a significant influence of the development of specific antibody responses. This is consistent with the peak shift hypothesis; where higher transmission intensity results in exposure to a greater antigenic diversity (Fulford *et al.*, 1992).

Older age is also a significant predictor of protection against reinfection, thus providing further evidence in support of an association between antigen exposure and the development of protective antibody responses. This cumulative exposure to parasite antigens has historically been linked to the release of previously cryptic proteins as adult worms die (Woolhouse and Hagan, 1999). Worm death has also been proposed as the mechanism through which IgG₁ associated anti-fecundity immunity develops during *S. haematobium* infection (Mitchell *et al.*,

2012), but in order for the immune response to act on worm fecundity, it follows that the target protein must be accessible whilst the worm is still viable. Proteomic analysis of potential targets of an anti-fecundity immune response were explored in Chapter 5.

Identifying the antigenic targets of IgG1-mediated anti-fecundity immunity in *S. haematobium*

A number of previous proteomic studies have investigated antibody specific serological responses to *S. haematobium* SWA preparation, separated by two-dimensional Western blot (Mutapi *et al.*, 2008, 2011). Nevertheless, the analysis presented in Chapter 5 is unique in that it explores the differential recognition of SWA proteins by individuals in whom reductions in worm fecundity have explicitly been observed, via paired circulating anodic antigen titres, as a measure of live worm burden, and urinary egg counts.

The anti-fecundity immune response has previously been associated with IgG₁ antibodies specific to schistosome adult worm antigens (Wilson *et al.*, 2014); yet knowledge of the schistosome proteins targeted by IgG₁-driven protection, and the mechanism by which naturally acquired anti-fecundity immunity occurs remains unclear. Traditional proteomic methods were used in Chapter 5 to identify SWA proteins that invoke an IgG₁ response in individuals with reduced fecundity scores compared to matched controls.

The proteomic analysis identified several candidate proteins that are closely associated with the parasite-host interface. These included excreted/secreted proteins, tegumental proteins and several gut proteins that have previously been identified in digestive regurgitant, suggesting that transient exposure of proteins at the parasite-host interface during the lifetime of the worm may be the stimulus for development of the delayed anti-fecundity immune response. Furthermore, the analysis of specific IgG₁ to anti-fecundity target proteins pre- and post-treatment did not find evidence to support an association with the development of a protective immune response against worm fecundity following worm death, as proposed by previous mathematical modelling studies (Mitchell *et al.*, 2012).

Serological analysis of IgG₁ isotype responses to recombinant anti-fecundity target proteins again found that specific antibody responses were strongly associated with age and village of residence, as a proxy for the force of transmission. This provides evidence to support the link between cumulative antigen exposure and the development of protective antibodies in an anti-fecundity immune response, similar to the development of anti-infection immunity.

The development of specific antibody responses is dependent on the degree of antigen exposure

The association between the development of protective antibody responses and exposure is of particular importance to the planning and implementation of control programmes, whether this be exposure to cercariae in the environment, cumulative antigen exposure with increasing host age, or exposure with varying force of transmission.

MDA has been shown to be effective in reducing the prevalence of high-intensity infections in several regions in sub-Saharan Africa (Deol *et al.*, 2019; Kittur *et al.*, 2019) and may have the potential to reduce the force of transmission at the community level (Anderson and May, 1985; French *et al.*, 2010). The analysis presented in this thesis suggests that the development of protective antibody responses is determined by the profile of exposure within a certain population. The reduction in transmission that may result from long-term intervention with MDA may therefore reduce the rate at which protective immunity is acquired. This may have important implications for the rebound of schistosome infection should control interventions be reduced as the WHO targets for schistosomiasis control are met.

Conversely, evidence from field studies suggests that MDA programmes may augment the development of immunological responses (Black *et al.*, 2010). This is likely to result from a boost in antigen exposure following worm death (Woolhouse and Hagan, 1999). Whilst evidence presented in Chapter 4 supports a post-treatment boost in antibody responses against reinfection, Chapter 5 indicates that this is not the case for anti-fecundity immunity, instead specific antibodies associated with reduced fecundity may be stimulated by, and active against, the live adult worm.

This thesis has highlighted the importance of understanding the combined contribution of antigenic exposure and the acquisition of immunity in explaining schistosome infection profiles. Further development of the mathematical model presented in Chapter 3, to improve the immunity function and include MDA intervention strategies will help elucidate ways in which control interventions, primarily MDA, may influence antigenic exposure. The work presented in this thesis forms part of a systems-based approach in understanding both the heterogenous nature of schistosome transmission and ultimately how best to control this devastating parasite.

References

- Aalberse, R. C. *et al.* (2009) 'Immunoglobulin G4: An odd antibody', *Clinical and Experimental Allergy*, 39(4), pp. 469–477. doi: 10.1111/j.1365-2222.2009.03207.x.
- Aalberse, R. C. and Platts-Mills, T. A. E. (2004) 'How do we avoid developing allergy: modifications of the TH2 response from a B-cell perspective.', *The Journal of allergy and clinical immunology*. United States, 113(5), pp. 983–986. doi: 10.1016/j.jaci.2004.02.046.
- Abath, F. G. and Werkhauser, R. C. (1996) 'The tegument of *Schistosoma mansoni*: functional and immunological features.', *Parasite immunology*. England, 18(1), pp. 15–20. doi: 10.1046/j.1365-3024.1996.d01-6.x.
- Acosta, D. *et al.* (2008) 'Fasciola hepatica leucine aminopeptidase, a promising candidate for vaccination against ruminant fasciolosis.', *Molecular and biochemical parasitology*. Netherlands, 158(1), pp. 52–64. doi: 10.1016/j.molbiopara.2007.11.011.
- Agnew, A. *et al.* (1995) 'The relationship between worm burden and levels of a circulating antigen (CAA) of five species of *Schistosoma* in mice.', *Parasitology*, 111, pp. 67–76. doi: 10.1017/S0031182000064611.
- Agnew, A. *et al.* (1996) 'Age-dependent reduction of schistosome fecundity in *Schistosoma haematobium* but not *Schistosoma mansoni* infections in humans.', *The American journal of tropical medicine and hygiene*, 55(3), pp. 338–43.
- Agnew, Alison *et al.* (1996) 'Age-dependent reduction of schistosome fecundity in *Schistosoma haematobium* but not *Schistosoma mansoni* infections in humans', *American Journal of Tropical Medicine and Hygiene*, 55(3), pp. 338–343.
- Akbar, A. N. and Henson, S. M. (2011) 'Are senescence and exhaustion intertwined or unrelated processes that compromise immunity?', *Nature reviews. Immunology*. England, 11(4), pp. 289–295. doi: 10.1038/nri2959.
- Alexa, A. and Rahnenführer, J. (2019) 'Gene set enrichment analysis with topGO'. Bioconductor. Available at:
<https://bioconductor.org/packages/release/bioc/vignettes/topGO/inst/doc/topGO.pdf>.

- Alexander, N. (2012) 'Review: Analysis of parasite and other skewed counts', *Tropical Medicine and International Health*, 17(6), pp. 684–693. doi: 10.1111/j.1365-3156.2012.02987.x.
- Allen, E. J. and Victory, H. D. (2003) 'Modelling and simulation of a schistosomiasis infection with biological control', *Acta Tropica*, 87(2), pp. 251–267. doi: 10.1016/S0001-706X(03)00065-2.
- Anderson, R. M. *et al.* (2015) 'What is required in terms of mass drug administration to interrupt the transmission of schistosome parasites in regions of endemic infection?', *Parasites & Vectors*. *Parasites & Vectors*, pp. 1–11. doi: 10.1186/s13071-015-1157-y.
- Anderson, R. M. *et al.* (2016) *Studies of the Transmission Dynamics, Mathematical Model Development and the Control of Schistosome Parasites by Mass Drug Administration in Human Communities*, *Advances in Parasitology*. Elsevier Ltd. doi: 10.1016/bs.apar.2016.06.003.
- Anderson, Roy M and May, R. M. (1985) 'Helminth infections of humans: mathematical models, population dynamics, and control', *Advances in parasitology*, 24(November), pp. 1–101. doi: [http://dx.doi.org/10.1016/S0065-308X\(08\)60561-8](http://dx.doi.org/10.1016/S0065-308X(08)60561-8).
- Anderson, R M and May, R. M. (1985) 'Herd immunity to helminth infection and implications for parasite control.', *Nature*, 315(June), pp. 493–6. doi: 10.1038/315493a0.
- Anderson, R. M. and May, R. M. (1991) 'Biology of host-macroparasite associations', in Anderson, R.M.; May, R. M. (ed.) *Infectious Disease of Humans, Dynamics and Control*. Oxford: Oxford University Press, pp. 433–466.
- Armstrong, J. C. (1965) 'Mating behavior and development of schistosomes in the mouse.', *The Journal of parasitology*. United States, 51, pp. 605–616.
- Ashburner, M. *et al.* (2000) 'Gene Ontology : tool for the unification of biology', *Nature Genetics*, 25(may), pp. 25–29.
- de Assis, R. R. *et al.* (2016) 'A next-generation proteome array for *Schistosoma mansoni*.', *International journal for parasitology*. England, 46(7), pp. 411–415. doi: 10.1016/j.ijpara.2016.04.001.
- Atassi, M. Z. and Smith, J. A. (1978) 'A proposal for the nomenclature of antigenic sites in peptides and proteins.', *Immunochemistry*. England, 15(8), pp. 609–610. doi: 10.1016/0161-5890(78)90016-0.
- Balkom, B. W. M. Van *et al.* (2005) 'Mass Spectrometric Analysis of the *Schistosoma mansoni* Tegumental Sub-proteome research articles', *Journal of Proteome Research*, pp. 958–966.
- Barbour, A. D. (1978) 'Macdonald's model and the transmission of bilharzia', *Transactions of the*

Royal Society of Tropical Medicine and Hygiene, 72(1), pp. 6–15. doi: 10.1016/0035-9203(78)90290-0.

Barbour, A. D. (1985) 'The importance of age and water contact patterns in relation to schistosoma haematobium infection', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 79(2), pp. 151–153. doi: 10.1016/0035-9203(85)90319-0.

Bärenbold, O. *et al.* (2017) 'Estimating sensitivity of the Kato-Katz technique for the diagnosis of *Schistosoma mansoni* and hookworm in relation to infection intensity', *PLoS Neglected Tropical Diseases*. Public Library of Science, 11(10), p. e0005953. Available at: <https://doi.org/10.1371/journal.pntd.0005953>.

Basch, P. F. (1990) 'Why do schistosomes have separate sexes?', *Parasitology Today*, 6(5), pp. 160–163. doi: 10.1016/0169-4758(90)90339-6.

Basch, P. F. and Basch, N. (1984) 'Intergeneric reproductive stimulation and parthenogenesis in *Schistosoma mansoni*', *Parasitology*. 2009/04/06. Cambridge University Press, 89(2), pp. 369–376. doi: DOI: 10.1017/S0031182000001372.

Basch, P. F. and Nicolas, C. (1989) 'Schistosoma mansoni: pairing of male worms with artificial surrogate females.', *Experimental parasitology*. United States, 68(2), pp. 202–207. doi: 10.1016/0014-4894(89)90098-2.

Berhe, N. *et al.* (2004) 'Variations in helminth faecal egg counts in Kato-Katz thick smears and their implications in assessing infection status with *Schistosoma mansoni*.', *Acta tropica*. Netherlands, 92(3), pp. 205–212. doi: 10.1016/j.actatropica.2004.06.011.

Berridge, M. J., Lipp, P. and Bootman, M. D. (2000) 'The versatility and universality of calcium signalling.', *Nature reviews. Molecular cell biology*. England, 1(1), pp. 11–21. doi: 10.1038/35036035.

Berrow, N. S., Alderton, D. and Owens, R. J. (2009) 'The Precise Engineering of Expression Vectors Using High-Throughput In-Fusion™ PCR Cloning', in Doyle, S. A. (ed.) *High Throughput Protein Expression and Purification. Methods in Molecular Biology*, vol 498. New York: Humana Press.

Black, C. L. *et al.* (2010) 'Influence of exposure history on the immunology and development of resistance to human *Schistosomiasis mansoni*', *PLoS Neglected Tropical Diseases*, 4(3). doi: 10.1371/journal.pntd.0000637.

Bleay, C. *et al.* (2007) 'Density-dependent immune responses against the gastrointestinal nematode *Strongyloides ratti*', *International Journal for Parasitology*, 37(13), pp. 1501–1509. doi: <https://doi.org/10.1016/j.ijpara.2007.04.023>.

Booth, M. *et al.* (2003) 'The influence of sampling effort and the performance of the Kato-Katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Côte d'Ivoire.', *Parasitology*. England, 127(Pt 6), pp. 525–531. doi: 10.1017/s0031182003004128.

Booth, M. *et al.* (2004) 'Hepatosplenic morbidity in two neighbouring communities in Uganda with high levels of *Schistosoma mansoni* infection but very different durations of residence.', *Transactions of the Royal Society of Tropical Medicine and Hygiene*. England, 98(2), pp. 125–136. doi: 10.1016/s0035-9203(03)00018-x.

Boulanger, D. *et al.* (1991) 'Immunization of mice and baboons with the recombinant Sm28GST affects both worm viability and fecundity after experimental infection with *Schistosoma mansoni*.', *Parasite immunology*, 13(5), pp. 473–90. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1956696>.

Bourke, C. D. *et al.* (2013) 'Integrated analysis of innate, Th1, Th2, Th17, and regulatory cytokines identifies changes in immune polarisation following treatment of human schistosomiasis', *Journal of Infectious Diseases*, 208(1), pp. 159–169. doi: 10.1093/infdis/jis524.

Bradley, D. J. and May, R. M. (1978) 'Consequences of helminth aggregation for the dynamics of schistosomiasis', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 72(3), pp. 262–273. doi: 10.1016/0035-9203(78)90205-5.

Braschi, S., Borges, W. C. and Wilson, R. A. (2006) 'Proteomic analysis of the schistosome tegument and its surface membranes', *Memorias do Instituto Oswaldo Cruz*, 101(SUPPL. 1), pp. 205–212. doi: S0074-02762006000900032 [pii].

Brown, S. P. and Grenfell, B. T. (2001) 'An unlikely partnership: parasites, concomitant immunity and host defence.', *Proceedings. Biological sciences / The Royal Society*, 268(1485), pp. 2543–9. doi: 10.1098/rspb.2001.1821.

Brynjolfsson, S. F. *et al.* (2018) 'Long-Lived Plasma Cells in Mice and Men', *Frontiers in Immunology*, 9, p. 2673. doi: 10.3389/fimmu.2018.02673.

Busemeyer, J. R. *et al.* (2015) 'Model Comparison and the Principle of Parsimony'. Oxford University Press.

Bushara, H. O. *et al.* (1980) 'Observations on cattle schistosomiasis in the Sudan, a study in comparative medicine. II. Experimental demonstration of naturally acquired resistance to *Schistosoma bovis*.', *The American journal of tropical medicine and hygiene*. United States, 29(3), pp. 442–451. doi: 10.4269/ajtmh.1980.29.442.

Bushara, H. O. *et al.* (1983) 'Observations on cattle schistosomiasis in the Sudan, a study in comparative medicine. IV. Preliminary observations on the mechanism of naturally acquired

resistance.’, *The American journal of tropical medicine and hygiene*. United States, 32(5), pp. 1065–1070. doi: 10.4269/ajtmh.1983.32.1065.

Bushara, H. O. *et al.* (1993) ‘Suppression of *Schistosoma bovis* egg production in cattle by vaccination with either glutathione S-transferase or keyhole limpet haemocyanin’, *Parasite immunology*, 15(7), pp. 383–390. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8414642>.

Butterworth, A. *et al.* (1988) ‘Immunity in human schistosomiasis mansoni: cross-reactive IgM and IgG2 anti-carbohydrate antibodies block the expression of immunity.’, *Biochimie*. France, 70(8), pp. 1053–1063. doi: 10.1016/0300-9084(88)90268-4.

Butterworth, A. E. *et al.* (1984) ‘Immunity after treatment of human schistosomiasis mansoni. I. Study design, pretreatment observations and the results of treatment.’, *Transactions of the Royal Society of Tropical Medicine and Hygiene*. England, 78(1), pp. 108–123. doi: 10.1016/0035-9203(84)90190-1.

Butterworth, A. E. *et al.* (1987) ‘Immunity in human schistosomiasis mansoni: prevention by blocking antibodies of the expression of immunity in young children.’, *Parasitology*, 94 (Pt 2), pp. 281–300.

van Buuren, S. and Groothuis-Oudshoorn, K. (2011) ‘mice: Multivariate imputation by chained equations in R’, *Journal of Statistical Software*, 45(3), pp. 1–67. doi: 10.18637/jss.v045.i03.

Capron, A. *et al.* (1994) ‘Development of a vaccine strategy against human and bovine schistosomiasis. Background and update’, *Trop Geogr Med*, 46(4), pp. 242–246.

Carson, J. *et al.* (2018) ‘Molecular & Biochemical Parasitology The tegumental allergen-like proteins of *Schistosoma mansoni*: A biochemical study of SmTAL4-TAL13’, *Molecular & Biochemical Parasitology*. Elsevier, 221(January), pp. 14–22. doi: 10.1016/j.molbiopara.2018.02.002.

Castresana, J. (2000) ‘Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis.’, *Molecular biology and evolution*. United States, 17(4), pp. 540–552. doi: 10.1093/oxfordjournals.molbev.a026334.

Castro-Borges, W., Simpson, D. M., *et al.* (2011) ‘Abundance of tegument surface proteins in the human blood fluke *Schistosoma mansoni* determined by QconCAT proteomics’, *Journal of Proteomics*. Elsevier B.V., 74(9), pp. 1519–1533. doi: 10.1016/j.jprot.2011.06.011.

Castro-Borges, W., Dowle, A., *et al.* (2011) ‘Enzymatic shaving of the tegument surface of live schistosomes for proteomic analysis: A rational approach to select vaccine candidates’, *PLoS Neglected Tropical Diseases*, 5(3). doi: 10.1371/journal.pntd.0000993.

Chan, M. S. *et al.* (1995) ‘The development of an age structured model for schistosomiasis

transmission dynamics and control and its validation for *Schistosoma mansoni*.', *Epidemiol Infect*, 115(2), pp. 325–344. doi: 10.1017/S0950268800058453.

Chan, M. S. *et al.* (1996) 'Dynamic aspects of morbidity and acquired immunity in schistosomiasis control', *Acta Tropica*, 62(2), pp. 105–117. doi: 10.1016/S0001-706X(96)00039-3.

Chan, M. S. (1996) 'The consequences of uncertainty for the prediction of the effects of schistosomiasis control programmes.', *Epidemiology and infection*, 117(3), pp. 537–50. doi: 10.1017/S0950268800059239.

Chan, M. S. *et al.* (2000) 'Stochastic simulation and the detection of immunity to schistosome infections.', *Parasitology*, 120 (Pt 2(September 2000), pp. 161–169. doi: 10.1017/S003118209900534X.

Chan, M. S. and Bundy, D. A. (1997) 'Modelling the dynamic effects of community chemotherapy on patterns of morbidity due to *Schistosoma mansoni*.', *Trans R Soc Trop Med Hyg*, 91(2), pp. 216–220. doi: 10.1016/S0035-9203(97)90231-5.

Chan, M. S., Guyati, H. L. and Kingdom, U. (1996) 'Development Infection Disease Late Disease Development', 55(1), pp. 52–62.

Chan, M. S. and Isham, V. S. (1998) 'A stochastic model of schistosomiasis immuno-epidemiology', *Mathematical Biosciences*, 151(2), pp. 179–198. doi: 10.1016/S0025-5564(98)10014-7.

Chandiwana, S. K. and Woolhouse, M. E. J. (1991) 'Heterogeneities in water contact patterns and the epidemiology of schistosoma haematobium', *Parasitology*. 2009/04/06. Cambridge University Press, 103(3), pp. 363–370. doi: DOI: 10.1017/S0031182000059874.

Cheesbrough, M. (2005) *District Laboratory Practice in Tropical Countries*. 2nd edn. Cambridge: Cambridge University Press. doi: DOI: 10.1017/CBO9780511581304.

Cheever, A. W. *et al.* (1994) 'Kinetics of egg production and egg excretion by *Schistosoma mansoni* and *S. japonicum* in mice infected with a single pair of worms', *American Journal of Tropical Medicine and Hygiene*, 50(3), pp. 281–295.

Cheever, A. W. and Andrade, Z. A. (1967) 'Pathological lesions associated with *Schistosoma mansoni* infection in man.', *Transactions of the Royal Society of Tropical Medicine and Hygiene*. England, 61(5), pp. 626–639. doi: 10.1016/0035-9203(67)90125-3.

Churcher, T. S., Filipe, J. A. N. and Basáñez, M. G. (2006) 'Density dependence and the control of helminth parasites', *Journal of Animal Ecology*, 75(6), pp. 1313–1320. doi: 10.1111/j.1365-2656.2006.01154.x.

- Civitello, D. J. and Rohr, J. R. (2014) 'Disentangling the effects of exposure and susceptibility on transmission of the zoonotic parasite *Schistosoma mansoni*', *Journal of Animal Ecology*, 83(6), pp. 1379–1386. doi: 10.1111/1365-2656.12222.
- Clegg, J. A., Smithers, S. R. and Terry, R. J. (1971) 'Concomitant immunity and host antigens associated with schistosomiasis', *International Journal for Parasitology*, 1(1), pp. 43–49. doi: 10.1016/0020-7519(71)90045-2.
- Coakley, G., Buck, A. H. and Maizels, R. M. (2016) 'Host parasite communications-Messages from helminths for the immune system: Parasite communication and cell-cell interactions.', *Molecular and biochemical parasitology*, 208(1), pp. 33–40. doi: 10.1016/j.molbiopara.2016.06.003.
- Coakley, G., Maizels, R. M. and Buck, A. H. (2015) 'Exosomes and Other Extracellular Vesicles: The New Communicators in Parasite Infections.', *Trends in parasitology*, 31(10), pp. 477–489. doi: 10.1016/j.pt.2015.06.009.
- Cohen, J. E. (1973) 'Selective Host Mortality in a Catalytic Model Applied to Schistosomiasis', *The American Naturalist*. [University of Chicago Press, American Society of Naturalists], 107(954), pp. 199–212. Available at: <http://www.jstor.org/stable/2459794>.
- Colley, D. G. *et al.* (2014) 'Human schistosomiasis.', *Lancet*. Elsevier, 383(9936), pp. 2253–64. doi: 10.1016/S0140-6736(13)61949-2.
- Colley, D. G. and Evan Secor, W. (2004) 'Immunoregulation and World Health Assembly resolution 54.19: Why does treatment control morbidity?', *Parasitology International*, 53(2), pp. 143–150. doi: 10.1016/j.parint.2004.01.005.
- Colley, D. G. and Secor, W. E. (2014) 'Immunology of human schistosomiasis', *Parasite Immunology*, 36, pp. 347–357. doi: 10.1111/pim.12087.
- Conover, W. J. and Iman, R. L. (1979) *On multiple-comparisons procedures. Technical Report LA-7677-MS*.
- Cornford, E. M. (1986) 'Influence of mating on surface nutrient exchange in schistosomes', *Journal of Chemical Ecology*, 12(8), pp. 1777–1796. doi: 10.1007/BF01022383.
- Corstjens, P. L. A. M. *et al.* (2014) 'Tools for diagnosis, monitoring and screening of *Schistosoma* infections utilizing lateral-flow based assays and upconverting phosphor labels.', *Parasitology*, 141(14), pp. 1841–1855. doi: 10.1017/S0031182014000626.
- Cox, D. R. (1972) 'Regression Models and Life-Tables', *Journal of the Royal Statistical Society. Series B (Methodological)*. [Royal Statistical Society, Wiley], 34(2), pp. 187–220. Available at: <http://www.jstor.org/stable/2985181>.

- Coyne, M. J. and Smith, G. (1991) 'The regulation of mortality and fecundity in *Schistosoma mattheei* following a single experimental infection in sheep.', *International journal for parasitology*. England, 21(8), pp. 877–882. doi: 10.1016/0020-7519(91)90161-y.
- Croote, D. *et al.* (2018) 'High-affinity allergen-specific human antibodies cloned from single IgE B cell transcriptomes', *Science*, 362(6420), pp. 1306 LP – 1309. doi: 10.1126/science.aau2599.
- Crosnier, C. *et al.* (2019) 'Systematic screening of 96 *Schistosoma mansoni* cell-surface and secreted antigens does not identify any strongly protective vaccine candidates in a mouse model of infection', *Wellcome open research*. F1000 Research Limited, 4, p. 159. doi: 10.12688/wellcomeopenres.15487.1.
- Curwen, R. S. *et al.* (2006) 'Identification of novel proteases and immunomodulators in the secretions of schistosome cercariae that facilitate host entry.', *Molecular & cellular proteomics : MCP*. United States, 5(5), pp. 835–844. doi: 10.1074/mcp.M500313-MCP200.
- D'Udine, F. *et al.* (2015) 'Migration and Conservation in the Bale Mountains Ecosystem', *International Institute for Sustainable Development*. Winnipeg, (August), pp. 1–26.
- Da'dara, A. A. and Krautz-Peterson, G. (2014) 'New insights into the reaction of *Schistosoma mansoni* cercaria to the human complement system', *Parasitology research*. 2014/07/17, 113(10), pp. 3685–3696. doi: 10.1007/s00436-014-4033-3.
- Dall'antonia, F. *et al.* (2014) 'Structure of allergens and structure based epitope predictions.', *Methods (San Diego, Calif.)*, 66(1), pp. 3–21. doi: 10.1016/j.ymeth.2013.07.024.
- van Dam, G. J. *et al.* (2004) 'Diagnosis of schistosomiasis by reagent strip test for detection of circulating cathodic antigen.', *Journal of clinical microbiology*, 42(12), pp. 5458–5461. doi: 10.1128/JCM.42.12.5458-5461.2004.
- Van Dam, G. J. *et al.* (1996) 'Antibody response patterns against *Schistosoma mansoni* in a recently exposed community in Senegal', *Journal of Infectious Diseases*, 173(5), pp. 1232–1241. doi: 10.1093/infdis/173.5.1232.
- Deelder, A. M. *et al.* (1989) 'Sensitive determination of circulating anodic antigen in *Schistosoma mansoni* infected individuals by an enzyme-linked immunosorbent assay using monoclonal antibodies.', *The American journal of tropical medicine and hygiene*. United States, 40(3), pp. 268–272. doi: 10.4269/ajtmh.1989.40.268.
- Deelder, A. M. *et al.* (1994) 'Quantitative diagnosis of *Schistosoma* infections by measurement of circulating antigens in serum and urine.', *Tropical and geographical medicine*. Netherlands, 46(4 Spec No), pp. 233–238.
- Deelder, M. *et al.* (1989) 'Sensitive determination of circulating anodic antigen in *Schistosoma*

mansoni infected individuals by an enzyme-linked immunosorbent assay using monoclonal antibodies', 40(3), pp. 268–272.

Delcroix, M. *et al.* (2007) 'Proteomic analysis of adult *S. mansoni* gut contents', *Molecular and biochemical parasitology*. 2007/03/20, 154(1), pp. 95–97. doi: 10.1016/j.molbiopara.2007.03.008.

Demeure, C. E., Rihet, P., Abel, L., Ouattara, M., Bourgois, A., *et al.* (1993) 'Resistance to *Schistosoma mansoni* in humans: influence of the IgE/IgG4 balance and IgG2 in immunity to reinfection after chemotherapy.', *The Journal of infectious diseases*. United States, 168(4), pp. 1000–1008. doi: 10.1093/infdis/168.4.1000.

Demeure, C. E., Rihet, P., Abel, L., Ouattara, M., Bourgois, a, *et al.* (1993) 'Resistance to *Schistosoma mansoni* in humans: influence of the IgE/IgG4 balance and IgG2 in immunity to reinfection after chemotherapy.', *The Journal of infectious diseases*, 168(4), pp. 1000–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7690821>.

