Chapter 4: Oligonucleotide Microarray Analysis of Strain-Associated Gene Expression in the Human Blood Fluke, *Schistosoma japonicum*.

4.1 Introduction

The Chinese (SJC) and Philippine (SJP) strains of *S. japonicum* exhibit a number of morphological and other phenotypic differences as listed in Chapter 1 (section 1.4.). Although no major genetic variations have been previously shown, a recent study has reported significant DNA microsatellite genetic variation between strains and isolates of *S. japonicum* [62]. As previously stated in Chapter 3, based on known morphological and other phenotypic differences it was hypothesised that there may also be major differences at the expression levels between the strains of *S. japonicum*. Here I have utilised a custom made 60mer Schistosome microarray (Chapter 3) to profile the strain-associated differential gene expression between SJC and SJP.

4.2 Materials and Methods

4.2.1 Parasites

Oncomelania hupensis hupensis snails, infected with Chinese (Anhui) *S. japonicum*, were imported from the Institute of Parasitic Diseases, Shanghai. *O. hupensis quadrasi* snails, harbouring the Philippine (Sorsogon, Luzon) strain (laboratory isolate, >14 years) were maintained at the Queensland Institute of Medical Research (QIMR). Australian Research Council (ARC) Swiss mice were anaesthetized by an

intraperitoneal (IP) injection (100 µl per 10 g body weight) of a 1:1 Ketamine 100 mg/ml (Provet, Banyo, Australia), Xylazine (Provet) 20 mg/ml together with 12 ml H_2O (Sigma-Aldrich, Sydney, Australia). The mice were infected percutaneously with 40 cercariae each that were shed from snails. Adult worms were perfused in phosphate-buffered saline (PBS) 6 weeks after challenge from infected mice, which were pre-IP injected with 100 ml Heparin (5,000 IU in 5 ml) (Central Pharmacy, NewMarket, Australia) and euthanased by CO_2. After perfusion, schistosomes were counted, kept as pairs or sex-separated, washed in chilled PBS, and stored in liquid nitrogen until used.

4.2.2 Total RNA isolation

Total RNA was isolated from adult pairs or single sex schistosomes (100-200 worms) using published protocols [110]. Visualisation of total RNA quality/quantity was assessed using the Bioanalyzer RNA Nano LabChip (Bioanalyzer) (Agilent, Santa Clara, USA) (Supplementary File 1). The quantification of total RNA was also determined using a BioMate 3 Thermo Spectronic spectrometer (Thermo, New York, USA). A 500 ng aliquot of total RNA from each sample was used to create fluorophore-labelled cRNA using Cyanine 3-Cytidine 5′-triphosphate (CY3c) or Cyanine 5-Cytidine 5′-triphosphate (CY5c) (PerkinElmer, Boston, USA). A Qiagen RNeasy kit (Qiagen, Melbourne, Australia) was used to purify each sample. All Cyanine-labelled cRNA samples were analysed by the Bioanalyzer, after which, samples were measured at A*260*, A*550* and A*650* nm in a BioMate 3 Thermo Spectronic spectrometer to determine yield, concentration, amplification efficiency and abundance of Cyanine fluorophore. Once the concentration had been determined, 0.75 µg aliquots of CY3c and CY5c were combined in a fresh tube together and "10 X control targets" [Targets for positive probes included on the chip] (Agilent) added.

4.2.3 Microarray hybridization

Sample fluorophore-labelled cRNA (target) was applied to a gasket slide that was pre-positioned in a hybridization chamber (Agilent). The microarray slide was placed probe side towards the target. The chamber was assembled and placed in a hybridization oven (SHEL LAB 1012AG, Cornelius, USA) and incubated for 17 h at 60˚C rotating at 4-rpm. After hybridization, the chamber was opened and the microarray slide was washed for one minute in a 6× sodium chloride sodium phosphate and ethylenediaminetetraacetic acid (SSPE) / N-Lauroylsarcosine solution [4]. The slide was then washed for another minute in a 0.06× SSPE/N-Lauroylsarcosine solution before drying using a Wash 3 Solution [Stabilization and Drying Solution (active component acetonitrile) (Agilent)].

4.2.4 Feature extraction and analysis of data

Microarray slides were scanned by a Microarray Scanner (B version) (Agilent), 21 x 61 scan-region at 10 µm scan-resolution. The "tag image format files" (tiff) produced by the scanner were loaded into the image analysis program Feature Extraction 7.5.1 (Agilent Technologies, Santa Clara, USA) to produce standardised data for statistical analysis. All microarray slides were checked for background evenness by viewing the tiff image on Feature Extraction. Differential expression was analysed using Rosetta Resolver (Rosetta Biosoftware, Seattle, USA). Image files were automatically processed according to default settings contained in the XML design file that describes the microarray physical layout, content and statistical parameters for normalisation by both Linear and LOWESS to correct for systematic errors and to model technology-

specific random errors. The software produced log-ratios from the processed data, with log ratio error and associated probability (p)-values generated for every feature according to the Agilent-Rosetta error model for *in situ* inkjet synthesised oligonucleotide DNA microarrays (Chapter 3.3.3). The data generated from these files was output in Gene Expression Mark-up Language (GEML), MicroArray and Gene Expression Mark-up Language (MAGE-ML) and tab-delimited text formats. Gene Ontology functional categories were analysed using GeneSpring (Agilent Technologies) and various quality indicators were reviewed using Microsoft Excel (Microsoft Corporation, Redmond, USA).

