This copy has been supplied by the Library of the University of Otago on the understanding that the following conditions will be observed:

1. To comply with s56 of the Copyright Act 1994 [NZ], this thesis copy must only be used for the purposes of research or private study.

2. The author's permission must be obtained before any material in the thesis is reproduced, unless such reproduction falls within the fair dealing guidelines of the Copyright Act 1994. Due acknowledgement must be made to the author in any citation.

3. No further copies may be made without the permission of the Librarian of the University of Otago.
Approaches to Identify Candidate Genes for Resistance to Facial Eczema Disease in Sheep

Elizabeth Jenness Duncan

A thesis submitted for the degree of Doctor of Philosophy at the University of Otago, Dunedin, New Zealand

June 2007
What use is knowledge if there is no understanding?

Aristotle (5th Century)
Abstract

Facial eczema disease (FE) is a secondary photosensitisation disease of ruminants caused by exposure to the mycotoxin sporidesmin. Resistance to FE has a significant genetic component and previous research has included a whole genome scan and investigation of candidate genes. The aim of this study was to use multiple approaches to identify genes associated with resistance to FE.

ABC transporters have been considered as putative candidate genes for FE since the yeast ABC transporter, PDR5, was found to modulate sensitivity to sporidesmin in *Saccharomyces cerevisiae*. A previous study had shown that hepatic expression of the ovine ABC transporter, ABCB1, was induced following exposure to sporidesmin but only in resistant animals (Longley (1998) PhD Thesis, University of Otago). In the present study, using qRT-PCR, a difference in the expression of ABCB1 between resistant and susceptible animals was not confirmed. It is concluded that ABCB1 is not likely to be a candidate gene for FE.

As the full genome sequences for several mammalian species are now available, phylogenetic analyses were used to identify the most likely mammalian ortholog of the yeast PDR5 protein. This analysis found that the yeast PDR5 protein was most closely related to the mammalian ABCG sub-family. The human ABCG sub-family has five members one of which, ABCG2, is a known xenobiotic transporter. Comparative mapping of ABCG2 indicated that it co-localised to a region of the sheep genome weakly associated with resistance to FE. The full-length sequence of ovine ABCG2 was determined and two synonymous polymorphisms were found. These two polymorphisms, together with an intronic SNP were genotyped across a panel of selection-line animals. The allele frequencies of the intronic SNP were found to be significantly different between the selection lines, providing evidence for the association of ABCG2 with resistance to FE. The hepatic expression of ABCG2 was examined but no differential expression between the selection-lines was observed.

Global gene expression profiling via microarray analysis was undertaken as a novel approach to identify candidate genes. Differences in gene expression were examined between naïve and sporidesmin-dosed resistant and susceptible animals using a bovine cDNA microarray. A small number of differentially expressed genes were identified. Follow-up studies found that there were a relatively high number of errors in EST identity. Eight differentially expressed genes were selected for confirmation by Northern analysis. Six of these genes were shown to be differentially expressed, but neither the patterns nor the magnitude of the differential expression reflected that observed on the microarray. One of the six genes identified as differentially expressed was catalase, which has previously been implicated in resistance to FE. This finding validates the approach taken using gene expression profiling to identify candidate genes.

The final approach used in this study necessitated the development and characterisation of an *in vitro* system for studying sporidesmin toxicity. The system
chosen was a human hepatoma cell line, HepG2. To date the only effective treatment for FE is the prophylactic administration of high levels of zinc sulphate. The mechanism of protection by zinc is unknown, but zinc is known to be a potent modulator of gene expression. Conceptually, any genes modulated by zinc are possible candidates for resistance to FE. It was shown that zinc pre-treatment could protect HepG2 cells against sporidesmin-induced cytotoxicity. Equivalent protection was provided by the addition of zinc in the presence of the transcriptional inhibitor actinomycin D, suggesting that the mechanism of zinc protection is independent of de novo gene transcription.

Overall, the goal of this project was to find genes to assist selection of sheep resistant to FE. Toward this goal, this research has identified several new candidate genes and avenues for investigation.
Acknowledgements

First and foremost, I would like to thank my supervisors Dr. Sin Phua and Dr. Mary Thompson, for their immeasurable patience, support and understanding. You have both taught me more than either of you will ever know, and there is no doubt in my mind without both of you this thesis would never have been finished.

Secondly, I would like to thank my PhD committee members: Dr. Chris Brown and Associate Professor Iain Lamont for their guidance and advice throughout this PhD.

This PhD project was supported by a MeatNZ fellowship obtained by Dr. Sin Phua and additional financial support was provided by AgResearch. In addition I was able to attend the European Societies of Toxicology meeting in Poland thanks to the financial support of: the NZSBMB, The Oxidative Stress in Health and Disease research theme, The Functional Genomics, Gene Expression and Proteomics theme, the Department of Biochemistry, The Otago Institute and The Vincent George House of Travel.

Many people have contributed to this project and in particular I would like to acknowledge Chris Morris for his considerable efforts in setting up and maintaining the FE selection lines. I would like to thank Dianne Hyndman and Theresa Wilson for synthesising the microarray slides used in chapter five, and for their helpful advice. I would also like to thank David Baird for performing the microarray statistical analysis, Dr. Ken Dodds for helping me out with a number of statistical queries and for performing the peddrift analysis presented in chapter four and Benoit Auvray for the haplotype estimation also in chapter four. Finally, I would like to thank Hannah Henry for her help in performing the genotyping presented in chapter four and for her continued moral support and guidance.

I would also like to thank all the members of the AgResearch Molecular Biology Unit at the University of Otago (both past and present), you have been a fantastic group of people to work with and I have learnt a lot from you all.

On a personal note, I would like to express my gratitude to my family, my Mother and Grandmother specifically: You are both a continued source of inspiration to me.

I would also like to acknowledge my “science-geek” friends who have provided an enormous amount of support, understanding, encouragement, scientific discussion and most importantly the odd drink or two: Dr. Belinda Bray, Dr. Berni Chapman, Dr. Andrew Cridge, Dr. Peter Dearden, Julie Hagen, Dr. Orla Keane, Dr. Veronique Hermann, Dr. Robert McLaren, Justine Murrell, Dr. Tina Summerfield and Tony Zaharic.

Thanks also to my “non-science” or “ex-science” friends whose encouragement and support is perhaps all the more remarkable: Kirsty Dickson, Chetna Patel, Petra Dearden, Samantha Eastman and Corey Moir.

Thank you.
Table of Contents

Abstract ........................................................................................................................... i
Acknowledgements ...................................................................................................... iii
Table of Contents ........................................................................................................... v
List of Tables ................................................................................................................ ix
List of Figures ............................................................................................................. xi
List of Abbreviations .................................................................................................. xiii

CHAPTER ONE: INTRODUCTION .................................................................................. 1
  1.1. AETIOLOGY OF FACIAL ECZEMA ........................................................... 2
  1.2. ECOLOGY AND DISTRIBUTION OF P. CHARTARUM ........................................ 2
  1.3. SPORIDESMIN CHEMISTRY AND TOXICOLOGY ........................................... 4
      1.3.1. ETP mycotoxins ................................................................................... 4
      1.3.2. Sporidesmin biosynthesis ..................................................................... 6
      1.3.3. Sporidesmin metabolism ...................................................................... 8
      1.3.4. Sporidesmin toxicity ............................................................................. 9
  1.4. PATHOLOGY AND PATHOPHYSIOLOGY OF FACIAL ECZEMA ...
      ......................................................................................................................... 14
  1.5. CONTROL OF FACIAL ECZEMA .................................................................. 18
  1.6. GENETIC RESISTANCE TO FACIAL ECZEMA ........................................... 21
      1.6.1. Resistant and Susceptible selection lines ............................................. 21
      1.6.2. Breed specific differences (Ovis aries) .............................................. 22
      1.6.3. Species specific differences ................................................................... 23
  1.7. RESEARCH INTO RESISTANCE TO FACIAL ECZEMA ................................ 24
      1.7.1. Genome scan ......................................................................................... 25
      1.7.2. Candidate gene approach .................................................................... 26
  1.8. OBJECTIVES OF THIS RESEARCH .......................................................... 29