Deol, A. K. *et al.* (2019) 'Schistosomiasis - Assessing Progress toward the 2020 and 2025 Global Goals', *The New England journal of medicine*. Massachusetts Medical Society, 381(26), pp. 2519–2528. doi: 10.1056/NEJMoal812165.

deWalick, S. *et al.* (2011) 'The proteome of the insoluble *Schistosoma mansoni* eggshell skeleton', *International Journal for Parasitology*, 41(5), pp. 523–532. doi: <https://doi.org/10.1016/j.ijpara.2010.12.005>.

Dickinson, H. A. (2012) *The tegumental allergen-like proteins of Schistosoma haematobium : developmental expression and human antibody responses*. University of Cambridge.

Diemert, D. J. *et al.* (2012) 'Generalized urticaria induced by the Na-ASP-2 hookworm vaccine: implications for the development of vaccines against helminths.', *The Journal of allergy and clinical immunology*. United States, 130(1), pp. 169–76.e6. doi: 10.1016/j.jaci.2012.04.027.

Dinno, A. (2017) 'Conover-Iman Test of Multiple Comparisons Using Rank Sums', p. 7. Available at: <https://cran.r-project.org/web/packages/conover.test/conover.test.pdf>.

Doehring, E., Feldmeier, H. and Daffalla, A. A. (1983) 'Day-to-day variation and circadian rhythm of egg excretion in urinary schistosomiasis in the Sudan.', *Annals of tropical medicine and parasitology*. England, 77(6), pp. 587–594. doi: 10.1080/00034983.1983.11811757.

Driguez, P. *et al.* (2010) 'Schistosomiasis vaccine discovery using immunomics.', *Parasit Vectors*, 3, p. 4. doi: 1756-3305-3-4 [pii]\r10.1186/1756-3305-3-4.

Driguez, P. *et al.* (2015) 'Protein microarrays for parasite antigen discovery.', *Methods in molecular biology (Clifton, N.J.)*. United States, 1201, pp. 221–233. doi: 10.1007/978-1-4939-1438-

Dunn, O. J. (1964) 'Multiple Comparisons Using Rank Sums', *Technometrics*. Taylor & Francis, 6(3), pp. 241–252. doi: 10.1080/00401706.1964.10490181.

Dunne, D. W. *et al.* (1992) 'Immunity after treatment of human schistosomiasis: association between IgE antibodies to adult worm antigens and resistance to reinfection', *European Journal of Immunology*, 22, pp. 1483–1494.

Edgar, R. C. (2004) 'MUSCLE: a multiple sequence alignment method with reduced time and space complexity.', *BMC bioinformatics*, 5, p. 113. doi: 10.1186/1471-2105-5-113.

Emlen, S. T. and Oring, L. W. (1977) 'Ecology, sexual selection, and the evolution of mating systems.', *Science (New York, N.Y.)*. United States, 197(4300), pp. 215–223. doi: 10.1126/science.327542.

Erazo, A. *et al.* (2007) 'Unique maturation program of the IgE response in vivo.', *Immunity*, 26(2), pp. 191–203. doi: 10.1016/j.immuni.2006.12.006.

Etard, J. F., Audibert, M. and Dabo, A. (1995) 'Age-acquired resistance and predisposition to reinfection with *Schistosoma haematobium* after treatment with praziquantel in Mali.', *The American journal of tropical medicine and hygiene*. United States, 52(6), pp. 549–558. doi: 10.4269/ajtmh.1995.52.549.

Van Etten, L. *et al.* (1997) 'Day-to-day variation of egg output and schistosome circulating antigens in urine of *Schistosoma haematobium*-infected school children from Gabon and follow-up after chemotherapy.', *The American journal of tropical medicine and hygiene*, 57(3), pp. 337–41. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9311646> <http://hinari-gw.who.int/whalecomwww.ajtmh.org/whalecom0/content/57/3/337.full.pdf+html?frame=sidebar>.

European Bioinformatics Institute (2018) *MEROPS - the peptidase database, Release 12.1*. Available at: <https://www.ebi.ac.uk/merops/cgi-bin/speccards?sp=sp002458;type=peptidase> (Accessed: 4 May 2020).

Fajtová, P. *et al.* (2015) 'Prolyl Oligopeptidase from the Blood Fluke *Schistosoma mansoni*: From Functional Analysis to Anti-schistosomal Inhibitors', *PLoS Neglected Tropical Diseases*. Public Library of Science, 9(6), p. e0003827. Available at: <https://doi.org/10.1371/journal.pntd.0003827>.

Farnell, E. J. *et al.* (2015) 'Known allergen structures predict *Schistosoma mansoni* IgE-binding antigens in human infection', *Frontiers in Immunology*, 6(FEB). doi: 10.3389/fimmu.2015.00026.

Feldmeier, H. and Poggensee, G. (1993) 'Diagnostic techniques in schistosomiasis control. A

- review.’, *Acta tropica*. Netherlands, 52(4), pp. 205–220. doi: 10.1016/0001-706x(93)90009-z.
- Feng, Z., Eppert, A., Milner, F.A., *et al.* (2004) ‘Estimation of parameters governing the transmission dynamics of schistosomosomes’, *Applied Mathematics Letters*, 17, pp. 1105–1112. doi: 10.1016/j.aml.2004.02.002.
- Feng, Z., Eppert, A., Milner, F. A., *et al.* (2004) ‘Estimation of parameters governing the transmission dynamics of schistosomosomes’, *Applied Mathematics Letters*, 17(10), pp. 1105–1112. doi: 10.1016/j.aml.2004.02.002.
- Figueiredo, B. C. P. *et al.* (2015) ‘Kicking in the guts: *Schistosoma mansoni* digestive tract proteins are potential candidates for vaccine development’, *Frontiers in Immunology*, 6(JAN), pp. 1–7. doi: 10.3389/fimmu.2015.00022.
- Filipe, A. N. *et al.* (2005) ‘Human infection patterns and heterogeneous exposure in river blindness’, 102(42).
- Finkelman, F. D. *et al.* (1988) ‘IL-4 is required to generate and sustain in vivo IgE responses.’, *The Journal of Immunology*, 141(7), pp. 2335 LP – 2341. Available at: <http://www.jimmunol.org/content/141/7/2335.abstract>.
- Fitzsimmons, C. M. *et al.* (2004) ‘Human IgE response to the *Schistosoma haematobium* 22. 6 kDa antigen’, *Parasite Immunology*, 26(9-Aug), pp. 371–376. Available at: isi:000226654600009.
- Fitzsimmons, C. M., Jones, F. M., Pinot de Moira, A., *et al.* (2012) ‘Progressive cross-reactivity in IgE responses: An explanation for the slow development of human immunity to Schistosomiasis?’, *Infection and Immunity*, 80(12), pp. 4264–4270. doi: 10.1128/IAI.00641-12.
- Fitzsimmons, C. M., Jones, F. M., Stearn, A., *et al.* (2012) ‘The *Schistosoma mansoni* tegumental-allergen-like (TAL) protein family: Influence of developmental expression on human igE responses’, *PLoS Neglected Tropical Diseases*, 6(4). doi: 10.1371/journal.pntd.0001593.
- Flemming, A. (2019) ‘Class-switch recombination revised’, *Nature Reviews Immunology*, 19(10), pp. 596–597. doi: 10.1038/s41577-019-0214-8.
- Floudas, A. *et al.* (2017) ‘Composition of the *Schistosoma mansoni* worm secretome: Identification of immune modulatory Cyclophilin A’, *PLoS Neglected Tropical Diseases*, 11(10). doi: 10.1371/journal.pntd.0006012.
- Francis, P. and Bickle, Q. (1992) ‘Cloning of a 21.7-kDa vaccine-dominant antigen gene of *Schistosoma mansoni* reveals an EF hand-like motif.’, *Molecular and biochemical parasitology*. Netherlands, 50(2), pp. 215–224. doi: 10.1016/0166-6851(92)90218-9.

- French, M. D. *et al.* (2010) 'Observed reductions in *Schistosoma mansoni* transmission from large-scale administration of praziquantel in Uganda: a mathematical modelling study.', *PLoS neglected tropical diseases*, 4(11), p. e897. doi: 10.1371/journal.pntd.0000897.
- French, M. D. *et al.* (2015) 'Estimation of changes in the force of infection for intestinal and urogenital schistosomiasis in countries with schistosomiasis control initiative-assisted programmes.', *Parasites & vectors*, 8(1), p. 558. doi: 10.1186/s13071-015-1138-1.
- Fulford, A. J. *et al.* (1996) 'Water contact observations in Kenyan communities endemic for schistosomiasis: methodology and patterns of behaviour.', *Parasitology*, 113 (Pt 3, pp. 223–41. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8811848>.
- Fulford, A. J. *et al.* (1992) 'On the use of age-intensity data to detect immunity to parasitic infections, with special reference to *Schistosoma mansoni* in Kenya.', *Parasitology*, 105 (Pt 2(1988), pp. 219–27. doi: 10.1017/S003118200007414X.
- Fulford, A. J. *et al.* (1995) 'A statistical approach to schistosome population dynamics and estimation of the life-span of *Schistosoma mansoni* in man.', *Parasitology*, 110 (Pt 3, pp. 307–16.
- Fulford, A. J. C. *et al.* (1995) 'A statistical approach to schistosome population dynamics and estimation of the life-span of *Schistosoma mansoni* in man', *Parasitology*, 110(Pt 3), pp. 307–316.
- Fulford, A. J. C. *et al.* (1998) 'Puberty and age-related changes in susceptibility to schistosome infection', *Parasitology Today*, 14(1), pp. 23–26. doi: 10.1016/S0169-4758(97)01168-X.
- Fulford, A. J. C. (2000) *Analysis of the Patterns of Exposure to, and Infection by, Schistosoma mansoni*. Department. University of Cambridge.
- Galanti, S. E., Huang, S. C. C. and Pearce, E. J. (2012) 'Cell death and reproductive regression in female *Schistosoma mansoni*', *PLoS Neglected Tropical Diseases*. doi: 10.1371/journal.pntd.0001509.
- Garthwaite, P. H., Fan, Y. and Sisson, S. A. (2016) 'Adaptive optimal scaling of Metropolis–Hastings algorithms using the Robbins–Monro process', *Communications in Statistics - Theory and Methods*. Taylor & Francis, 45(17), pp. 5098–5111. doi: 10.1080/03610926.2014.936562.
- Gaze, S. *et al.* (2014) 'An Immunomics Approach to Schistosome Antigen Discovery: Antibody Signatures of Naturally Resistant and Chronically Infected Individuals from Endemic Areas', *PLoS Pathogens*, 10(3), p. e1004033. doi: 10.1371/journal.ppat.1004033.
- Gould, H. J. and Wu, Y.-C. B. (2018) 'IgE repertoire and immunological memory: compartmental regulation and antibody function', *International immunology*. Oxford University Press, 30(9), pp. 403–412. doi: 10.1093/intimm/dxy048.

- Grabski, A., Mehler, M. and Drott, D. (2005) 'The Overnight Express Autoinduction System: High-density cell growth and protein expression while you sleep', *Nature Methods*, 2(3), pp. 233–235. doi: 10.1038/nmeth0305-233.
- Gray, D. J. *et al.* (2011) 'Diagnosis and management of schistosomiasis', *BMJ (Clinical research ed.)*. BMJ Publishing Group Ltd., 342, pp. d2651–d2651. doi: 10.1136/bmj.d2651.
- Green, M. and Sambrook, J. (2012) *Molecular Cloning: A Laboratory Manual*. Fourth Edi. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Griffin, M., Coutinho, F. A. and Thomas, J. D. (1988) 'A schistosomiasis transmission model incorporating concomitant immunity.', *Revista brasileira de biologia*. Brazil, 48(3), pp. 553–563.
- Gryseels, B. *et al.* (1994) 'Epidemiology, immunology and chemotherapy of *Schistosoma mansoni* infections in a recently exposed community in Senegal.', *Tropical and geographical medicine*. Netherlands, 46(4 Spec No), pp. 209–219.
- Gryseels, B. *et al.* (2006) 'Human schistosomiasis.', *Lancet (London, England)*, 368(9541), pp. 1106–18. doi: 10.1016/S0140-6736(06)69440-3.
- Grzych, J. *et al.* (1993) 'IgA antibodies to a protective antigen in human Schistosomiasis mansoni.', *Journal of immunology*, 150(2), pp. 527–535.
- Gupta, B. C. and Basch, P. F. (1987) 'The Role of *Schistosoma mansoni* Males in Feeding and Development of Female Worms', *The Journal of Parasitology*, 73(3), pp. 481–486. Available at: <http://www.jstor.org/stable/pdf/3282125.pdf>.
- Gurarie, D. *et al.* (2015) 'Modelling control of *Schistosoma haematobium* infection : predictions of the long-term impact of mass drug administration in Africa', *Parasites & Vectors*. Parasites & Vectors, pp. 1–14. doi: 10.1186/s13071-015-1144-3.
- Gurarie, D. and King, C. H. (2005) 'Heterogeneous model of schistosomiasis transmission and long-term control: the combined influence of spatial variation and age-dependent factors on optimal allocation of drug therapy.', *Parasitology*, 130(Pt 1), pp. 49–65. doi: 10.1017/S0031182004006341.
- Gurarie, D. and King, C. H. (2014) 'Population biology of *Schistosoma* mating, aggregation, and transmission breakpoints: more reliable model analysis for the end-game in communities at risk.', *PLoS One*, 9(12), p. e115875. doi: 10.1371/journal.pone.0115875.
- Gurarie, D., King, C. H. and Wang, X. (2010) 'A new approach to modelling schistosomiasis transmission based on stratified worm burden.', *Parasitology*, 137(13), pp. 1951–1965. doi: 10.1017/S0031182010000867.

Guyatt, H. L. *et al.* (1994) 'Aggregation in schistosomiasis: comparison of the relationships between prevalence and intensity in different endemic areas.', *Parasitology*. England, 109 (Pt 1, pp. 45–55. doi: 10.1017/s0031182000077751.

Hackett, L. W. (1937) *Malaria in Europe. An Ecological Study*. Oxford Univ. Pr.; London H. Milford.

Hagan, P *et al.* (1991) 'Human IgE, IgG4 and resistance to reinfection with *Schistosoma haematobium*.' *Nature*, 349(6306), pp. 243–5. doi: 10.1038/349243a0.

Hagan, Paul *et al.* (1991) 'Human IgE, IgG4 and resistance to reinfection with *Schistosoma haematobium*', *Nature*, 349(6306), pp. 243–245. doi: 10.1038/349243a0.

Hagan, P. (1992) 'Reinfection, exposure and immunity in human schistosomiasis', *Parasitology Today*, 8(1), pp. 12–16. doi: 10.1016/0169-4758(92)90303-J.

Hairston, N. G. (1965) 'On the mathematical analysis of schistosome populations.', *Bulletin of the World Health Organization*, 33(1), pp. 45–62.

Hall, S. L. *et al.* (2011) 'Insights into blood feeding by schistosomes from a proteomic analysis of worm vomitus', *Molecular and Biochemical Parasitology*. Elsevier B.V., 179(1), pp. 18–29. doi: 10.1016/j.molbiopara.2011.05.002.

Han, Z.-G. *et al.* (2009) 'Schistosoma genomics: new perspectives on schistosome biology and host-parasite interaction.', *Annual review of genomics and human genetics*. United States, 10, pp. 211–240. doi: 10.1146/annurev-genom-082908-150036.

Haniuda, K. *et al.* (2016) 'Autonomous membrane IgE signaling prevents IgE-memory formation', *Nature Immunology*, 17(9), pp. 1109–1117. doi: 10.1038/ni.3508.

Hatta, T. *et al.* (2006) 'Identification and characterisation of a leucine aminopeptidase from the hard tick *Haemaphysalis longicornis*', *International Journal for Parasitology*, 36(10), pp. 1123–1132. doi: <https://doi.org/10.1016/j.ijpara.2006.05.010>.

He, J.-S. *et al.* (2015) 'Biology of IgE Production: IgE Cell Differentiation and the Memory of IgE Responses', in: Springer, Cham, pp. 1–19. doi: 10.1007/978-3-319-13725-4_1.

He, J.-S. *et al.* (2017) 'IgG1 memory B cells keep the memory of IgE responses.', *Nature communications*, 8(1), p. 641. doi: 10.1038/s41467-017-00723-0.

Heesterbeek, H., Anderson, R., *et al.* (2015) 'Modeling infectious disease dynamics in the complex landscape of global health', *Science*, 347(6227), p. aaa4339. doi: 10.1126/science.aaa4339.Modeling.

Heesterbeek, H., Anderson, R. M., *et al.* (2015) 'Modeling infectious disease dynamics in the

complex landscape of global health', *Science*, 347(6227). Available at: <http://science.sciencemag.org/content/347/6227/aaa4339.abstract>.

Hewitson, J. P., Grainger, J. R. and Maizels, R. M. (2009) 'Helminth immunoregulation: The role of parasite secreted proteins in modulating host immunity', *Molecular and Biochemical Parasitology*. Elsevier B.V., 167(1), pp. 1–11. doi: 10.1016/j.molbiopara.2009.04.008.

Hockley, D. J. and McLaren, D. J. (1973) 'Schistosoma mansoni: Changes in the outer membrane of the tegument during development from cercaria to adult worm', *International Journal for Parasitology*, 3(1), pp. 13–20. doi: 10.1016/0020-7519(73)90004-0.

Hoffmann, K. F. and Strand, M. (1997) 'Molecular characterization of a 20.8-kDa Schistosoma mansoni antigen. Sequence similarity to tegumental associated antigens and dynein light chains.', *The Journal of biological chemistry*. United States, 272(23), pp. 14509–14515. doi: 10.1074/jbc.272.23.14509.

Hollingsworth, T. D. *et al.* (2015) 'Seven challenges for modelling indirect transmission: vector-borne diseases, macroparasites and neglected tropical diseases.', *Epidemics*, 10, pp. 16–20. doi: 10.1016/j.epidem.2014.08.007.

Jamieson, B. G. M. (2016) *Schistosoma: Biology, Pathology and Control*. First edit. Boca Raton: CRC Press.

Janeway, C. A. *et al.* (2001) *Immunobiology: The Immune System in Health and Disease*. 5th Editio. New York: Garland Publishing.

Jimenez-Saiz, R. *et al.* (2017) 'Lifelong memory responses perpetuate humoral T H 2 immunity and anaphylaxis in food allergy', *Journal of Allergy and Clinical Immunology*, 140(6), pp. 1604–1615.e5. doi: 10.1016/j.jaci.2017.01.018.

Jiz, M. *et al.* (2009) 'Immunoglobulin E (IgE) Responses to Paramyosin Predict Resistance to Reinfection with Schistosoma japonicum and Are Attenuated by IgG4', *Infection and Immunity*, 77(5), pp. 2051–2058. doi: 10.1128/IAI.00012-09.

Jones, M. K., Jamieson, B. G. M. and Justine, J. L. (2017) 'Reproductive System of Schistosoma', in Jamieson, B. G. M. (ed.) *Schistosoma Biology, Pathology and Control*. Boca Raton: CRC Press, p. 280.

De Jonge, N. *et al.* (1989) 'Presence of the schistosome circulating anodic antigen (CAA) in urine of patients with Schistosoma mansoni or S. haematobium infections.', *The American journal of tropical medicine and hygiene*. United States, 41(5), pp. 563–569. doi: 10.4269/ajtmh.1989.41.563.

Joshua, G. W. (2001) 'Functional analysis of leucine aminopeptidase in Caenorhabditis elegans.', *Molecular and biochemical parasitology*. Netherlands, 113(2), pp. 223–232. doi: 10.1016/s0166-

6851(01)00221-3.

Karanja, D. M. *et al.* (2002) 'Resistance to reinfection with *Schistosoma mansoni* in occupationally exposed adults and effect of HIV-1 co-infection on susceptibility to schistosomiasis: a longitudinal study', *The Lancet*, 360(9333), pp. 592–596. doi: 10.1016/S0140-6736(02)09781-7.

Kariuki, H. C. *et al.* (2013) 'Long term study on the effect of mollusciciding with niclosamide in stream habitats on the transmission of schistosomiasis mansoni after community-based chemotherapy in Makueni District, Kenya', *Parasites and Vectors*, 6(1). doi: 10.1186/1756-3305-6-107.

Katz, N., Chaves, A. and Pellegrino, J. (1972) 'A simple device for quantitative stool thick-smear technique in Schistosomiasis mansoni.', *Revista do Instituto de Medicina Tropical de Sao Paulo*. Brazil, 14(6), pp. 397–400.

Keymer, A. (1982) 'Density-dependent mechanisms in the regulation of intestinal helminth populations', *Parasitology*. 2009/04/06. Cambridge University Press, 84(3), pp. 573–587. doi: DOI: 10.1017/S0031182000052847.

Keymer, A. E. and Slater, A. F. G. (1987) 'Helminth fecundity: Density dependence or statistical illusion?', *Parasitology Today*, 3(2), pp. 56–58. doi: 10.1016/0169-4758(87)90215-8.

Kildemoes, A. O. *et al.* (2017) 'Rapid clearance of *Schistosoma mansoni* circulating cathodic antigen after treatment shown by urine strip tests in a Ugandan fishing community - Relevance for monitoring treatment efficacy and re-infection.', *PLoS Neglected Tropical Diseases*, 11(11), p. e0006054. doi: 10.1371/journal.pntd.0006054.

Kim, Y.-J. *et al.* (2012) 'Identification and characterization of a novel 21.6-kDa tegumental protein from *Clonorchis sinensis*', *Parasitology Research*, 110(5), pp. 2061–2066. doi: 10.1007/s00436-011-2681-0.

King, C. H. (2017) 'The evolving schistosomiasis agenda 2007-2017—Why we are moving beyond morbidity control toward elimination of transmission', *PLoS Neglected Tropical Diseases*. Public Library of Science, 11(4), p. e0005517. Available at: <https://doi.org/10.1371/journal.pntd.0005517>.

King, C. H., Dickman, K. and Tisch, D. J. (2005) 'Reassessment of the cost of chronic helminthic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis.', *Lancet (London, England)*, 365(9470), pp. 1561–9. doi: 10.1016/S0140-6736(05)66457-4.

King, C. L. *et al.* (1991) 'Frequency analysis of IgE-secreting B lymphocytes in persons with normal or elevated serum IgE levels.', *Journal of immunology (Baltimore, Md. : 1950)*. United

States, 146(5), pp. 1478–1483.

Kittur, N. *et al.* (2019) 'Persistent Hotspots in Schistosomiasis Consortium for Operational Research and Evaluation Studies for Gaining and Sustaining Control of Schistosomiasis after Four Years of Mass Drug Administration of Praziquantel.', *The American journal of tropical medicine and hygiene*, 101(3), pp. 617–627. doi: 10.4269/ajtmh.19-0193.

Kjetland, E. F. *et al.* (2006) 'Association between genital schistosomiasis and HIV in rural Zimbabwean women.', *AIDS (London, England)*. England, 20(4), pp. 593–600. doi: 10.1097/01.aids.0000210614.45212.0a.

Knopp, S. *et al.* (2015) 'Sensitivity and Specificity of a Urine Circulating Anodic Antigen Test for the Diagnosis of *Schistosoma haematobium* in Low Endemic Settings', *PLoS Neglected Tropical Diseases*, 9(5), p. e0003752. doi: 10.1371/journal.pntd.0003752.

Krauth, S. J. *et al.* (2012) 'An In-Depth Analysis of a Piece of Shit: Distribution of *Schistosoma mansoni* and Hookworm Eggs in Human Stool', *PLoS Neglected Tropical Diseases*, 6(12). doi: 10.1371/journal.pntd.0001969.

Krauth, S. J. *et al.* (2019) 'A Call for Systems Epidemiology to Tackle the Complexity of Schistosomiasis, Its Control, and Its Elimination.', *Tropical medicine and infectious disease*, 4(1). doi: 10.3390/tropicalmed4010021.

Kubo, S. *et al.* (2003) 'Long term maintenance of IgE-mediated memory in mast cells in the absence of detectable serum IgE.', *Journal of immunology (Baltimore, Md. : 1950)*. United States, 170(2), pp. 775–780. doi: 10.4049/jimmunol.170.2.775.

Kück, P. *et al.* (2010) 'Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees', *Frontiers in Zoology*, 7(1), p. 10. doi: 10.1186/1742-9994-7-10.

de la Torre-Escudero, E. *et al.* (2012) 'Molecular cloning, characterization and diagnostic performance of the *Schistosoma bovis* 22.6 antigen.', *Veterinary parasitology*. Netherlands, 190(3–4), pp. 530–540. doi: 10.1016/j.vetpar.2012.06.023.

Lawrence, J. D. (1973) 'The ingestion of red blood cells by *Schistosoma mansoni*', *The Journal of Parasitology*, 59(1), pp. 60–63. Available at: http://www.jstor.org/stable/pdf/3278572.pdf?_=1464347640493.

Lawrence, M. G. *et al.* (2017) 'Half-life of IgE in serum and skin: Consequences for anti-IgE therapy in patients with allergic disease', *Journal of Allergy and Clinical Immunology*. Elsevier Inc., 139(2), pp. 422–428.e4. doi: 10.1016/j.jaci.2016.04.056.

Lessler, J. *et al.* (2015) 'Seven challenges for model-driven data collection in experimental and

- observational studies.’, *Epidemics*, 10, pp. 78–82. doi: 10.1016/j.epidem.2014.12.002.
- Li, E. Y. *et al.* (2019) ‘Improving public health control of schistosomiasis with a modified WHO strategy: a model-based comparison study’, *The Lancet Global Health*. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license, 7(10), pp. e1414–e1422. doi: 10.1016/S2214-109X(19)30346-8.
- von Lichtenberg, F. (1975) ‘Schistosomiasis as a worldwide problem: pathology.’, *Journal of toxicology and environmental health*. United States, 1(2), pp. 175–184. doi: 10.1080/15287397509529319.
- Liu, P. *et al.* (2012) ‘Schistosoma Japonicum UDP-Glucose 4-Epimerase Protein Is Located on the Tegument and Induces Moderate Protection against Challenge Infection’, *PLoS One*. Public Library of Science, 7(7), p. e42050. Available at: <https://doi.org/10.1371/journal.pone.0042050>.
- Looney, T. J. *et al.* (2016) ‘Human B-cell isotype switching origins of IgE.’, *The Journal of allergy and clinical immunology*, 137(2), pp. 579–586.e7. doi: 10.1016/j.jaci.2015.07.014.
- Lopata, A. L., Kleine-Tebbe, J. and Kamath, S. D. (2016) ‘Allergens and molecular diagnostics of shellfish allergy: Part 22 of the Series Molecular Allergology’, *Allergo journal international*. 2016/11/02. Springer Medizin, 25(7), pp. 210–218. doi: 10.1007/s40629-016-0124-2.
- LoVerde, P. T. *et al.* (2004) ‘Schistosoma mansoni male–female interactions’, *Canadian Journal of Zoology*. NRC Research Press, 82(2), pp. 357–374. doi: 10.1139/z03-217.
- Lu, Z. *et al.* (2016) ‘Schistosome sex matters: a deep view into gonad-specific and pairing-dependent transcriptomes reveals a complex gender interplay’, *Scientific Reports*, 6(1), p. 31150. doi: 10.1038/srep31150.
- Lu, Z. *et al.* (2017) ‘A gene expression atlas of adult Schistosoma mansoni and their gonads’, *Scientific Data*, 4(1), p. 170118. doi: 10.1038/sdata.2017.118.
- Lu, Z. *et al.* (2019) ‘Males, the Wrongly Neglected Partners of the Biologically Unprecedented Male–Female Interaction of Schistosomes’, *Frontiers in Genetics*, 10(September), pp. 1–9. doi: 10.3389/fgene.2019.00796.
- Ludolf, F. *et al.* (2014) ‘Serological Screening of the Schistosoma mansoni Adult Worm Proteome’, *PLoS Neglected Tropical Diseases*. Public Library of Science, 8(3), p. e2745. Available at: <https://doi.org/10.1371/journal.pntd.0002745>.
- Lunn, T. J. *et al.* (2019) ‘Dose-response and transmission: the nexus between reservoir hosts, environment and recipient hosts.’, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 374(1782), p. 20190016. doi: 10.1098/rstb.2019.0016.