Two lists of differentially expressed schistosome sequences meeting a p-value threshold of $\leq$0.001 were generated; one for up-regulated probes, the other for down-regulated probes. Probe populations were then further partitioned according to signal intensity to identify probes that failed to perform in hybridizations with *S. mansoni* or *S. japonicum* mRNA. Experiments included one or more dye swaps (replicates), of which any bias was removed by the linear and LOWESS normalisation to produce extensive lists of differentially expressed probes, which were further filtered through p-value ($\leq$0.001) and flag error exclusion. Considerable care was taken to ensure that the results were MIAME compliant [79] (Table 4.1).

Table 4.1 MIAME compliant GEO submitted Materials and Methods

Experiment 2	**SJPX V SJCX**			
GEO Samples	GSM115029	GSM115030	GSM115031	GSM115032
Chip Barcode	251260210055	251260210056	251260210057	251260210064
Article URL:	http://dx.doi.org/ doi:10.1016/j.mcp.2006.02.002			
Experiment 3	**SJPM V SJCM**			
GEO Samples	GSM116391	GSM116392		
Chip Barcode	251260210058	251260210059		
Article URL:	http://dx.doi.org/ doi:10.1016/j.mcp.2006.02.002			
Experiment 4	**SJCF V SJPF**			
GEO Samples	GSM116395	GSM116397		
Chip Barcode	251260210062	251260210063		
Article URL:	http://dx.doi.org/ doi:10.1016/j.mcp.2006.02.002			

Key: SJP = *S. japonicum* Philippine, SJC = *S. japonicum* Chinese, X = Mixed sex M = Males, F = Females, V = versus.

All of the microarray material, methods and results were submitted to NCBI Gene Expression Omnibus. The GEO sample number relates to the accession display of the relevant microarray experiment which may be viewed online: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE5234.

4.2.5 Real time PCR

Gene expression patterns determined by microarray analysis were validated using real time PCR. Complementary DNA was synthesized from total RNA using a modified SuperScript™ III protocol (Invitrogen, Melbourne, Australia) with $p(dT)_{15}$ primers (Roche, Sydney, Australia). Forward and reverse primers (Sigma-Aldrich) were designed from *S. japonicum* contigs (Table 4.2). All cDNA samples synthesised from aliquots of the same total RNA used for the microarray experiments were diluted to 1 µg/µl, quantified by a ND-1000 spectrophotometer (Nano Drop, Wilmington, USA), and then 1 µl aliquots were combined with 10 µl of SYBER® Green (Applied Biosystems, Foster City, USA), 3 µl of water (Sigma) and 2 µl (5 pmol) of the forward

and reverse primers in a 0.1 ml tube (Corbett Research, Sydney, Australia). All reactions were performed on a Rotor-Gene (3000) real time PCR (Corbett Research) and analysed by Rotor Gene 6 Software (Corbett Research). In order to minimise indiscriminate binding of double-stranded DNA, which can produce readings in the "no template" controls, separate reverse transcription and PCR steps were used [125]. Parameters were set by determination of primer melting temperature and addition of a melt curve to show primer viability. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a house-keeping gene for all real time PCR studies: the DNA sequence of this gene is identical in the Chinese and Philippine strains of *S. japonicum* [126].

Table 4.2 List of primers used for real time PCR to validate microarray results.

Contig Name	Primer Pair Sequence (5'-3')	
	Forward	Reverse
Contig 1347	TTCAAGCGAATGAACACGTC	CCGATTTACACAGTGCCGTA
Contig 204	AGCCGCAAACCTATCAACAG	GTGAAACTCGCCTGGATTGT
Contig 2213	TGTGGAAAGTGGTGGTTCAC	CCCCGGTTGCATTTTTAACT
Contig 2522	TGGCCTTGACTCACTCACTG	GTGGAGGGTACAGGTCAAG
Contig 2733	CGCCCTCGTCTCTGTTACAT	CGCTGCTAAAGAATGCCTCT
Contig 2935	CCCACGTTTGGTTCCTGTAT	TTGGTGCAAAACGTGGTACT
Contig 346	AAAGTGACGTTGGGAGATGG	GCAGGAAAATCCGCTACTGT
Contig 3466	CGATCAGATTCCGACGACTT	CTGGGCCAAGTACTGAGAG
Contig 3518	AGGTCCAAATGGTGGACTTG	CGAAAATCCCATTTCCTTTG
Contig 3984	ACAGGCAGACCAGGTATTGG	AGGCGAGCTAGTGAAGCTCT
Contig 4489	TGCTGACCTCCAACAGGAAT	CTGAAGCTTGGAGGATGGT
Contig 5020	CGAAATCAACATCACCACCA	GGGCTGATTTGGAAATTGTG
Contig 5336	TTCAACGGCGACATTATCAG	TCAGTTCCATGAGTGCATGA
Contig 6432	CGTCAAAGCTGCGTAACAAA	GCGTTGACGAAGTTTCTGTG
Contig 6442	CCATGTCAGGAACATCCACA	TGTCATTGATTCGCCTTGAA
Contig 6582	CTGAACGAATGGCTCAGTTG	GGTATTGAGTCACGCACGAA
Contig 6760	TCATGGTGGCGGTATAGCTT	AAGCGTGAATCTTGCCATTA
Contig 6953	CGATCACGTGGTGTTGGTAG	ACCCTCATGTTGATGTCGTG
Contig 7060	ACCACGAAGACGAAGAACA	GGGGAAGACAATCACATTGG
Contig 7411	ACGAGGGTGATTTGTTGCAT	GTGAGGTCGTGGACGTTTCT
Contig 8365	ATCCAACGTGTCCAGGAGAA	TTAACGACGTCCAGTGGTCA
Contig 8664	CAATTTATGAACGGGCAGGT	CGTGACAGTGACGGCAATAC
Contig 8689	TTGTCGTGGATTCGGTCTAA	GCCACCGCTCTTACTTCATC

