5. Discussion

Two populations of *B. glabrata* Campinas strain, one susceptible and one resistant, were exposed to the miracidia of *S. mansoni*. The hepatopancreas, the gastropod organ where sporocysts reproduce asexually and release cercariae (Negrão-Corrêa *et al.*, 2007), was dissected out. Gels created by 2DE, displaying the proteins in susceptible and resistant samples were compared and contrasted. Following analysis proteins found to be unique to the resistant gel and therefore likely to be biologically significant were examined more closely.

The fragmentation spectra produced in this project were of poor quality so only spot 201 was able to be analysed by this method. The results from the spectra produced for this spot (Figure 6) indicate the peptide detected was most likely an autolysis product, trypsin (Table 4). The MS appeared to be working as it was detecting the peptide trypsin, but as it was just autolysis products detected, it was concluded there was not enough protein in the sample so new gels were run with five times increase in protein concentration incorporated into the protocol.

When spots were analysed from the new gels the spectra produced were still poor. Once again the mass spectrometer seemed to be working, but the peptides for some reason did not appear to be charging. Most peptides in the spectra were singly charged. It is possible some of the peptides were sticking to the plastic, although the method did incorporate re-suspending the supernatant in TFA to reduce this.

To gain more information from the MS results, although not technically correct, it was decided to treat the data as if it had gone through MALDI ToF, generating peak lists from PMF's.

5.1 Overview of Comparisons of Susceptible Gel Compared to Resistant Gel

The total number of protein spots on each gel was relatively similar, with 575 spots on the resistant gel compared with 536 on the susceptible gel (Table 5). These total values however do not indicate anything about changes in regulation. Slight alterations in environmental conditions during the running of the gel may account for the minor difference in the total number of spots, although unlikely as the two samples were run at the same time, under the same conditions.

110 and 105 of the matched spots showed up-regulation and down-regulation of more than two fold respectively (Table 5), possibly indicating significant change in the expression of resistant snail proteins compared with the susceptible snail proteins. These proteins may be involved in the immune response by the resistant snails against *S. mansoni*. These proteins found to be up- or down-regulated twofold or more, and therefore likely to be biologically significant (Barrett *et al.*, 2005) can be seen in Table 6.

Proteins found to be up- or down-regulated by a factor of less than two could be suggested to show little variation and they could be involved in other biological functions which are necessary in both susceptible and resistant snails.

B. glabrata expresses diverse parasite reactive haemolymph lectins in response to infection, for instance 13 different FREP subfamilies (Loker *et al.*, 2004) due to diversity to point mutations and alternative splicing. Such diverse

parasite reactive haemolymph proteins account for the differences in expression profiles between the resistant and susceptible snails.

The resistant sample contains more unique spots (273) than the susceptible sample (15) (Table 5). Unique spots may indicate resistant snails consistently expressing proteins which are not expressed in susceptible snails, some of which maybe involved in an immune response against *S. mansoni*.

It has been suggested (Barrett *et al.*, 2005) that changes in protein volume must be at least two-fold for the result to be statistically 'significant', recommended as estimates of the coefficient of variation for relative protein volume range from 22-55% for biological variation and 16-39% for analytical variation. Unique protein spots or spots displaying 'significant' difference in volume could putatively be associated with an immune response or be the result of suppression as a result of secretory/excretory products from an invading parasite.

Three spots of interest, spots 1880, 1906 and 2017, were identified to be analysed in more detail by MS, that were unique to the resistant gel. These spots where chosen as they had large normalised volumes (Table 7) providing relatively large quantities of protein for MS detection, and where in optimum positions on the gel where they could be picked minimising risk of cutting neighbouring spots (Figures 10a, 11a, 12a). However first two spots of interest from a previous study (Murray unpublished, 2008) were analysed by MS (spot 594 and 201) that had shown up/down regulation of more than two fold, as clues had been provided into their putative identifications, however their true identifications by MS remained to be determined.

5.2 PMF

PMF search results from the MASCOT database (Table 8) are ordered on a probability basis, the MOWSE score (probability molecular weight search). The MOWSE scoring algorithm is described by Pappin, Horjrup & Bleasby (1993).

The total score for each protein is given by $-10 \times \log_{10}(P)$, where P is the absolute probability that the match is random. There is a low probability that the match is random (i.e. a good match) if the protein has a high score.

MOWSE is the name of the E-value (expectation value) in MASCOT. It is directly equivalent to the E-value in a BLAST search result. The lower the E-value the more significant the result. It is the number of matches with equal or better scores that are expected to occur by chance alone (i.e. if the database was probed with random fragments). Generally values less than 0.02 are probably significant, 1 - 0.02 cannot be completely ruled out, and greater than one could expect this good a match by chance alone (Barrett *et al.*, 2005).