CHAPTER TWO: MATERIALS AND METHODS ............................................................... 31
  2.1. APPROVALS AND PERMITS ...................................................................... 31
  2.2. ANIMALS AND ANIMAL TISSUES ......................................................... 32
  2.3. MATERIALS AND EQUIPMENT .................................................................. 34
  2.4. DNA METHODS ............................................................................................ 34
      2.4.1. Genomic DNA isolation ................................................................. 34
      2.4.2. Oligonucleotide design and synthesis ............................................ 35
      2.4.3. Polymerase chain reaction (PCR) .................................................... 35
      2.4.4. Agarose gel electrophoresis ......................................................... 37
      2.4.5. Purification of PCR products ......................................................... 38
      2.4.6. Phenol extraction and ethanol precipitation of DNA .................... 38
      2.4.7. Cloning of DNA fragments ............................................................ 39
      2.4.8. Plasmid DNA isolation .................................................................... 43
      2.4.9. Restriction endonuclease digestion ............................................... 44
      2.4.10. Determination of nucleic acid concentration .............................. 44
      2.4.11. DNA sequencing ............................................................................. 45
      2.4.12. SNP genotyping .............................................................................. 46
  2.5. RNA METHODS ............................................................................................. 48
      2.5.1. RNA isolation .................................................................................... 48
      2.5.2. DNase treatment of RNA ............................................................... 50
      2.5.3. Formaldehyde agarose gel electrophoresis ................................. 50
      2.5.4. cDNA synthesis ............................................................................... 51
Table of Contents

2.5.5. Northern blotting .................................................. 52
2.5.6. TaqMan qRT-PCR .................................................. 54
2.5.7. SYBR green and melt curve analysis ......................... 57
2.5.8. Microarray methods .................................................. 59

2.6. CELL CULTURE METHODS ........................................... 63
2.6.1. Cells, media and consumables ......................................... 63
2.6.2. Passage of cells .......................................................... 63
2.6.3. Treatment conditions .................................................... 64
2.6.4. Assays for cell death ..................................................... 66

2.7. COMPUTATIONAL METHODS AND STATISTICAL ANALYSES ...... 67
2.7.1. General statistical analyses ............................................ 67
2.7.2. Graphing ................................................................. 68
2.7.3. Phylogenetic analyses ................................................... 68
2.7.4. Comparative mapping ................................................... 70
2.7.5. Linkage mapping ......................................................... 71

CHAPTER THREE: QUANTIFICATION OF ABCB1 EXPRESSION .......... 73
3.1. INTRODUCTION .................................................................. 73
3.1.1. ABC Transporters .......................................................... 74
3.1.2. PDR5 (Plieiotrophic drug resistance protein 5, STS1, suppressor of sporidesmin toxicity 1) ........................................... 76
3.1.3. ABCB1 (P-glycoprotein, MDR1) ......................................... 76
3.1.4. ABCB gene family ........................................................ 78
3.1.5. Aims and rationale ......................................................... 78

3.2. RESULTS ........................................................................... 79
3.2.1. TaqMan Assay Design ................................................... 79
3.2.2. TaqMan Assay Optimisation ............................................ 86
3.2.3. Quantification of ABCB1 expression using the TaqMan Assay ... 88
3.2.4. Characterisation of the Competitive RT-PCR product .......... 93
3.2.5. Northern blot using competitive RT-PCR product as a probe ....... 98

3.3. DISCUSSION ..................................................................... 102

CHAPTER FOUR: SECTION A: Identification of the closest mammalian ortholog of the S. cerevisiae ABC transporter PDR5 ................................................................. 109
4.1. INTRODUCTION ................................................................. 109
4.1.1. ABC transporters ......................................................... 109
4.1.2. Functional inference from comparative genomics and phylogenetics .. 111
4.1.3. Comparative Mapping .................................................... 112
4.1.4. Aims and rationale ......................................................... 113

4.2. RESULTS ........................................................................... 114
4.2.1. Identifying the mammalian homologs of the yeast PDR5 protein ... 114
4.2.2. Defining the ABCG family in sheep ................................. 125
4.2.3. Collation of information concerning the ABCG sub-family ...... 128
4.2.4. In silico mapping (comparative mapping) ......................... 131

SECTION B: Evaluation of ABCG2 as a candidate gene in facial eczema disease... ............................................. 133

4.3. INTRODUCTION ................................................................. 133
4.3.1. ABCG2 ........................................................................ 133
4.3.2. Linkage Mapping ......................................................... 136
CHAPTER FIVE: cDNA MICROARRAYS AS A TOOL FOR CANDIDATE GENE IDENTIFICATION

5.1. INTRODUCTION.................................................................................. 161
  5.1.1. cDNA microarray technology ........................................................ 161
  5.1.2. Microarrays and complex disease traits ........................................... 162
  5.1.3. Toxicogenomics ............................................................................ 163
  5.1.4. Biomarkers .................................................................................... 163
  5.1.5. Aims and Rationale ....................................................................... 165

5.2. RESULTS .......................................................................................... 166
  5.2.1. Description of the Microarray Slide ................................................ 166
  5.2.2. Tissue isolation and RNA extraction .............................................. 167
  5.2.3. Experimental design ....................................................................... 167
  5.2.4. Microarray hybridisations and post-processing ................................ 168
  5.2.5. Statistical analysis ........................................................................ 169
  5.2.6. Bioinformatics ............................................................................. 179
  5.2.7. Re-sequencing of EST clones ..................................................... 188
  5.2.8. Confirmation of differential expression by Northern hybridisation. 190
  5.2.9. Evaluation of the accuracy of the microarray data .......................... 211

5.3. DISCUSSION ..................................................................................... 214

CHAPTER SIX: AN IN VITRO SYSTEM TO EXAMINE THE MECHANISM OF ZINC PROTECTION AGAINST SPORIDESMIN TOXICITY

6.1. INTRODUCTION ................................................................................. 223
  6.1.1. Epidithiodioxopiperazine (ETP) toxicity ....................................... 223
  6.1.2. Zinc, copper and facial eczema disease ........................................... 224
  6.1.3. HepG2 cells as an in vitro model system ....................................... 226
  6.1.4. Aims and rationale ....................................................................... 228

6.2. RESULTS ............................................................................................ 230
  6.2.1. Sporidesmin cytotoxicity: a comparison of HepG2 with HeLa cells 230
  6.2.2. HepG2 cells and sporidesmin toxicity .......................................... 232
  6.2.3. HepG2 cells and metal toxicity (zinc and copper) .......................... 236
  6.2.4. Zinc protection of HepG2 cells from sporidesmin toxicity .......... 240
  6.2.5. The role of transcription in zinc protection from sporidesmin toxicity 242

6.3. DISCUSSION ..................................................................................... 246
Table of Contents

CHAPTER SEVEN: GENERAL DISCUSSION AND FUTURE WORK........253
7.1. Project Overview.................................................................253
7.2. Quantification of ABCB1 Gene Expression .........................254
7.3. Identification And Evaluation Of A Mammalian Ortholog Of The Yeast PDR5 Protein ..................................................256
7.4. cDNA Microarrays as a Tool for Candidate Gene Identification....260
7.5. HepG2 Cells as an In Vitro System for Studying Sporidesmin Toxicity....264
7.6. Final Remarks ......................................................................267