4.3 Results and Discussion

A 22,575 feature 60-mer oligonucleotide microarray was used to investigate gene expression patterns between the SJC and SJP strains of *S. japonicum.* It is important to identify any differences in the quality of the isolated total RNA, since this can compromise the measurement of gene expression. Extreme care was taken to ensure that all samples of total RNA were of high quality and quantity (A*260*/A*280* nm $\geq$1.7 in nuclease-free water [127]) with minimal degradation, as recommended by Bustin and Nolan [125]. The design of the microarray incorporating two closely related species, while providing a powerful tool, does require care in the interpretation of analyses. The

sequence divergence between interrelated species could affect hybridization, but since mRNA used was from two strains of the same species, this was not an issue. Pooling 100 or more worms, strain-specific differential expression in *S. japonicum* was readily identified.

4.3.1 Gene expression differences between the Chinese and Philippine strains of *S. japonicum*

The majority of previous research has demonstrated only minor genetic differences between geographical strains of *S. japonicum* despite the distinct phenotypic differences reported in terms of fecundity, drug sensitivity and pathology [61, 128]. Studies of nuclear and mitochondrial genetic markers reflected the high levels of conservation between the two strains [129] Additionally, enzyme electrophoresis and restriction fragment length polymorphism (RFLP) analysis also confirmed the remarkable similarity between SJC and SJP at the genetic level [61]. Shrivastava *et al*. [62] showed that there was 'very great' genetic differentiation of DNA microsatellite markers between the Chinese (SJC) and Philippine (SJP) strains of *S. japonicum*. DNA microsatellite markers, commonly known as simple sequence repeats, are statistically significant clusters of short sequences [130]. The majority of which are found in the non-coding regions of genomes. The coding regions do contain a small amount of these repeats, which may reflect in the function of some proteins [130]. The work reported by Shrivastava *et al*. [62] suggests that this differentiation is also evident between the Chinese mainland isolates. In the study reported here, it was postulated that differences in gene expression levels may be the basis for the phenotypic differences evident. In order to investigate this, I used a microarray covering the majority of the schistosome transcriptome to compare transcription levels between the two strains.

In this study, probes identified as having a p-value of $\leq$0.001 were considered differentially expressed between the strains of adult Philippine and Chinese *S. japonicum*. Four microarrays (including two dye swaps) of labelled mixed sex *S. japonicum* cRNA showed that 593 probes were differentially expressed between the SJP and SJC strains (Figs. 4.1 & 4.2). In single-sex strain comparison experiments, two microarrays each (including one dye swap) where used to identify 664 and 426 probes as differentially expressed between SJC and SJP males and between SJC and SJP females, respectively. When the overlap of individual probes up-regulated in more than one comparison was investigated, some probes were differentially over-expressed in the same strain in multiple comparisons. Probes which were differentially expressed in mixed sex and both single sex comparisons were identified as "common probes". A graphical representation of the number of probes that were up-regulated in more than one comparison is presented as a Venn diagram in Fig. 4.1. An example of a "common probe" is Contig 812 (Fig. 4.1; Table 4.3). The low number of "common probes" is due to contrast expression between the sexes, flagged material and possible splice site variation. Complete lists of all probes differentially expressed between SJC and SJP are presented in supplementary Tables 6-8.