- Macdonald, G. (1965) 'The dynamics of helminth infections, with special reference to schistosomes', *Transactions of the Royal Society of Tropical Medicine & Hygiene*, 59(5), pp. 489–506.
- Maizels, R. M. *et al.* (1991) *Parasite Antigens, Parasite Genes: A Laboratory Manual for Molecular Parasitology*. Cambridge: Cambridge University Press.
- Maldonado, J. F. and Acosta-Matienzo, J. (1948) 'Biological Studies on the Miracidium of *Schistosoma mansoni*', *The American Journal of Tropical Medicine and Hygiene*. The American Society of Tropical Medicine and Hygiene, 1-28(5), pp. 645–657. doi: 10.4269/ajtmh.1948.sl-28.645.
- Marcilla, A. *et al.* (2012) 'Extracellular Vesicles from Parasitic Helminths Contain Specific Excretory/Secretory Proteins and Are Internalized in Intestinal Host Cells', *PLoS ONE*. Public Library of Science, 7(9), p. e45974. Available at: <https://doi.org/10.1371/journal.pone.0045974>.
- Marcilla, A. *et al.* (2014) 'Extracellular vesicles in parasitic diseases', *Journal of Extracellular Vesicles*, 3(1). doi: 10.3402/jev.v3.25040.
- Margolis, L. and Sadovsky, Y. (2019) 'The biology of extracellular vesicles: The known unknowns', *PLoS Biology*. Public Library of Science, 17(7), p. e3000363. Available at: <https://doi.org/10.1371/journal.pbio.3000363>.
- Martin, A. *et al.* (2015) 'Chapter 20 - Somatic Hypermutation: The Molecular Mechanisms Underlying the Production of Effective High-Affinity Antibodies', in Alt, F. W. *et al.* (eds). London: Academic Press, pp. 363–388. doi: <https://doi.org/10.1016/B978-0-12-397933-9.00020-5>.
- May, R. M. (1977) 'Togetherness among Schistosomes: its effects on the dynamics of the infection', *Mathematical Biosciences*, 35(3–4), pp. 301–343. doi: 10.1016/0025-5564(77)90030-X.
- Mbanefo, E. C. *et al.* (2015) 'Immunogenicity and anti-fecundity effect of nanoparticle coated glutathione S-transferase (SjGST) DNA vaccine against murine *Schistosoma japonicum* infection.', *Parasitology international*. Netherlands, 64(4), pp. 24–31. doi: 10.1016/j.parint.2015.01.005.
- McCarthy, E. *et al.* (2004) 'Leucine aminopeptidase of the human blood flukes, *Schistosoma mansoni* and *Schistosoma japonicum*', *International Journal for Parasitology*, 34(6), pp. 703–714. doi: <https://doi.org/10.1016/j.ijpara.2004.01.008>.
- McSorley, H. J., Hewitson, J. P. and Maizels, R. M. (2013) 'Immunomodulation by helminth parasites: defining mechanisms and mediators.', *International journal for parasitology*. England, 43(3–4), pp. 301–310. doi: 10.1016/j.ijpara.2012.11.011.

- Medhat, a *et al.* (1998) 'Increased interleukin-4 and interleukin-5 production in response to *Schistosoma haematobium* adult worm antigens correlates with lack of reinfection after treatment.', *The Journal of infectious diseases*, 178(2), pp. 512–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9697734>.
- Medley, G. F. and Bundy, D. A P (1996) 'Dynamic modeling of epidemiologic patterns of schistosomiasis morbidity', *American Journal of Tropical Medicine and Hygiene*, 55(5 SUPPL.), pp. 149–158.
- Medley, G. F. and Bundy, D. A.P. (1996) 'Dynamic modeling of epidemiologic patterns of schistosomiasis morbidity', *American Journal of Tropical Medicine and Hygiene*, 55(5 SUPPL.), pp. 149–158. doi: 10.4269/ajtmh.1996.55.149.
- Michaels, R. M. (1969) 'Mating of *Schistosoma mansoni* in Vitro', *Experimental Parasitology*, 25, pp. 58–71. doi: [https://doi.org/10.1016/0014-4894\(69\)90052-6](https://doi.org/10.1016/0014-4894(69)90052-6).
- Mitchell, K. M. *et al.* (2012) 'Protective immunity to *Schistosoma haematobium* infection is primarily an anti-fecundity response stimulated by the death of adult worms', *Proceedings of the National Academy of Sciences*, 109(33), pp. 13347–13352. doi: 10.1073/pnas.1121051109.
- Mitchell, K. M. *et al.* (2014) 'Predicted impact of mass drug administration on the development of protective immunity against *Schistosoma haematobium*.', *PLoS neglected tropical diseases*, 8(7), p. e3059. doi: 10.1371/journal.pntd.0003059.
- Moncrief, N. D., Kretsinger, R. H. and Goodman, M. (1990) 'Evolution of EF-hand calcium-modulated proteins. I. Relationships based on amino acid sequences.', *Journal of molecular evolution*. Germany, 30(6), pp. 522–562. doi: 10.1007/BF02101108.
- Morris, G. P. (1968) 'Fine structure of the gut epithelium of *Schistosoma mansoni*.', *Experientia*. Switzerland, 24(5), pp. 480–482. doi: 10.1007/BF02144405.
- Moutsoglou, D. M. and Dreskin, S. C. (2016) 'B cells establish, but do not maintain, long-lived murine anti-peanut IgE(a).', *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology*, 46(4), pp. 640–653. doi: 10.1111/cea.12715.
- Mutapi, F. *et al.* (1998) 'Chemotherapy accelerates the development of acquired immune responses to *Schistosoma haematobium* infection.', *The Journal of infectious diseases*, 178(1), pp. 289–93.
- Mutapi, F. *et al.* (1998) 'Chemotherapy Accelerates the Development of Acquired Immune Responses to *Schistosoma haematobium* Infection', *Journal of Infectious Diseases*, 178(1), pp. 289–293. doi: 10.1086/517456.
- Mutapi, F. *et al.* (2005) 'Praziquantel treatment of individuals exposed to *Schistosoma*

haematobium enhances serological recognition of defined parasite antigens.', *The Journal of infectious diseases*, 192(i), pp. 1108–1118. doi: 10.1086/432553.

Mutapi, F. *et al.* (2008) 'Age-Related and Infection Intensity-Related Shifts in Antibody Recognition of Defined Protein Antigens in a Schistosome-Exposed Population', *Journal of Infectious Diseases*, 198(2), pp. 167–175. doi: 10.1086/589511.

Mutapi, F. *et al.* (2011) 'Differential recognition patterns of *Schistosoma haematobium* adult worm antigens by the human antibodies IgA, IgE, IgG1 and IgG4', *Parasite Immunology*, 33(3), pp. 181–192. doi: 10.1111/j.1365-3024.2010.01270.x.

Naus, C. W. *et al.* (1998) 'Human IgE, IgG subclass, and IgM responses to worm and egg antigens in schistosomiasis haematobium: a 12-month study of reinfection in Cameroonian children.', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 26(5), pp. 1142–1147. doi: 10.1086/520310.

Naus, C. W. A. *et al.* (2003) 'The relationship between age, sex, egg-count and specific antibody responses against *Schistosoma mansoni* antigens in a Ugandan fishing community', *Tropical Medicine and International Health*, 8(6), pp. 561–568. doi: 10.1046/j.1365-3156.2003.01056.x.

Neves, L. X. *et al.* (2015) 'What's in SWAP? Abundance of the principal constituents in a soluble extract of *Schistosoma mansoni* revealed by shotgun proteomics', *Parasites & Vectors*, 8(1), p. 337. doi: 10.1186/s13071-015-0943-x.

Norbury, L. J. *et al.* (2018) 'Intranasal delivery of a formulation containing stage-specific recombinant proteins of *Fasciola hepatica* cathepsin L5 and cathepsin B2 triggers an anti-fecundity effect and an adjuvant-mediated reduction in fluke burden in sheep', *Veterinary Parasitology*, 258, pp. 14–23. doi: <https://doi.org/10.1016/j.vetpar.2018.05.008>.

Nowacki, F. *et al.* (2015) 'Protein and small non-coding RNA-enriched extracellular vesicles are released by the pathogenic blood fluke *Schistosoma mansoni*.', *Journal of extracellular vesicles*. Taylor & Francis, 4, p. 28665. doi: 10.3402/jev.v4.28665.

Nowacki, F. C. *et al.* (2015) 'Protein and small non-coding RNA-enriched extracellular vesicles are released by the pathogenic blood fluke *Schistosoma mansoni*', *Journal of Extracellular Vesicles*. Taylor & Francis, 4(1), p. 28665. doi: 10.3402/jev.v4.28665.

Obeng, B. B. *et al.* (2008) 'Application of a circulating-cathodic-antigen (CCA) strip test and real-time PCR, in comparison with microscopy, for the detection of *Schistosoma haematobium* in urine samples from Ghana', *Annals of Tropical Medicine & Parasitology*. Taylor & Francis, 102(7), pp. 625–633. doi: 10.1179/136485908X337490.

Ochodo, E. A. *et al.* (2015) 'Circulating antigen tests and urine reagent strips for diagnosis of

active schistosomiasis in endemic areas.’, *The Cochrane database of systematic reviews*, 3(3), p. CD009579. doi: 10.1002/14651858.CD009579.pub2.

Oettgen, H. C. (2016) ‘Fifty years later: Emerging functions of IgE antibodies in host defense, immune regulation, and allergic diseases’, *Journal of Allergy and Clinical Immunology*. Elsevier, 137(6), pp. 1631–1645. doi: 10.1016/j.jaci.2016.04.009.

Oettle, R. C. and Wilson, S. (2017) ‘The Interdependence between Schistosome Transmission and Protective Immunity’, *Tropical Medicine and Infectious Disease* . doi: 10.3390/tropicalmed2030042.

Orr, R. *et al.* (2012) ‘FhCaBP4: a *Fasciola hepatica* calcium-binding protein with EF-hand and dynein light chain domains’, *Parasitology Research*, 111(4), pp. 1707–1713. doi: 10.1007/s00436-012-3010-y.

Pacífico, L. G. G. *et al.* (2006) ‘Immunization with *Schistosoma mansoni* 22.6 kDa antigen induces partial protection against experimental infection in a recombinant protein form but not as DNA vaccine.’, *Immunobiology*. Netherlands, 211(1–2), pp. 97–104. doi: 10.1016/j.imbio.2005.06.004.

Paterson, S. and Viney, M. E. (2002) ‘Host immune responses are necessary for density dependence in nematode infections’, *Parasitology*. 2003/01/16. Cambridge University Press, 125(3), pp. 283–292. doi: DOI: 10.1017/S0031182002002056.

Pearson, M. S. *et al.* (2015) ‘Of Monkeys and Men: Immunomic Profiling of Sera from Humans and Non-Human Primates Resistant to Schistosomiasis Reveals Novel Potential Vaccine Candidates’, *Frontiers in Immunology*, 6(May). doi: 10.3389/fimmu.2015.00213.

Pelley, J. W. (2012) ‘17 - Protein Synthesis and Degradation’, in Pelley, J. W. B. T.-E. I. R. B. (Second E. (ed.). Philadelphia: W.B. Saunders, pp. 149–160. doi: <https://doi.org/10.1016/B978-0-323-07446-9.00017-9>.

Peralta, J. M. and Cavalcanti, M. G. (2018) ‘Is POC-CCA a truly reliable test for schistosomiasis diagnosis in low endemic areas? The trace results controversy’, *PLoS neglected tropical diseases*. Public Library of Science, 12(11), pp. e0006813–e0006813. doi: 10.1371/journal.pntd.0006813.

Pérez-Sánchez, R. *et al.* (2008) ‘A proteomic approach to the identification of tegumental proteins of male and female *Schistosoma bovis* worms’, *Molecular and Biochemical Parasitology*, 161(2), pp. 112–123. doi: <https://doi.org/10.1016/j.molbiopara.2008.06.011>.

Peters, P. A., Warren, K. S. and Mahmoud, A. A. (1976) ‘Rapid, accurate quantification of schistosome eggs via nuclepore filters.’, *The Journal of parasitology*. United States, 62(1), pp. 154–155.

- Pinot de Moira, A. *et al.* (2010) 'Analysis of complex patterns of human exposure and immunity to Schistosomiasis mansoni: The influence of age, sex, ethnicity and IgE', *PLoS Neglected Tropical Diseases*, 4(9). doi: 10.1371/journal.pntd.0000820.
- Pinot De Moira, A. *et al.* (2007) 'Microgeographical and tribal variations in water contact and Schistosoma mansoni exposure within a Ugandan fishing community', *Tropical Medicine and International Health*, 12(6), pp. 724–735. doi: 10.1111/j.1365-3156.2007.01842.x.
- Pinot De Moira, A. *et al.* (2010) 'Analysis of Complex Patterns of Human Exposure and Immunity to Schistosomiasis mansoni: The Influence of Age, Sex, Ethnicity and IgE'. doi: 10.1371/journal.pntd.0000820.
- Pinot De Moira, A. *et al.* (2013) 'Effects of treatment on IgE responses against parasite allergen-like proteins and immunity to reinfection in childhood schistosome and hookworm coinfections', *Infection and Immunity*, 81(1), pp. 23–32. doi: 10.1128/IAI.00748-12.
- Pokharel, D. R. and Rathaur, S. (2008) 'Purification and characterization of a leucine aminopeptidase from the bovine filarial parasite Setaria cervi.', *Acta tropica*. Netherlands, 106(1), pp. 1–8. doi: 10.1016/j.actatropica.2007.12.009.
- Polman, K. *et al.* (2000) 'Relating serum circulating anodic antigens to faecal egg counts in Schistosoma mansoni infections: a modelling approach.', *Parasitology*, 121 Pt 6, pp. 601–610. doi: 10.1017/S0031182000006843.
- Popiel, I. (1986) 'Male-stimulated female maturation in Schistosoma: A review', *Journal of Chemical Ecology*, 12(8), pp. 1745–1754. doi: 10.1007/BF01022380.
- Popiel, I. and Basch, P. F. (1984) 'Reproductive Development of Female Schistosoma mansoni (Digenea: Schistosomatidae) Following Bisexual Pairing of Worms and Worm Segments', *THE JOURNAL OF EXPERIMENTAL ZOOLOGY*, 232, pp. 141–150.
- Potter, K. N. and Capra, J. D. (1998) 'Diversity, Generation of', in Delves, P. J. (ed.) *Encyclopedia of Immunology*. 2nd edn. Academic Press, pp. 764–770.
- Poulsen, L. K. and Sørensen, T. B. (1993) 'Elimination of viral infection risk from blood samples for allergy testing.', *Allergy*, 48(3), pp. 207–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8506990>.
- Qian, Z. L. and Deelder, A. M. (1983) 'Schistosoma japonicum: immunological characterization and detection of circulating polysaccharide antigens from adult worms.', *Experimental parasitology*. United States, 55(2), pp. 168–178. doi: 10.1016/0014-4894(83)90011-5.
- R Core Team (2015) 'R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria'. Vienna, Austria: R Foundation for Statistical

Computing. Available at: <https://www.r-project.org/>.

Raina, O. K. *et al.* (2011) 'Lack of protective efficacy in buffaloes vaccinated with *Fasciola gigantica* leucine aminopeptidase and peroxiredoxin recombinant proteins.', *Acta tropica*. Netherlands, 118(3), pp. 217–222. doi: 10.1016/j.actatropica.2011.02.008.

Rawlings, N. D. *et al.* (2018) 'The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database', *Nucleic Acids Research*. Oxford University Press, 46(D1), pp. D624–D632. doi: 10.1093/nar/gkx1134.

Restif, O. *et al.* (2012) 'Model-guided fieldwork: Practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics', *Ecology Letters*, 15(10), pp. 1083–1094. doi: 10.1111/j.1461-0248.2012.01836.x.

El Ridi, R. *et al.* (2014) 'Cysteine peptidases as schistosomiasis vaccines with inbuilt adjuvanticity.', *PLoS One*, 9(1), p. e85401. doi: 10.1371/journal.pone.0085401.

Rihet, P. and Demeure, C. E. (1991) 'Evidence for an association between human resistance to *Schistosoma mansoni* and high anti-larval IgE level s', *European journal of immunology*, 21(c), pp. 2679–2686.

Rinaldi, G. *et al.* (2009) 'RNA interference targeting leucine aminopeptidase blocks hatching of *Schistosoma mansoni* eggs.', *Molecular and biochemical parasitology*, 167(2), pp. 118–126. doi: 10.1016/j.molbiopara.2009.05.002.

Ritz, C. *et al.* (2015) 'Dose-response analysis using R', *PLoS One*, 10(12), pp. 1–13. doi: 10.1371/journal.pone.0146021.

Ritz, C. and Strebig, J. C. (2016) 'Package "drc": Analysis of Dose-Response Curves', *R Project*, p. 149. Available at: <https://cran.r-project.org/web/packages/drc/drc.pdf>.

Riveau, G. *et al.* (1998) 'Glutathione S-transferases of 28kDa as major vaccine candidates against schistosomiasis', *Memórias do Instituto Oswaldo Cruz*. scielo, 93, pp. 87–94. Available at: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02761998000700012&nrm=iso.

Riveau, G. *et al.* (2012) 'Safety and immunogenicity of rSh28GST antigen in humans: Phase 1 randomized clinical study of a vaccine candidate against urinary schistosomiasis', *PLoS Neglected Tropical Diseases*, 6(7), pp. 1–8. doi: 10.1371/journal.pntd.0001704.

Riveau, G. *et al.* (2018) 'Safety and efficacy of the rSh28GST urinary schistosomiasis vaccine: A phase 3 randomized, controlled trial in Senegalese children', *PLoS Neglected Tropical Diseases*, 12(12), pp. 1–22. doi: 10.1371/journal.pntd.0006968.

- Roberts, G. O., Gelman, A. and Gilks, W. R. (1997) 'Weak convergence and optimal scaling of random walk Metropolis algorithms', *Annals of Applied Probability*, 7(1), pp. 110–120. doi: 10.1214/aoap/1034625254.
- Robichon, C. *et al.* (2011) 'Engineering Escherichia coli BL21 (DE3) Derivative Strains To Minimize E . coli Protein Contamination after Purification by Immobilized Metal Affinity Chromatography †', 77(13), pp. 4634–4646. doi: 10.1128/AEM.00119-11.
- Rosenberg, H. F., Dyer, K. D. and Foster, P. S. (2013) 'Eosinophils: changing perspectives in health and disease', *Nature Reviews Immunology*, 13(1), pp. 9–22. doi: 10.1038/nri3341.
- Ruth Lawson, J. and Wilson, R. A. (1980) 'The survival of the cercariae of *Schistosoma mansoni* in relation to water temperature and glycogen utilization', *Parasitology*. 2009/04/06. Cambridge University Press, 81(2), pp. 337–348. doi: DOI: 10.1017/S0031182000056079.
- Samoil, V. *et al.* (2018) 'Vesicle-based secretion in schistosomes: Analysis of protein and microRNA (miRNA) content of exosome-like vesicles derived from *Schistosoma mansoni*', *Scientific Reports*. Springer US, 8(1), pp. 1–16. doi: 10.1038/s41598-018-21587-4.
- Samuelson, J. C. and Caulfield, J. P. (1982) 'Loss of covalently labeled glycoproteins and glycolipids from the surface of newly transformed schistosomula of *Schistosoma mansoni*.', *The Journal of cell biology*, 94(2), pp. 363–369. doi: 10.1083/jcb.94.2.363.
- Sanneh, B. *et al.* (2017) 'Field evaluation of a schistosome circulating cathodic antigen rapid test kit at point-of-care for mapping of schistosomiasis endemic districts in The Gambia', *PloS One*. Public Library of Science, 12(8), pp. e0182003–e0182003. doi: 10.1371/journal.pone.0182003.
- Santiago, M. L. *et al.* (1998) 'Identification of the *Schistosoma japonicum* 22.6-kDa antigen as a major target of the human IgE response: similarity of IgE-binding epitopes to allergen peptides.', *International archives of allergy and immunology*. Switzerland, 117(2), pp. 94–104. doi: 10.1159/000023995.
- Satayathum, S. A. *et al.* (2006) 'Factors affecting infection or reinfection with *Schistosoma haematobium* in coastal Kenya: survival analysis during a nine-year, school-based treatment program.', *The American journal of tropical medicine and hygiene*, 75(1), pp. 83–92.
- Satti, M. Z. *et al.* (1996) 'Clinical, parasitological and immunological features of canal cleaners hyper-exposed to *Schistosoma mansoni* in the Sudan', *Clinical and experimental immunology*. Blackwell Science Inc, 104(3), pp. 426–431. doi: 10.1046/j.1365-2249.1996.00051.x.
- Saunders, S. P. *et al.* (2019) 'Non-classical B Cell Memory of Allergic IgE Responses', *Frontiers in immunology*, 10(April), p. 715. doi: 10.3389/fimmu.2019.00715.
- Schad, G. A. and Anderson, R. M. (1985) 'Predisposition to hookworm infection in humans.',

Science (New York, N.Y.). United States, 228(4707), pp. 1537–1540. doi: 10.1126/science.4012307.

Senawong, G. *et al.* (2012) 'Cloning, expression, and characterization of a novel *Opisthorchis viverrini* calcium-binding EF-hand protein', *Parasitology International*, 61(1), pp. 94–100. doi: <https://doi.org/10.1016/j.parint.2011.07.012>.

Shaw, J. R., Marshall, I. and Erasmus, D. A. (1977) 'Schistosoma mansoni: in vitro stimulation of vitelline cell development by extracts of male worms.', *Experimental parasitology*. United States, 42(1), pp. 14–20. doi: 10.1016/0014-4894(77)90056-x.

Simes, R. J. (1986) 'Biometrika Trust An Improved Bonferroni Procedure for Multiple Tests of Significance Author(s) An improved Bonferroni procedure for multiple tests of significance', *Biometrika*, 73(3), pp. 751–754. Available at: <http://www2.math.uu.se/~thulin/mm/HW2-Simes.pdf>⁰<http://www.jstor.org/stable/2336545>⁰<http://www.jstor.org/>⁰[http://www.jstor.org/](http://www.jstor.org/action/showPublisher?publisherCode=bio.)⁰<http://www.jstor.org/>

Skelly, P. J. *et al.* (2014) 'Schistosome feeding and regurgitation.', *PLoS pathogens*, 10(8), p. e1004246. doi: 10.1371/journal.ppat.1004246.

Skelly, P. J., Stein, L. D. and Shoemaker, C. B. (1993) 'Expression of *Schistosoma mansoni* genes involved in anaerobic and oxidative glucose metabolism during the cercaria to adult transformation', *Molecular and Biochemical Parasitology*, 60(1), pp. 93–104. doi: [https://doi.org/10.1016/0166-6851\(93\)90032-S](https://doi.org/10.1016/0166-6851(93)90032-S).

Smithers, S. R. and Terry, R. J. (1967) 'Resistance to experimental infection with *Schistosoma mansoni* in rhesus monkeys induced by the transfer of adult worms.', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 61(4), pp. 517–33.

Smithers, S. R. and Terry, R. J. (1969) 'Immunity in Schistosomiasis', *Annals of the New York Academy of Sciences*, 160(2), pp. 826–40. Available at: <https://doi.org/10.1111/j.1749-6632.1969.tb15904.x>.

Sotillo, J. *et al.* (2016) 'Extracellular vesicles secreted by *Schistosoma mansoni* contain protein vaccine candidates', *International Journal for Parasitology*. Australian Society for Parasitology Inc., 46(1), pp. 1–5. doi: 10.1016/j.ijpara.2015.09.002.

Sotillo, J. *et al.* (2019) 'In-depth proteomic characterization of *Schistosoma haematobium*: Towards the development of new tools for elimination', *PLoS Neglected Tropical Diseases*. Public Library of Science, 13(5), p. e0007362. Available at: <https://doi.org/10.1371/journal.pntd.0007362>.

Sow, S. *et al.* (2011) 'The contribution of water contact behavior to the high *Schistosoma mansoni* Infection rates observed in the Senegal River Basin', *BMC Infectious Diseases*, 11. doi:

10.1186/1471-2334-11-198.

Stavnezer, J., Guikema, J. E. J. and Schrader, C. E. (2008) 'Mechanism and regulation of class switch recombination', *Annual review of immunology*, 26, pp. 261–292. doi: 10.1146/annurev.immunol.26.021607.090248.

Steinauer, M. L. (2009) 'The sex lives of parasites: Investigating the mating system and mechanisms of sexual selection of the human pathogen *Schistosoma mansoni*', *International Journal for Parasitology*, 39(10), pp. 1157–1163. doi: 10.1016/j.ijpara.2009.02.019.

Stelma, F. F. *et al.* (1993) 'Epidemiology of *Schistosoma mansoni* infection in a recently exposed community in Northern Senegal', *American Journal of Tropical Medicine and Hygiene*, 49(6), pp. 701–706. doi: 10.4269/ajtmh.1993.49.701.

Stothard, J. R. *et al.* (2011) 'Closing the praziquantel treatment gap: new steps in epidemiological monitoring and control of schistosomiasis in African infants and preschool-aged children', *Parasitology*, 138(12), pp. 1593–1606. doi: 10.1017/S0031182011001235.

Strait, R. T., Morris, S. C. and Finkelman, F. D. (2006) 'IgG-blocking antibodies inhibit IgE-mediated anaphylaxis in vivo through both antigen interception and Fc gamma RIIB cross-linking.', *The Journal of clinical investigation*, 116(3), pp. 833–841. doi: 10.1172/JCI25575.

Stroehlein, A. J. *et al.* (2019) 'High-quality *Schistosoma haematobium* genome achieved by single-molecule and long-range sequencing', *GigaScience*, 8(9). doi: 10.1093/gigascience/giz108.

Sturrock, R. F. (2001) 'The schistosomes and their intermediate hosts', in Mahmoud, A. A. F. (ed.) *Schistosomiasis*. London: Imperial College Press.

Subpipattana, P., Grams, R. and Vichasri-Grams, S. (2012) 'Analysis of a calcium-binding EF-hand protein family in *Fasciola gigantica*', *Experimental Parasitology*, 130(4), pp. 364–373. doi: <https://doi.org/10.1016/j.exppara.2012.02.005>.

Supek, F. *et al.* (2011) 'REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms', *PLoS One*. Public Library of Science, 6(7), p. e21800. Available at: <https://doi.org/10.1371/journal.pone.0021800>.

Talavera, G. and Castresana, J. (2007) 'Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments.', *Systematic biology*. England, 56(4), pp. 564–577. doi: 10.1080/10635150701472164.

Tan, G. *et al.* (2015) 'Current Methods for Automated Filtering of Multiple Sequence Alignments Frequently Worsen Single-Gene Phylogenetic Inference.', *Systematic biology*, 64(5), pp. 778–791. doi: 10.1093/sysbio/syv033.