REFERENCES .............................................................................269

APPENDICES
APPENDIX A: Oligonucleotide primer sequences..........................297
APPENDIX B: Microarray data......................................................301
APPENDIX C: Publications arising from this work..........................303
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Structure and chemical formula of known variants of sporidesmin</td>
</tr>
<tr>
<td>1.2</td>
<td>Relative resistance and susceptibility of various species to sporidesmin</td>
</tr>
<tr>
<td>2.1</td>
<td>Details of resistant selection-line animals used in this study</td>
</tr>
<tr>
<td>2.2</td>
<td>Details of susceptible selection-line animals used in this study</td>
</tr>
<tr>
<td>4.1</td>
<td>The number of sequences and domains that were extracted from each species using PDR5 (NA_014796) as a query sequence</td>
</tr>
<tr>
<td>4.2</td>
<td>Sequences that grouped with the mammalian ABCG sub-family</td>
</tr>
<tr>
<td>4.3</td>
<td>Collation of information about the mammalian ABCG proteins</td>
</tr>
<tr>
<td>4.4</td>
<td>Comparative mapping of the mammalian ABCG family</td>
</tr>
<tr>
<td>4.5</td>
<td>Comparison of map positions with existing QTL data</td>
</tr>
<tr>
<td>4.6</td>
<td>Two point linkage analysis of the SNP marker ss65824075:G&gt;A</td>
</tr>
<tr>
<td>4.7</td>
<td>Allele distributions of ABCG2 SNP markers in FE selection lines</td>
</tr>
<tr>
<td>4.8</td>
<td>Haplotype distribution for the three ABCG2 SNP markers</td>
</tr>
<tr>
<td>5.1</td>
<td>Summary of the number and distribution of differentially expressed ESTs</td>
</tr>
<tr>
<td>5.2</td>
<td>ESTs differentially expressed in sporidesmin-dosed resistant sheep</td>
</tr>
<tr>
<td>5.3</td>
<td>ESTs differentially expressed in sporidesmin-dosed susceptible sheep</td>
</tr>
<tr>
<td>5.4</td>
<td>ESTs differentially expressed in naïve resistant sheep</td>
</tr>
<tr>
<td>5.5</td>
<td>ESTs differentially expressed in naïve susceptible sheep</td>
</tr>
<tr>
<td>5.6</td>
<td>ESTs differentially expressed in response to sporidesmin exposure</td>
</tr>
<tr>
<td>5.7</td>
<td>ESTs differentially expressed in response to genetic-line</td>
</tr>
<tr>
<td>5.8</td>
<td>Putative biomarkers for either resistance to FE or sporidesmin exposure</td>
</tr>
<tr>
<td>5.9</td>
<td>GO terms significantly associated with the differentially expressed genes in resistant sheep</td>
</tr>
<tr>
<td>5.10</td>
<td>GO terms significantly associated with the differentially expressed genes in susceptible sheep</td>
</tr>
<tr>
<td>5.11</td>
<td>Table of EST clones that were re-sequenced</td>
</tr>
<tr>
<td>5.12</td>
<td>Pearson correlation coefficients for the microarray data as compared with the Northern expression data</td>
</tr>
<tr>
<td>5.13</td>
<td>Summary of the trends of differential expression across the microarray and Northern hybridisation data</td>
</tr>
</tbody>
</table>
6.1 Effect of actinomycin D on zinc protection of HepG2 cells against sporidesmin toxicity ......................................................................................................................................................243
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Generic structure of an oxidised ETP ............................................................... 4</td>
</tr>
<tr>
<td>1.2</td>
<td>Structure of gliotoxin (A) and sirodesmin (B) .................................................. 5</td>
</tr>
<tr>
<td>1.3</td>
<td>The reduction of sporidesmin disulphide to sporidesmin dithiol .................... 10</td>
</tr>
<tr>
<td>1.4</td>
<td>The proposed reaction scheme for the redox cycling of sporidesmin ............. 11</td>
</tr>
<tr>
<td>1.5</td>
<td>The proposed mechanism for zinc coordination of sporidesmin ..................... 12</td>
</tr>
<tr>
<td>1.6</td>
<td>The proposed mechanism of mixed disulphide formation by sporidesmin .... 13</td>
</tr>
<tr>
<td>1.7</td>
<td>Clinical manifestation of FE disease in sheep ............................................. 15</td>
</tr>
<tr>
<td>1.8</td>
<td>Haematoxylin and eosin stained sheep liver sections ...................................... 16</td>
</tr>
<tr>
<td>2.1</td>
<td>Timeline for the experimental treatments of HepG2 cells .............................. 65</td>
</tr>
<tr>
<td>3.1</td>
<td>Schematic of ABC transporters ....................................................................... 73</td>
</tr>
<tr>
<td>3.2</td>
<td>Schematic of the <em>S. cerevisiae</em> ABC transporter PDR5 ................................ 76</td>
</tr>
<tr>
<td>3.3</td>
<td>Diagram illustrating TaqMan chemistry ....................................................... 79</td>
</tr>
<tr>
<td>3.4</td>
<td>Neighbour-joining tree for the human ABCB sub-family ................................ 80</td>
</tr>
<tr>
<td>3.5</td>
<td>Sequence similarity of human ABCB1 and ABCB4 ....................................... 81</td>
</tr>
<tr>
<td>3.6</td>
<td>Testing of the ABCB1 qRT-PCR primers ....................................................... 82</td>
</tr>
<tr>
<td>3.7</td>
<td>Multiple sequence alignment of ovine GAPDH sequences.................................. 83</td>
</tr>
<tr>
<td>3.8</td>
<td>Testing of the GAPDH qRT-PCR primers ....................................................... 85</td>
</tr>
<tr>
<td>3.9</td>
<td>TaqMan qRT-PCR assay optimisation .................................................................. 87</td>
</tr>
<tr>
<td>3.10</td>
<td>qRT-PCR standard curves for GAPDH and ABCB1 .......................................... 89</td>
</tr>
<tr>
<td>3.11</td>
<td>Graphs show the relative levels of ABCB1 and GAPDH mRNAs .......................... 90</td>
</tr>
<tr>
<td>3.12</td>
<td>Expression of ABCB1 normalised to GAPDH for individual animals ............... 92</td>
</tr>
<tr>
<td>3.13</td>
<td>Expression of ABCB1 normalised to GAPDH for pooled animals ...................... 93</td>
</tr>
<tr>
<td>3.14</td>
<td>Comparison of oligonucleotide primers used for competitive RT-PCR ............ 94</td>
</tr>
<tr>
<td>3.15</td>
<td>Agarose gel showing restricted competitive RT-PCR product .......................... 95</td>
</tr>
<tr>
<td>3.16</td>
<td>SYBR Green Melt curve analysis for competitive RT-PCR products ............... 97</td>
</tr>
<tr>
<td>3.17</td>
<td>Northern analysis of the competitive RT-PCR product ................................ 99</td>
</tr>
<tr>
<td>4.1</td>
<td>Schematic of the <em>S. cerevisiae</em> ABC transporter PDR5 .............................. 110</td>
</tr>
<tr>
<td>4.2</td>
<td>Simplified neighbour-joining tree of ABC domain sequences ........................ 117</td>
</tr>
<tr>
<td>4.3</td>
<td>Detailed neighbour-joining tree of ABC domain sequences ........................... 119</td>
</tr>
<tr>
<td>4.4</td>
<td>Neighbour-joining analysis of the ABCG sub-family of ABC domains .......... 122</td>
</tr>
<tr>
<td>4.5</td>
<td>Maximum parsimony analysis of the ABCG sub-family of ABC domains .......... 123</td>
</tr>
</tbody>
</table>
### List of Figures

4.6 Simplified phylogeny of Eutherian mammals ............................................125  
4.7 Stylised syntenic map of the ABCG2 locus ..................................................127  
4.8 Nucleotide and predicted amino acid sequence of the sheep ABCG2 cDNA 141  
4.9 The hydropathy plot of the deduced ovine ABCG2 protein sequence ..........142  
4.10 DNA sequence of the ovine ABCG2 intron 4 ............................................145  
4.11 Localisation of the ABCG2 locus on ovine chromosome 6 ..........................146  
4.12 QTL analysis of ovine chromosome 6 ........................................................147  
4.13 Association studies of markers in FE selection lines ................................149  
4.14 Northern analysis of ovine ABCG2 ..........................................................150  
5.1 Schematic of the microarray experimental design .......................................167  
5.2 Pseudo-coloured image of a hybridised microarray slide .............................168  
5.3 Northern hybridisation of CAT mRNA in ovine liver ..................................193  
5.4 Northern hybridisation of LECT2 mRNA in ovine liver ...............................194  
5.5 Northern hybridisation of MATR3 mRNA in ovine liver ..............................196  
5.6 Northern hybridisation of MT2 mRNA in ovine liver ..................................198  
5.7 Northern hybridisation of RAC2 mRNA in ovine liver ................................203  
5.8 Northern hybridisation of CRP mRNA in ovine liver ..................................205  
5.9 Northern hybridisation of FGA mRNA in ovine liver ..................................206  
5.10 Northern hybridisation of GSTA2 mRNA in ovine liver ..............................208  
5.11 Northern hybridisation of EPAS1 mRNA in ovine liver ..............................209  
5.12 Northern hybridisation of KBRAS2 mRNA in ovine liver ............................210  
6.1 Cytotoxic effect of sporidesmin on HepG2 cells .........................................230  
6.2 Cytotoxic effect of sporidesmin on HeLa cells ...........................................231  
6.3 Time course of sporidesmin toxicity in HepG2 cells ...................................233  
6.4 Cytopathological effects of sporidesmin on HepG2 cells ............................235  
6.5 Cytotoxic effect of zinc sulphate on HepG2 cells .......................................237  
6.6 Effect of copper sulphate on HepG2 cells .................................................238  
6.7 Cytotoxic effect of mixtures of zinc and copper on HepG2 cells ....................239  
6.8 Effect of zinc pre-treatment on sporidesmin toxicity ..................................241  
6.9 Effect of zinc pre-treatment time on sporidesmin toxicity ...........................242  
6.10 Effect of actinomycin D on zinc protection against sporidesmin toxicity ....244
Abbreviations

All standard abbreviations are defined in the “instructions to authors” of the Biochemical Journal (2007). Non-standard abbreviations are defined below.