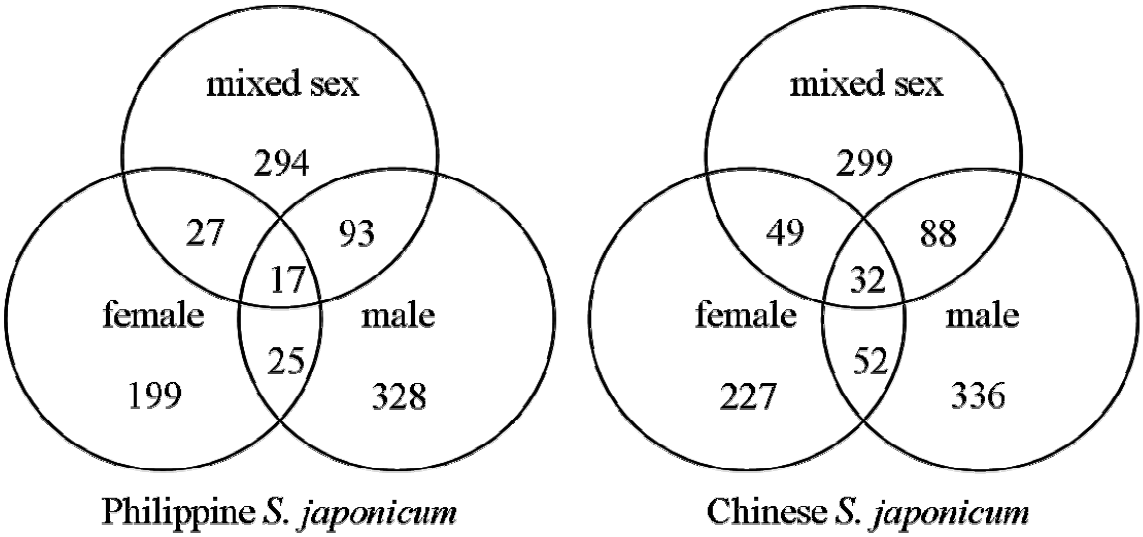

Figure 4.1 Venn diagram of expressed probes of Philippine and Chinese strains of *S. japonicum*. Both diagrams represent the probes up-regulated in comparison of experiments of SJP or SJC (p-value $\leq$0.001). The central number of the three overlapping circles represents "common probes" which comprise of probes which are consistently up-regulated in all three comparative experiments which are: Mixed sex: comparison of mixed sex SJP and SJC. Female: comparison of female SJP and SJC. Male: comparison of male SJP and SJC. Complete lists of these probes are presented in supplementary Table 9.

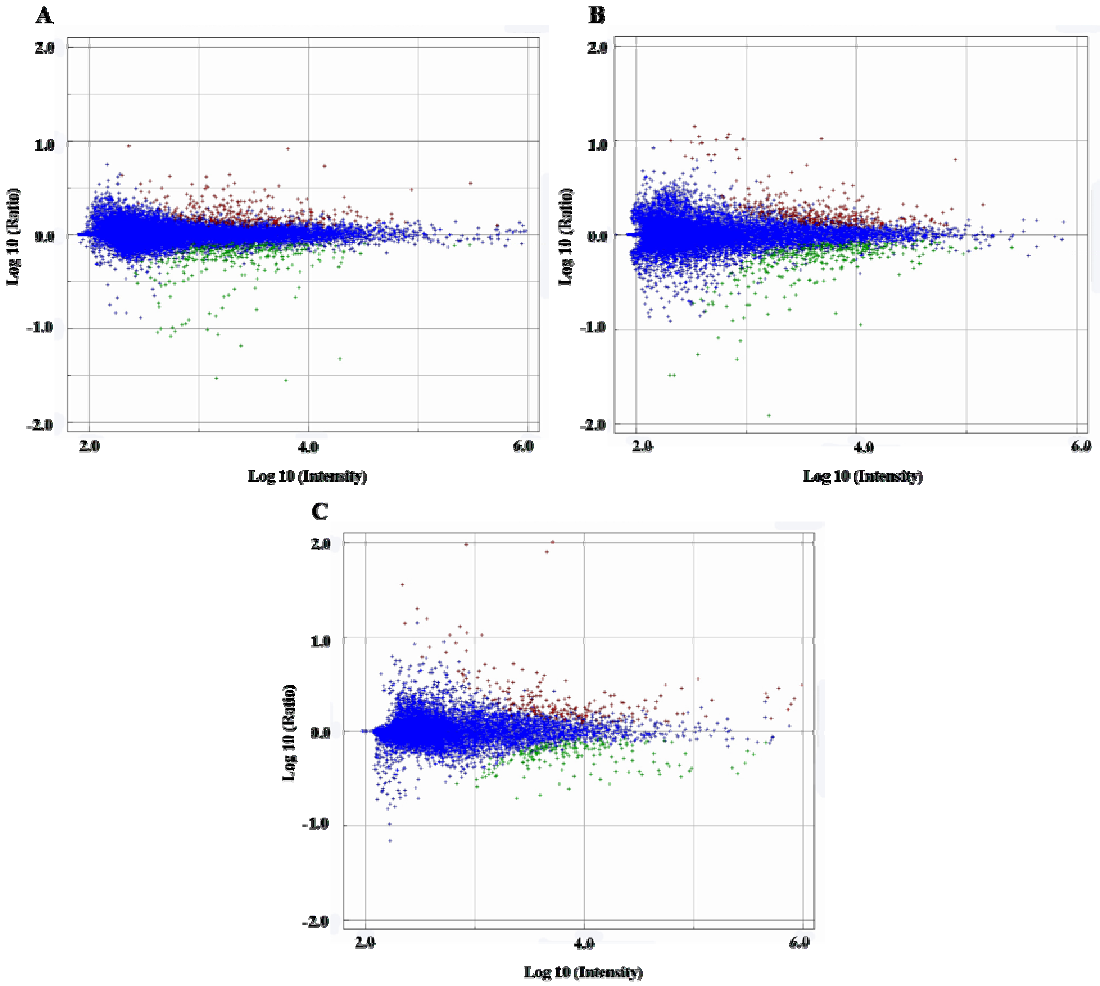

Figure 4.2 Scatter plots showing the distribution of fold changes between Philippine and Chinese strains of *S. japonicum*. Scatter plots generated by Rosetta Luminator demonstrating strain variation in: (**A**) mixed sex comparison; 294 up-regulated SJP compared to 299 down-regulated SJC, (**B**) male comparison; 328 up-regulated SJP compared to 336 down-regulated SJC and (**C**) female comparison; 227 up-regulated SJC compared to 199 down-regulated SJP. Each figure represents $\log_{10}$ intensity [X] versus $\log_{10}$ Ratio [Y] scatter plot. Blue (expanding from $\log_{10}$ Ratio [0.0]) indicates no differential expression (p-value $\leq$0.001). Up-regulated probes are shown in red while down-regulated probes are shown in green.