- Tebeje, B. M. *et al.* (2016) 'Schistosomiasis vaccines: where do we stand?', *Parasites & Vectors*, 9(1), p. 528. doi: 10.1186/s13071-016-1799-4.
- Thomas, C. M. *et al.* (2014) 'Comparative biochemical analysis of three members of the *Schistosoma mansoni* TAL family: Differences in ion and drug binding properties', *Biochimie*. Elsevier B.V, 108, pp. 40–47. doi: 10.1016/j.biochi.2014.10.015.
- Tinghino, R. *et al.* (2002) 'Molecular, structural, and immunologic relationships between different families of recombinant calcium-binding pollen allergens', *Journal of Allergy and Clinical Immunology*. Elsevier, 109(2), pp. 314–320. doi: 10.1067/mai.2002.121528.
- Trowsdale, J. and Parham, P. (2004) 'Mini-review: defense strategies and immunity-related genes.', *European journal of immunology*. Germany, 34(1), pp. 7–17. doi: 10.1002/eji.200324693.
- Truscott, J. E. *et al.* (2017) 'A comparison of two mathematical models of the impact of mass drug administration on the transmission and control of schistosomiasis', *Epidemics*. Elsevier B.V., 18, pp. 29–37. doi: 10.1016/j.epidem.2017.02.003.
- Turner, J. D. *et al.* (2005) 'Allergen-specific IgE and IgG4 are markers of resistance and susceptibility in a human intestinal nematode infection', *Microbes and Infection*, 7(7–8), pp. 990–996. doi: 10.1016/j.micinf.2005.03.036.
- Valenta, R. *et al.* (1998) 'Calcium-binding allergens: from plants to man', *International Archives of Allergy and Immunology*, 117(3), pp. 160–166.
- Vennervald, B. J. *et al.* (2004) 'Detailed clinical and ultrasound examination of children and adolescents in a *Schistosoma mansoni* endemic area in Kenya: hepatosplenic disease in the absence of portal fibrosis.', *Tropical medicine & international health : TM & IH*. England, 9(4), pp. 461–470. doi: 10.1111/j.1365-3156.2004.01215.x.
- Vennervald, B. J. and Polman, K. (2009) 'Helminths and malignancy.', *Parasite immunology*. England, 31(11), pp. 686–696. doi: 10.1111/j.1365-3024.2009.01163.x.
- Verdino, P. *et al.* (2008) 'Three-Dimensional Structure of the Cross-Reactive Pollen Allergen Che a 3: Visualizing Cross-Reactivity on the Molecular Surfaces of Weed, Grass, and Tree Pollen Allergens', *The Journal of Immunology*, 180(4), pp. 2313 LP – 2321. doi: 10.4049/jimmunol.180.4.2313.
- Vieira, P. and Rajewsky, K. (1988) 'The half-lives of serum immunoglobulins in adult mice', *European Journal of Immunology*, 18(2), pp. 313–316. doi: 10.1002/eji.1830180221.
- de Vlas, S. J. *et al.* (1992) 'A model for variations in single and repeated egg counts in *Schistosoma mansoni* infections.', *Parasitology*, 104(3), pp. 451–60. doi: 10.1017/s003118200006371x.

- Walter, K. *et al.* (2006) 'Increased human IgE induced by killing *Schistosoma mansoni* in vivo is associated with pretreatment Th2 cytokine responsiveness to worm antigens.', *Journal of immunology (Baltimore, Md. : 1950)*, 177(8), pp. 5490–8. doi: 10.1002/jimr.10000 [pii].
- Wang, J. *et al.* (2015) 'Intake of erythrocytes required for reproductive development of female *Schistosoma japonicum*', *PLoS One*, 10(5), pp. 1–17. doi: 10.1371/journal.pone.0126822.
- Warren, K. S. (1973) 'Regulation of the prevalence and intensity of schistosomiasis in man: immunology or ecology?', *The Journal of infectious diseases*. United States, 127(5), pp. 595–609. doi: 10.1093/infdis/127.5.595.
- Webster, M. *et al.* (1996) 'Human Immunoglobulin E Responses to a Recombinant 22.6-Kilodalton Antigen from *Schistosoma mansoni* Adult Worms Are Associated with Low Intensities of Reinfection after Treatment', *INFECTION AND IMMUNITY*, 64(10), pp. 4042–4046.
- Weiss, W., Weiland, F. and Görg, A. (2009) 'Protein detection and quantitation technologies for gel-based proteome analysis.', *Methods in molecular biology (Clifton, N.J.)*. United States, 564, pp. 59–82. doi: 10.1007/978-1-60761-157-8_4.
- Whitfield, P. J. *et al.* (2003) 'Age-dependent survival and infectivity of *Schistosoma mansoni* cercariae', *Parasitology*. 2003/10/09. Cambridge University Press, 127(1), pp. 29–35. doi: DOI: 10.1017/S0031182003003263.
- Wiegand, R. E. *et al.* (2017) 'A Persistent Hotspot of *Schistosoma mansoni* Infection in a Five-Year Randomized Trial of Praziquantel Preventative Chemotherapy Strategies.', *The Journal of infectious diseases*, 216(11), pp. 1425–1433. doi: 10.1093/infdis/jix496.
- Wilkins, H. A. *et al.* (1984) 'Dynamics of *Schistosoma haematobium* infection in a Gambian community. III. Acquisition and loss of infection', *Transactions of the Royal Society of Tropical Medicine & Hygiene*, 78(2), pp. 227–232.
- Williams, C. B. (1937) 'The use of logarithms in the interpretation of certain entomological problems', *The Annals of Applied Biology*, 24, pp. 404–414.
- Wilson, R. A. *et al.* (2007) 'Oming in on schistosomes: prospects and limitations for post-genomics', *Trends in Parasitology*, 23(1), pp. 14–20. doi: 10.1016/j.pt.2006.10.002.
- Wilson, S. *et al.* (2013) 'Rapidly boosted Plasma IL-5 induced by treatment of human Schistosomiasis haematobium is dependent on antigen dose, IgE and eosinophils.', *PLoS neglected tropical diseases*, 7(3), p. e2149. doi: 10.1371/journal.pntd.0002149.
- Wilson, S. *et al.* (2014) 'Human *Schistosoma haematobium* antifecundity immunity is dependent on transmission intensity and associated with immunoglobulin G1 to worm-derived antigens.',

- The Journal of infectious diseases*, 210(12), pp. 2009–16. doi: 10.1093/infdis/jiu374.
- Woof, J. M. and Kerr, M. A. (2004) 'IgA function--variations on a theme.', *Immunology*, 113(2), pp. 175–177. doi: 10.1111/j.1365-2567.2004.01958.x.
- Woolhouse, Mark E. J.; Hasibeder, G.; Chandiwana, S. K. (1996) 'On estimating the basic reproduction number for *Schistosoma haematobium*', *Tropical Medicine and International Health*, 1(4), pp. 456–463.
- Woolhouse, M. E. (1991) 'On the application of mathematical models of schistosome transmission dynamics. I. Natural transmission', *Acta Trop*, 49(4), pp. 241–270. doi: 10.1016/0001-706X(91)90077-W.
- Woolhouse, M. E. *et al.* (1998) 'Heterogeneities in schistosome transmission dynamics and control.', *Parasitology*, 117 (Pt 5, pp. 475–482. doi: 10.1017/S003118209800331X.
- Woolhouse, M. E. . (1998) 'Patterns in Parasite Epidemiology: The Peak Shift', *Parasitology Today*, 14(10), pp. 428–434. doi: 10.1016/S0169-4758(98)01318-0.
- Woolhouse, M. E. and Hagan, P. (1999) 'Seeking the ghost of worms past', *Nat Med*, 5(11), pp. 1225–1227. doi: 10.1038/15169.
- Woolhouse, M. E. J. *et al.* (1991) 'Acquired immunity and epidemiology of *Schistosoma haematobium*', *Nature*, 351(June), pp. 757–759.
- Woolhouse, M. E. J. (1991) 'On the application of mathematical models of schistosome transmission dynamics. I. Natural transmission', *Acta Tropica*, 49(4), pp. 241–270. doi: [https://doi.org/10.1016/0001-706X\(91\)90077-W](https://doi.org/10.1016/0001-706X(91)90077-W).
- World Health Organisation (2017) *Crossing the billion. Lymphatic filariasis, onchocerciasis, schistosomiasis, soil-transmitted helminthiasis and trachoma: preventive chemotherapy for neglected tropical diseases*. Geneva: World Health Organisation. doi: Licence: CC BY-NC-SA 3.0 IGO.
- World Health Organization (2006) *Preventive chemotherapy in human helminthiasis Preventive chemotherapy in human helminthiasis*.
- World Health Organization (2012) *Accelerating Work to Overcome the Global Impact of Neglected Tropical Diseases: A Roadmap for Implementation* (World Health Organizaiton, Geneva, Switzerland, 2012). Geneva. Available at: http://www.who.int/neglected_diseases/NTD_RoadMap_2012_Fullversion.pdf.
- World Health Organization (2020a) *Ending the Neglect to Attain the Sustainable Development Goals - A road map for neglected tropical diseases 2021–2030*. Available at:

<http://apps.who.int/bookorders>.

World Health Organization (2020b) Schistosomiasis: Key Facts. Available at: <https://www.who.int/news-room/fact-sheets/detail/schistosomiasis> (Accessed: 3 August 2020).

Wu, Zhenyu et al. (2018) 'Extracellular Vesicle-Mediated Communication Within Host-Parasite Interactions.', *Frontiers in immunology*, 9, p. 3066. doi: 10.3389/fimmu.2018.03066.

Xiao, J. ying et al. (2013) 'Molecular Cloning and Characterization of Taurocyamine Kinase from *Clonorchis sinensis*: A Candidate Chemotherapeutic Target', *PLoS Neglected Tropical Diseases*, 7(11), pp. 1–12. doi: 10.1371/journal.pntd.0002548.

Xiong, H. et al. (2012) 'Sequential class switching is required for the generation of high affinity IgE antibodies.', *The Journal of experimental medicine*, 209(2), pp. 353–364. doi: 10.1084/jem.20111941.

Yang, Z., Sullivan, B. M. and Allen, C. D. C. (2012) 'Fluorescent In Vivo Detection Reveals that IgE+ B Cells Are Restrained by an Intrinsic Cell Fate Predisposition', *Immunity*, 36(5), pp. 857–872. doi: <https://doi.org/10.1016/j.immuni.2012.02.009>.

You, H. et al. (2015) 'Suppression of the Insulin Receptors in Adult *Schistosoma japonicum* Impacts on Parasite Growth and Development: Further Evidence of Vaccine Potential', *PLoS Neglected Tropical Diseases*. Public Library of Science, 9(5), p. e0003730. Available at: <https://doi.org/10.1371/journal.pntd.0003730>.

Young, N. D. et al. (2012) 'Whole-genome sequence of *Schistosoma haematobium*', *Nature Genetics*, 44(2), pp. 221–225. doi: 10.1038/ng.1065.

Appendix 1

Buffer compositions

A1.1 Buffer for ShSWA sample preparation

Rehydration solution

8 M Urea

4 % Chaps

0.2 % DTT

1 % IPG buffer 3-10NL (GE Healthcare Life Science)

(add DTT and IPG just before use)

0.002% bromophenol blue

A1.2 Buffers for SDS-PAGE one-dimensional gel electrophoresis

Sample buffer

NuPAGE™ lithium dodecyl sulfate (LDS) Sample Buffer (4x)

β-mercaptoethanol (ME)

Colloidal Coomassie blue stain

Dissolve 1 g of Coomassie Brilliant Blue (Bio-Rad) in 1 L of the following solution:

Methanol (50 % [v/v])

Glacial acetic acid (10 % [v/v])

H₂O (40 %)

filter through Whatman filter paper

Coomassie de-stain

10 % acetic acid

50 % methanol

40 % H₂O

A1.3 Buffers for two-dimensional gel electrophoresis

Chaps/Thiourea buffer

6 M Urea

2 M Thiourea

4 % Chaps

5 mM Mg Acetate

10 mM Tris pH 8.5

2 x sample buffer Chaps/Urea

8 M Urea

4 % Chaps

2 % DTT
2 % IPG buffer 3-10NL GE Healthcare Life Science
(add DTT and IPG just before use)

SDS-PAGE gel

12% gel (30 ml for 13 cm gel with 1 mm spacer)
9.9 ml dH₂O
12 ml 30 % acrylamide mix
7.5 ml 1.5 M Tris (pH 8.8)
Filter solution through 0.45 µm filter
300 µl 10 % SDS
300 µl 10 % ammonium persulfate
12 µl TEMED

Equilibration buffer-reducing

6 M Urea
75 mM Tris pH 8.8
30 % Glycerol
2 % SDS
1 % DTT (add fresh just before use)

Equilibration buffer-alkylation

6 M Urea
75 mM Tris pH 8.8
30 % Glycerol
2 % SDS
2.5 % iodoacetamide (add fresh just before use)

Sealing agarose

0.5 % agarose
0.002 % Bromophenol blue
dissolved in SDS running buffer

Running Buffer

1x TGS (Tris/Glycine/SDS) (Bio-Rad)

Colloidal Coomassie fix solution

40 % methanol
7 % acetic acid

Silver staining solution (1 Litre)

2 g AgNO₃
0.2 g or 540 µL of a 37% w/v solution Formaldehyde 37%

Silver development solution

3 % Na₂CO₃

0.0005 % $\text{Na}_2\text{S}_2\text{O}_3$

0.05 % Formaldehyde (add fresh before use: 135 μl /100ml)

A1.4 Buffers for Western blot

Transfer buffer

20x NuPage transfer buffer

20 % Methanol

A1.5 Buffers for trypsin digest

Silver destain solution (freshly prepared)

98.7 mg potassium ferricyanide

10 ml HPLC water

248.2 mg sodium thiosulphate

10 ml HPLC water

Digest reduction buffer (freshly prepared)

10 mM DTT in 100 mM NH_4HCO_3

1.5 mg dithiothreitol in 1 ml 100 mM NH_4HCO_3

15.4 mg in 10 ml 100 mM NH_4HCO_3

Digest alkylation buffer (freshly prepared)

55 mM iodoacetamide in 100 mM NH_4HCO_3

10 mg iodoacetamide in 1 ml 100 mM NH_4HCO_3

102 mg in 10 ml 100 mM NH_4HCO_3

Trypsin mixture

20 μg trypsin (1x vial) in 200 μl of 50 mM NH_4HCO_3

A1.6 Buffers for mass spectrometry

Solvent A

0.1 % formic acid in dH_2O

Solvent B

80 % acetonitrile

20 % dH_2O

0.1 % formic acid

A1.7 Buffers for protein production

2x YT Medium

31 g/L 2x YT Microbial Medium Powder (Sigma) was added to dH_2O .
sterilised by autoclaving at 121°C for 15 min.

Ampicillin selective plates

1.5 % Agar in 2x YT
+ 100 µg/ml ampicillin stock

A1.8 Anti-fecundity buffers**TBE buffer (Tris/Borate/EDTA)**

54 g of Tris base
27.5 g of boric acid
20 ml of 0.5 M EDTA (pH 8.0)
Adjust pH to 8.3 by HCl

DNA Loading Buffer

0.25 % w/v Bromophenol Blue
in 30% v/v Glycerol

LB Agar plates

Carbenicillin (Cb) 50mg/ml, 1:1000
20 % X-Gal [in DMF], 1:1000
IPTG 1 M stock; 1:1000
In warm agar.

Overnight Express™ Instant Terrific Broth (TBONX) (300ml)

18 g media
3 ml glycerol

NPI-10 (Binding/lysis buffer for native conditions, 1 L):

50 mM NaH₂PO₄ 6.90 g NaH₂PO₄·H₂O (MW 137.99 g/mol)
300 mM NaCl 17.54 g NaCl (MW 58.44 g/mol)
10 mM imidazole 0.68 g imidazole (MW 68.08 g/mol)
0.05 % v/v Tween 20
Adjust pH to 8.0 using NaOH and sterile filter (0.2 or 0.45 µm).

NPI-20 (Wash buffer for native conditions, 1 L)

50 mM NaH₂PO₄ 6.90 g NaH₂PO₄·H₂O (MW 137.99 g/mol)
300 mM NaCl 17.54 g NaCl (MW 58.44 g/mol)
20 mM imidazole 1.36 g imidazole (MW 68.08 g/mol)
0.05 % v/v Tween 20
Adjust pH to 8.0 using NaOH and sterile filter (0.2 or 0.45 µm).

NPI-75 (Wash buffer for native conditions, 1 L)

50 mM NaH₂PO₄ 6.90 g NaH₂PO₄·H₂O (MW 137.99 g/mol)
300 mM NaCl 17.54 g NaCl (MW 58.44 g/mol)
75 mM imidazole 1.36 g imidazole (MW 68.08 g/mol)
0.05 % v/v Tween 20
Adjust pH to 8.0 using NaOH and sterile filter (0.2 or 0.45 µm).

NPI-150 (Wash buffer for native conditions, 1 L)

50 mM NaH_2PO_4 6.90 g $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (MW 137.99 g/mol)

300 mM NaCl 17.54 g NaCl (MW 58.44 g/mol)

75 mM imidazole 1.36 g imidazole (MW 68.08 g/mol)

0.05 % v/v Tween 20

Adjust pH to 8.0 using NaOH and sterile filter (0.2 or 0.45 μm).

NPI-250 (Elution buffer for native conditions, 1 L)

50 mM NaH_2PO_4 6.90 g $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (MW 137.99 g/mol)

300 mM NaCl 17.54 g NaCl (MW 58.44 g/mol)

250 mM imidazole 17.0 g imidazole (MW 68.08 g/mol)

0.05 % v/v Tween 20

Adjust pH to 8.0 using NaOH and sterile filter (0.2 or 0.45 μm).

DNAse Type I stock solution (100,000 Kunitz units/ml)

5 g ($\geq 2,000,000$ Kunitz units) bottle of DNAse I (Sigma DN-25)

add 5 ml of UHQ Sterile Water.

Aliquot into 100 μl aliquots, store at -20°C

Add one aliquot to the 25 ml of Buffer NPI-10-Tween buffer to be used for lysis (final concentration 400 Kunitz units/ml).

SDS PAGE gel loading buffer

100 mM Tris, pH 6.8

4 % w/v SDS

0.2 % w/v Bromophenol blue

10 % v/v β - mercaptoethanol

20 % v/v glycerol

A1.9 Buffers for ELISA**Coating buffer (pH 9.6)**

0.795 g Na_2CO_3

1.465 g NaHCO_3

dissolved in 500ml of distilled water

Wash buffer

PBS + 0.03 % Tween 20

Incubation buffer

PBS + 0.05 % Tween 20

0.1 % Marvel

Blocking buffer

PBS + 0.05 % Tween 20

1 % Marvel (skimmed milk powder, Premier International Foods (UK) Limited, Spalding, UK)

OPD substrate

10 ml distilled water

4.96 ml 0.1 M citric acid

5.04 ml 0.2 M Na_2HPO_4

20 μl 30 % H_2O_2

0.2 ml OPD stock (1 x 30 mg OPD tablet (Sigma-Aldrich Company Ltd) in 3 ml dH_2O)

Appendix 2

Anti-fecundity target PCR primers

Protein	Forward primer (5'→3')*	Reverse primer (5'→3')†
ENA_1 KGB37341 Alpha-adducin	TCTGGTAACCAATAAGCAAGGCG	GTGTTTCTTCTTTCTGGAAAAATG
ENA_1 KGB33608 ATP-dependent RNA helicase eIF4A	CAACGCATTGGTTCCTATCTAAC	CAGAAAAATCAACAATGTCATCAGG
ENA_1 KGB34274 COP9 signalosome complex subunit 4	GGCGAGAAATCTAAGTACTATTCTTTC	CACTGGTGGATCTATAAACCATTTCTG
ENA_1 KGB40482 Dihydropyrimidinase-related protein 4	TATGAAAAAGTCCCGTCAAAACAG	CCATAAGGCGTTTGACGCCCC
ENA KGB39377 Fragile X mental retardation protein 1-like protein	TGTCACTGTACAATTAAAGTCGG	ACTTGCTCTTGCCCTTATTGTGG
ENA KGB37895 Glutamine tRNA ligase	TTAGTGAATGGTGACAGTGATG	AATTGTTTTACCAGGATCGGC
ENA_1 KGB40610 Glycogenin-1	CTAGATATTATGGGGTCTGTATTG	TTTTTTTGATTTTTTGATTCCAGAAGATTTTG
ENA_1 KGB36367 hypothetical protein	AACAATTATACTGGACAATATGGATAC	CGACGATTCCCGCATGTGTATCCTC
ENA_1 KGB36836 hypothetical protein	ACAACTGATTTACCTGAAGTAACTGG	GTAGCCAAAGTGGATAACCGCATTTGG
ENAKGB31588_1 partial hypothetical protein	AGTGTGTGTGTTGTGGATAGTG	ATGATCAGGGTTTCACTGTCTTC
ENA_1 KGB38612 Kinesin light chain	TCTGTACAGTCTGGCAAAGTCAG	ATGATTCAAACCTGTTGTTTAAAGATG
ENA_1 KGB31407 L-lactate dehydrogenase A chain	GTCGGAATGGCTGCAGCATTC	CCACTGGATTTTTCGTTATAACTTC
EMBOSS_001_1 Lethal(2) giant larvae protein	GCTCACCGGGCATTGTTTTC	TATTCCAGTGGCAATTGTGATCG
ENA_1 KGB36326 Lysosomal protective protein	ATGTCTGTGGTTCATCTTAACACTC	AAAAGGTTTGAAGTCAATGAACAGAC
ENA_1 KGB35184 Major egg antigen	GACAACTACAGAACGGAAATAAACATAC	AACTGTGCATAATAATGGCGCTTCAAC

KGB31714.1.1 partial Major vault protein	AGTCCCCATCACTCATCACG	CGTTTCAATAGATAAGGCTTCTGC
ENA_1 KGB33726 NADP-dependent malic enzyme	ACTTTGTTTCAGGGAAC TGCAATTC	TACAATGGTACATTGTTGTGTGGGAG
A_05241_1 gene=A_05241 Peroxiredoxin -1	TTATTACCAAATCAGCCTGCTC	GTGCACAGACGAGAGAGTAGG
ENA KGB41475 Phosphoglucomutase	TTGACAAATCGATATAAAGCCAAACC	TGTGATGACATTTTGGAGCTACTC
ENA KGB36476 Phosphoglycerate kinase	TACATAGGAAATAGTTTATTGATGC	GTGAGCATCCGTAAAGTGCAAC
ENA KGB33558 Phosphorylase b kinase gamma catalytic chain, skeletal muscle/heart isoform	CGTTGGCGGTCTTTTCGTTTG	ATAATCAATTCTAAATCGGTATCCTG
KGB39067.1.1 Plastin	TCAATAAATCAATTAAC TCAAGAACAAATG	TTCAGTTCATTAAATTTTATTGTCTGTC
ENA KGB33720 Prolyl endopeptidase	AAAACGTTGTCAGGACCCAC	TTTATTTTGCTTGTAACCTTACTGATTTTG
ENA_1 KGB38962 Protein SET	GCGACCCCAACCTAAAGTACCC	AACATCCGCTACACCACTTTC
KGB39995.1.1 Putative aminopeptidase W07G4.4	TCTACCAATGTTGTTACTCCCG	CTTAAAACTTAGTCGTGGTAAAAATATATC
ENA KGB41647 putative protein disulfide-isomerase ER-60	ATGTCTCGAGCGGGACCTG	TAGATCACTTTTCTTTGGATTGCC
ENA_1 KGB36581 putative UDP-glucose 4-epimerase	CAGAAAGGCAATAAAGGAGTTG	ATTATTAGAATGTGTTGAAGTGAG
ENA KGB37900 KGB37900.1 putative serine--tRNA ligase, cytoplasmic	ACCATGTTCTTTACACAGGAAAC	ATCCGATTGTTTCTTCTTTGGTTTTC
ENA KGB40310 KGB40310.1 Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	GCAGGGGATGATAAGGTGAAC	TAATTTCCCTTTTCGCTTTTAGCCCC
ENA KGB39342 KGB39342.1 Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial	GAGTTGTTAAAAAAGTACGAAATACC	ATAAGGCAGTTCGAATTTTACATTTATAG
ENA KGB33452 KGB33452.1 T-complex protein 1 subunit eta	GATAAGCTTATAATAGACGATAAAGG	ATGCCTCATCCCCAGGAGGCC
ENA KGB37298 KGB37298.1 Taurocyamine kinase	CAGGTTGAAAGCCTCCACAATC	TAGTCCCTTTTCAATTTCAATCATTTTAG
ENA KGB41386 KGB41386.1 Thimet oligopeptidase	TCTGTTGCTTCAGGTTGTCGTC	TGGTAAATTAACATTTTAATAATTGAAG
ENA KGB34110 KGB34110.1 Thioredoxin domain-containing protein C06A6.5	GATTTATTTAAAAAAGTCGCTGC	GAATTCATCATGTATATAATTGTATATC
A_04057 gene=A_04057 threonyl tRNA synthetase	GCCAGTAATCACCAGTGTGG	AAATTTCTTGTCTGCTCCAGC

B_00064 gene=B_00064 transaldose	GCTATTAAAAAGTATTACCAATTGATTC	TTGTTGTTCCATACGTGTTTAAATTAATC
ENA KGB38058 KGB38058.1 Tropomodulin	TCGAATAAGTTATTATTGGTAAAGC	ATTTTATTACGCCGTAGTTGCC
ENA KGB39286 KGB39286.1 Troponin T	TTTGGTCTTGTTGTCCTTATTATTTC	TTCAACAGGAACCTCAGTTGCAG
ENA KGB35524 KGB35524.1 Transketolase	GCGTACGTCGGGAAATATATTG	TTTGCCAAAGTAATGCTTTTACAGC
ENA KGB36313 KGB36313.1 Annexin A7	GATCCGAATCTTTTCAGATCTTTTCAG	TGCTCCGATAAGCGCCAGTAATAG
ENA KGB33031 KGB33031.1 Glycogen phosphorylase, liver form	TTGTCTGATGCTGAGAAAAAGGAAG	TTTCTTTTCTATAGCTGGTTCAAATGG
ENA KGB33009 KGB33009.1 Cystathionine beta-synthase	ACGCGCATTTTGTTGTACATTTTG	CTGGAAAGGTTTCACCTTACTTTC
ENA KGB31695 KGB31695.1 partial hypothetical protein	GAACGTGCGAAACGTACATTG	GTCGACCTCCTCAATCGTTG
ENA KGB33210 KGB33210.1 Tropomyosin	GATGGAATTAAAAAGAAAAATGATTGC	GATCGAATGAAATTCATTGATCATTG
ENA KGB38978 partial cAMP-dependent protein kinase type II regulatory subunit	GCTTCAGTATTAAACATCACTTGC	TCCAAGTAATTCAGTTAATTGTTTTC
ENA KGB37594 Dipeptidyl peptidase 2	GTGCATCAAAACCCCTGTTTTC	TAAATCAAGCCATGAACGAATTG
ENA_1 KGB39908 Elongation factor 2	GATTTTAAAGCGGAATATCCGC	TAGTTTATCCAAAGAAGTTATCTAAGG
ENA KGB38992 Glutamine synthetase	CCGTTAACATCACGTTTCACC	AAGACATGTTGTTTCTACTAACATTTC
ENA KGB38015 Adenylylcyclase-associated protein 1	GAAAGACTTGATTCTCTACTTTGTC	TGAATTCTGTTTCATCTGCAAAATC

* The sequence AAGTTCTGTTTCAGGGCCCG, complementary to the pOPIN vector, was added to the 5' end of the forward primer

† The sequence ATGGTCTAGAAAGCTTTA, complementary to the pOPIN vector, was added to the 5' end of the reverse primer