All gene names are given in full on their first usage and then abbreviated on subsequent usages. Gene names that are used commonly in the thesis are defined below. All gene names and symbols are the standard names as defined by NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene).

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ATP binding cassette</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>ANT</td>
<td>Adenine nucleotide transporter</td>
</tr>
<tr>
<td>BMP</td>
<td>Bone morphogenetic protein</td>
</tr>
<tr>
<td>BTA</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>CEL</td>
<td>Caenorhabditis elegans</td>
</tr>
<tr>
<td>[α-32P]dCTP</td>
<td>2'-deoxyctydine 5'-[α-32P]dCTP triphosphate</td>
</tr>
<tr>
<td>CFA</td>
<td>Canis familiaris</td>
</tr>
<tr>
<td>CFU</td>
<td>Colony forming units</td>
</tr>
<tr>
<td>cM</td>
<td>Centimorgan</td>
</tr>
<tr>
<td>C&lt;sub&gt;T&lt;/sub&gt;</td>
<td>Threshold cycle</td>
</tr>
<tr>
<td>DEPC</td>
<td>Diethylpropyl carbonate</td>
</tr>
<tr>
<td>DME</td>
<td>Drosophila melanogaster</td>
</tr>
<tr>
<td>EASE</td>
<td>Expression analysis systematic explorer</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed sequence tag</td>
</tr>
<tr>
<td>ETP</td>
<td>Epidithiodioxopiperazine toxin</td>
</tr>
<tr>
<td>FE</td>
<td>Facial eczema</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GDH</td>
<td>Glutamate dehydrogenase</td>
</tr>
<tr>
<td>GGT</td>
<td>Gamma glutamyltransferase</td>
</tr>
<tr>
<td>GO</td>
<td>Gene ontology</td>
</tr>
<tr>
<td>HSA</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>LB</td>
<td>Luria – Bertani medium</td>
</tr>
<tr>
<td>LC</td>
<td>Lethal concentration</td>
</tr>
<tr>
<td>LD</td>
<td>Linkage disequilibrium</td>
</tr>
<tr>
<td>LDH</td>
<td>Lactate dehydrogenase</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>Description</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>LMW</td>
<td>Low molecular weight</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>Matrix-assisted laser desorption/ionization - time of flight mass spectrometry</td>
</tr>
<tr>
<td>MCS</td>
<td>Multiple cloning site</td>
</tr>
<tr>
<td>MMU</td>
<td>Mus musculus</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-Morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>NBD</td>
<td>Nucleotide binding domain</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>NOEL</td>
<td>No observed effect level</td>
</tr>
<tr>
<td>OAR</td>
<td>Ovis aries</td>
</tr>
<tr>
<td>PEG</td>
<td>Poly(ethylene glycol)</td>
</tr>
<tr>
<td>PIPES</td>
<td>Piperazine-N,N'-bis[2-ethane-sulfonic acid]</td>
</tr>
<tr>
<td>PMT</td>
<td>Photomultiplier tube</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative (real time) reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>QTL</td>
<td>Quantitative trait loci</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RNO</td>
<td>Rattus norvegicus</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>SCE</td>
<td>Saccharomyces cerevisiae</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SOC</td>
<td>Super Optimal broth with Catabolite repression (glucose)</td>
</tr>
<tr>
<td>SSC</td>
<td>Standard saline citrate</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris acetate EDTA buffer</td>
</tr>
<tr>
<td>TE</td>
<td>Tris EDTA buffer</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris-(hydroxymethyl)-amino methane</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>v/v</td>
<td>Volume in volume</td>
</tr>
<tr>
<td>w/v</td>
<td>Weight by volume</td>
</tr>
<tr>
<td>X-Gal</td>
<td>5-bromo-4-chloro-3-indoyl-β-D-galactoside</td>
</tr>
</tbody>
</table>
Chapter One
General Introduction

Facial eczema (FE) is a secondary photosensitisation disease caused by exposure to sporidesmin, the toxin produced by the saprophytic fungus *Pithomyces chartarum*. FE affects most farmed ruminants including sheep (*Ovis aries*), cattle (*Bos taurus*), alpacas (*Vicugna pacos*), and fallow deer (*Dama dama L.*) (Mortimer and Ronaldson, 1983).

In the short term FE can be fatal to the animal. However, even if the animal recovers there are significant long term production losses associated with this disease. These losses include decreased weight and weight gain (Smeaton *et al.*, 1985), a reduction in carcass and offal weight, as well as decreased milk, wool and pelt production (Pinto *et al.*, 2005). Animals also show a decrease in reproductive success (Moore *et al.*, 1983; Smeaton *et al.*, 1985; Morris *et al.*, 1991) and longevity (Morris *et al.*, 2002). These effects may persist for some years following exposure to sporidesmin and the animals may be at a greater risk from subsequent toxic challenges (Mortimer and Ronaldson, 1983). The severe consequences of sporidesmin exposure results in an annual cost to the New Zealand agricultural industry of between $63 and $126 million (Anonymous, 1990), and FE is the second most important production disease in New Zealand (behind parasite infection) (Brunsdon, 1988).

Several approaches have been taken to control the growth of the fungus, including bio-competitive exclusion (Collin and Towers, 1995a) and fungicide use (Towers, 1986), neither of which have been particularly successful. The most effective treatment for FE to date is the prophylactic administration of zinc sulphate (Smith *et al.*, 1977). However, the mechanism by which zinc protects animals against sporidesmin toxicity is unknown, and for various reasons the long-term treatment of animals with zinc is impractical. Given that resistance to FE has a significant genetic
component (refer to 1.6) it is thought that the best long-term management strategy for FE in New Zealand is using marker assisted selection to exploit the genetic heritability of resistance (Campbell et al., 1981) to breed animals resistant to FE.

1.1 AETIOLOGY OF FACIAL ECZEMA

FE was first observed in New Zealand in the early 20\textsuperscript{th} century as a condition resulting in exudation and swelling of the skin on the head and ears of afflicted sheep. It was quickly realised that the facial lesions were actually secondary to the primary disease-causing lesion, which was liver injury. This liver injury was proposed to cause photosensitisation due to the build up of phyloerythrin, which is a photodynamic breakdown product of chlorophyll. The symptoms and pathology of FE were first comprehensively described in 1942 and the liver was determined to be the major organ system affected (Cunningham et al., 1942) (refer section 1.4).

Because of the non-transmissible nature of FE and the seasonal incidence it was thought that the disease may be caused by the ingestion of an abnormal metabolite present in rapidly growing ryegrass. Dried ryegrass fed to guinea pigs (Cavia porcellus) induced FE symptoms, confirming that a compound associated with the ryegrass was the causative agent. Subsequently it was found that exposure of laboratory animals to P. chartarum spores was also able to induce the symptoms of disease and the causative compounds were found to be the sporidesmins (Table 1.1) (Percival and Thornton, 1958; Thornton and Percival, 1959).