Table 4.3 Examples* of probes differentially expressed between Philippine and Chinese strains of *S. japonicum*.

Probe Name	Description/Annotation	Up-regulated in	Compared to	Fold Change	p-value
Contig2213	*Homo sapiens* BAC clone RP11-362J16 from Y.	SJC male	SJP male	82.03	0.00E+0
Contig4561	*S. mansoni* amidase	SJP female	SJC female	66.80	6.58E-04
Contig7851	*S. japonicum* T-complex protein mRNA, partial cds	SJP male	SJC female	10.91	4.56E-09
Contig8086	*S. japonicum* serine protease inhibitor serpin (SPI) mRNA, SPI-yyt allele	SJC female	SJP female	4.62	8.08E-14
Contig5188	*S. japonicum* structural muscle protein paramyosin mRNA	SJP mixed sex	SJC mixed sex	1.82	2.00E-05
		SJP male	SJC male	1.96	1.22E-27
TC16805	Triosephosphate isomerase (BC 5.3.1) (TIM). [Blood fluke]	SJP male	SJC male	1.47	9.72E-04
Contig8817	*S. japonicum* glutathione-S-Transferase mRNA	SJC mixed sex	SJP mixed sex	1.30	2.00E-05
TC10879	Similar to thioredoxin 2 (*M. undulates*)	SJP mixed sex	SJC mixed sex	3.12	8.12E-07
Contig812	*S. japonicum* retrotransposon SjR2 polyprotein gene	SJC mixed sex	SJP mixed sex	11.74	0.00E+0
		SJC male	SJP male	8.12	6.19E-21
		SJC female	SJP female	8.53	3.24E-10
Contig5261	*S. japonicum* clone G29A non-LTR retrotransposon SjR2-like	SJC mixed sex	SJP mixed sex	1.42	3.00E-05
		SJC male	SJP male	1.25	2.15E-04
		SJC female	SJP female	1.7	3.60E-04

*Selected on the basis of fold change (p-value $\leq$0.001) and/or annotation (See supplementary Table 1 for more details on individual probes).

4.3.2 Differentially expressed probes

The microarray findings indicate that substantial differences in gene expression occur between SJC and SJP, with many of the differentially expressed probes having no known function or sequence similarity with genes in public databases. One particular

group of annotated genes that were consistently up-regulated in SJC are retrotransposons (Table 4.3). Retrotransposons are suspected to influence genomic evolution, since they make up at least 20% of the entire schistosome genome and respond to environmental stress by shaping genome size and stability through effects on retrotransposon replication rates [131, 132]. For example an increase in genome size in wild barley is a result of environmental stress which results in the induction of *BARE-1* retrotransposon replication [133]. Cellular stress has also been shown to activate long interspersed nuclear elements (L1 retrotransposons) in mammalian genomes [134]. Schistosomes also encounter environmental stress, either from host immune responses or from the external environment, which may promote mobilization of retrotransposons [132], such as Contig 812 (*S. japonicum* retrotransposon SjR2 polyprotein gene, 8.12 – 11.74 fold) or Contig5261 (*S. japonicum* clone G29A non-LTR retrotransposon SjR2-like, 1.25-1.7 fold) both up-regulated in SJC. This variation in retrotransposon expression may reflect the fact that the Chinese isolate of *S. japonicum* used in the current study was obtained directly from the field and was probably under increased external selection pressures [135, 136] compared with the Philippine isolate that has been maintained for a considerable period (over 14 years) in the laboratory. Since the differences in up-regulation of the SjR2-like retrotransposons may be related to copy number at the genomic level, Southern blot analysis may be used in further experiments to exclude multi-copy contribution to gene expression levels.

The gene encoding the enzyme amidase [Contig 4561] was shown in the microarray analysis to be up-regulated in SJP (Table 4.3). In eukaryotes, amidases break down proteins during catabolism, stabilize proteins by an N-terminal amidation reaction, and additionally, in bacterial infections they play an important role in immune-defence by cleaving bacterial amide compounds [136]. It has been further postulated

that N-acetymuramylalanine amidase suppresses both the beneficial and harmful properties of lysozyme [137]. In humans, lysozymes contribute to the bactericidal and bacteriolytic activity of serum and to the breakdown of bacteria engulfed by macrophages [137]. Amidase in SJP *S. japonicum* may play a similar role as a novel immuno-defence mechanism, possibly preventing bacterial infection of the parasite in the mammalian host.