Appendix 3

R Code for *S. mansoni* age- and sex-structured transmission model

A3.1 Transmission model code

```
# Loading In Required Packages
library(deSolve); library(ggplot2); library(gridExtra); library(tictoc)

## immunity function (depends on shape parameters & constant linking number of dying worms)
## therefore dynamically updated during numerical integration
#a1=c,a2=d,a3=b,a4=e
imm.a<-function(parameters, immuno.stimulus) {
  with(as.list(c(parameters)), {
    j = seq(1,na)
    a = (j-1)*amax/na + amax/(2*na)
    da <- amax/na
    p.estab <- a1+(a2-a1)/(1+exp(a3*(log(immuno.stimulus)-log(a4))))
    #p.estab <- exp(-a1*immuno.stimulus) +
    p.estab# this returns the probability of establishment
    # which varies between 1 and 0.1 (this will need to be revised)
  })
}

## demographic age structure
pi.a <- function(parameters) {
  with(as.list(c(parameters)), {
    j = seq(1,na)
    a = (j-1)*amax/na + amax/(2*na)
    da <- amax/na
    da*mu.h*exp(-mu.h*a)/(1-(exp(-mu.h*amax)))
  })
}

## cercarial exposure by age
rho.a<- function(parameters, m) {
  with(as.list(c(parameters)), {
    j = seq(1,na)
    a = (j-1)*amax/na + amax/(2*na)
    if(m==1) {
      return( (a*c*exp(-(beta*a)) +d)*rel.exp.m)
    } else {
      return( (a*c*exp(-(beta*a)) +d) )
    }
  })
}
```



```

## Transmission model
model <- function(times, y, parameters){
  with(as.list(c(y,parameters)),{

    ## state variable matrices structured by sex
    ymat <- matrix(y, ncol=2*na, nrow=(nw+2))

    dymat <- matrix(0, ncol=2*na, nrow=(nw+2))

    ## indicator variables for elements of i corresponding to live, dead and optical density IgE
    live.worms.id <- seq(1:nw)
    dead.worms.id <- nw+1
    IgE.id <- nw + 2 # this compartment tracks the IgE optical density in different age groups
    tot.comp <- IgE.id # total number of compartments
    #age step
    da <- 1/(amax/na)
    # female column indicators
    fem.cols <- seq(1,na,1)
    # male column indicators
    mal.cols <- seq(na+1,2*na, 1)

    ## proportion male and female in the population
    p.f <- sr
    p.m <- 1-sr

    # IgE OD, either a function of dying worms or cumulative experience of dead worms
    IgE.a.f <- ymat[IgE.id, fem.cols]
    IgE.a.m <- ymat[IgE.id, mal.cols]

    #establishment probability by each age group and by sex
    probest.a.f <- imm.a(parameters, IgE.a.f)
    probest.a.m <- imm.a(parameters, IgE.a.m)

    # age structure (assumed same for males and females)
    pia <- pi.a(parameters)

    ### exposure function structured by age and sex
    rhoa.f <- rho.a(parameters, m=0) # exposure - females
    rhoa.m <- rho.a(parameters, m=1) # exposure - males

    Wtot.f <- colSums(ymat[live.worms.id,fem.cols]) #Total number of worms - females
    Wtot.m <- colSums(ymat[live.worms.id,mal.cols]) #Total number of worms - males

    ## force of infection - females
    lambda.a.f <- (mu.w*R0*rhoa.f*probest.a.f*(sum(pia*rhoa.f*Wtot.f)*p.f + sum(pia*rhoa.m*Wtot.m)*p.m))/
      (sum(pia*rhoa.f^2)*p.f + sum(pia*rhoa.m^2)*p.m )

    ## force of infection - males

```



```

lambda.a.m <- (mu.w*R0*rhoa.m*probest.a.m*(sum(pia*rhoa.f*Wtot.f)*p.f
+ sum(pia*rhoa.m*Wtot.m)*p.m))/
  (sum(pia*rhoa.f^2)*p.f + sum(pia*rhoa.m^2)*p.m )

lambda.a <- c(lambda.a.f, lambda.a.m)

### MALE and FEMALES together
for (j in 1:(2*na)) { #'for' Loop (first age group to total number (na
))
  if (j==1) { #first age group (female)
    for (i in 1:tot.comp) { #for worm compartment (i) in worm compartm
ents W0:D (within the age loop)
      if (i==1) { #i.e. W0
        #Transmission equation for W0: foi - worm progression from thi
s compartment - aging from this age strata
        dymat[i,j] = lambda.a[j]-(nw*mu.w + da)*ymat[i,j]
      } else if (i<dead.worms.id) { #i.e. W1..Wi
        #Transmission equation for W1..Wi worm progression from previo
us compartment - worm progression from this compartment - aging from this
age strata
        dymat[i,j] = nw*mu.w*ymat[i-1,j] - (nw*mu.w + da)*ymat[i,j]
      } else if (i==dead.worms.id) { #i.e. D
        #Transmission equation for D: gain from previous worm compartm
ent - aging from this age strata
        dymat[i,j] = nw*mu.w*ymat[nw,j] - da*ymat[i,j]
      }
      else if (i==IgE.id) { #i.e. Imm
        #Transmission equation for immunity: product of dying worms an
d a constant
        dymat[i,j] = gain.imm*ymat[nw,j]*nw*mu.w - ymat[i,j]*loss.imm
- ymat[i,j] *da
      }
    }
  }
  #previously = else if (j>1), changed to (j>1&j<na)
  else if ( (j>1 & j<na) ) { #for age classes j+1..na-1 or na+
    for (i in 1:tot.comp) {
      if (i==1) { #i.e. W0
        #Transmission equation for W0: foi - worm progression - aging
from this age strata + aging from previous age strata
        dymat[i,j] = lambda.a[j]-(nw*mu.w +da)*ymat[i,j] + ymat[i,j-1]
*da
      } else if (i<dead.worms.id) { #i.e. W1..Wi
        #Transmission equation for W1..Wi: worm progression from previ
ous compartment - worm progression from this compartment - aging from this
compartment + aging from previous age strata
        dymat[i,j] = nw*mu.w*ymat[i-1,j] - nw*mu.w*ymat[i,j] + (ymat[i
,j-1] - ymat[i,j])*da
      } else if (i==dead.worms.id) { #i.e. D
        #Transmission equation for D: worm progression from previous c
ompartment - aging from this compartment + aging from previo
us age strata
        dymat[i,j] = nw*mu.w*ymat[nw,j] + (ymat[i,j-1] - ymat[i,j])*da
      }
      else if (i==IgE.id) { #i.e. Imm
        #Transmission equation for immunity: product of dead worm comp

```



```

artment and constant
    #this component is not used
    dymat[i,j] = gain.imm*ymat[nw,j]*nw*mu.w - ymat[i,j]*loss.imm
+ (ymat[i,j-1] - ymat[i,j])*da
    }
    }
    #previously = else if (j>na), changed to (j==na) - could also be j
ust 'else'?
    } else if (j==na ) { #i.e. eldest age class
        for (i in 1:tot.comp) {
            if (i==1) { #i.e. W0
                #Transmission equation for W0: foi - worm progression + aging
from previous age strata
                dymat[i,j] = lambda.a[j]-(nw*mu.w)*ymat[i,j] + (ymat[i,j-1] -
ymat[i,j])*da
            } else if (i<dead.worms.id) { #i.e. W1..Wi
                #Transmission equation for W1..Wi: worm progression from previ
ous compartment - worm progression from this compartment + aging from previ
ous age strata
                dymat[i,j] = nw*mu.w*ymat[i-1,j] - nw*mu.w*ymat[i,j] + (ymat[i
,j-1] - ymat[i,j])*da
            } else if (i==dead.worms.id) { #i.e. D
                #Transmission equation for D: worm progression from previous c
ompartment + aging from previous age strata
                dymat[i,j] = nw*mu.w*ymat[nw,j] + (ymat[i,j-1] - ymat[i,j])*da
            }
            else if (i==IgE.id) { #i.e. Imm
                #Transmission equation for immunity: product of dead worm comp
artment and constant
                # this component is not used
                dymat[i,j] = gain.imm*ymat[nw,j]*nw*mu.w - ymat[i,j]*loss.imm
+ (ymat[i,j-1] - ymat[i,j])*da
            }
        }
    } else if (j==(na+1)) {
        for (i in 1:tot.comp) { #for worm compartment (i) in worm compartm
ents W0:D (within the age loop)
            if (i==1) { #i.e. W0
                #Transmission equation for W0: foi - worm progression from thi
s compartment - aging from this age strata
                dymat[i,j] = lambda.a[j]-(nw*mu.w + da)*ymat[i,j]
            } else if (i<dead.worms.id) { #i.e. W1..Wi
                #Transmission equation for W1..Wi worm progression from previo
us compartment - worm progression from this compartment - aging from this
age strata
                dymat[i,j] = nw*mu.w*ymat[i-1,j] - (nw*mu.w + da)*ymat[i,j]
            } else if (i==dead.worms.id) { #i.e. D
                #Transmission equation for D: gain from previous worm compartm
ent - aging from this age strata
                dymat[i,j] = nw*mu.w*ymat[nw,j] - da*ymat[i,j]
            }
            else if (i==IgE.id) { #i.e. Imm
                #Transmission equation for immunity: product of dying worms an
d a constant
                dymat[i,j] = gain.imm*ymat[nw,j]*nw*mu.w - ymat[i,j]*loss.imm
- ymat[i,j] *da
            }
        }
    }
}

```



```

    }
  }
  } else if ( j>(na+1) & j<(2*na) ) {
    for (i in 1:tot.comp) {
      if (i==1) { #i.e. W0
        #Transmission equation for W0: foi - worm progression - aging
        from this age strata + aging from previous age strata
        dymat[i,j] = lambda.a[j]-(nw*mu.w +da)*ymat[i,j] + ymat[i,j-1]
      } else if (i<dead.worms.id) { #i.e. W1..Wi
        #Transmission equation for W1..Wi: worm progression from previ
        ous compartment - worm progression from this compartment - aging from this
        compartment + aging from previous age strata
        dymat[i,j] = nw*mu.w*ymat[i-1,j] - nw*mu.w*ymat[i,j] + (ymat[i
        ,j-1] - ymat[i,j])*da
      } else if (i==dead.worms.id) { #i.e. D
        #Transmission equation for D: worm progression from previous c
        ompartment - aging from this compartment +
        aging from previo
        us age strata
        dymat[i,j] = nw*mu.w*ymat[nw,j] + (ymat[i,j-1] - ymat[i,j])*da
      }
      else if (i==IgE.id) { #i.e. Imm
        #Transmission equation for immunity: product of dead worm comp
        artment and constant
        #this component is not used
        dymat[i,j] = gain.imm*ymat[nw,j]*nw*mu.w - ymat[i,j]*loss.imm
        + (ymat[i,j-1] - ymat[i,j])*da
      }
    }

    } else if (j==(2*na) ) {
      for (i in 1:tot.comp) {
        if (i==1) { #i.e. W0
          #Transmission equation for W0: foi - worm progression + agin
          g from previous age strata
          dymat[i,j] = lambda.a[j]-(nw*mu.w)*ymat[i,j] + (ymat[i,j-1]
          - ymat[i,j])*da
        } else if (i<dead.worms.id) { #i.e. W1..Wi
          #Transmission equation for W1..Wi: worm progression from pre
          vious compartment - worm progression from this compartment + aging from pr
          evius age strata
          dymat[i,j] = nw*mu.w*ymat[i-1,j] - nw*mu.w*ymat[i,j] + (ymat
          [i,j-1] - ymat[i,j])*da
        } else if (i==dead.worms.id) { #i.e. D
          #Transmission equation for D: worm progression from previous
          compartment + aging from previous age strata
          dymat[i,j] = nw*mu.w*ymat[nw,j] + (ymat[i,j-1] - ymat[i,j])*
          da
        }
        else if (i==IgE.id) { #i.e. Imm
          #Transmission equation for immunity: product of dead worm co
          mpartment and constant
          # this component is not used
          dymat[i,j] = gain.imm*ymat[nw,j]*nw*mu.w - ymat[i,j]*loss.i
          mm + (ymat[i,j-1] - ymat[i,j])*da
        }
      }
    }
  }
}

```



```

    }

    }
}

## return relevant variables alongside matrix output
worms.a <- Wtot.f*p.f + Wtot.m*p.m
worms.f <- weighted.mean(Wtot.f, pia)
worms.m <- weighted.mean(Wtot.m, pia)

worms.mn <- weighted.mean(worms.a, pia)

## egg output (no density dependence - this may need to be revised)
epg.a.f <- Wtot.f*0.5*eggrate # the 0.5 is for even worm sex ratio
epg.a.m <- Wtot.m*0.5*eggrate
epg.a <- epg.a.f*p.f + epg.a.m*p.m

epg.mn <- weighted.mean(epg.a, pia)

dead.worms.a.f <- ymat[dead.worms.id,fem.cols]
dead.worms.a.m <- ymat[dead.worms.id,mal.cols]
dead.worms.a <- dead.worms.a.f*p.f + dead.worms.a.m*p.m

dead.worms.mn <- weighted.mean(dead.worms.a, pia)

probest.mn.f <- weighted.mean(probest.a.f, pia)
probest.mn.m <- weighted.mean(probest.a.m, pia)
probest.mn <- probest.mn.f*p.f + probest.mn.m*p.m

IgE.mn.f <- weighted.mean(IgE.a.f, pia)
IgE.mn.m <- weighted.mean(IgE.a.m, pia)
IgE.mn <- IgE.mn.f*p.f + IgE.mn.m*p.m
IgE.a <- IgE.a.f*p.f + IgE.a.m*p.m

j = seq(1,na)
age = (j-1)*amax/na + amax/(2*na)

return(list(y=rbind(dymat), age=age,
             worms.by.a=worms.a, worms.mn=worms.mn, IgE.a.f = IgE.a.f,
             IgE.a.m=IgE.a.m, IgE.mn=IgE.mn,
             dead.worms.mn = dead.worms.mn,
             epg.a.f = epg.a.f , epg.a.m = epg.a.m,
             epg.mn = epg.mn,
             worms.a.f = Wtot.f, worms.a.m = Wtot.m,
             worms.mn.f = worms.f, worms.mn.m = worms.m
             ))
})
}

runmod <- function(parameters) {
  y <- rep(1, (parameters["nw"]+2)*parameters["na"]*2)
  ## stepsize and times
  stepsize=60/365

```



```

## 50 years to equilibrium
times<-seq(0,50, stepsize)
out<- deSolve::rk4(y=y, times=times, func=model, parms=parameters)
out <- as.data.frame(out)
out
}

## function to extract age dependency at equilibrium
extract <- function(out, parameters) {
  tmp <- t(out[nrow(out), grep("worms.by.a", colnames(out),fixed=T)])
  tmp <- cbind(tmp, t(out[nrow(out), grep("age", colnames(out))]))
  tmp <- as.data.frame(cbind(tmp,
                             t(out[nrow(out), grep("epg.a.f", colnames(out)
))]),
                             t(out[nrow(out), grep("epg.a.m", colnames(out)
))]),
                             t(out[nrow(out), grep("worms.a.f", colnames(o
ut))]),
                             t(out[nrow(out), grep("worms.a.m", colnames(o
ut))]),
                             t(out[nrow(out), grep("IgE.a.f", colnames(out)
))]),
                             t(out[nrow(out), grep("IgE.a.m", colnames(out)
))]))
  colnames(tmp) <- c("worms.a", "age", "epg.a.f", "epg.a.m", "worms.a.f",
"worms.a.m",
                    "IgE.a.f", "IgE.a.m")
  tmp
}

```

A3.2 MH-MCMC functions

```

# Prior Function
prior_function <- function(parameters_vector) {

  R0 <- parameters_vector["R0"]
  worm_mortality_rate <- parameters_vector["mu.w"]
  immunity_gain <- parameters_vector["gain.imm"]
  immunity_loss <- parameters_vector["loss.imm"]
  egg_prod_rate <- parameters_vector["eggrate"]
  sex_exposure_scaling_factor <- parameters_vector["rel.exp.m"]

  R0_prior_value <- dunif(R0, min = 0.5, max = 8, log = TRUE)
  worm_mortality_rate_prior_value <- dunif(worm_mortality_rate, min = 0.06
, max = 0.29, log = TRUE) # change this so that it isn't the rate but the
actual lifespan
  immunity_gain_prior_value <- dunif(immunity_gain, min = 0, max = 1, log
= TRUE)
  immunity_loss_prior_value <- dunif(immunity_loss, min = 0.01, max = 18,
log = TRUE)
}

```



```

    egg_prod_rate_prior_value <- dunif(egg_prod_rate, min = 1, max = 10, log
= TRUE)
    sex_exposure_scaling_factor_prior_value <- dunif(sex_exposure_scaling_fa
ctor, min = 5, max = 45, log = TRUE)

    prior_output <- R0_prior_value + worm_mortality_rate_prior_value + immun
ity_gain_prior_value + immunity_loss_prior_value +
                    egg_prod_rate_prior_value + sex_exposure_scaling_factor_
prior_value

    return(unname(prior_output))
}

#Likelihood Function
loglikelihood_function <- function(proposed_parameter_values, parameters_v
ector, breakdown) {

  # Ensuring the Proposed Parameter Values Have the Correct Names
  names(proposed_parameter_values) <- c("R0", "mu.w", "gain.imm", "loss.im
m", "eggrate", "rel.exp.m")

  # Input proposed parameters into parameters vector
  parameters_vector["R0"] <- proposed_parameter_values[1]
  parameters_vector["mu.w"] <- proposed_parameter_values[2]
  parameters_vector["gain.imm"] <- proposed_parameter_values[3]
  parameters_vector["loss.imm"] <- proposed_parameter_values[4]
  parameters_vector["eggrate"] <- proposed_parameter_values[5]
  parameters_vector["rel.exp.m"] <- proposed_parameter_values[6]

  # Running the model
  model_output <- runmod(parameters_vector)
  processed_output <- extract(model_output, parameters_vector)

  # Processing ReadySchisto Data
  male_subsetter <- Ready_Schisto$sex == 1
  female_subsetter <- Ready_Schisto$sex == 0
  male_ages <- Ready_Schisto$age[male_subsetter]
  female_ages <- Ready_Schisto$age[female_subsetter]
  male_epg <- round(Ready_Schisto$epg_base[male_subsetter])
  female_epg <- round(Ready_Schisto$epg_base[female_subsetter])
  male_ige <- Ready_Schisto$TAL1_IgE_base[male_subsetter] + 0.03
  female_ige <- Ready_Schisto$TAL1_IgE_base[female_subsetter] +0.03

  # Extracting the Male and Female Specific Model Outputs
  male_IgE_model_output <- processed_output$IgE.a.m[male_ages]
  female_IgE_model_output <- processed_output$IgE.a.f[female_ages]
  male_epg_model_output <- round(processed_output$epg.a.m[male_ages])
  female_epg_model_output <- round(processed_output$epg.a.f[female_ages])

  # Calculate the Loglikelihoods

  male_IgE_mean <- male_IgE_model_output
  male_IgE_variance <- 0.03253145
  male_gamma_shape <- (male_IgE_mean^2)/male_IgE_variance

```



```

male_gamma_scale <- male_IgE_variance/male_IgE_mean

female_IgE_mean <- female_IgE_model_output
female_IgE_variance <- 0.03253145
female_gamma_shape <- (female_IgE_mean^2)/female_IgE_variance
female_gamma_scale <- female_IgE_variance/female_IgE_mean

# Calculate the Loglikelihoods
loglik_epg_male <- sum(dnbinom(male_epg, mu = male_epg_model_output, size = 0.4097107, log = TRUE))
loglik_epg_female <- sum(dnbinom(female_epg, mu = female_epg_model_output, size = 0.4097107, log = TRUE))
loglik_IgE_male <- sum(dgamma(male_ige, shape = male_gamma_shape, scale = male_gamma_scale, log = TRUE))
loglik_IgE_female <- sum(dgamma(female_ige, shape = female_gamma_shape, scale = female_gamma_scale, log = TRUE))

if (breakdown == TRUE) {
  print(c("Male Epg Loglik is: ", loglik_epg_male))
  print(c("Female Epg Loglik is: ", loglik_epg_female))
  print(c("Male IgE Loglik is: ", loglik_IgE_male))
  print(c("Female IgE Loglik is: ", loglik_IgE_female))
}

# Loglik_IgE_male <- sum(dgamma(male_IgE_model_output, shape = 0.1868786, scale = 0.8995443, log = TRUE))
# Loglik_IgE_female <- sum(dgamma(female_IgE_model_output, shape = 0.1868786, scale = 0.8995443, log = TRUE))

# Return overall Loglikelihood
overall_loglikelihood <- loglik_epg_male + loglik_epg_female + loglik_IgE_male + loglik_IgE_female
return(overall_loglikelihood)
}

# Proposal Function
proposal_function <- function(current_parameters, covariance_matrix) {
  proposed_parameter_values <- mvrnorm(1, current_parameters, covariance_matrix)
  for (i in 1:length(proposed_parameter_values)) {
    if (proposed_parameter_values[i] <= 0) {
      proposed_parameter_values[i] <- 0.001
    }
  }
  return(proposed_parameter_values)
}

# Posterior Function
posterior_function <- function(proposed_parameter_values, parameters_vector){

  parameters_vector["R0"] <- proposed_parameter_values[1]
  parameters_vector["mu.w"] <- proposed_parameter_values[2]

```



```

parameters_vector["gain.imm"] <- proposed_parameter_values[3]
parameters_vector["loss.imm"] <- proposed_parameter_values[4]
parameters_vector["eggrate"] <- proposed_parameter_values[5]
parameters_vector["rel.exp.m"] <- proposed_parameter_values[6]

posterior <- loglikelihood_function(proposed_parameter_values, parameter
s_vector, FALSE) + prior_function(parameters_vector)
return(posterior)
}

# MCMC Function
MCMC_running_adaptive <- function(number_of_iterations, parameters_vector,
sd_proposals, start_covariance_adaptation) {

  # Scaling Factor Adaptation Parameters
  number_of_parameters <- 6
  required_acceptance_probability <- 0.234 # required acceptance probabili
ty
  alpha_r <- -qnorm(required_acceptance_probability/2) # same as -qtrunc(p
=p_r.accept/2, spec="norm", a = -Inf, b = Inf)
  r <- array(NA, dim = c(number_of_iterations + 1, 1))
  c_r <- array(NA, dim = c(number_of_iterations, 1))
  r[1, ] <- rexp(n = 1, rate = 1) # generating a random value to begin wit
h
  acceptance_tracker <- array(0, dim = c(number_of_iterations + 1, 1))

  # Storage for Output
  MCMC_output <- matrix(nrow = number_of_iterations + 1, ncol = 6)
  MCMC_output[1, ] <- c(parameters_vector["R0"], parameters_vector["mu.w"]
, parameters_vector["gain.imm"],
                        parameters_vector["loss.imm"], parameters_vector["
eggrate"], parameters_vector["rel.exp.m"])
  colnames(MCMC_output) = c("R0", "mu.w", "gain.imm", "loss.imm", "eggrate
", "rel.exp.m")
  Acceptances <- vector(length = number_of_iterations + 1)
  Acceptance_Ratio <- vector(length = number_of_iterations + 1)
  Logposterior_storage <- vector(length = number_of_iterations + 1)

  # Generating the Covariance Matrix
  sigma <- matrix(0, nrow = 6, ncol = 6)
  for (i in 1:6) {
    for (j in 1:6) {
      if (i == j) {
        sigma[i, j] <- sd_proposals[i]
      }
    }
  }

  # Calculating the Posterior Probability for the Initial Values
  initial_values <- vector(length = 6)
  initial_values[1] <- parameters_vector["R0"]
  initial_values[2] <- parameters_vector["mu.w"]
  initial_values[3] <- parameters_vector["gain.imm"]
  initial_values[4] <- parameters_vector["loss.imm"]
  initial_values[5] <- parameters_vector["eggrate"]

```



```

initial_values[6] <- parameters_vector["rel.exp.m"]

current_posterior <- posterior_function(initial_values, parameters_vector)

# Running the Actual MCMC
for (i in 1:number_of_iterations){

  proposed_parameter_values <- proposal_function(MCMC_output[i, ], r[i,
] * sigma)
  proposed_posterior <- posterior_function(proposed_parameter_values, parameters_vector)

  likelihood_ratio <- exp(proposed_posterior - current_posterior);

  if(runif(1) < likelihood_ratio) {

    MCMC_output[i + 1, ] <- proposed_parameter_values
    Acceptances[i] <- 1
    Logposterior_storage[i] <- proposed_posterior
    current_posterior <- proposed_posterior
    acceptance_tracker[i, ] <- 1

  } else{

    MCMC_output[i + 1, ] <- MCMC_output[i, ]
    Acceptances[i] <- 0
    Logposterior_storage[i] <- current_posterior
    acceptance_tracker[i, ] <- 0

  }

  # Adaptation Step Involves the Robins-Munro Adaptation of r, the SD of the Proposal Distribution
  c_r[i, ] <- r[i, ] * (1 - (1/number_of_parameters)) * ((2 * pi ^ 0.5 * exp((alpha_r ^ 2)/2)) / (2 * alpha_r)) + 1/(number_of_parameters * required_acceptance_probability * (1 - required_acceptance_probability))

  # Last iter was rejection if accept[i-1] is equal to
  r[i + 1, ] <- abs(ifelse((acceptance_tracker[i, ] == 0), r[i, ] - (c_r[i, ] * required_acceptance_probability)/max(100, i/number_of_parameters),
    r[i, ] + (c_r[i, ] * ((1 - required_acceptance_probability)/max(100, i/number_of_parameters))))))

  # Covariance Matrix Adaptation
  if (i > start_covariance_adaptation) {
    if (i == start_covariance_adaptation + 1) {
      previous_mean <- as.matrix(apply(MCMC_output[1:start_covariance_adaptation, ], 2, mean), nrow = number_of_parameters, ncol = 1)
      sigma <- cov(MCMC_output[1:start_covariance_adaptation, ])
    }
    new_mean <- as.matrix((1/i) * ((i-1) * previous_mean + MCMC_output[i + 1, ]), nrow = number_of_parameters, ncol = 1)
  }
}

```



```

    sigma <- (((i-2)/(i-1)) * sigma) + (previous_mean %*% t(previous_mean)) - ((i/(i-1)) * (new_mean %*% t(new_mean))) + ((1/(i-1)) * (as.matrix(MCMC_output[i + 1, ]) %*% t(as.matrix(MCMC_output[i + 1, ]))))
    previous_mean <- new_mean
  }

  Acceptance_Ratio[i] <- sum(Acceptances)/i

  if(i %% 100 == 0) {
    print(c("The iteration number is", i))
    print(c("The acceptance ratio is", Acceptance_Ratio[i]))
    print(c("The logposterior is", current_posterior))
    print(MCMC_output[i + 1, ])
    print(sigma)
  }
}
list <- list()
list[["MCMC_Output"]] <- MCMC_output
list[["Acceptances"]] <- Acceptances
list[["Posterior_Output"]] <- Logposterior_storage
return(list)
}

```


Appendix 4

Supplementary tables for Chapter 4

Table A4.1 Geometric mean odds ratios describing the association between post-treatment IgE responses to recombinant SmTAL1 domain antigens and intensity of reinfection at 2 years post-treatment.

	rSmTAL1 EF hand		rSmTAL1 DLC	
	GM ratio (95% CI)	P-value	GM ratio (95% CI)	P-value
rSmTAL1 domain [†]	0.89 (0.41–1.95)	ns	1.13 (0.50 – 2.52)	ns
Age (years) [‡]				
10–15	1.27 (0.42–3.83)	ns	1.20 (0.41–3.52)	ns
16 – 24	0.52 (0.15–1.80)	ns	0.48 (0.15–1.56)	ns
25 – 34	0.30 (0.09–1.00)	*	0.26 (0.09–0.80)	*
35 – 60	0.15 (0.05–0.49)	**	0.14 (0.05–0.43)	***
9wk egg count [§]	1.00 (1.00 – 1.00)	***	1.00 (1.00–1.00)	***

***p<0.001, **p<0.01, *p<0.05, ns non-significant;

[†] values for EF hand or DLC, as indicated;

[‡] values are compared to 7- to 9-year-old age group;

[§] proxy for treatment efficacy;

Table A4.2. Association between pre-treatment ShTAL-specific IgE responses and baseline CAA, as a measure of worm burden, adjusted for age, sex, village and ShTAL-specific IgG4. Results are displayed for A) high intensity and B) moderate intensity villages, independently.