1.2 ECOLOGY AND DISTRIBUTION OF P. CHARTARUM

P. chartarum is a ubiquitous fungus found in countries as ecologically diverse as Zambia, Australia, India and the United States (Atherton et al., 1974). The fungus has been found associated with various substrates including soil and human foodstuffs. However, its most common habitat is plant debris, and P. chartarum thrives in the litter layer of pasture (Mortimer and Ronaldson, 1983). The fungus sporulates freely in warm conditions (10 - 30°C) with high humidity (Brook, 1963).
This is consistent with the observed increase in FE outbreaks when periods of hot dry weather are interspersed with warm rain and high humidity (Mitchell et al., 1959). Several factors can influence the incidence and severity of FE including microclimatic and environmental conditions as well as differences in past and present grazing pressures. These differences mean that FE incidence can vary between pastures on the same farm as well as between farms in the same area (Smith et al., 1987). In general, FE is commonly found in the North Island and upper South Island of New Zealand and usually appears in late summer through to the end of autumn (January through to June). In low lying areas of the North Island serious outbreaks of FE are experienced on sheep, beef and dairy farms approximately every three years, although this is heavily dependent on weather patterns (Morris et al., 2004).

Strains of *P. chartarum* can be either toxigenic (sporidesmin producing) or non-toxigenic (unable to produce sporidesmin), and broadly the production of sporidesmin is related to the sporulating activity of the individual strains (di_Menna et al., 1970). Although there are a few exceptions i.e. cultures that sporulate freely but do not produce sporidesmin; these isolates are generally assumed to have mutations or deficiencies in the sporidesmin biosynthetic pathway (Halder et al., 1981; Brewer et al., 1989). The proportion of toxigenic strains varies with geographical location, for example 2% of the *P. chartarum* strains in Brazil are toxigenic compared with 67% in Australia and 95% in New Zealand (Collin et al., 1998). A recent report found that 100% of the isolates examined from the Azores Islands in Portugal were toxigenic, consistent with severe outbreaks of FE seen in cattle and sheep between 1999 and 2001 (Pinto et al., 2005). The geographical variation observed in the toxigenicity of *P. chartarum* may explain why New Zealand is more severely affected by FE than any other country.
1.3 SPORIDESMIN CHEMISTRY AND TOXICOLOGY

1.3.1 ETP mycotoxins

Sporidesmin is an epidithiodioxopiperazine (ETP) mycotoxin. The ETP compounds are a group of toxins that are characterised by the presence of a diketopiperazine ring and an internal disulphide bridge (Fig. 1.1, Fig. 1.2, Table 1.1). This disulphide bridge imparts the toxicity to these ETP compounds (Trown and Bilello, 1972; Mullbacher et al., 1986).

To date there are 14 ETP compounds known, and the majority of research has been focussed on the first ETP compound to be discovered, gliotoxin (Fig. 1.2A) (Weindling and Emerson, 1936; Weindling, 1941) and more recently on sirodesmin (Fig. 1.2B) (Gardiner and Howlett, 2005; Gardiner et al., 2005a; Gardiner et al., 2005b).

Gliotoxin was first identified from Gliocladium fimbriatum (Johnson et al., 1943), but is also produced by a number of Penicillium and Aspergillus species (Richard et al., 1994) as well as Thermoascus (Waring et al., 1987) and Candidia (Shah and Larsen, 1991). Gliotoxin is thought to be involved in the aetiology of invasive Aspergilloses infections in avian species (Richard et al., 1996) and humans (Sutton et al., 1994; Richard et al., 1996; Sutton et al., 1996; Tomee and Kauffman, 2000) particularly in immuno-compromised patients such as cancer patients (Bodey et al., 1992) and possibly AIDS patients (Eichner and Mullbacher, 1984). Gliotoxin is also proposed to play a role in gut barrier dysfunction resulting from fungal overgrowth (Upperman et al., 2003).

There has been substantial interest in gliotoxin as a therapeutic agent as it prevents viral RNA replication, via inhibition of reverse transcriptase (Rightsel et al., 1964; De Clercq et al., 1978). Gliotoxin has also been examined for therapeutic potential as

![Figure 1.1: Generic structure of an oxidised ETP. In some compounds the sulphur bridge contains one, three or four sulphur atoms; these compounds are usually co-produced with those containing two. R=any atom or group.](image-url)
Chapter One: Introduction

Figure 1.2: Structure of two ETP compounds. A) gliotoxin, which is produced by *Aspergillus fumigatus*, *Trichoderma virens*, *Penicillium* spp., and *Candida albicans*, B) sirodesmin, which is produced by *Leptosphaeria maculans* and *Sirodesmium diversum*.

an immunosuppressive agent in allogenic transplantation (Lissing et al., 1988; Mullbacher et al., 1988; Tuch et al., 1988; Sutton et al., 1995a; Sutton et al., 1995b). Additionally, gliotoxin has been shown to effectively protect against autoimmune diabetes in a susceptible mouse model (Larsen et al., 2000; Liu et al., 2000). The immunomodulatory properties of gliotoxin may be mediated through its selective toxicity to cells of the immune system (Sutton et al., 1994; Sutton et al., 1995a; Shah et al., 1998; Suen et al., 2001). Alternatively, gliotoxin is also capable of modulating immune function by modifying the gene expression and interactions between immune cells (Yamada et al., 2000; Wichmann et al., 2002; Stanzani et al., 2005; Niide et al., 2006).

Gliotoxin causes apoptotic cell death in varicus cell types both in vitro (Waring et al., 1988; Piva, 1994; Waring et al., 1997; Zhou et al., 2000; Suen et al., 2001; Wright et al., 2001) and in vivo. In fact gliotoxin has been shown to selectively induce apoptosis in hepatic stellate cells (Kweon et al., 2003). This is of interest as hepatic stellate cells have a role in fibrosis following liver injury (Elsharkawy et al., 2005). It is thought that by causing apoptosis of hepatic stellate cells, gliotoxin may act as an anti-fibrogenic compound with therapeutic potential (Wright et al., 2001; Dekel et al., 2003; Orr et al., 2004).

Sirodesmin is produced by the Loculoascomycete *Leptosphaeria maculans* which causes blackleg in *Brassica napus* (rapeseed or canola) (Howlett et al., 2001). The role of sirodesmin in the pathogenesis of blackleg is unclear as sirodesmin is not required for lesion formation on leaves (Gardiner et al., 2004). It is possible that
sirodesmin could be involved in the necrotrophic phase of infection or may provide a competitive advantage against other micro-organisms during saprophytic growth (Gardiner et al., 2005a).

Despite the structural dissimilarity and the differences in the aetiology of illnesses they cause, ETP toxins are thought to mediate toxicity in similar ways (refer 1.3.4).

1.3.2 Sporidesmin Biosynthesis
Most fungal toxins are secondary metabolites that are not required for fungal growth or sporulation but may provide selective advantage to the fungus under particular conditions (Gardiner et al., 2005a). All ETP compounds are synthesised via condensation of at least two amino acids, one of which must be aromatic (phenylalanine, tryptophan or tyrosine) as the diketopiperazine ring is derived from a cyclic dipeptide. Sporidesmin is synthesised via condensation of L-tryptophan with L-alanine to form the tricyclic ring system followed by sulfation of the dipiperazine ring using a selection of sulphur donors (Kirby and Robins, 1980).

P. chartarum produces seven structural variants of sporidesmin (Table 1.1). The most abundant variant is sporidesmin A which accounts for approximately 80% of the sporidesmin produced (Atherton et al., 1974).