As mentioned previously, the Chinese strain of *S. japonicum* is more pathogenic than the Philippine. This virulence is mainly related to the secreted antigen that comes from the oocyte/miracidium within the egg that promotes a Th2 response in the host [29], and to the number of eggs that is laid by the adult female worm [138]. The microarray results presented here show no major egg probe expressional difference between the strains. The *S. mansoni* probes TC 7634 [homologue to p48 eggshell protein] and TC 7487 [similar to major egg antigen] were up-regulated 1.32 and 1.39 fold in Chinese and Philippine mixed sex adults respectively. Additionally TC 7485 [similar to major egg antigen *S. mansoni*] was up-regulated in Philippine females compared to Chinese females. This is in contrast to the gynaecophoral canal probes that were up-regulated in the Chinese strain in comparison to Philippine males only. The three gynaecophoral canal probes Contig 7087, TC 10854 and Contig 6643 were up-regulated 1.61, 1.43 and 1.37 fold in Chinese strain, respectively. The successful development of schistosomes depends on signaling between the male and female parasites while paring [139]. This may suggest that the developmental system of schistosomes requires, upon pairing, signaling from the male to activate female gene expression [139, 140]. The up-regulation of the gynaecophoral canal probes in the Chinese strain may therefore explain some of the virulence differences reported

between the strains, for unpaired female worms are stunted in size and sexually immature [140].

Another example of a gene that was differentially expressed between SJC and SJP was paramyosin. Schistosome paramyosin has been demonstrated as an Fc binding protein [141] and may interfere with host immuno-modulation by the inhibition of the complement cascade [101, 142]. Paramyosin plays an important role in the structure of muscles in invertebrates stabilizing thick myofilaments [143]. It is found within a variety of different muscle types in varying quantities [143]. Paramyosin has also being identified in cercarial secretions in which it may play a role in host immune evasion by trapping host immunoglobulins or complement components [144]. One probe to paramyosin (Contig 5188) was up-regulated in SJP compared with SJC in both mixed sex (1.82 fold) and male (1.96 fold) adult worm comparisons (Table 4.3). Paramyosin has been considered as a putative vaccine candidate against *S. mansoni*, *S. japonicum*, filariasis, and cysticercosis [145, 146]. More specifically, over 70% protection was elicited in mice vaccinated with biochemically purified, native SJP paramyosin, while vaccination of mice with SJC paramyosin resulted in lower levels of protective efficacy [147-149]. Zhang *et al*. [146] reported a significant reduction in worms, worm pairs and eggs in the livers of mice after vaccine/challenge of recombinant *Sjc-97* paramyosin fragments. Their report showed no apparent statistical difference of protection between the four recombinant *Sjc-97* proteins used in the experiments. This suggested that antibodies elicited by the proteins may provide a contributory role in protection [146]. The microarray data from the work described in this thesis tends to corroborate these findings in that the more susceptible SJP had higher levels of paramyosin expression than SJC, suggesting more antigenic release to the host immune system resulting in an increased immune response and the higher vaccine efficacy reported.

4.3.3 Gene-ontological comparison of the Chinese and Philippine strains of *S. japonicum*

As defined in Chapter 3 (section 3.3.4), many of the probes that are present on the microarray constructed for these experiments have associated gene ontology (GO). In order to present an overview of the entire strain GO comparison, the final data from the mixed sex strain experiments (Supplementary Table 6) were combined with GO definitions (Supplementary Table 2) for molecular and biological function (Table 4.4), with definitions listed in Tables 4.5 and 4.6. Generally, the GO defined probes were up-regulated in the SJC compared to the SJP mixed sex (Table 4.4). This may be due to the increased virulence of the Chinese worms when compared to the Philippine strain [61]. The up-regulation of 'binding' probes shows an increased interaction of molecules within the Chinese worms; this together with the 'transporter activity' is reflected in the up-regulation of probes related to 'metabolism'. It is noteworthy there was a slight increase in the number of up-regulated SJP probes, GO defined as 'regulation of the cell'. Since this ontology involves gene expression and protein modification, the slight increase may be related to some morphological difference in the lab maintained isolate. This could be further explored by the characterisation of these eight probes (described as ribosomal and undefined genes) (Supplementary Table 10). Additionally, if a field isolate of the SJP strain could be obtained, real time PCR or microarray experiments could be run concurrently with the current laboratory strain to investigate the differential expression of these and other probes. These experiments together with Southern blot and sequence analysis may show if there was any difference due to genetic bottlenecking of isolates as previously suggested in other studies of schistosomes [135].

Table 4.4 Number of probes with associated gene ontology within the mixed sex SJP and SJC strains

Molecular function	SJPX	SJCX
Binding	10	15
Catalytic Activity	13	15
Chaperone Activity	0	2
Enzyme Regulator Activity	0	1
Signal Transducer Activity	1	1
Structural Molecule Activity	2	5
Translation Regulator Activity	1	0
Transporter Activity	1	7

Biological Function	SJPX	SJCX
Cell Communication	4	4
Cell Cycle	0	5
Cell Differentiation	3	2
Cell Organisation	6	14
Membrane Fusion	1	3
Metabolism	14	18
Physiological Process	15	20
Regulation of the Cell	8	7
Transport	0	3

Probes that were not differentially expressed within the *S. japonicum* Chinese (SJC) and *S. japonicum* Philippine (SJP) mixed sex (X) strain comparisons at p-value $\geq$0.001 were not included in the Table.