The sequence coverage (%) is the proportion of the theoretical protein which is covered by the peptides. 20% coverage or higher is likely to be significant, although it depends on the size of the protein. Small proteins are more likely to yield higher % coverage (Barrett *et al.*, 2005).

The lowest MOWSE from all the spots analysed is 3.2 from spot 594, a possible transposase. The next lowest MOWSE value is 6.8 found for both spots 594 and 1880. All other MOWSE values in the table are greater than these. It has been suggested by Barrett *et al.*, that any proteins with values greater than one

could expect a match by chance alone. None of the peptides queried had values less than 0.02 which would be considered to be significant.

Table 8 describes that a large number of potential protein matches, although not particularly significant are immune related (e.g. antigen T cell receptor, immunoglobulin heavy chain variable region, bone marrow macrophage cDNA).

Figure 12B shows the histogram of the protein score distribution. Protein scores greater than 71 are significant ($p<0.05$). None of the proteins in the table have scores greater than 71. For spot 594 the transposase has a score of 54, while the relaxin-3 precursor has a lower score of 50 indicating the chance the protein is a transposase is less random. The transposase however only has 5% sequence coverage, while the relaxin-3 precursor however has 19% with 20% coverage or higher is likely to be significant.

The relaxins p*I* (6.56) is similar to the estimate seen on the gel (Figure 4a) although it's predicted MW (15 kDa) differs substantially from that of spot 594 (~40 kDa) (Figure 4a). The difference in mass could, at least partly, be due to post translational modifications, such as glycosylation.

Murray (unpublished 2008), suggests it is possible the spot 594 (Figure 4), which was shown to be up regulated more than two fold, and therefore likely to be biologically significant, visualized on the gel is a member of the FREP 2 family, as it has been identified as having a MW of 44 kDa (Table 3) and a similar p*I* (Figure 4a), to FREP 2 who's predicted p*I* was 5.71, with a MW of 42 kDa.

FREPs are immunologically relevant proteins and their up regulation has been identified in FREP 2 and 4 in resistant strains of *B. glabrata* by a 57 and 4.5 fold increase respectively (Hertel, Adema & Locker *et al.,* 2005).

No FREPs were identified in the PMF analysis for spot 594 as suggested by Murray (unpublished 2008). Examination of the literature revealed that relaxin-3 like sequences could be present in invertebrates (Georges, Viguiermartinez & Poirier, 1990) such as *B. glabrata*, and may have some immune functions, explaining its difference in regulation here (Piccinni, Bani, Beloni, Manuelli *et al.,* 1999).

Relaxin-3 belongs to the insulin superfamily (Wilkinson, Speed, Tregear & Bathgate, 2005), which is composed of peptides with diverse sequences held together by characteristic disulphide links connecting A and B peptides chains (Liu, Bonaventure, Sutton *et al.,* 2004). The family's evolution has been contentious, with high sequence variability seen between closely related species, and distantly related species showing high similarity (Wilkinson *et al*, 2005). Relaxin-3 is likely to be the ancestral relaxin (Wilkinson *et al.*, 2005).

In 1983 relaxin-like activity was first detected in *Tetrahymena pyriformis* a protozoa (Schwabe, Leroith, Thompson, Shiloach *et al.,* 1983). A protein molecule was detected that behaves like relaxin in regard to solubility, charge, size, immunological activity, and sensitivity to reducing agents (Schwabe *et al.*, 1983). Invertebrate relaxin, a hormone with "relaxin-like" properties, has since been reported in other studies in ascidians (*Herdmania momus*) and tunicates (*Ciona intestinalis*) (Georges *et al.,* 1990). A cDNA and peptide sequence with almost 100% similarity to porcine relaxin was isolated from *C. intestinalis* (Georges & Schwabe, 1999; Wilkinson et al., 2005).

Relaxins are thought to play a role in the neuropeptide signalling process. In higher vertebrates it plays multifunctional roles including reproductive tissue growth, uterus relaxation and collagen remodeling in females, although it has also been reported to play important roles in nonreproductive functions including cardiac protection, allergic responses, and wound healing (Liu *et al.*, 2004).

It has been shown relaxin exhibits additional, multiple effects on organs other than reproductive ones (Bani, 1997), such as modulating the biological activity of bone marrow-derived cells, such as platelets, granulocytes mast cells, as well as reducing the allergic asthma-like reaction elicited by antigen inhalation in sensitized guinea pigs (Piccinni *et al.*, 1999).

Relaxin is induced by matrix metalloproteinases (MMPs) secreted by macrophages when they populate wound sites, and plays an important role in tissue remodeling and wound repair (Ho, Yan & Bagnell, 2007).

It has been demonstrated (Piccinni *et al.*, 1999) that relaxin can favour the *in vitro* development of human antigen specific $CD4^+$ T cells into Th1-like effectors. Their results suggest relaxin may contribute to the regulation of the immune homeostasis during pregnancy, and that it can directly influence both the differentiation and function of $CD4^+$ effector lymphocytes.