A.		ShTALI		ShTALI, 3		ShTALI, 3, 5		ShTALI, 3, 5, II	
High intensity		GM ratio (95% CI)	p-value	GM ratio (95% CI)	p-value	GM ratio (95% CI)	p-value	GM ratio (95% CI)	p-value
Age (years) [†]									
7-9		0.68 (0.20-2.25)	ns	0.68 (0.21-2.19)	ns	0.72 (0.22-2.36)	ns	0.64 (0.19-2.10)	ns
10-14		0.95 (0.30-3.06)	ns	0.77 (0.25-2.36)	ns	0.81 (0.26-2.56)	ns	0.74 (0.24-2.35)	ns
15 - 24		0.32 (0.10-1.01)	.	0.31 (0.10-0.94)	*	0.32 (0.10-0.98)	*	0.31 (0.10-0.98)	*
25 - 29		0.12 (0.04-0.38)	***	0.11 (0.03-0.33)	***	0.12 (0.04-0.36)	***	0.11 (0.04-0.36)	***
30+		0.10 (0.02-0.41)	**	0.09 (0.02-0.37)	**	0.09 (0.02-0.38)	**	0.09 (0.02-0.38)	**
Sex [‡]		0.44 (0.22-0.88)	*	0.47 (0.25-0.92)	*	0.46 (0.24-0.90)	*	0.48 (0.24-0.93)	*
B.									
Moderate intensity									
Age (years) [†]									
7-9		9.41 (2.79-31.74)	***	10.31 (3.12-34.05)	***	10.00 (1.88-11.48)	***	8.63 (2.63-28.31)	***
10-14		25.76 (6.53-101.60)	***	29.25 (7.60-112.65)	***	31.94 (3.69-25.21)	***	27.12 (7.22-101.88)	***
15 - 24		21.96 (5.20-92.81)	***	20.06 (4.68-85.91)	***	24.12 (1.97-14.41)	***	9.99 (2.14-46.65)	**
25 - 29		13.08 (3.73-45.91)	***	12.13 (3.35-43.98)	***	13.97 (1.20-7.52)	***	9.81 (2.77-34.75)	***
30+		7.23 (1.65-31.66)	**	7.41 (1.71-32.12)	**	11.60 (0.84-7.72)	**	6.60 (1.57-27.75)	*
Sex [‡]		-	ns	-	ns	-	ns	-	ns
C.									
TALx-IgE		1.31 (0.57-3.03)	ns	0.51 (0.14-1.83)	ns	0.55 (0.10-2.96)	ns	1.59 (0.08-33.71)	ns
TALI-IgG4		2.65 (0.85-8.21)	.	-	ns	-	ns	-	ns
TAL3-IgG4				3.99 (1.35-11.79)	*	-	ns	-	ns
TAL5-IgG4							*	-	ns
TALII-IgG4						8.62 (1.49-49.98)	*	-	ns
								4.54 (1.67-12.32)	**

***p<0.001, **p<0.01, *p<0.05, . p<0.1, ns non-significant; [†] values are compared to 5- to 6-year-old age group; [‡] values are provided for males compared to females;

Table A4.3. Association between 9-weeks post-treatment ShTAL-specific IgE responses and reinfection status at two years. Binomial model adjusted for age, sex, village and ShTAL-specific IgG4, non-significant terms removed. Results are displayed for the high transmission and moderate transmission villages separately.

	ShTAL	P-value	ShTAL, 3	P-value	ShTAL, 3, 5	P-value	ShTAL, 3, 5, II	P-value
High intensity (n = 98)	ShTALx-IgE	1.16 (0.06–9.24)	ns	0.83 (0.10–5.50)	ns	0.24 (0.03–1.63)	0.73 (0.09–15.55)	ns
	Age ≥10 [†]	1.42 (0.52–4.30)	ns	0.05 (0.02–0.14)	***	0.05 (0.02–0.14)	0.05 (0.02–0.12)	***
Moderate intensity (n = 106)	ShTALx-IgE	0.04 (0.02–0.11)	ns	0.41 (0.13–1.20)	ns	0.44 (0.12–1.46)	1.07 (0.14–5.56)	ns

***p<0.001, ns non-significant; [†] values compared to age <10 years.

Appendix 5

Chapter 5 Supplementary Information

Table A5.1 Proteins identified from Mascot searches of LC-MS/MS data against the *S. haematobium* protein database (version 1; Bioproject PRJNA78265, 13073 sequences; 5666088 residues)

Accession	Gene Name	Size (aa)	MW (kDa)	Score*	Peptides* (unique)	Spot ID
A_00547	14-3-3 protein I(MS3_03976)	242	27.2	329	6(6)	392, 894
A_04333	14-3-3 protein I(MS3_03977)	252	28.3	401	7(7)	894
B_00937	26S protease regulatory subunit 8(MS3_11217)	271	30.2	141	3(3)	841
A_05901	26S proteasome non-ATPase regulatory subunit II(MS3_04609)	415	46.8	213	5(5)	842
A_01943	26S proteasome non-ATPase regulatory subunit I2(MS3_07874)	611	72.1	607	12(12)	821
A_07972	26S proteasome non-ATPase regulatory subunit 6(MS3_07520)	347	40.1	106	3(3)	842
A_01124	6-phosphogluconate dehydrogenase, decarboxylating(MS3_06507)	478	52.8	854	17(17)	821, 211
A_08205	60 kDa heat shock protein, mitochondrial(MS3_08926)	572	61.1	751	13(13)	821, 785, 881
C_00478	78 kDa glucose-regulated protein(MS3_10713)	648	71.3	639	13(12)	143, 392, 881
A_04688	97 kDa heat shock protein(MS3_02688)	569	64	750	12(12)	142, 143
A_00100	Actin-1(MS3_01922)	149	16.4	234	3(3)	143, 841, 211, 392, 881, 456
A_00382	Actin-1(MS3_01933)	149	16.4	601	10(10)	881
A_00561	Actin-1(MS3_02465)	215	23.5	156	3(3)	108,699, 143, 158, 785, 841, 211, 794, 796, 392, 881, 456, 108, 699, 142, 714, 143, 821, 158, 785, 841, 211, 794, 842, 796, 392, 797, 881, 456, 328
A_01049	Actin-2(MS3_07374)	376	41.7	1399	24(24)	821
A_06911	Adenosylhomocysteinase A(MS3_04449)	445	49.1	792	15(15)	821
A_00362	Adenylosuccinate synthetase(MS3_05283)	405	44.3	190	4(4)	821
A_00946	Adenylyl cyclase-associated protein I(MS3_06380)	458	50.4	368	6(6)	821, 211, 456

A_05823	ADP-ribosylation factor 4(MS3_03443)	181	20.7	102	2(2)	392
A_01588	Alanine aminotransferase 2(MS3_00805)	493	55.3	138	4(4)	821, 211
A_00552	Alanine--tRNA ligase, cytoplasmic(MS3_04103)	924	103	119	2(2)	799
A_05634	Alcohol dehydrogenase class-3(MS3_08826)	363	38.97	381	8(8)	841
A_04318	ALG-2 interacting protein X(MS3_04592)	712	82.7	524	11(11)	108
A_07016	Alpha-adducin(MS3_05669)	575	62.5	403	9(9)	821
A_02297	Alpha-centractin(MS3_04014)	375	42.5	575	10(10)	841, 842
A_03954	Aminoacylase-I(MS3_08008)	429	48.9	552	11(11)	821
A_04216	Ankyrin repeat domain-containing protein 13D(MS3_03267)	650	72.5	365	9(9)	108
A_00951	Ankyrin-3(MS3_00592)	2350	259.4	731	13(13)	108, 803, 142, 143, 796, 797
B_00989	Ankyrin-3(MS3_04585)	1406	155.9	419	7(7)	796, 797, 799
A_06610	Annexin A7(MS3_01952)	284	31.8	125	3(3)	894
A_05991	Annexin A7(MS3_04598)	548	60.4	123	3(3)	803
A_04779	Asparagine--tRNA ligase, cytoplasmic(MS3_05222)	592	66.8	196	3(3)	143, 799
A_05293	Aspartate aminotransferase, cytoplasmic(MS3_00663)	230	26	117	2(2)	841
A_00071	ATP synthase subunit alpha, mitochondrial(MS3_07464)	451	49.2	166	3(3)	392
A_02920	ATP synthase subunit beta, mitochondrial(MS3_03295)	312	34.1	180	2(2)	392
A_00203	ATP synthase subunit beta, mitochondrial(MS3_03301)	196	21.1	121	2(2)	392
B_00447	ATP-dependent RNA helicase eIF4A(MS3_01778)	123	13.7	145	3(3)	881
A_04685	ATP-dependent RNA helicase eIF4A(MS3_01780)	286	32.7	674	14(13)	881
A_07985	Bifunctional glutamate/proline--tRNA ligase(MS3_01412)	1341	152.1	600	13(13)	729, 799
A_04102	cAMP-dependent protein kinase type II regulatory subunit(MS3_07384)	336	37.7	427	9(9)	881
A_06121	CAP-Gly domain-containing linker protein 1(MS3_05296)	454	51.9	99	2(2)	894
A_00672	Cell division cycle and apoptosis regulator protein 1(MS3_00097)	856	96.6	134	2(2)	881
A_03354	COP9 signalosome complex subunit 2(MS3_08425)	461	53.9	418	9(9)	821, 211
A_08296	COP9 signalosome complex subunit 4(MS3_02487)	437	49	840	16(16)	881
A_00035	Cystathionine beta-synthase(MS3_01163)	788	88.6	248	5(5)	803
B_00683	Cytoplasmic dynein 1 intermediate chain 2(MS3_09957)	510	56.7	291	6(6)	821
A_06634	Cytosolic purine 5'-nucleotidase(MS3_02593)	386	44	225	4(4)	699
A_02231	Dihydrolipoyl dehydrogenase, mitochondrial(MS3_02752)	461	48.9	363	7(7)	821, 211

A_01645	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial(MS3_07158)	420	45.4	270	4(4)	821
A_00281	Dihydropyrimidinase-related protein 4(MS3_08954)	463	50.6	791	16(16)	803, 142, 143, 729, 158, 794, 796, 797, 799
B_00335	Dihydropyrimidinase(MS3_02950)	524	57.3	353	7(7)	142, 143, 785
A_04268	Dipeptidyl peptidase 2(MS3_05939)	434	49.4	248	5(5)	142, 158, 785, 794
A_07925	DNA repair protein RAD51 I(MS3_03659)	338	36.7	628	11(11)	108, 699, 803, 142, 714, 143, 729, 821, 158, 785, 841, 211, 794, 842, 796, 861, 797, 881, 456, 799, 894
A_04791	Dynamin-I(MS3_03369)	909	102	122	4(4)	158
A_03038	EH domain-containing protein 1(MS3_01309)	544	62	248	6(6)	803
A_04199	Elongation factor 1-alpha 1(MS3_08479)	465	50.9	246	4(4)	143, 841, 211, 861, 392, 456, 328
A_03435	Elongation factor 2(MS3_08361)	849	94.5	513	9(9)	699, 143, 821, 796, 392, 797, 799
A_06659	Elongation factor Tu, mitochondrial(MS3_03615)	438	48.6	427	8(8)	841
A_08172	Enolase(MS3_02425)	434	46.7	1635	28(28)	142, 143, 821, 211, 842, 796, 392, 881, 799
B_00771	Eukaryotic peptide chain release factor subunit 1(MS3_10821)	125	14.2	135	3(3)	821
A_06789	Eukaryotic translation initiation factor 3 subunit E(MS3_00094)	382	45.5	73	2(2)	821
A_01832	Filamin-A(MS3_01744)	2483	268.3	1511	28(28)	108, 803, 142, 143, 158, 785, 794, 796, 797, 799
A_01136	Flotillin-1(MS3_09536)	383	42.8	375	8(8)	821
A_02036	Fragile X mental retardation protein 1-like protein(MS3_07796)	770	83.4	120	2(2)	158, 785, 794
A_06709	Fructose-bisphosphate aldolase(MS3_02856)	363	39.6	1362	22(22)	821, 841, 211, 842, 881

A_00541	Glucose-6-phosphate 1-dehydrogenase(MS3_09439)	163	19.3	406	8(8)	158785, 794
A_03341	Glutamine synthetase(MS3_07403)	413	46.1	1028	16(16)	842, 881
A_06665	Glutamine--tRNA ligase(MS3_06258)	562	64.5	1240	24(24)	142, 796, 797, 799
B_00393	Glutathione S-transferase class-mu 28 kDa isozyme(MS3_06482)	211	23.9	229	4(4)	803, 729
A_07113	Glycine--tRNA ligase(MS3_08104)	1242	139.6	288	6(6)	108
A_05311	Glycogen phosphorylase, liver form(MS3_01181)	841	97	211	3(2)	803, 729, 821
A_06781	Glycogenin-1(MS3_09086)	282	32.3	94	2(2)	881
A_00361	Glyoxalase domain-containing protein 4(MS3_04550)	294	32.8	693	11(11)	392
A_06188	Golgin subfamily A member 4(MS3_02962)	1706	199.8	208	5(5)	894
A_02496	GTP-binding nuclear protein Ran(MS3_08533)	216	24.9	189	4(4)	392
A_06572	Guanine nucleotide exchange factor VAV2(MS3_06721)	1080	122.4	461	10(10)	821
D_00312	Heat shock 70 kDa protein-like protein(MS3_11293)	276	30.4	435	7(7)	108, 143, 158, 785, 842, 392, 881
D_00354	Heat shock protein 83(MS3_11059)	283	32.1	421	6(6)	894
A_04778	Hexokinase(MS3_10478)	451	50.3	1666	29(29)	821, 211
A_01979 (+2)	Histone H3 (MS3_07735)			75	2(2)	
A_03330	Histone H4(MS3_07685)	103	11.4	250	6(6)	108, 699, 714, 143, 729, 785, 796, 392, 456, 799
A_04334	Hsc70-interacting protein(MS3_02787)	322	36.5	382	7(7)	881, 894
A_01880	hypothetical protein(MS3_00626) - BLAST hypothetical	915	102.8	509	12(12)	108
A_01128	hypothetical protein(MS3_00869) - BLAST major tegumental antigen SMI5	955	108.8			392, 881
A_08087	hypothetical protein(MS3_02084) - BLAST hypothetical	1039	117.2	158	3(3)	881
A_00125	hypothetical protein(MS3_04659) - BLAST zinc finger	581	68.3	624	11(11)	881
A_06542	hypothetical protein(MS3_05150) - BLAST EF hand protein	342	40.5	557	11(11)	881
A_00950	hypothetical protein(MS3_05715) - BLAST expressed conserved protein	480	54.2	142	3(3)	842
A_08269	hypothetical protein(MS3_07333) - BLAST cell adhesion molecule	237	27.1	94	2(2)	842
A_04731	hypothetical protein(MS3_09614) - BLAST ?	556	63.8	261	4(4)	803
D_00097	hypothetical protein(MS3_11248) - BLAST septin	101	11.9	96	2(2)	821

D_00394	hypothetical protein(MS3_11411) - BLAST HSP70	372	40.6	272	6(6)	108, 142, 143, 158, 785, 842, 392, 881
C_00979 (+1)	hypothetical protein(MS3_11537) - BLAST MF3 Sj egg shell	205	22.4	90	2(2)	
A_04829	Inactive tyrosine-protein kinase 7(MS3_04080)	1495	166	375	6(6)	803
A_04232	Isocitrate dehydrogenase [NADP], mitochondrial(MS3_05465)	165	18.4	144	2(2)	841
A_02280	Kinesin heavy chain(MS3_01335)	1521	173.5	508	7(7)	108, 821
A_06413	Kinesin light chain(MS3_07007)	693	79.3	1178	23(23)	158785, 794
A_00866	Kyphoscoliosis peptidase(MS3_04531)	875	99.5	207	5(5)	108, 799
D_00386	L-lactate dehydrogenase A chain(MS3_10510)	301	32.8	948	16(16)	841, 861, 894, 328
B_00767	Lethal(2) giant larvae protein 1(MS3_10379)	1358	147.3	415	9(9)	158, 785, 794
A_04811	Lissencephaly-1-like protein(MS3_01543)	387	43	187	4(4)	211
B_00563	Lysine--tRNA ligase(MS3_11149)	186	21.8	201	5(5)	142, 143, 796, 797
A_01815	Lysosomal protective protein(MS3_04617)	500	57	166	3(3)	796
A_02474	Major egg antigen(MS3_03417)	467	52.3	365	7(7)	158, 785, 794
A_04279	Major egg antigen(MS3_03420)	361	41.4	611	10(10)	842
B_00663	Major vault protein(MS3_11226)	393	44.4	129	3(3)	821
B_00584	Major vault protein(MS3_11386)	156	17.4	383	6(6)	821, 881
A_06969	Malate dehydrogenase, cytoplasmic(MS3_07730)	280	31	447	8(8)	861, 894
B_00187	Malate dehydrogenase, mitochondrial(MS3_02136)	199	21.4	134	2(2)	861, 328
A_04837	Mannose-1-phosphate guanylttransferase alpha- A(MS3_09733)	372	41.6	226	4(4)	211
A_03883	Mannose-1-phosphate guanylttransferase beta- A(MS3_10791)	204	22.9	134	2(2)	842
A_02116	Methionine aminopeptidase 1(MS3_09103)	531	59.6	329	6(6)	211
A_05404	Methionine aminopeptidase 2(MS3_06239)	463	52.1	203	4(4)	842
A_05628	Microtubule-associated protein 1A(MS3_02964)	980	106.6	91	2(2)	328
A_00875	Microtubule-associated protein futsch(MS3_02923)	2241	253.5	411	9(9)	821
A_06910	Mitochondrial-processing peptidase subunit beta(MS3_06348)	199	22.6	595	9(9)	821
A_07550	NADP-dependent malic enzyme(MS3_01897)	565	63.9	391	9(9)	142, 143, 796, 797
B_00348	Nuclear protein localization protein 4-like protein(MS3_01219)	630	70.1	170	5(5)	803, 796, 797, 799
A_01472	Obg-like ATPase 1(MS3_05870)	396	44.8	500	10(10)	841

A_08044	Ornithine aminotransferase, mitochondrial(MS3_07231)	316	34.5	317	7(7)	841
A_08297	Peptidyl-prolyl cis-trans isomerase FKBP4(MS3_08019)	384	43.1	302	5(5)	821
A_07033	Peroxioredoxin-2(MS3_07531)	219	24.9	50	2(2)	456
A_02332	Phenylalanine--tRNA ligase alpha subunit(MS3_08204)	497	56.3	203	5(5)	729
A_01521	Phenylalanine--tRNA ligase beta subunit(MS3_06796)	526	58.9	388	8(8)	143, 796, 797
A_03579	Phosphoglucomutase-1(MS3_09994)	575	63.2	1756	28(28)	108, 803, 142, 143, 821, 158, 785, 794, 796, 797
A_05407	Phosphoglucomutase-2(MS3_06719)	518	58.4	196	5(5)	799
A_03905	Phosphoglycerate kinase(MS3_04778)	168	17.7	777	13(13)	821, 841, 211, 842, 881
A_08431	Phosphoglycerate kinase(MS3_05999)	269	29.4	1081	21(21)	821, 841, 211, 842, 392, 881
A_04622	Phosphorylase b kinase gamma catalytic chain, skeletal muscle/heart isoform(MS3_01719)	445	51.1	148	3(3)	881
A_02645	Plastin-2(MS3_07481)	651	74.1	446	9(9)	803, 729, 158, 794, 799
A_02310	Prolyl endopeptidase(MS3_01899)	555	62.7	518	11(11)	108
A_05539	Prolyl endopeptidase(MS3_01906)	104	12.7	136	3(3)	108
A_06245	Proteasomal ATPase-associated factor 1(MS3_09672)	211	23.1	156	3(3)	803
C_00456	Proteasome subunit beta type-1-A(MS3_03070)	212	23.2	99	2(2)	456
A_04404	Protein NDRG3(MS3_00196)	430	48.3	297	6(6)	821
A_02739	Protein SET(MS3_07372)	254	29.3	164	4(4)	894
A_02810	Putative aminopeptidase W07G4.4(MS3_01749)	532	58.6	131	4(4)	803, 729, 796, 797, 799
A_04283	Putative aminopeptidase W07G4.4(MS3_08450)	534	57.7	2092	35(35)	699, 803, 729, 821, 158, 785, 794, 796, 797, 799
A_05679	putative mannose-6-phosphate isomerase(MS3_06680)	477	53.3	341	7(7)	842
A_02429	putative protein disulfide-isomerase ER-60(MS3_10187)	327	36.8	97	2(2)	158, 785, 794, 881
A_00749	putative serine--tRNA ligase, cytoplasmic(MS3_06265)	557	63.5	500	9(9)	142, 158, 785, 794
A_04952	putative serine/threonine-protein phosphatase PP2A regulatory subunit(MS3_02928)	516	56.2	324	6(6)	785, 794
A_01177	putative UDP-glucose 4-epimerase(MS3_04887)	395	43.9	94	2(2)	894

A_06975	Pyruvate kinase PKM(MS3_08937)	564	62.1	2317	36(36)	699, 803, 729, 821, 785, 799
A_08135	Rab GDP dissociation inhibitor alpha(MS3_05530)	450	50.7	597	10(10)	821, 841, 842
A_04335	Rab GDP dissociation inhibitor alpha(MS3_05531)	422	48	420	9(9)	821
A_06034	Ras-related protein Rab- 5C(MS3_05910)	198	21.9	128	2(2)	392
A_00217	Ras-related protein Rab- 8B(MS3_00353)	205	23.2	116	2(2)	392
A_06331	Ras-related protein RABA2a(MS3_02050)	150	17.1	86	2(2)	392
D_00031	Septin-7(MS3_10902)	365	41.2	309	5(5)	821
A_07737	Serine/threonine-protein kinase OSR1(MS3_03201)	552	60.4	108	3(3)	328
B_01003	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform(MS3_11054)	244	26.9	353	6(6)	785, 794
A_00105	Serine/threonine-protein phosphatase PPI-gamma catalytic subunit(MS3_08776)	292	33.2	148	4(4)	392
A_07707	Severin(MS3_03933)	699	80.5	189	3(3)	881
A_02392	Severin(MS3_03937)	364	41.7	1777	31(31)	881
A_01662	Spliceosome RNA helicase DDX39B(MS3_07944)	396	45.8	714	13(13)	821, 841, 392
A_04087	Striatin(MS3_08222)	846	90.9	242	6(6)	842
A_05946	Succinyl-CoA ligase [ADP- forming] subunit beta, mitochondrial(MS3_07763)	416	45.3	749	15(15)	881
A_04008	Succinyl-CoA ligase [GDP- forming] subunit beta, mitochondrial(MS3_09608)	365	39.2	216	4(4)	881
A_01627	Syntaxin-binding protein 2(MS3_03372)	823	92.7	218	4(4)	797
A_02220	T-complex protein 1 subunit alpha(MS3_10572)	520	56.5	1512	24(24)	108, 142, 143, 785, 796, 797, 799
B_00656	T-complex protein 1 subunit beta(MS3_07556)	429	45.8	812	12(12)	158, 785, 794
A_08125	T-complex protein 1 subunit delta(MS3_06928)	547	59.2	535	10(10)	803
A_00711	T-complex protein 1 subunit epsilon(MS3_03054)	502	55.3	1783	31(31)	108, 142, 143, 796, 392, 797, 799
A_01496	T-complex protein 1 subunit eta(MS3_01627)	497	54.4	1780	32(32)	158, 785, 794
A_00917	T-complex protein 1 subunit gamma(MS3_00785)	427	47.5	1385	24(24)	142, 143, 796, 797, 799
A_07163	T-complex protein 1 subunit theta(MS3_08399)	510	55.7	1336	22(22)	108, 142, 143, 796
A_08009	T-complex protein 1 subunit zeta(MS3_06669)	425	46.9	636	12(12)	796, 797, 799

A_02556	Taurocyamine kinase(MS3_05631)	716	80.4	2993	54(54)	108, 714, 841
A_08162	Tegument antigen(MS3_08446)	190	22.6	508	9(9)	456
A_00696	Thimet oligopeptidase(MS3_02331)	683	78.4	1944	39(39)	108
A_08102	Thimet oligopeptidase(MS3_09903)	664	75.9	2166	44(44)	108
A_08148	Thioredoxin domain-containing protein C06A6.5(MS3_02305)	230	26.4	182	2(2)	211
A_01294	Titin(MS3_00472)	587	65.1	119	3(3)	894
A_01686	Titin(MS3_07693)	761	83.3	852	13(13)	894
A_05934	Titin(MS3_09412)	1849	209.2	140	2(2)	881
B_00064	Transaldose(MS3_0020028)	327	37.5	845	13(13)	821, 881
A_07461	Transitional endoplasmic reticulum ATPase(MS3_03802)	775	86.5	130	4(4)	108, 392
A_03666	Transketolase(MS3_03775)	497	53.7	257	6(6)	799
A_06499	Tropomodulin(MS3_06431)	357	40.5	93	2(2)	881
A_02365	Tropomyosin-2(MS3_01764)	568	65.3	157	3(3)	894
A_00073	Tropomyosin(MS3_01364)	348	40.6	1534	29(29)	894
A_00746	Troponin T(MS3_07702)	304	35.4	357	6(6)	881
A_00703	Tryptophan--tRNA ligase, cytoplasmic(MS3_07724)	362	41.4	439	8(8)	841
A_05492	Tubulin beta chain(MS3_03381)	444	49.8	945	15(15)	392, 881
A_05674	Ubiquinone biosynthesis protein COQ7-like protein(MS3_04827)	421	49.5	373	6(6)	894
A_03654	Uridine 5'-monophosphate synthase(MS3_01755)	1496	168.5	90	2(2)	211
A_05092	UTP--glucose-1-phosphate uridylyltransferase(MS3_03534)	514	57.6	60	2(2)	699
A_06795	WD repeat and FYVE domain-containing protein 1(MS3_06935)	356	39.9	97	2(2)	841
A_05842	WD repeat-containing protein 47(MS3_06367)	1118	123.8	258	4(4)	841
A_05917	Zinc phosphodiesterase ELAC protein 1(MS3_04039)	436	48.7	491	7(7)	842

*For proteins identified in multiple spots, the highest MASCOT Score and peptide sequence match figures are given



Figure A5.1. Biological process GO terms significantly enriched in 191 identified proteins compared to *Schistosoma haematobium* whole genome annotation. Bubble intensity indicates p-value, where darker bubbles represent p-values of greater significance (p-values for each term can be found in Chapter 5, Table 5.2). Edges link nodes representing highly similar GO terms, width of line indicates degree of similarity. Plotted in REVIGO (Supek *et al.*, 2011).