The biosynthetic gene cluster for sirodesmin has been identified in L. maculans using fungal gene-disruption (Gardiner et al., 2004). These data have been used to predict the corresponding biosynthetic gene pathway for Gliotoxin in A. fumigatus (Gardiner et al., 2005b). Although no sequence data exist for P. chartarum key features of the biosynthetic pathway are likely to be conserved.
Table 1.1: Structure and chemical formula of known variants of sporidesmin (% yield = the % of total sporidesmin yield when P. chartarum is grown on rye-grain medium (Atherton et al., 1974))

<table>
<thead>
<tr>
<th>Variant</th>
<th>% Yield</th>
<th>Formula</th>
<th>Mw (Da)</th>
<th>Chemical Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpdA</td>
<td>82.17</td>
<td>C₁₆H₂₀CIN₅O₂S₂</td>
<td>473.946</td>
<td><img src="image1.png" alt="Chemical Structure" /></td>
</tr>
<tr>
<td>SpdB</td>
<td>8.28</td>
<td>C₁₂H₁₆CIN₄O₂S₂</td>
<td>457.947</td>
<td><img src="image2.png" alt="Chemical Structure" /></td>
</tr>
<tr>
<td>SpdD</td>
<td>8.28</td>
<td>C₁₆H₂₀CIN₅O₂S₂</td>
<td>504.015</td>
<td><img src="image3.png" alt="Chemical Structure" /></td>
</tr>
<tr>
<td>SpdE</td>
<td>8.28</td>
<td>C₁₆H₂₀CIN₅O₂S₃</td>
<td>506.006</td>
<td><img src="image4.png" alt="Chemical Structure" /></td>
</tr>
<tr>
<td>SpdF</td>
<td></td>
<td>C₁₂H₁₆CIN₄O₂S</td>
<td>455.913</td>
<td><img src="image5.png" alt="Chemical Structure" /></td>
</tr>
<tr>
<td>SpdG</td>
<td>0.1</td>
<td>C₁₆H₂₀CIN₅O₂S₄</td>
<td>538.066</td>
<td><img src="image6.png" alt="Chemical Structure" /></td>
</tr>
<tr>
<td>SpdH</td>
<td>0.1</td>
<td>C₁₆H₂₀CIN₅O₂S₄</td>
<td>441.948</td>
<td><img src="image7.png" alt="Chemical Structure" /></td>
</tr>
</tbody>
</table>
1.3.3 Sporidesmin Metabolism

Little is known about the way in which sporidesmin is metabolised by ruminants. The usual route of exposure to sporidesmin is via the gastrointestinal tract. Following exposure, unconjugated sporidesmin can be rapidly (within 2 hours) detected in the plasma, urine and bile. The concentration detected in the bile is 100 times greater than the plasma and 10 times greater than in urine (Mortimer and Stanbridge, 1968).

Using $^{35}$S labelled sporidesmin, maximum biliary excretion of sporidesmin was observed between 1 and 3 hours after exposure and basal levels were restored after approximately 24 hours (Fairclough and Smith, 1983). This differs from the previously reported maximum of 1 – 8 hours (Mortimer and Taylor, 1962), but the earlier study (Mortimer and Taylor, 1962) used only 5 animals and also a relatively insensitive cytotoxicity assay to determine sporidesmin concentrations. Both of these assays (cytotoxicity and $^{35}$S) only detect unmodified sporidesmin, which is problematic as the most abundant sporidesmin metabolites are mono- and di-hydroxy derivatives (Fairclough et al., 1978), in these derivatives the two sulphur atoms have been replaced with a hydroxyl group and a hydrogen atom (Fairclough et al., 1978). These modifications negate the toxicity of the compound (Mullbacher et al., 1986) and eliminate the radionlabelled sulphur. Urinary excretion times were determined to be between 1 and 8 hours (Mortimer and Taylor, 1962; Fairclough et al., 1978), but the concentrations observed were verging on undetectable and this estimate is likely to be erroneous. A more recent study using a sensitive and reliable immunoassay for sporidesmin found maximum urinary excretion of sporidesmin between 15 and 30 hours after exposure (Smith et al., 1999).

The high concentrations of sporidesmin observed in the bile may account for some of the liver pathology seen associated with FE (refer 1.4) particularly the pericholangitis and obstructive biliary hepatopathy (Mortimer and Stanbridge, 1969). Although the observed concentration of sporidesmin in the urine was relatively low it may account for some of the rarer presentations of FE including urinary bladder cytotoxicity and haemorrhagic cystitis (Mortimer and Taylor, 1962).
Delineation of the *in vivo* metabolism of sporidesmin is crucial to understanding sporidesmin toxicity and how this toxicity contributes to disease associated pathology. Using isolated hepatic microsomes at least 8 metabolites of sporidesmin have been identified, including the mono- and dihydroxy- derivatives (Fairclough *et al.*, 1978). However, the microsomal drug detoxification system primarily consists of the cytochrome P450 dependent monooxygenase enzymes, and other enzyme systems may also play an important role in sporidesmin metabolism.

A further four metabolites were identified in the urine of animals exposed to sporidesmin. These metabolites were all significantly more polar than sporidesmin and none had an absorbance at 218 nm suggesting metabolic modification of the chromophore or that the levels of metabolite were below the limits of UV detection (Smith *et al.*, 1999).

### 1.3.4 Sporidesmin Toxicity

While it is known that the bridged disulphide moiety of the diketopiperazine is required for the toxicity of sporidesmin and other ETP compounds (Trown and Bilello, 1972; Mullbacher *et al.*, 1986), the underlying cellular events that mediate this toxicity are not well defined. There are two prevailing theories regarding the toxic mechanism of sporidesmin and other ETP compounds. The first mechanism involves the redox cycling of sporidesmin to generate reactive oxygen species (ROS) that are capable of mediating cell damage. The second mechanism involves the modification of cellular proteins via reactive thiol residues. In this way, sporidesmin may alter protein function, turnover or possibly induce cell death. There is substantial experimental evidence for both of these mechanisms and they are discussed in more detail in sections 1.3.4.1 and 1.3.4.2.

Both mechanisms rely on the reduction of the ETP disulphide bridge to the dithiol moiety with a thiolate anion as the intermediate step (Fig. 1.3). This reduction is thought to be mediated by intracellular glutathione and is required for the intracellular retention of ETP compounds (Bemardo *et al.*, 2003).
Figure 1.3: The reduction of sporidesmin disulphide to sporidesmin dithiol. This reaction occurs in the presence of glutathione and generates the reactive sporidesmin thiolate anion intermediate.

1.3.4.1 Redox cycling of sporidesmin

There has been a great deal of research concerning the auto-oxidation of thiol containing compounds (Munday, 1982; Munday, 1984a; Munday, 1984b; Munday, 1985; Munday, 1987). It has been shown that the disulphide bond of sporidesmin chemically reacts in a glutathione linked (Munday, 1982), metal-catalysed cycle of reduction and oxidation to produce the toxic free radical superoxide and a cascade of other free radical species (Fig. 1.4) (Munday, 1984a).

Initially, sporidesmin disulphide is reduced to sporidesmin dithiol by glutathione (Fig. 1.3 and 1.4 (reactions 1, 2)). The intermediate in this reduction is the thiolate anion of sporidesmin. This thiolate anion can interact with a suitable metal catalyst such as copper to generate the thiy1 radical and the superoxide radical (reaction 3a). The thiolate anion can also participate in two further reactions that generate the thiy1 radical (reactions 3b,c). In reaction 3b the superoxide radical generated in reaction 3a is dismutated to hydrogen peroxide also generating a thiy1 radical. In reaction 3c the hydroxyl radical is converted to the hydroxyl ion also generating the thiy1 radical. The thiy1 radical can then be recycled back to sporidesmin disulphide by interacting with either hydrogen peroxide (reaction 4a) or molecular oxygen (reaction 4b) to generate the hydroxyl ion or superoxide, respectively. The hydroxyl ion may also be generated through metal-catalysed interactions between superoxide and hydrogen peroxide (Munday, 1987). If endogenously formed hydrogen peroxide is not detoxified to water, formation of the hydroxyl radical may be formed by Haber Weiss or Fenton chemistry.

Intracellular ROS and other free radicals are proposed to play a role in cell signalling (Kamata and Hirata, 1999) and modulation of gene expression (Haddad, 2002).
However, under standard conditions the intracellular production of ROS is small and it has been suggested that sporidesmin treatment could induce massive production of ROS. This wave of ROS production could overwhelm the cellular detoxification systems and cause widespread cellular damage thereby accounting for some of the cytotoxicity associated with sporidesmin exposure.

While all ROS are potentially damaging it is thought that the hydroxyl radical in particular is responsible for inducing the cytopathological changes induced by sporidesmin (Kappus and Sies, 1981). The hydroxyl radical is particularly damaging as it is very unstable and cellular exposure can lead to protein carbonylation, membrane peroxidation and DNA damage (Imlay, 2003). These widespread changes to normal cell function may lead to cell death and ultimately liver dysfunction.