Spot 201 (Figure 5), was up-regulated greater than two-fold and was identified as having a MW of ≈ 24 kDa (Table 3). Baker (unpublished 2008), suggested it could possibly be Manganese superoxide dismutase (MnSOD) potentially showing up-regulation as a result of cellular defence mechanisms against oxidative stress, as Jung, Nowak, Zhang, Hertel *et al.*, (2005) calculated such a protein as having a molecular weight of 24 kDa and a p*I* of 8.7.

MnSOD is an antioxidant enzyme present in *B. glabrata* was found to have an increased expression in the susceptible snail strain following infection with *S. mansoni.* Up-regulation of the protein was hypothesized to be part of a stress response fond in susceptible M line strains of *B. glabrata* (Jung *et al.*, 2005). The tissue used in Jung's experiment was from the head foot region, as opposed to Baker (unpublished 2008) which was from the hepatopancreas. Baker noted non-linear IEF strips were used in her experiment, pH 3-10, making estimating p*I* difficult and very inaccurate without mass spectrometry conformation.

MnSOD was not identified in the PMF analysis of spot 201. The most probably match was a 2 days pregnant adult female oviduct cDNA (Table 8), although this is not considered significant, as it has an E-value of 7.4 and protein score of only 30. The match however does have a predicted MW of 25 kDa, close to the 24 kDa of spot 201, and a p*I* of 8.35, similar to the p*I* of 8.7 from Jungs experiment.

5.3 Limitations of the Study & Future Work

B. glabrata is not a genome verified organism, therefore protein identification by database searching becomes more problematical (Barrett *et al.*, 2005). A number of approaches to cross species protein identification have been suggested, but if the organism being studied is only distantly related to any organism with a sequenced genome then the likelihood of protein identification remains small (Barrett *et al.*, 2005).

When searching databases potential matches can be limited by using species or tissue-specific databases, although specialist databases may not contain all of the relevant protein sequences, so failure to find a match in a species-specific database for example does not mean that a significant match will not be found in a more general database. When searching databases such as MASCOT, if they do not contain the unknown protein, then the aim is to identify those entries which exhibit the closest homology, often equivalent proteins from related species. Proteomic research on the hepatopancreas of *B. glabrata* is sparse, therefore proteins in databases are sparse so to compensate the metazoans database at NCBI was searched using MASCOT.

Within the hepatopancreas, epithelial lining immune proteins would expected to be present, that are also present throughout the snail, although there is also the expectancy of organ specific immune responses taking place which must be considered when investigating one organ alone. The analysis of organ specific immune responses could take place with more accurate identification of observed proteins, in turn improving our understanding of sporocyst-snail interactions during development within the intermediate host. Vergote *et al.*, highlighted gene expression in the albumin gland as possibly possessing underlying susceptibility/resistance characteristics to parasite invasion in *B. glabrata* (Vergote, Bouchut, Sautiere, Roger *et al.*, 2005).

It is important to consider the effect of post-translational modifications, which can be detected due to the high resolution capacity of 2DE, unlike studying the transcriptome where post-translational modifications are not apparent. Post-translational modifications such as the addition of carbohydrate or lipid side groups, for example lectins which bind to carbohydrate epitopes on the

surface of *S. mansoni*, can affect the migration of proteins altering the position of proteins on the gel. The increase or decrease in MW and p*I* due to carbohydrate or lipid side groups, of protein would hinder the identification of such immune proteins (Barrett *et al.*, 2005), if these values where used for identification alone, as previous projects (Baker unpublished, 2008; Murray unpublished, 2008). The use of mass spectrometry reduces the chance of misidentification.

Estimating the p*I* and MW on gels is a crude science. This method of protein identification is relatively erroneous, as the margin for estimating MWs can be 20% inaccurate (Barrett *et al.*, 2005). Determining the MW of specific spots was difficult as some of the molecular markers migrated out along a large area of the gel, reducing the probability of correctly identifying the actual MW of a spot.

It has been shown that the range of immune proteins from *B. glabrata* include proteins that are higher than 99 kDa, which would be present on a 12% gel, and lower than 14.3 kDa. To allow for the identification of lower molecular weight proteins, a higher percentage resolving gel could be used, decreasing the size of the pores and allow proteins with lower MWs to be identified.

It has been shown (Spray & Granath, 1990) that bands of haemolymph proteins of 90-210 kDa can be been found on polyacrylamide gels. Creating a lower percentage resolving gel would increase the size of the pores and allow proteins of higher MW than 99kDa to migrate through the gel and be identified. A range of resolving gels with different percentages would give gels with more representative proteins involved in the immune response, allowing for a more accurate comparison and analysis between resistant and susceptible snails.