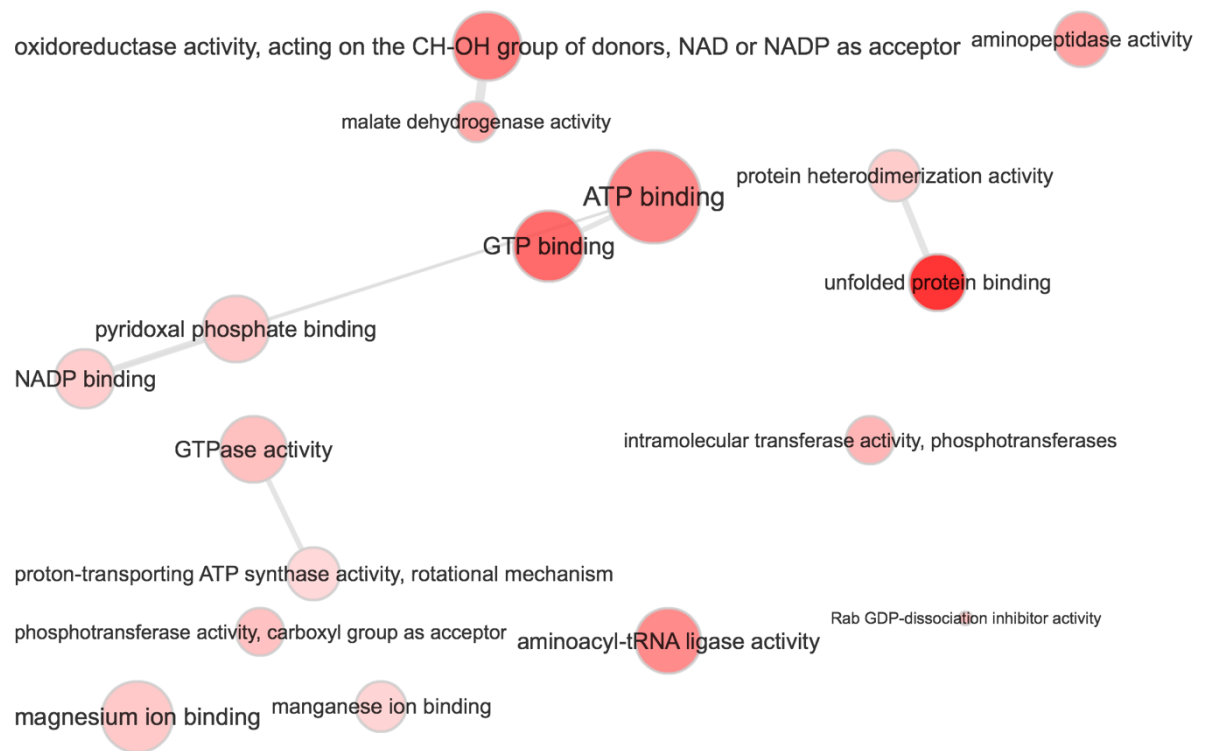


Figure A5.2. Molecular function GO terms significantly enriched in 191 identified proteins compared to *Schistosoma haematobium* whole genome annotation. Bubble intensity indicates p-value, where darker bubbles represent p-values of greater significance (p-values for each term can be found in Chapter 5, Table 5.4). Edges link nodes representing highly similar GO terms, width of line indicates degree of similarity. Plotted in REVIGO (Supek *et al.*, 2011).

Table A5.2 EV proteins identified by Nowacki *et al.* (2015) that were identified in *S. mansoni* (version 5.2) MASCOT search of immunogenic anti-fecundity case proteins

Accession number	Protein name	UniProt ID	SWAP abundance*
Smp_024110	enolase	G4VQ58	1
Smp_194770	taurocyamine kinase; creatine kinase; arginine kinase	PI6641	30
Smp_042160.2	fructose-bisphosphate aldolase	G4VJT9	2
Smp_106930	heat shock protein 70	G4V8L4	7
Smp_056970.1	glyceraldehyde 3 phosphate dehydrogenase	G4V5I2	4
Smp_046690	polyubiquitin C	G4VCP4	-
Smp_008660.1	gelsolin	G4VIJ1	101
Smp_037230	fimbrin	G4LXR0	95
Smp_094420	rab GDP dissociation inhibitor beta	G4VC25	154
Smp_099870	eukaryotic translation elongation factor 1 alpha	G4VAD2	19
Smp_202970.1	cytoplasmic actin	G4VLN1	80
Smp_009760	14-3-3 protein homolog 1	G4LUI9	9
Smp_054160	Glutathione S-transferase 28 kDa (GST 28) (GST class-mu), putative	P09792	6
Smp_014010	adenylyl cyclase associated protein 1	G4LVI8	-
Smp_005670	rab11	G4LZE2	98
Smp_009780.1	hypothetical protein (14-3-3, putative [V.4.0])	G4LUI7	61
Smp_040570.1	ras protein Rab 5A	G4VB69	345
Smp_045500	annexin	G4VI66	-

*Rank of relative abundance amongst 633 proteins identified in SWAP (Neves *et al.*, 2015)

Table A5.3 Proteins found to be common to candidate anti-fecundity targets and proteomic analysis of *S. mansoni* vomitus (Hall *et al.*, 2011)

Primary cellular function	Accession #	Identity
Lysosome	Smp_071610	Dipeptylpeptidase II
Enzyme inhibitor	Smp_089670	Alpha-2-macroglobulin
Tegument	Smp_045200	Sm22.6
	Smp_074140	Annexin
	Smp_017730	Sm200 surface protein
Cytosolic, stress	Smp_054160	GST28
Chaperone	Smp_106930	HSP70
	Smp_062420	HSP interacting protein
Glycolysis/energy	Smp_024110	Enolase
	Smp_042160	Fructose bisphosphate aldolase
	Smp_052810	Glucose 6 phosphate dehydrogenase
Structural/motor	Smp_040680	Dynein
	Smp_008660	Gelsolin
	Smp_029500	Thimet oligopeptidase K
	Smp_127820	Unknown

Table A5.4 Proteins identified within the MEROPS peptidase database
(<https://www.ebi.ac.uk/merops/index.shtml>)

Accession	Identity	Peptidase	MEROPS accession
A_01815	Lysosomal protective protein(MS3_04617)	serine carboxypeptidase A	MER0941858
A_05404	Methionine aminopeptidase 2(MS3_06239)	subfamily M24A unassigned peptidases	MER1137300
A_04404	Protein NDRG3(MS3_00196)	family S33 non-peptidase homologues	MER1137410
A_00866	Alpha-adducin(MS3_05669)	family C110 non-peptidase homologues	MER0626867
A_03954	Aminoacylase-1(MS3_08008)	subfamily M20F non-peptidase homologues	MER0853298
A_00897	Carboxypeptidase E(MS3_06451)	subfamily M14B non-peptidase homologues	MER0836381
A_02116	Cullin-3(MS3_08348)	subfamily M24A unassigned peptidases	MER1137090
A_04268	Dipeptidyl peptidase 2(MS3_05939)	family S28 unassigned peptidases	MER1053534
A_06910	Mitochondrial-processing peptidase subunit beta(MS3_06348)	subfamily M16B non-peptidase homologues	MER1137239
A_02310	Prolyl endopeptidase(MS3_01899)	subfamily S9A unassigned peptidases	MER0991507
A_06500	Plasminogen(MS3_04920)	subfamily S1A unassigned peptidases	MER0957331
C_00456	Proteasome subunit beta type-1-A(MS3_03070)	subfamily T1A unassigned peptidases	MER1091633
A_02810	Putative aminopeptidase W07G4.4(MS3_01749)	PwLAP aminopeptidase	MER0929461
A_04283	Putative aminopeptidase W07G4.4(MS3_08450)	family M17 unassigned peptidases	MER0929572
B_00335	Dihydropyrimidinase(MS3_02950)	family M38 non-peptidase homologues	MER1137408

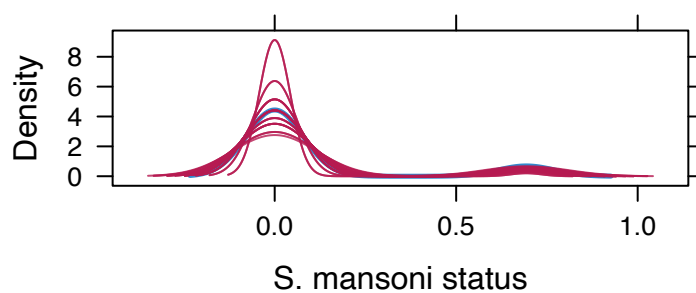


Fig. A5.3 Multiple imputation by chained equations (MICE) density plot. The density plot demonstrates the observed density distribution of *S. mansoni* infection status from the Malian data (blue) and the results of 25 imputations (red). Details of the imputation method can be found in Chapter 2, section 2.16.11.

Appendix 6

Publications

Oettle, R. C. and Wilson, S. (2017) 'The Interdependence between Schistosome Transmission and Protective Immunity', *Tropical Medicine and Infectious Disease* . doi: 10.3390/tropicalmed2030042.



Review

The Interdependence between Schistosome Transmission and Protective Immunity

Rebecca C. Oettle  and Shona Wilson *

Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK;
ro291@cam.ac.uk

* Correspondence: sw320@cam.ac.uk; Tel.: +44-1223-333338

Received: 17 July 2017; Accepted: 8 August 2017; Published: 23 August 2017

Abstract: Mass drug administration (MDA) for control of schistosomiasis is likely to affect transmission dynamics through a combination of passive vaccination and reduction of local transmission intensity. This is indicated in phenomenological models of immunity and the impact of MDA, yet immunity parameters in these models are not validated by empirical data that reflects protective immunity to reinfection. There is significant empirical evidence supporting the role of IgE in acquired protective immunity. This is proposed to be a form of delayed concomitant immunity, driven at least in part by protective IgE responses to the tegument allergen-like (TAL) family of proteins. Specific questions have arisen from modeling studies regarding the strength and duration of the protective immune response. At present, field studies have not been specifically designed to address these questions. There is therefore a need for field studies that are explicitly designed to capture epidemiological effects of acquired immunity to elucidate these immunological interactions. In doing so, it is important to address the discourse between theoretical modelers and immuno-epidemiologists and develop mechanistic models that empirically define immunity parameters. This is of increasing significance in a climate of potential changing transmission dynamics following long-term implementation of MDA.

Keywords: schistosomes; IgE; immunity; transmission dynamics

The age-infection intensity curve for all three major human-infecting schistosomes follows a pattern of increasing intensity through cumulative exposure until adolescence. After this infection intensities drop, irrespective of water exposure patterns [1]. It is therefore proposed that protective immunity is acquired with cumulative experience of infection. Once an individual has been exposed to a threshold of antigen, an effective immune response is mounted. Experience of infection is determined by either duration of infection or intensity of exposure. Since individuals in high-transmission areas are exposed to a higher level and greater diversity of antigens earlier in their lifetime, peak infection intensity is greater and occurs at an earlier age, when compared to areas of lower infection intensity [2]. This observation is termed the ‘peak shift’ [3,4]. Treatment with the antihelminthic praziquantel has been the cornerstone of morbidity control since the 1980s. Praziquantel kills adult schistosome worms, thereby providing passive immunisation through an antigenic stimulus delivered as adult worms die. Chemotherapy affects the development of individual protective immunity [4–6], but may also have the potential to considerably affect transmission at the community level [7].

Theoretical models are useful for policy making and can be used in both planning and evaluating the impact of control interventions on a broad epidemiological scale. For theoretical models of transmission to be useful they must have sufficient biological basis. The challenge of developing a schistosomiasis transmission model is in striking the balance between reality and parsimony. Schistosomiasis is highly focal, demonstrating significant spatial heterogeneity both between village and within village. This has been empirically demonstrated when studying the social, behavioural and cultural determinants of transmission on a microgeographical scale [8]. Likewise, immunological

processes are studied on a microgeographical scale, relevant to the mechanisms and pathways that drive the development of immunity within populations studied. A model that aims to independently quantify the contribution of each modifying factor will be complex and prove challenging to draw useful analytical conclusions from [9,10]. Furthermore, model validity is likely to be restricted to the population against which the model is tested. Population and environmental heterogeneities may, however, play a significant role in dictating transmission dynamics [11].

The computational predictions of mathematical models are only as good as the data against which they are calibrated and validated. Models are based upon the existing knowledge of biological processes and the data available to describe and quantify these interactions [12]. Recent studies continue to base the dynamics of infection solely on the explanation of variable exposure with age [13–15], despite considerable field evidence to support the role of acquired immunity. It is important that models evolve alongside the emergence of experimental evidence supporting model parameterization [16]. When immunological concepts have been modeled, immune processes have typically been described phenomenologically, with interactions based on assumptions with limited empirical backing [12]. A shortage of field data to inform quantitative estimates for parameter values and limits to theoretical understanding have frequently been cited as reasons for not exploring the contribution of immunity to transmission dynamics in more detail [13,17]. Furthermore, immuno-epidemiological studies have not historically been designed to provide empirical evidence for parameter values directly [16,17].

1. Required Immunological Parameters

“A closer collaboration between biometricians and parasitologists, and a better acquaintanceship of each with the methods of the other, is one of the most useful things we can work for today” [18,19].

When studying the dynamics of infectious diseases interdisciplinary collaboration is essential. This is particularly important in cases of such complex social and ecological interactions as schistosomiasis. It is therefore important to identify specific areas where there is a greater need for quantitative estimates of model parameters. This is especially important with the recent clear shift from a focus on control towards discussion of elimination efforts. Subsequently, a number of recent models have attempted to predict the requirements and impacts of mass drug administration (MDA) in reaching proposed elimination targets [13,20,21].

It has been argued that models that do not incorporate the dynamics of acquired immunity lack power in reproducing age-intensity profiles [13,22] and may overestimate the positive impact of control achieved by MDA [23,24]. This would occur if reduced transmission delays the development of immunity. Existing transmission models incorporating protective immunity suggest that worm lifespan and the rate of immune decay significantly contribute to the magnitude of specific antibody response [5,20]. Moreover, as many communities have now experienced access to mass drug administration with praziquantel, to a varying degree [25], fluctuations in antibody magnitude are likely to be occurring. There is therefore a need for field studies that are explicitly designed to capture epidemiological effects of acquired immunity. This will aid model development through provision of parameter estimates for transmission models [9].

The following gaps in our theoretical understanding and data requirements have been highlighted in phenomenological examination of immune responses, some of which relate to immuno-epidemiological study design, some specifically to immune mechanisms:

- Information on the demographic profile of communities from which immunological cohorts are drawn [15].
- Longitudinal studies of age-specific correlations between parasite burden and immune responses, incorporating individual data from a full range of age classes, in addition to comparison of communities with differing overall levels of endemic infection [13,26].
- Better knowledge of patterns of exposure experienced by immuno-epidemiological cohorts [26], requirement for water contact studies including cultural and sociological analysis [8].

- Specific analysis of age-dependent immunological parameters, including immunological responsiveness (strength of protective response) and loss of immunological memory [5,20,26], with particular reference to the parameterisation of protective immunity.
- Effects of inappropriate responses such as blocking antibodies and parasite-induced immunomodulation [26].
- The impact of immune responses that affect parasite fecundity and host excretion of eggs [20,26].

2. Delayed Concomitant Immunity

Mechanisms that confer partial immunity to schistosomes are not fully understood due to reliance on immuno-epidemiological evidence, with the murine experimental model being largely inappropriate for modelling human resistance to infection. However, levels of IgE specific to schistosome antigens, particularly antigens derived from the adult worm (SWA), are negatively associated with reinfection in studies examining all three major human-infecting schistosomes [27–29].

While historically a switch in the balance between SWA-IgE and the blocking antibody SWA-IgG4 was proposed as the key to immunity, with adults having a greater SWA-IgE:IgG4 ratio than children [30,31], the current leading hypothesis of the mechanism of immunity is that exposure to worm death is required for the development of immunity [2,6]. This has been referred to as delayed concomitant immunity [32], a development on the concomitant immunity first described by Smithers and Terry, in which established live worms mediate an immune response that prevents establishment of further parasites [33]. In the model of delayed concomitant immunity, the source of antigen to which the IgE is raised is the adult worms, upon worm death, while the focus of immune attack remains the early schistosomule. From an evolutionary point of view, this represents a stand-off between the parasite and the host—sufficient time is awarded to the parasite to produce the next generation, while the host is protected against the build-up of an excessive worm burden. Biologically this hypothesis is also compelling, with the adult worm being a master of immune evasion, successfully living within the blood stream for an estimated 3.5–10.5 years depending on species [34,35], while the transformation undergone during the transition from free-living cercaria to host-dwelling schistosomule provides a point of vulnerability within the parasite's life history for the host to attack.

This hypothesis of delayed concomitant immunity is best described through the responses to members of the *Schistosoma mansoni*-tegumental allergen-like (SmTALs) protein family. SmTAL1 (previously Sm22.6), the first of the family to be described, is the major IgE antigen detected on western blots probed with sera from endemic populations [29]. SmTAL1-IgE levels rise with age (Figure 1a), commensurate with the drop in infection intensities in adolescence, and have been shown to be negatively associated with reinfection intensities [36,37]. SmTAL1 is expressed from the 24 h schistosomule onwards, mainly within the tegument of the adult worm, where it is not substantially exposed to the host by live worms. IgE levels to SmTAL1 increase in the weeks after praziquantel treatment, confirming increased exposure of the antigen to the host during worm death [38]. SmTAL3, also expressed within the adult worm, follows a similar increase in detectable IgE with age, but with SmTAL3-IgE responders being a subset of SmTAL1-IgE responders (Figure 1b) [32,37]; thus, it is proposed that a greater exposure to dying worms is required for a SmTAL3-IgE response. If true, this would be illustrative of an antigen-threshold effect occurring for IgE responses to antigens of this family.

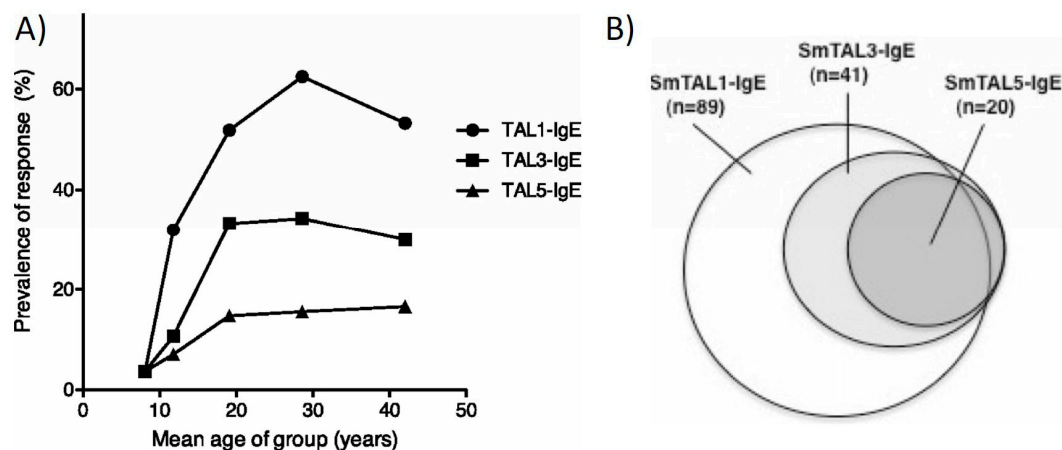


Figure 1. IgE Responses to SmTAL1, SmTAL3 and SmTAL5. Shown are the 9-week post-treatment IgE responses of a cohort of males, aged 7–60 years, resident in a Ugandan fishing village of high endemicity. (A) Prevalence of detectable IgE specific to SmTAL1, SmTAL3 and SmTAL5 by age; (B) Venn diagram displaying sequentiality of detectable of IgE specific to SmTAL1, SmTAL3 and SmTAL5. Figures are reproduced with permission from [32].

Responsiveness to SmTAL2, which is expressed throughout the life cycle, offers an illustrative counterpoint to SmTAL1/3 responses. An IgE response to this protein is measurable, but only amongst the very youngest members of populations in endemic areas, ultimately being regulated by IgG4 [37,39]. The constant exposure to SmTAL2 from dying eggs within the tissue, in contrast to the intermittent exposure to SmTAL1 and SmTAL3 from the adults, makes it imperative that the host down-regulates the potentially damaging IgE response to SmTAL2, a down-regulation that is not required for the members of the family cryptically expressed in adults. However, with SmTAL1 and 3 not expressed in the earliest schistosomule stages, as would be required if this stage is the target of immunity, IgE response to SmTAL1 and SmTAL3 will not be explanatory of immunity, and indeed after controlling for age, IgE responses to these antigens are not significantly associated with immunity [32].

Immunity via exposure to dying adult worms is not easily reproduced in murine models, with long-term egg deposition altering vasculature, and hence migratory pathways of worms infecting from trickle exposures. Immuno-epidemiological studies suggest a slow-developing IgE-mediated mechanism, rather than sentinel myeloid cell-derived regulatory immune responses, as seen in a multiple (pre-egg laying) cercarial challenge murine model [40]. It is therefore hypothesised that skin resident mast cells would be the first responders in immune individuals [41]. This will dramatically alter the immune profile of the response upon cercarial challenge from one of regulation to one characterised by IgE mediated by type-1 hypersensitivity reactions. Potential targets include SmTAL5, for which age-dependent IgE responsiveness is observed (Figure 1a) [37]. Expressed in the early schistosomules, SmTAL5 displays IgE cross-reactivity with SmTAL3, with SmTAL3 appearing to be the antigenic source for this IgE response [32]. It is this cross-reactivity between the adult worm expressed SmTAL3 and the early schistosomule expressed SmTAL5 that is proposed to exert the delayed concomitant immunity, though it is unlikely that this is the sole cross-reactive antigen that is involved in the development of partial immunity.

Eosinophil infiltration has also been implicated in protective human IgE responsiveness. In vitro eosinophils can mediate killing of schistosomules in an IgE-mediated manner [42]. In *S. haematobium* infection, eosinophils from Gambian individuals with the greatest expansion in eosinophil number have the greatest capability to kill schistosomules [43], and in adult Kenyan car-washers infected with *S. mansoni*, both eosinophil numbers, and their expression of the tetrameric high-affinity IgE receptor on their surface, are associated with immunity to reinfection [44]. It is tempting to hypothesise that cross-linking of IgE on mast cells leads to chemotactic infiltration of eosinophils, which then

release their granular contents to kill the schistosomules, but without a valid model of IgE-armed skin responses this remains a hypothesis.

3. IgE Responses and Praziquantel Treatment

The proposed reliance on worm death for development of protective immunity to schistosomes suggests that treatment with praziquantel could provide a 'passive vaccination' effect, boosting the development of immunity. This passive vaccination effect is eloquently illustrated in the studies of car-washers conducted in Kisumu, Kenya, in which praziquantel treatment was administered after every detectable *S. mansoni* reinfection. As the number of treatments administered increased, the longer the observed duration to reinfection [45]. In comparative studies enrolling longer-term exposed sand-harvesters, this immunity was observable after two treatments, but an extended process of 7 infections and treatment was required for car-washers who were more recently exposed to infection. In agreement with our hypothesis, at baseline the greater exposed sand-harvesters had greater circulating levels of SmTAL1-IgE, indicating that this group had moved further along the path towards immunity prior to the commencement of the study [46]. During the reinfection-treatment cycles, elements of the immune system involved in IgE production, in particular soluble and B cell membrane-bound levels of CD23, the low-affinity IgE receptor, became elevated in comparison to baseline [47]. The elevation of these immune elements was also observed in a cohort of schoolchildren treated four times annually [48].

Immunologically, elevated levels of SWA and SmTAL1-, 3- and 5-specific IgE are observed 3–9 weeks post-treatment. These post-treatment IgE levels are better correlates of resistance to reinfection as measured 1–2 years post-treatment, than pre-treatment antibody levels [32,38,49]. This boost in IgE to adult worm-derived antigens, and antigens cross-reactive with worm-derived antigens, is accompanied by an increase in histamine release upon antigen stimulation of basophils, indicating that the increased IgE is biologically active [50]. The elevated specific IgE levels are also accompanied by an increase in immune cell proliferation [51] and production of type-2 immune cytokines, such as IL-4, 5 and 13 by peripheral blood mononuclear cell and whole blood cultures [52,53]. IL-5, an eosinophil maturation factor, in particular, is elevated in post-treatment supernatants of SWA stimulated immune cells [54,55]. For Ugandan children residing in high-transmission fishing villages, the pre-treatment strength of this IL-5 cytokine response is associated with levels of SWA-IgE measured post-treatment with praziquantel [56]. Mirroring this ex vivo IL-5 response are circulating 24 h post-treatment IL-5 levels measured in the plasma [57]. For *S. haematobium* infection, the IL-5 levels detected 24 h post-treatment are associated with pre-treatment SWA-IgE and eosinophil levels. Furthermore, post-treatment IL-5 is associated with pre-treatment infection intensity, indicating that both the primed state of the immune system, or prior exposure to dying worms, and the amount of antigen the system is exposed to upon treatment, is crucial in generating this IL-5 response. In turn, 24 h post-treatment levels of IL-5 are positively associated with protective post-treatment SWA-IgE levels and elevated circulating eosinophil numbers [58].

4. Rate of Decay of Protective Immunity

What is clear from the above immunological evidence is that treatment with praziquantel can give a boost to proposed elements of a protective immune response; that is, type 2 cytokine responsiveness, levels of specific IgE and eosinophil numbers. Moreover, this could be providing a 'passive vaccination' effect, allowing individuals in endemic areas to develop their immunity at an earlier age. While this is encouraging, as stated earlier, when conducting reinfection studies immuno-epidemiologists work on a microgeographical scale and tend to seek out study sites with high transmission, or particularly highly-exposed occupational cohorts, as this allows us to examine correlates of immunity. Little is known about whether a significant vaccination effect occurs if transmission is low or, as is the aim with multifaceted control strategies, driven to be being very low. Successful control programs have, over time, altered the distribution of schistosomiasis. French et al. [59,60], noted a change in force

of infection following repeated MDA campaigns in five countries where the Schistosomiasis Control Initiative has been working for the past decade. In an era where control is primarily based on upscaling chemotherapy, it is increasingly important to understand the impact of MDA on transmission dynamics. Are the levels of antigen exposure induced by annual treatment going to be sufficient to drive a vaccinating effect? Or indeed, by driving transmission to a point when it is very low, but continuing, will antigen exposure be limited to the point that it takes into late adolescence or early adulthood for antigen thresholds to be passed and immunity to have fully developed? Early modeling studies indicate that mass chemotherapy may reduce immunity at the community level, to the extent that, following cessation of control, parasite load in the older age classes may rise above the pre-control burden [7]. The latter will have implications for school-based strategies if the focus of control moves further from public health morbidity control strategies towards interruption of transmission.

Mitchell et al. [20] related protective immunity to reduced egg production, a phenomenon observed in *S. haematobium* field studies, but not seen for *S. mansoni* infections [61,62]. It was concluded that MDA will initially boost protective immunity, but antibody levels could decline below pre-treatment levels during or after cessation of MDA. The model is calibrated on short-term pre- and post-treatment studies of *S. haematobium* infection, but does support the hypothesis that MDA may disrupt the build-up of protective immunity. They also proposed that if reduced transmission led to reduced antigen exposure, antibody levels will be strongly influenced by immune decay rates and demonstrated a rebound in egg excretion following cessation of MDA activities [20]. It is important to note that this paper does not examine protective response against reinfection. Anti-fecundity and anti-infection immunity are likely mediated by different antibody classes [62], therefore have differing mechanisms of induction and preservation by the memory component of the immune system.

In relation to the immune decay rate of IgE, biologically-active IgE may be sustained for months after exposure to the initial stimuli of production due to long-lived IgE plasma cells maintaining IgE production [63]. Yet, when we conducted a longitudinal study looking at immune factors when annual praziquantel treatment of Kenyan school-aged children was combined with biannual mollusciciding of the one source of transmission in the area, the levels of circulating SWA-IgE were significantly lower one and two years after the baseline survey (unpublished results). This suggests that antigen exposure is required to maintain production of SWA-IgE at pre-treatment levels. Basophil bioassays were not conducted so we do not know whether SWA-IgE was reduced to a level sufficiently low to prevent cross-linking of IgE bound to receptors on effector cells. At the same time-points, however, whole blood cultures produced elevated SWA-specific type-2 cytokines, indicating that T cell memory is robust and maintained over that extended period [64]. Although it is thought that B cells switched to IgE are very poorly recruited to the memory compartment of the immune system [63,65], it has recently been shown in murine models of peanut allergy that a memory response is crucial to re-activation of clinical IgE reactivity, with IL-4 from T cells driving IgG memory B cells to switch and mature into IgE plasmablasts [66]. In humans, grass pollen allergic rhinitis sufferers have higher numbers of circulating memory B cells that proliferate upon allergen exposure than control subjects [67]. It is therefore plausible that if transmission is still occurring, the memory component of the immune response could maintain or re-establish protective IgE-mediated immunity upon intermittent exposure to worm death, the long-term dynamics of which will depend on the local transmission patterns.