Table 4.5 Defined gene ontologies for molecular functions

Molecular Function	Description
Antioxidant Activity	Inhibition of the reactions brought about by dioxygen (O_2) or peroxides. Usually the antioxidant is effective because it can itself be more easily oxidized than the substance protected. The term is often applied to components that can trap free radicals, thereby breaking the chain reaction that normally leads to extensive biological damage.
Binding	The selective, often stoichiometric interaction of a molecule with one or more specific sites on another molecule.
Catalytic Activity	Catalysis of a biochemical reaction at physiological temperatures. In biologically catalysed reactions, the reactants are known as substrates, and the catalysts are naturally occurring macromolecular substances known as enzymes. Enzymes possess specific binding sites for substrates, and are usually composed wholly or largely of protein, but RNA that has catalytic activity (ribosome) is often also regarded as enzymatic.
Chaperone Activity	Modulates the activity of a molecular chaperone.
Enzyme Regulator Activity	Modulates the activity of an enzyme.
Motor Activity	Catalysis of movement along a polymeric molecule such as a microfilament or microtubule, coupled to the hydrolysis of a nucleoside triphosphate.
Protein Tagging Activity	Acting as an indicator or marker to facilitate recognition by other molecules in the cell. Recognition of the tag, which can be covalently attached to the target molecule, may result in modification, sequestration, transport or degradation of the molecule in question.
Signal Transducer Activity	Mediates the transfer of a signal from the outside to the inside of a cell by means other than the introduction of the signal molecule itself into the cell.
Structural Molecule Activity	The action of a molecule that contributes to the structural integrity of a complex or assembly within or outside a cell.
Transcription Regulator Activity	Plays a role in regulating transcription; may bind a promoter or enhancer DNA sequence or interact with a DNA-binding transcription factor.
Translation Regulator Activity	Any substance involved in the initiation, activation, perpetuation, repression or termination of polypeptide synthesis at the ribosome.
Transporter Activity	Enables the directed movement of substances (such as macromolecules, small molecules, ions) into, out of, within or between cells.

Key GO terminology ‘Adapted from AmiGO’:

http://www.godatabase.org/cgi-bin/amigo/go.cgi

Table 4.6 Defined gene ontologies for biological functions

Biological Function	Description
Behaviour	The specific actions or reactions of an organism in response to external or internal stimuli. Patterned activity of a whole organism in a manner dependent upon some combination of that organism's internal state and external conditions.
Cell Communication	Any process that mediates interactions between a cell and its surroundings. Encompasses interactions such as signalling or attachment between one cell and another cell, between a cell and an extracellular matrix, or between a cell and any other aspect of its environment.
Cell Cycle	The progression of biochemical and morphological phases and events that occur in a cell during successive cell replication or nuclear replication events. Canonically, the cell cycle comprises the replication and segregation of genetic material followed by the division of the cell, but in endocycles or syncytial cells nuclear replication or nuclear division may not be followed by cell division.
Cell Differentiation	The process whereby relatively unspecialised cells, e.g. embryonic or regenerative cells, acquire specialized structural and/or functional features that characterize the cells, tissues, or organs of the mature organism or some other relatively stable phase of the organism's life history. Differentiation includes the processes involved in commitment of a cell to a specific fate.
Cell Organisation	The processes involved in the assembly and arrangement of cell structures, including the plasma membrane and any external encapsulating structures such as the cell wall and cell envelope.
Cell Proliferation	The multiplication or reproduction of cells, resulting in the rapid expansion of a cell population.
Homeostasis	Any of the processes involved in the maintenance of an internal equilibrium within an organism or cell.
Membrane Fusion	The joining of two lipid bilayers to form a single membrane.
Metabolism	Processes that cause many of the chemical changes in living organisms, including anabolism and catabolism. Metabolic processes typically transform small molecules, but also include macromolecular processes such as DNA repair and replication, and protein synthesis and degradation. Note that metabolic processes do not include single functions or processes such as protein-protein interactions, protein-nucleic acids, nor receptor-ligand interactions.
Physiological Process	Those processes specifically pertinent to the functioning of integrated living units: cells, tissues, organs, and organisms.
Regulation of the Cell	Any process that modulates the frequency, rate or extent of a biological process. Biological processes are regulated by many means; examples include the control of gene expression, protein modification or interaction with a protein or substrate molecule.
Transport	The directed movement of substances (such as macromolecules, small molecules, ions) into, out of, within or between cells.

Key GO terminology ‘Adapted from AmiGO’:

http://www.godatabase.org/cgi-bin/amigo/go.cgi

4.3.4 Drug sensitivity

Variation in sensitivity to praziquantel has been reported among isolates of mainland Chinese *S. japonicum* [150]. Hope *et al*. [61] reported minimal variation in drug sensitivity between SJC and SJP strains, stating that three doses of praziquantel were more effective against the Philippine strain [61]. Praziquantel is the main drug used in the treatment of schistosomiasis, its mode of action has not been elucidated upon, but may be related to the disruption of Ca^{2+} homeostasis in adult worms [151]. Gobert G N, McInnes R, **Moertel L P**, Nelson C, Jones M K, Hu W and McManus D P [105] showed some probes related to calcium modulation were up-regulated in the *S. japonicum* species (Chinese) over *S. mansoni*. The study reported in this chapter showed that there was no expressional difference between the Chinese and Philippine strains relating to these probes, with the exception being Contig 8411 [*S. japonicum* calcium-binding protein Sj66 precursor] which was up-regulated 1.49 fold in SJC compared to SJP males only. Further, no other calcium related probes present on the chip were up-regulated between the strains. The activity of the enzyme lactate dehydrogenase has been shown to be inhibited by praziquantel and artemether *in vitro* [152, 153]. A search for lactate dehydrogenase annotated probes revealed four probes, all of which were not differentially expressed between the Chinese and Philippine strains. Although viable in their own right, these results do not explain the differences of praziquantel sensitivity between SJC and SJP as reported by Hope *et al*. [61]. Further research would be needed to investigate such differences and would possibly be benefited by further characterization of the unknown probes on the microarray mentioned here.