The hepatopancreas protein samples used had been stored in protease inhibitor and PMSF in a -80^{o}C freezer for a long time. Protease inhibitors can not completely inhibit proteases, with a degree of proteolysis still occurring even when samples are stored at -80^{o}C.

A method for improving the specificity of a PMF has been proposed by James, Quadroni, Carafoli & Gonnet (1994), suggesting additional digests using different proteases. Observing the same protein with a high score in two independent digests provides a similar degree of confidence to seeing multiple peptide matches in an MS-MS ions search.

PMF's main drawback is the ambiguity in protein identification due to peptide mass redundancy e.g. the peptide sequence PWFIL will have the same mass as FILWP, PWFIL etc. therefore a sufficient number of peptide must be analyse to provide the specificity required, and consequently PMF is not vey successful when used with protein mixtures (Barrett *et al.*, 2005).

2DE remains an imperfect technique, despite technical improvements, and geometric distortions of protein patterns are constitutional in creating and running the gels (Barrett *et al.*, 2005), therefore gels are not perfectly reproducible, causing problems in the matching step in gel analysis. Future work could involve dual labelling protein samples with covalently linked fluorescent dyes prior to electrophoresis, combined with multi-spectral imaging enabling the resistant and susceptible proteomes to be compared on the same gel, reducing problems associated with reproducibility and matching (Spibey, Jackson & Herick, 2001).

Although 2DE and MS have been very successfully employed to identify proteins, several papers recently have emphasised the 'pitfalls' of 2DE

experiments, especially in relationship to experimental design, poor statistical treatment and the high rate of 'false positive' results with regard to protein identification (Biron, Brun, Lefever, Lebarbenchon *et al.*, 2006).

2D gel analysis is tiring, complex and can be subjective. For accurate results editing has to be consistent everyday. Adequately trained individuals will analyse gels slightly differently.

Replicate gels were run in previous projects (Baker unpublished, 2008; Murray unpublished, 2008), while this project was only concerned with creating preparative gels to identify spots of interest. If more gels were created then the use of statistical analysis could have been included, such as analysis of variance (ANOVA) and the students t test, to measure the significance of the spots which are down and up regulated more than two fold such. It is recommended for one way ANOVA and the students t test that a maximum of five replicates are produced for each treatment (Biron, Brun, Lefever, Lebarbenchon *et al.*, 2006). As only one resistant and one susceptible gel was run the up regulation and down regulation of two-fold, considered to be biologically significant (Barrett *et al.*, 2005) was applied to the experiment analysis.

5.4 Conclusions

This project has examined the differences in the proteome of the hepatopancreas of susceptible and resistant *B. glabrata*, however due to limitations of MS no significant proteins were identified. Proteins with possible immune functions were identified by PMF through MASCOT, although none of the matches were significant. Future work would involve further MS to yield

more accurate results leading to a better understanding of the components of the proteome of the hepatopanceas. The project has highlighted many points of interest to investigate and has emphasized the many gaps in our understanding.

Further investigation could also see the production of more gels, allowing better familiarization with equipment, and reducing error as protocols used are complex and convoluted. The production of more gels would allow for statistical analysis such as ANOVA and students t test.

The identification of immune proteins will be enhanced with the completion of sequencing the complete genome of *B. glabrata* (Bouchut, Sautiere, Coustau & Mitta, 2006). Before this occurs however, expressed sequence tag (EST) gene discovery strategies such as ESTs (ORESTES) can be employed to identify transcripts which may be involved in snail-schistosome interactions (Lockyer, Spinks, Walker, Kane *et al.,* 2007).

It is difficult to discuss gene expression in *B. glabrata* without also giving some attention to those genes expressed by the invading parasite (Knight *et al.*, 2000). Further investigation will also be required to identify the factors involved in *S. mansoni* for compatibility, as well as host factors of *B. glabrata*, considering the hypothesis that success or failure of infection does not depend on susceptible or resistant status of the snail, but on matched' or 'mismatched' status of the host and parasite phenotypes. Efforts must be made to identify the molecular determinant of self/nonself recognition and the potential variability (Theron & Coustau, 2005).

Even when more is understood about the molecular mechanisms of host parasite compatibility, genetic transformation of snails with parasite resistance genes for a genetic control technique will still present several technical problems,

such as delivery system for the efficient transformation and production of stable transformants, and an adequate marker for selecting stable transformants (Knight *et al.*, 2000).

If further investigation does lead to a 'genetic control' technique, it will still need to be regarded as one component, of integrated schistosomiasis control programmes, that will complement existing strategies, including chemotherapy and health education. Although debatable, praziquantel resistance is here or on the horizon, the need for this 'genetic control' is more pressing than ever.

Section	**No. of Words**
Introduction and Aims and Objectives	3661
Materials and Methods	1360
Results	757
Disscussion	3091
Total	**8869**