5. Variation in Estimates of Immune Parameters

An important consideration is that numerical analyses are necessarily based upon a restricted set of parameter combinations [26]. These broad parameter estimates are likely to vary between locations [68], with authors highlighting the importance of using locality-specific values to generate predictions from models [69]. The generalisability of these models from one situation to another is consequently restricted [12]. In particular reference to immunological parameters, Chan et al. [68] argue that immuno-epidemiological studies on the individual community level fail to capture the true dynamics of transmission as a result of variability in parasitological examination and antibody

titres. They therefore proposed the use of mathematical models to inform the design of field studies, a concept further discussed in relation to ecological modelling by Restif et al. [16]. Chan et al. [68] advocate that model-guided design of fieldwork can strengthen the power of data analysis and inferences. They also suggest that the village is the smallest unit considered when studying patterns of transmission. Immunologically, this leads to the technical hurdle of scale; however, advancing technology for sample processing, including liquid handling capabilities and associated reductions in sample volume requirements, are now such that the potential to conduct serologically-based immunological studies on a wider scale does exist.

We believe, given the field-based evidence, that more mechanistic modelling of immunity in future schistosome transmission models is required. We have provided discussion, from an immunological perspective, on what is known from existing field data, but acknowledge that previous immunology studies, through scale and design, may fail to capture the parameters that modelers identify as important. However, with technological advances, an integrated approach between theoretical modelers and field immunologists is now possible and of utmost importance if transmission models are to truly capture schistosome infection dynamics within endemic areas.

Acknowledgments: Rebecca C. Oettle is funded by the UK Medical Research Council (DTG PhD). Shona Wilson is funded by the Royal Society (Grant AA130107) and the UK Medical Research Council (Grant MR/M019780/1).

Author Contributions: R.C.O. and S.W. wrote the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kabatereine, N.B.; Vennervald, B.J.; Ouma, J.H.; Kemijumbi, J.; Butterworth, A.E.; Dunne, D.W.; Fulford, A.J. Adult resistance to schistosomiasis mansoni: Age-dependence of reinfection remains constant in communities with diverse exposure patterns. *Parasitology* **1999**, *118*, 101–105. [[CrossRef](#)] [[PubMed](#)]
2. Woolhouse, M.E.; Hagan, P. Seeking the ghost of worms past. *Nat. Med.* **1999**, *5*, 1225–1227. [[CrossRef](#)] [[PubMed](#)]
3. Woolhouse, M.E.J.; Taylor, P.; Matanhire, D.; Chandiwana, S.K. Acquired immunity and epidemiology of *Schistosoma haematobium*. *Nature* **1991**, *351*, 757–759. [[CrossRef](#)] [[PubMed](#)]
4. Fulford, A.; Butterworth, A.E.; Sturrock, R.F.; Ouma, J.H. On the use of age-intensity data to detect immunity to parasitic infections, with special reference to *Schistosoma mansoni* in Kenya. *Parasitology* **1992**, *105 Pt 2*, 219–227. [[CrossRef](#)] [[PubMed](#)]
5. Chan, M.S.; Anderson, R.M.; Medley, G.F.; Bundy, D.A.P. Dynamic aspects of morbidity and acquired immunity in schistosomiasis control. *Acta Trop.* **1996**, *62*, 105–117. [[CrossRef](#)]
6. Mutapi, F.; Ndhlovu, P.D.; Hagan, P.; Spicer, J.T.; Mduluzi, T.; Turner, C.M.; Chandiwana, S.K.; Woolhouse, M.E. Chemotherapy accelerates the development of acquired immune responses to *Schistosoma haematobium* infection. *J. Infect. Dis.* **1998**, *178*, 289–293. [[CrossRef](#)] [[PubMed](#)]
7. Anderson, R.M.; May, R.M. Herd immunity to helminth infection and implications for parasite control. *Nature* **1985**, *315*, 493–496. [[CrossRef](#)]
8. De Moira, A.P.; Fulford, A.J.C.; Kabatereine, N.B.; Kazibwe, F.; Ouma, J.H.; Dunne, D.W.; Booth, M. Microgeographical and tribal variations in water contact and *Schistosoma mansoni* exposure within a Ugandan fishing community. *Trop. Med. Int. Health* **2007**, *12*, 724–735. [[CrossRef](#)] [[PubMed](#)]
9. Woolhouse, M.E. On the application of mathematical models of schistosome transmission dynamics. I. Natural transmission. *Acta Trop.* **1991**, *49*, 241–270. [[CrossRef](#)]
10. Woolhouse, M.E.J. On the application of mathematical models of schistosome transmission dynamics. II. Control. *Acta Trop.* **1992**, *50*, 189–204. [[CrossRef](#)]
11. Woolhouse, M.E.; Etard, J.F.; Dietz, K.; Ndhlovu, P.D.; Chandiwana, S.K. Heterogeneities in schistosome transmission dynamics and control. *Parasitology* **1998**, *117*, 475–482. [[CrossRef](#)] [[PubMed](#)]
12. Heesterbeek, H.; Anderson, R.; Andreasen, V.; Bansal, S.; De, D.; Dye, C.; Eames, K.; Edmunds, J.; Frost, S.; Funk, S.; et al. Modeling infectious disease dynamics in the complex landscape of global health. *Science* **2015**, *347*, aaa4339. [[CrossRef](#)] [[PubMed](#)]

13. Anderson, R.M.; Turner, H.C.; Farrell, S.H.; Yang, J.; Truscott, J.E. What is required in terms of mass drug administration to interrupt the transmission of schistosome parasites in regions of endemic infection? *Parasit Vectors* **2015**, *8*, 553. [[CrossRef](#)] [[PubMed](#)]
14. Gurarie, D.; King, C.H.; Yoon, N.; Li, E. Refined stratified-worm-burden models that incorporate specific biological features of human and snail hosts provide better estimates of *Schistosoma* diagnosis, transmission, and control. *Parasit Vectors* **2016**, *9*, 428. [[CrossRef](#)] [[PubMed](#)]
15. Truscott, J.E.; Gurarie, D.; Alsallaq, R.; Toor, J.; Yoon, N.; Farrell, S.H.; Turner, H.C.; Phillips, A.E.; Aurelio, H.O.; Ferro, J.; et al. A comparison of two mathematical models of the impact of mass drug administration on the transmission and control of schistosomiasis. *Epidemics* **2017**, *18*, 29–37. [[CrossRef](#)] [[PubMed](#)]
16. Restif, O.; Hayman, D.T.S.; Pulliam, J.R.C.; Plowright, R.K.; George, D.B.; Luis, A.D.; Cunningham, A.A.; Bowen, R.A.; Fooks, A.R.; O'Shea, T.J.; et al. Model-guided fieldwork: Practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics. *Ecol. Lett.* **2012**, *15*, 1083–1094. [[CrossRef](#)] [[PubMed](#)]
17. Allen, E.J.; Victory, H.D. Modelling and simulation of a schistosomiasis infection with biological control. *Acta Trop.* **2003**, *87*, 251–267. [[CrossRef](#)]
18. Hackett, L.W. *Malaria in Europe: An Ecological Study*; Oxford University Press: London, UK, 1937.
19. Anderson, R.M.; Turner, H.C.; Farrell, S.H.; Truscott, J.E. Studies of the transmission dynamics, mathematical model development and the control of schistosome parasites by mass drug administration in human communities. *Adv. Parasitol.* **2016**, *94*, 199–246. [[CrossRef](#)] [[PubMed](#)]
20. Mitchell, K.M.; Mutapi, F.; Mduluzi, T.; Midzi, N.; Savill, N.J.; Woolhouse, M.E.J. Predicted impact of mass drug administration on the development of protective immunity against *Schistosoma haematobium*. *PLoS Negl. Trop. Dis.* **2014**, *8*, e3059. [[CrossRef](#)] [[PubMed](#)]
21. Gurarie, D.; Yoon, N.; Li, E.; Ndeffo-mbah, M.; Durham, D.; Phillips, A.E.; Aurelio, H.O.; Ferro, J.; Galvani, A.P.; King, C.H. Modelling control of *Schistosoma haematobium* infection: Predictions of the long-term impact of mass drug administration in Africa. *Parasit. Vectors* **2015**, *8*, 1–14. [[CrossRef](#)] [[PubMed](#)]
22. Yang, H.M.; Coutinho, F.A.; Massad, E. Acquired immunity on a schistosomiasis transmission model—Fitting the data. *J. Theor. Biol.* **1997**, *188*, 495–506. [[CrossRef](#)] [[PubMed](#)]
23. Gurarie, D.; King, C.H. Heterogeneous model of schistosomiasis transmission and long-term control: The combined influence of spatial variation and age-dependent factors on optimal allocation of drug therapy. *Parasitology* **2005**, *130*, 49–65. [[CrossRef](#)] [[PubMed](#)]
24. Gurarie, D.; King, C.H. Population biology of *Schistosoma* mating, aggregation, and transmission breakpoints: More reliable model analysis for the end-game in communities at risk. *PLoS ONE* **2014**, *9*, e115875. [[CrossRef](#)] [[PubMed](#)]
25. Mutapi, F.; Maizels, R.; Fenwick, A.; Woolhouse, M. Human schistosomiasis in the post-mass drug administration era. *Lancet Infect. Dis.* **2017**, *17*, e42–e48. [[CrossRef](#)]
26. Woolhouse, M.E.J. A theoretical framework for the immunoepidemiology of helminth infection. *Parasite Immunol.* **1992**, *14*, 563–578. [[CrossRef](#)] [[PubMed](#)]
27. Jiz, M.; Friedman, J.F.; Leenstra, T.; Jarilla, B.; Pablo, A.; Langdon, G.; Pond-Tor, S.; Wu, H.-W.; Manalo, D.; Olveda, R.; et al. Immunoglobulin E (IgE) responses to paramyosin predict resistance to reinfection with *Schistosoma japonicum* and are attenuated by IgG4. *Infect. Immun.* **2009**, *77*, 2051–2058. [[CrossRef](#)] [[PubMed](#)]
28. Hagan, P.; Blumenthal, U.J.; Dunn, D.; Simpson, A.J.G.; Wilkins, H.A. Human IgE, IgG4 and resistance to reinfection with *Schistosoma haematobium*. *Nature* **1991**, *349*, 243–245. [[CrossRef](#)] [[PubMed](#)]
29. Dunne, D.W.; Butterworth, A.E.; Fulford, A.J.C.; Curtis Kariuki, H.; Langley, J.G.; Ouma, J.H.; Capron, A.; Pierce, R.J.; Sturrock, R.F. Immunity after treatment of human schistosomiasis: Association between IgE antibodies to adult worm antigens and resistance to reinfection. *Eur. J. Immunol.* **1992**, *22*, 1483–1494. [[CrossRef](#)] [[PubMed](#)]
30. Butterworth, A.E.; Bensted-Smith, R.; Capron, A.; Capron, M.; Dalton, P.R.; Dunne, D.W.; Grzych, J.M.; Kariuki, H.C.; Khalife, J.; Koech, D. Immunity in human schistosomiasis mansoni: Prevention by blocking antibodies of the expression of immunity in young children. *Parasitology* **1987**, *94*, 281–300. [[CrossRef](#)] [[PubMed](#)]

31. Demeure, C.E.; Rihet, P.; Abel, L.; Ouattara, M.; Bourgois, A.; Dessein, A.J. Resistance to *Schistosoma mansoni* in humans: Influence of the IgE/IgG4 balance and IgG2 in immunity to reinfection after chemotherapy. *J. Infect. Dis.* **1993**, *168*, 1000–1008. [[CrossRef](#)] [[PubMed](#)]
32. Fitzsimmons, C.M.; Jones, F.M.; Pinot De Moira, A.; Protasio, A.V.; Khalife, J.; Dickinson, H.A.; Tukahebwa, E.M.; Dunne, D.W. Progressive cross-reactivity in IgE responses: An explanation for the slow development of human immunity to schistosomiasis? *Infect. Immun.* **2012**, *80*, 4264–4270. [[CrossRef](#)] [[PubMed](#)]
33. Smithers, S.R.; Terry, R.J. Resistance to experimental infection with *Schistosoma mansoni* in rhesus monkeys induced by the transfer of adult worms. *Trans. R. Soc. Trop. Med. Hyg.* **1967**, *61*, 517–533. [[CrossRef](#)]
34. Wilkins, H.A.; Goll, P.H.; Marshall, T.F.; Moore, P.J. Dynamics of *Schistosoma haematobium* infection in a Gambian community. III. Acquisition and loss of infection. *Trans. R. Soc. Trop. Med. Hyg.* **1984**, *78*, 227–232. [[CrossRef](#)]
35. Fulford, A.J.; Butterworth, A.E.; Ouma, J.H.; Sturrock, R.F. A statistical approach to schistosome population dynamics and estimation of the life-span of *Schistosoma mansoni* in man. *Parasitology* **1995**, *110 Pt 3*, 307–316. [[CrossRef](#)] [[PubMed](#)]
36. Webster, M.; Fulford, A.J.C.; Braun, G.; Ouma, J.H.; Kariuki, H.C.; Havercroft, J.C.; Gachuhi, K.; Sturrock, R.F.; Butterworth, A.E.; Dunne, D.W. Human immunoglobulin E responses to a recombinant 22.6-kilodalton antigen from *Schistosoma mansoni* adult worms are associated with low intensities of reinfection after treatment. *Infect. Immun.* **1996**, *64*, 4042–4046. [[PubMed](#)]
37. Fitzsimmons, C.M.; Jones, F.M.; Stearn, A.; Chalmers, I.W.; Hoffmann, K.F.; Wawrzyniak, J.; Wilson, S.; Kabatereine, N.B.; Dunne, D.W.; Oliveira, S.C. The *Schistosoma mansoni* tegumental-allergen-like (TAL) protein family: Influence of developmental expression on human IgE responses. *PLoS Negl. Trop. Dis.* **2012**, *6*, e1593. [[CrossRef](#)] [[PubMed](#)]
38. Pinot De Moira, A.; Jones, F.M.; Wilson, S.; Tukahebwa, E.; Fitzsimmons, C.M.; Mwatha, J.K.; Bethony, J.M.; Kabatereine, N.B.; Dunne, D.W. Effects of treatment on IgE responses against parasite allergen-like proteins and immunity to reinfection in childhood schistosome and hookworm coinfections. *Infect. Immun.* **2013**, *81*, 23–32. [[CrossRef](#)] [[PubMed](#)]
39. De Moira, A.P.; Sousa-Figueiredo, J.C.; Jones, F.M.; Fitzsimmons, C.M.; Betson, M.; Kabatereine, N.B.; Stothard, J.R.; Dunne, D.W. *Schistosoma mansoni* infection in preschool-aged children: Development of immunoglobulin E and immunoglobulin G4 responses to parasite allergen-like proteins. *J. Infect. Dis.* **2012**, *207*, 362–366. [[CrossRef](#)] [[PubMed](#)]
40. Cook, P.C.; Aynsley, S.A.; Turner, J.D.; Jenkins, G.R.; Van Rooijen, N.; Leeto, M.; Brombacher, F.; Mountford, A.P. Multiple helminth infection of the skin causes lymphocyte hypo-responsiveness mediated by Th2 conditioning of dermal myeloid cells. *PLoS Pathog.* **2011**, *7*, e1001323. [[CrossRef](#)] [[PubMed](#)]
41. Ganley-Leal, L.M.; Mwinzi, P.N.M.; Cetre-Sossah, C.B.; Andove, J.; Hightower, A.W.; Karanja, D.M.S.; Colley, D.G.; Secor, W.E. Higher percentages of circulating mast cell precursors correlate with susceptibility to reinfection with *Schistosoma mansoni*. *Am. J. Trop. Med. Hyg.* **2006**, *75*, 1053–1057. [[PubMed](#)]
42. Soussi Gounni, A.; Lamkhioued, B.; Ochiai, K.; Tanaka, Y.; Delaporte, E.; Capron, A.; Kinet, J.-P.; Capron, M. High-affinity IgE receptor on eosinophils is involved in defence against parasites. *Nature* **1994**, *367*, 183–186. [[CrossRef](#)] [[PubMed](#)]
43. Hagan, P.; Wilkins, H.A.; Blumenthal, U.J.; Hayes, R.J.; Greenwood, B.M. Eosinophilia and resistance to *Schistosoma haematobium* in man. *Parasite Immunol.* **1985**, *7*, 625–632. [[CrossRef](#)] [[PubMed](#)]
44. Ganley-Leal, L.M.; Mwinzi, P.N.; Cetre-Sossah, C.B.; Andove, J.; Hightower, A.W.; Karanja, D.M.S.; Colley, D.G.; Secor, W.E. Correlation between eosinophils and protection against reinfection with *Schistosoma mansoni* and the effect of human immunodeficiency virus type 1 coinfection in humans. *Infect. Immun.* **2006**, *74*, 2169–2176. [[CrossRef](#)] [[PubMed](#)]
45. Karanja, D.M.; Hightower, A.W.; Colley, D.G.; Mwinzi, P.N.; Galil, K.; Andove, J.; Secor, W.E. Resistance to reinfection with *Schistosoma mansoni* in occupationally exposed adults and effect of HIV-1 co-infection on susceptibility to schistosomiasis: A longitudinal study. *Lancet* **2002**, *360*, 592–596. [[CrossRef](#)]
46. Black, C.L.; Mwinzi, P.N.M.; Muok, E.M.O.; Abudho, B.; Fitzsimmons, C.M.; Dunne, D.W.; Karanja, D.M.S.; Secor, W.E.; Colley, D.G. Influence of exposure history on the immunology and development of resistance to human schistosomiasis mansoni. *PLoS Negl. Trop. Dis.* **2010**, *4*, e637. [[CrossRef](#)] [[PubMed](#)]

47. Mwinzi, P.N.M.; Ganley-Leal, L.; Black, C.L.; Secor, W.E.; Karanja, D.M.S.; Colley, D.G. Circulating CD23⁺ B cell subset correlates with the development of resistance to *Schistosoma mansoni* reinfection in occupationally exposed adults who have undergone multiple treatments. *J. Infect. Dis.* **2009**, *199*, 272–279. [[CrossRef](#)] [[PubMed](#)]
48. Black, C.L.; Muok, E.M.O.; Mwinzi, P.N.M.; Carter, J.M.; Karanja, D.M.S.; Secor, W.E.; Colley, D.G. Increases in levels of *Schistosoma*-specific immunoglobulin E and CD23⁺ B cells in a cohort of Kenyan children undergoing repeated treatment and reinfection with *Schistosoma mansoni*. *J. Infect. Dis.* **2010**, *202*, 399–405. [[CrossRef](#)] [[PubMed](#)]
49. Pinot De Moira, A.; Fulford, A.J.C.; Kabatereine, N.B.; Ouma, J.H.; Booth, M.; Dunne, D.W.; Quinnell, R.J. Analysis of complex patterns of human exposure and immunity to *Schistosomiasis mansoni*: The influence of age, sex, ethnicity and IgE. *PLoS Negl. Trop. Dis.* **2010**, *4*, e820. [[CrossRef](#)] [[PubMed](#)]
50. Pinot De Moira, A.; Fitzsimmons, C.M.; Jones, F.M.; Wilson, S.; Cahen, P.; Tukahebwa, E.; Mpairwe, H.; Mwatha, J.K.; Bethony, J.M.; Skov, P.S.; et al. Suppression of basophil histamine release and other IgE-dependent responses in childhood *Schistosoma mansoni*/hookworm coinfection. *J. Infect. Dis.* **2014**, *210*, 1198–1206. [[CrossRef](#)] [[PubMed](#)]
51. Colley, D.G.; Barsoum, I.S.; Dahawi, H.S.; Gamil, F.; Habib, M.; el Alamy, M.A. Immune responses and immunoregulation in relation to human schistosomiasis in Egypt. III. Immunity and longitudinal studies of in vitro responsiveness after treatment. *Trans. R. Soc. Trop. Med. Hyg.* **1986**, *80*, 952–957. [[CrossRef](#)]
52. Medhat, A.; Shehata, M.; Bucci, K.; Mohamed, S.; Dief, A.D.; Badary, S.; Galal, H.; Nafeh, M.; King, C.L. Increased interleukin-4 and interleukin-5 production in response to *Schistosoma haematobium* adult worm antigens correlates with lack of reinfection after treatment. *J. Infect. Dis.* **1998**, *178*, 512–519. [[CrossRef](#)] [[PubMed](#)]
53. Joseph, S.; Jones, F.M.; Walter, K.; Fulford, A.J.; Kimani, G.; Mwatha, J.K.; Kamau, T.; Kariuki, H.C.; Kazibwe, F.; Tukahebwa, E.; et al. Increases in human T helper 2 cytokine responses to *Schistosoma mansoni* worm and worm-tegument antigens are induced by treatment with praziquantel. *J. Infect. Dis.* **2004**, *190*, 835–842. [[CrossRef](#)] [[PubMed](#)]
54. Van den Biggelaar, A.H.J.; Borrmann, S.; Kremsner, P.; Yazdanbakhsh, M. Immune responses induced by repeated treatment do not result in protective immunity to *Schistosoma haematobium*: Interleukin (IL)-5 and IL-10 responses. *J. Infect. Dis.* **2002**, *186*, 1474–1482. [[CrossRef](#)] [[PubMed](#)]
55. Roberts, M.; Butlerworth, A.E.; Kimani, G.; Kamau, T.; Fulford, A.J.C.; Dunne, D.W.; Ouma, J.H.; Sturrock, R.F. Immunity after treatment of human schistosomiasis: Association between cellular responses and resistance to reinfection. *Infect. Immun.* **1993**, *61*, 4984–4993. [[PubMed](#)]
56. Walter, K.; Fulford, A.J.C.; McBeath, R.; Joseph, S.; Jones, F.M.; Kariuki, H.C.; Mwatha, J.K.; Kimani, G.; Kabatereine, N.B.; Vennervald, B.J.; et al. Increased human IgE induced by killing *Schistosoma mansoni* in vivo is associated with pretreatment Th2 cytokine responsiveness to worm antigens. *J. Immunol.* **2006**, *177*, 5490–5498. [[CrossRef](#)] [[PubMed](#)]
57. Fitzsimmons, C.M.; Joseph, S.; Jones, F.M.; Reimert, C.M.; Hoffmann, K.F.; Kazibwe, F.; Kimani, G.; Mwatha, J.K.; Ouma, J.H.; Tukahebwa, E.M.; et al. Chemotherapy for schistosomiasis in Ugandan fishermen: Treatment can cause a rapid increase in interleukin-5 levels in plasma but decreased levels of eosinophilia and worm-specific immunoglobulin E. *Infect. Immun.* **2004**, *72*, 4023–4030. [[CrossRef](#)] [[PubMed](#)]
58. Wilson, S.; Jones, F.M.; Fofana, H.K.M.; Doucouré, A.; Landouré, A.; Kimani, G.; Mwatha, J.K.; Sacko, M.; Vennervald, B.J.; Dunne, D.W. Rapidly boosted plasma IL-5 induced by treatment of human schistosomiasis haematobium is dependent on antigen dose, IgE and eosinophils. *PLoS Negl. Trop. Dis.* **2013**, *7*, e2149. [[CrossRef](#)] [[PubMed](#)]
59. French, M.D.; Churcher, T.S.; Gambhir, M.; Fenwick, A.; Webster, J.P.; Kabatereine, N.B.; Basáñez, M.G. Observed reductions in *Schistosoma mansoni* transmission from large-scale administration of praziquantel in Uganda: A mathematical modelling study. *PLoS Negl. Trop. Dis.* **2010**, *4*, e897. [[CrossRef](#)] [[PubMed](#)]
60. French, M.D.; Churcher, T.S.; Webster, J.P.; Fleming, F.M.; Fenwick, A.; Kabatereine, N.B.; Sacko, M.; Garba, A.; Toure, S.; Nyandindi, U.; et al. Estimation of changes in the force of infection for intestinal and urogenital schistosomiasis in countries with schistosomiasis control initiative-assisted programmes. *Parasit. Vectors* **2015**, *8*, 558. [[CrossRef](#)] [[PubMed](#)]

61. Agnew, A.; Fulford, A.J.; Mwanje, M.T.; Gachuhi, K.; Gutschmann, V.; Krijger, F.W.; Sturrock, R.F.; Vennervald, B.J.; Ouma, J.H.; Butterworth, A.E.; et al. Age-dependent reduction of schistosome fecundity in *Schistosoma haematobium* but not *Schistosoma mansoni* infections in humans. *Am. J. Trop. Med. Hyg.* **1996**, *55*, 338–343. [[CrossRef](#)] [[PubMed](#)]
62. Wilson, S.; Jones, F.M.; Van Dam, G.J.; Corstjens, P.L.; Riveau, G.; Fitzsimmons, C.M.; Sacko, M.; Vennervald, B.J.; Dunne, D.W. Human *Schistosoma haematobium* antifecundity immunity is dependent on transmission intensity and associated with immunoglobulin G1 to worm-derived antigens. *J. Infect. Dis.* **2014**, *210*, 2009–2016. [[CrossRef](#)] [[PubMed](#)]
63. He, J.-S.; Narayanan, S.; Subramaniam, S.; Ho, W.Q.; Lafaille, J.J.; Curotto de Lafaille, M.A. Biology of IgE production: IgE cell differentiation and the memory of IgE responses. *Curr. Top. Microbiol. Immunol.* **2015**, *388*, 1–19. [[CrossRef](#)] [[PubMed](#)]
64. Wilson, S.; Jones, F.M.; Kenty, L.-C.; Mwatha, J.K.; Kimani, G.; Kariuki, H.C.; Dunne, D.W. Posttreatment changes in cytokines induced by *Schistosoma mansoni* egg and worm antigens: Dissociation of immunity- and morbidity-associated Type 2 responses. *J. Infect. Dis.* **2014**, *209*, 1792–1800. [[CrossRef](#)] [[PubMed](#)]
65. Haniuda, K.; Fukao, S.; Kodama, T.; Hasegawa, H.; Kitamura, D. Autonomous membrane IgE signaling prevents IgE-memory formation. *Nat. Immunol.* **2016**, *17*, 1109–1117. [[CrossRef](#)] [[PubMed](#)]
66. Jimenez-Saiz, R.; Chu, D.K.; Mandur, T.S.; Walker, T.D.; Gordon, M.E.; Chaudhary, R.; Koenig, J.; Saliba, S.; Galipeau, H.J.; Utley, A.; et al. Lifelong memory responses perpetuate humoral TH 2 immunity and anaphylaxis in patients with food allergy. *J. Allergy Clin. Immunol.* **2017**. [[CrossRef](#)] [[PubMed](#)]
67. Wong, K.J.; Timbrell, V.; Xi, Y.; Upham, J.W.; Collins, A.M.; Davies, J.M. IgE⁺ B cells are scarce, but allergen-specific B cells with a memory phenotype circulate in patients with allergic rhinitis. *Allergy* **2015**, *70*, 420–428. [[CrossRef](#)] [[PubMed](#)]
68. Chan, M.S.; Mutapi, F.; Woolhouse, M.E.; Isham, V.S. Stochastic simulation and the detection of immunity to schistosome infections. *Parasitology* **2000**, *120 Pt 2*, 161–169. [[CrossRef](#)] [[PubMed](#)]
69. Chan, M.S.; Guyatt, H.L.; Bundy, D.A.; Booth, M.; Fulford, A.J.; Medley, G.F. The development of an age structured model for schistosomiasis transmission dynamics and control and its validation for *Schistosoma mansoni*. *Epidemiol. Infect.* **1995**, *115*, 325–344. [[CrossRef](#)] [[PubMed](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).