Other disulphides show a similar ability to redox cycle between reduced and oxidised compounds, but usually at significantly higher concentrations (Cordiner and Jordan, 1983). This indicates that the piperazine ring of sporidesmin provides a chemically reactive environment for the bridged disulphide.
Free radical formation requires the presence of a suitable metal catalyst such as copper. This may perhaps explain the protective effect of dosing animals with zinc, as zinc could inhibit free radical production via the formation of stable mercaptide linkages with sporidesmin (Munday, 1984b). Sporidesmin and zinc are known to form a stable complex in vitro (Woodcock et al., 2001a; Woodcock et al., 2001b), as shown in Fig. 1.5. The pro-oxidant nature of copper may also account for some of the species specific effects in susceptibility as ruminants absorb and store copper more readily than non-ruminant species and this is reflected in the high levels of copper found in ruminant livers (Underwood, 1978).

1.3.4.2 Modification of cellular thiols
A large number of proteins contain free thiols that can be modified by the formation of internal disulphides or by mixed disulphides with low-molecular-mass thiols. ETP compounds contain a reactive disulphide and have been shown to form mixed disulphides with free thiol groups (Waring and Beaver, 1996). The proposed mechanism for the formation of mixed disulphides is shown in Fig. 1.6. These mixed disulphide compounds could possibly have altered structure, activity or turnover compared to the native proteins.

ETPs do not have exclusive protein targets, exposing cells to radiolabelled gliotoxin results in many cellular proteins becoming labelled (Waring and Beaver, 1996). Gliotoxin has been shown to interact with several functionally diverse proteins and modulate their activity. For some of these proteins the cysteine residues that form the mixed disulphide bonds have been identified. For example, gliotoxin has been shown to be a potent inhibitor of viral RNA polymerase 3Dpol (Rodriguez and
Carrasco, 1992), creatine kinase (Hurne et al., 2000) and Ras farnesyltransferase (Hara and Han, 1995). Recently, sporidesmin has been shown to be a potent inhibitor of glutaredoxin via the formation of mixed disulphides between the protein and toxin (Srinivasan et al., 2006). The effect of mixed disulphides on protein activity has been most closely examined for the enzyme alcohol dehydrogenase (Waring et al., 1995). It has been shown that gliotoxin inactivates alcohol dehydrogenase by forming a 1:1 covalent complex with a cysteine residue, either 281 or 282 (Waring et al., 1995). In addition, gliotoxin has been shown to inhibit NFκB, probably via interaction with an essential thiol residue (Pahl et al., 1996). It is proposed that this inhibition of NFκB may be important for the pathophysiology of gliotoxin exposure, as it is critical for T-cell activation as well as the production of several pro-inflammatory cytokines (Abbas and Lichtman, 2003). Therefore the inhibition of NFκB may be responsible for gliotoxin’s immunosuppressive properties (refer 1.3.1).
Mitochondria are also targeted by ETP compounds. Sporidesmin has been shown to produce rapid and reversible swelling of mitochondria possibly by altering membrane permeability (Middleton, 1974) or by altering reactive thiol containing proteins (Middleton, 1974). Gliotoxin treatment causes the release of both calcium and magnesium from mitochondria (Schweizer and Richter, 1994; Salvi et al., 2004). Gliotoxin also causes a redox-dependent change in electrophoretic mobility of a component of the mitochondrial permeability transition pore, the adenine nucleotide transporter (ANT). This transition pore is an essential mediator of apoptosis modulated by mitochondria (Orr et al., 2004), the ANT itself is sensitive to oxidizing agents and has two cysteine residues which may be targeted by gliotoxin (McStay et al., 2002).

1.4 PATHOLOGY AND PATHOPHYSIOLOGY OF FACIAL ECZEMA

FE is a hepatogenous (secondary) photosensitisation disease. Secondary photosensitisation diseases occur as the result of impaired liver function. In FE the primary liver damage is caused by ingestion of sporidesmin. Ruminants grazing on P. chartarum infested pasture are exposed to sporidesmin via the gastrointestinal route. The sporidesmin is then absorbed from the intestine and transported to the liver by the portal venous blood. In the liver the sporidesmin is concentrated (Towers, 1970) and excreted, primarily in the bile (Mortimer and Stanbridge, 1968) (refer 1.3.3).

FE is characterised by necrosis and apoptosis of liver cells, bile duct destruction and inflammation of hepatic and portal vein branches (Cunningham et al., 1942). The damaged areas are infiltrated by granulation tissue in the process of normal repair and this causes obstruction of the bile ducts and restricts bile flow (obstructive cholestasis). The hepatocellular damage and repair results in severe liver damage as seen in Fig. 1.7A and 1.7B.

Obstruction to bile flow prevents excretion of metabolic by-products including phylloerythrin. Phylloerythrin is a photodynamic pigment derived from microbial
breakdown of plant chlorophyll in the gut. Phylloerythrin is usually conjugated in the liver and excreted in the bile. However, in FE the bile duct obstruction causes phylloerythrin to accumulate in the serum to levels exceeding 0.3 μM (un-affected sheep maintain phylloerythrin concentrations of less than 0.1 μM) (Scheie et al., 2003). When exposed to ultraviolet light the phylloerythrin undergoes photochemical reactions which cause the skin lesions that are seen in later stages of disease. The inflammation and scabbing is found on exposed areas of the animal, in particular the face (Fig. 1.7C) (Mortimer and Ronaldson, 1983). The onset of this photosensitisation is usually 10 - 20 days after sporidesmin exposure and several days after significant liver damage.

Once photosensitisation is observed the disease process is irreversible. Photosensitised animals actively seek shade and venture out only in overcast weather or in darkness. The animals exhibit behavioural changes consistent with the irritation

![Figure 1.7](image-url)

**Figure 1.7:** Clinical manifestation of FE disease in sheep. A) The liver of a susceptible sheep exposed to sporidesmin demonstrating the primary hepatic lesions with focal necrosis. B) Comparison of a healthy liver (lower) and two damaged livers (upper) of animals suffering from facial eczema. The damaged livers clearly demonstrate the extensive fibrosis and hepatic atrophy leading to microhepatica. C) A susceptible sheep exhibiting the clinical presentation of sporidesmin toxicity, that is the photosensitisation and scarring following exposure to UV light.
that the lesions are causing. The behavioural changes include head shaking, scratching and rubbing. The irritation and subsequent inflammation causes oedema of the ears, eyelids, face and lips and any other area of the animal not covered by wool. There can also be seepage of fluid through the oedematous skin which causes an increased risk of secondary infection. This photosensitisation can persist for several weeks (Cunningham *et al.*, 1942). There is no treatment for the photosensitisation and care of the animals is primarily palliative (refer 1.5).

When examined *post mortem* the liver displays extremely abnormal pathology. There are patches of discolouration which leaves the liver irregularly blotched with fine mottling over the entire surface, and there is additional evidence of necrosis and severe fibrosis (Fig. 1.7A,B). The damage is not uniformly distributed through the whole liver, usually the left lobe is more severely affected with the margin of the left lobe usually severely atrophied and fibrous. The difference in severity between lobes could possibly be explained by differences in the portal blood flow. The primary blood supply for the liver is the portal vein, which drains both the superior and inferior mesenteric arteries from the gastrointestinal tract. The inferior mesenteric

![Figure 1.8](image_url)

Figure 1.8: Haematoxylin and eosin stained sheep liver sections taken 6 weeks after sporidesmin exposure. (A) The sheep was dosed with sporidesmin at 0.3 mg/kg (live weight). The serum GGT level at 3 weeks was 67 U. The hepatic lobules are well defined and regular with unobstructed central veins and portal tracts, consistent with normal liver histology. However, in (B) the sheep was dosed with sporidesmin at 0.2 mg/kg (live weight) and the serum GGT level at 3 weeks was 618 U consistent with the abnormal histology seen in the tissue sections. The hepatic lobules have lost their regular defined shape, the central veins are irregularly shaped and the portal triads appear partially obstructed. Most noticeably however is the large degree of fibrosis (seen in the more purple stained sections), this fibrosis suggests recovery from hepatocellular damage. Photomicrographs were kindly provided by Dr. S. H. Phua.
artery supplies the colon and drains via the splenic vein into the portal vein. In humans, at the juncture of the inferior and superior mesenteric arteries with the portal vein there is uneven mixing of the blood. This results in the left lobe of the liver being supplied primarily by the inferior mesenteric artery (via the splenic vein) (Shiomi et al., 1996). This phenomenon has not been examined in sheep, but if conserved then compounds absorbed through the colon would concentrate in the left lobe of the liver. Additionally, the blood from the inferior mesenteric artery is more oxygenated and this oxygenation may exacerbate the toxicity of sporidesmin.