4.3.5 Real time PCR independent validation of fold change

In order to validate the microarray predictions detailed in this chapter, 36 probes were examined using real time PCR as an independent analysis of gene expression. Twenty three contigs were randomly selected and primers designed for use in all the real time PCR experiments detailed in Chapters 4 and 5 (Table 4.7). The real time PCR results shown are a calculated mean of 4 replicates from 2 separate experiments, in which 20 of the 27 strain associations showed similar expression patterns to the predicted cRNA microarray hybridizations (Table 4.7). Thirty of the thirty six (83%) probes were validated and the six that were not may have been due to mismatch bases causing erroneous hybridization to the microarray [69]. Indeed, some of the microarray sequences did contain mismatch bases when compared with previously published sequences of the same schistosome species or strain. These inconsistencies will be addressed in future by minor modification to the probes (Table 4.7) and when the complete genomic sequence for *S. japonicum* is available.

The selection of GAPDH was a good choice as a house-keeping gene for these real time experiments for it was shown to be equally expressed between adult worms of the same sex (mean confidence threshold (CT) 11.19 $\pm$3.7 and 11.39 $\pm$3.91 for the Chinese and Philippine isolates respectively, for all strain experiments). This is in comparison to the ribosomal 18s gene that would be a better choice for a house-keeping gene if multiple life stages were used [154]. This gene has been used in other studies with some success [107, 108], but notably the 18s may not be used as a reference gene if the cDNA is made using a oligo-dT primer or if mRNA is used as a template [155]. Additionally the 18s gene is highly expressed and may not be suitable for genes with low expression levels [155]. If the copy-number from the sample gene is low and the reference gene high, the data would show extremely low expression levels of the sample

gene. That is if sample data is too low it may not correctly reflect the expression levels of the individual genes.

Table 4.7 Real time PCR validation of microarray results based on fold change.

	Real Time PCR		Microarray		
Mixed-sex	**Direction**	**Fold Change**	**Direction**	**Fold change**	**p-value**
Contig 6760	Chinese	1.45	Chinese	1.62	3.20E-02
Contig 2522	Chinese	2.85	Chinese	2.13	4.35E-08
Contig 2935	***Chinese***	***1.35***	***Philippine***	***8.2***	***0.00E+00***
Contig 3518	Chinese	2.85	Chinese	2.97	1.71E-10
Contig 7411	Chinese	2.65	Chinese	1.19	1.35E-01
Contig 5020	***Chinese***	***1.7***	***Both***	***1***	***1.00E+00***
Contig 2213	Chinese	169.6	Chinese	34.19	0.00E+00
Contig 6432	Chinese	1.4	Chinese	1.25	1.18E-01
Contig 6953	Chinese	1.2	Chinese	35.92	0.00E+00
Female					
Contig 7060	Chinese	5.65	Chinese	3.28	2.05E-18
Contig 346	***Chinese***	***3.2***	***Philippine***	***2.07***	***1.31E-04***
Contig 8664	Philippine	2.2	Philippine	2.38	1.96E-26
Contig 2935	***Chinese***	***1.35***	***Philippine***	***5.18***	***4.54E-07***
Contig 3466	Philippine	4.4	Philippine	2.74	1.90E-34
Contig 204	Philippine	1.9	Philippine	1.88	2.00E-02
Contig 2733	Philippine	1.65	Philippine	3.68	0.00E+00
Contig 6582	***Chinese***	***2.7***	***Philippine***	***4.81***	***3.29E-30***
Contig 6442	Chinese	1.6	Chinese	78.58	0.00E+00
Male					
Contig 3984	Chinese	3.25	Chinese	2.03	8.56E-06
Contig 4489	Philippine	7.35	Philippine	11.52	2.06E-08
Contig 3518	Chinese	2.6	Chinese	3.67	4.38E-06
Contig 1347	Philippine	1.3	Philippine	2.59	2.04E-03
Contig 2522	Chinese	2.5	Chinese	3.07	1.12E-03
Contig 2935	***Chinese***	***1.1***	***Philippine***	***10.39***	***1.27E-11***
Contig 2733	Philippine	3.7	Philippine	3.14	1.00E-13
Contig 6432	Chinese	1.75	Chinese	1.41	4.45E-09
Contig 204	Chinese	2.05	Philippine	3.1	6.45E-11

The real time PCR results for both Philippine (SJP) and Chinese *S. japonicum* are means of 4 replicates from 2 separate experiments. The results in bold and italics indicate real time PCR results that did not agree with the microarray data. The right hand column shows p-values for correlating microarray fold change.

4.3.6 Conclusions

In summary, the study discussed in this chapter has demonstrated considerable differences in the gene expression profiles of the Chinese and Philippine strains of *S. japonicum*. This differential expression included probe sequences that may encode proteins that have the potential to influence the evolution of schistosomes and their virulence and in differential drug efficacy against these parasites. This information will provide new insights into the biology of *S japonicum* and provide a fulcrum for understanding the transcriptional basis of the phenotypic differences evident between the two strains. The strain-associated probes identified here serve as starting points for further research into better chemotherapeutic and vaccine targets against Philippine and Chinese *S. japonicum*.