Histologically, the portal area of the lobule is the most severely affected area of the liver and as a result there is fibrous thickening of the walls which surround the branches of the hepatic portal vein and hepatic artery (Fig. 1.8A,B). Examination of haematoxylin and eosin stained liver sections show that the lobules have lost their structural integrity, the central veins appear irregularly shaped and the portal triads appear partially obstructed. Additionally, there is a significant degree of fibrosis that could impair liver function (Fig. 1.8A,B). The intra-hepatic bile ducts have white, thickened fibrous walls and are often occluded, and the gall bladder is usually distended. The infiltration of polymorphonuclear leukocytes which surround the necrotic hepatocytes has also been observed (Cunningham et al., 1942).

While the most extensively affected organ in FE is the liver, post mortem analysis indicates abnormal pathology of several organ systems. In particular, there is discolouration and enlargement of the spleen (Cunningham et al., 1942), hypertrophy of the adrenal gland (Smith and Payne, 1991) and regression of the thymus (Cunningham et al., 1942; Smith and Payne, 1991). Both the spleen and thymus play essential roles in the immune system, the thymus in particular has a critical role in T-cell maturation (Abbas and Lichtman, 2003) and compromised function of these organs is likely to have a significant effect on the recovery and long term wellbeing of afflicted animals. It is not known if sporidesmin instigates these pathological changes or whether the changes are a result of prolonged liver dysfunction and cholestasis.

Another interesting, but rare pathology associated with FE is hepatic encephalopathy which accounts for some of the behavioural and neurological aspects of FE. It is
thought that this condition is a result of impaired liver function, in particular hyperammonaemia (Thompson et al., 1979).

Experiments performed in vitro have demonstrated that the earliest effects of sporidesmin on cultured liver cells are the reorganisation of cytoskeletal microfilaments and the loss of cell surface adhesiveness (Jordan and Pedersen, 1986). Other effects such as inhibition of phagocytosis and reduction in both bile flow and bile acid secretion are also observed (Bullock et al., 1974; Cordiner and Jordan, 1983; Cordiner et al., 1983).

1.5 CONTROL OF FACIAL ECZEMA

Once animals present with clinical symptoms the disease process is irreversible and the care of animals is primarily palliative. Animals must be removed from direct sunlight to keep the photosensitisation levels to a minimum. There is also some evidence that a high protein diet can aid recovery from sporidesmin toxicity (Wright, 1969; Anonymous, 1970), although this effect seems to be specific to casein, in particular phosphoserine (Hove and Wright, 1969).

As there is no effective treatment for FE there have been two different emphases on the development of control measures; the first has been on controlling the proliferation or toxigenicity of P. chartarum, and the second has been on modulating the animals’ response to the sporidesmin toxin.

Historically, FE control has been based on pasture monitoring and avoidance of toxic pastures, and to some degree this approach is still utilised today with weather surveillance forming an important part of farming practice in afflicted areas. Monitoring of pastures has been improved dramatically by the advent of fungal spore counts in the 1960s (Atherton et al., 1974). However, direct spore counts can be misleading due to the presence of both toxigenic and non-toxigenic strains (Atherton et al., 1974). More recently immunoassays have been developed that detect sporidesmin in pasture (Collin and Towers, 1995b) leading to more reliable and accurate determination of sporidesmin levels. However, the control still relies on
selective grazing and pasture management by moving stock or reducing grazing on infected fodder by using un-contaminated diet supplements.

Control of *P. chartarum* proliferation with fungicides has been popular since the late 1960s after it was shown that benzimidazole anthelmintics (which also have a fungicidal action) were capable of binding to *P. chartarum* spores and at concentrations of 140g / hectare was able to provide significant and prolonged protection against FE disease (Sinclair, 1967; Stutzenberger and Parle, 1972). However, due to the high cost, growing ineffectiveness and difficulty of administration fungicide use is less favoured (Towers, 1986). In addition there is a growing consumer resistance to overuse of such chemicals in agriculture.

A reasonably recent approach has involved bio-competitive exclusion of toxigeneic strains of *P. chartarum* with non-toxigenic strains. Although it is thought that sporidesmin production confers a selective advantage to toxigenic strains of *P. chartarum* it was found that a non-toxigenic strain originating from South Africa (Collin *et al*., 1996) performed well in competition assays on both agar and dried ryegrass and decreased the amount of sporidesmin produced (Collin and Towers, 1995a). However, in order for this approach to be successful the non-toxigenic strain of *P. chartarum* must be able to establish itself and successfully compete with endemic strains in the environment. A preliminary field study showed that pastures treated with non-toxigenic strains of *P. chartarum* (Collin and Towers, 1995a) resulted in a decrease of sporidesmin concentration by up to 80%. However, the untreated plots had spore levels that were barely approaching toxic levels and further field trials are required (Fitzgerald *et al*., 1998).

A few compounds have been shown to confer protection to sheep against sporidesmin in trials, for example hexachlorobenzene which is an inducer of several P450 isoforms (Mortimer *et al*., 1978). However, to date the only successful on-farm pharmacological treatment for FE is the prophylactic treatment with zinc (either in sulphate or oxide forms). Pre-treatment of both sheep (Smith *et al*., 1977) and cattle (Towers and Smith, 1978) with high levels of zinc sulphate (up to 50 mg/kg) has been shown to be very effective in protecting against FE disease. While there is some evidence that zinc treatment may be effective after animals are exposed to
sporidesmin (Rickard, 1975) zinc is most effective when administered prophylactically. Traditionally zinc was administered by drenching with a zinc oxide slurry or by the addition of zinc sulphate to drinking water. Administration of zinc has been greatly improved by the advent of the Time Capsule™ intra-ruminal device for the slow release of zinc oxide (Munday et al., 1997; Munday et al., 2001).

The mechanism by which zinc protects against sporidesmin and FE is currently unknown. However, it has been postulated that the zinc could protect against sporidesmin by decreasing hepatic copper concentrations thus preventing sporidesmin from auto-oxidation and redox cycling (refer section 1.3.4.1) (Munday, 1984b; Munday, 1985). This is supported by the finding that a number of iron compounds are known to also protect against sporidesmin (Munday and Manns, 1989) which suggests that zinc is not having an effect at the cellular level, as iron would be expected to exacerbate sporidesmin toxicity due to Fenton chemistry (Munday, 1989). Instead zinc may be acting to decrease the levels of hepatic copper (Prabowo et al., 1988) causing animals to become copper deficient. However, the antagonism seen between zinc and copper absorption is inconsistent dependent on age of the animal, and duration of zinc treatment (Rounce et al., 1998).

The third strategy that has been investigated is immunisation of the animals against sporidesmin. However this approach has been hampered by the fact that sporidesmin is a low molecular weight compound and is not antigenic (Jonas and Ronaldson, 1974; Jonas and Erasmuson, 1979). An approach which involved coupling the sporidesmin molecule to a protein (bovine thyroglobulin) or immunising with a structurally related compound (2-amino-5-chloro-3,4-dimethoxy benzyl alcohol) showed that immunised animals were actually more sensitive to liver damage induced by sporidesmin dosing (Faireclough et al., 1984).

By far the most promising and practical long term solution for control of FE is breeding for genetic resistance, as susceptibility to FE has a heritability of 0.42 (Campbell et al., 1981). This has been exploited to generate resistant and susceptible selection lines of Romney sheep (refer section 1.6.1) which have been used as a resource flock for research into the genetic basis of resistance to FE (refer section 1.7). Once genetic markers associated with resistance to FE have been developed,
marker-assisted selection would be used to breed resistance into commercial flocks. Currently selective breeding is based on performance testing, which is time consuming and relatively inefficient (Morris et al., 1994).

1.6 GENETIC RESISTANCE TO FACIAL ECZEMA

1.6.1 Resistant and Susceptible selection lines
Resistance to FE has a significant genetic component, and although heritability scores have only been estimated for Romney sheep the values obtained in separate studies are remarkably consistent. Specifically, using the liver injury score as an indicator trait FE was found to have a heritability of 0.42 (Campbell et al., 1981), and using serum GGT (gamma-glutamyltransferase) activity as an indicator trait the heritability estimate was 0.45 (Morris et al., 1989; Morris et al., 1995b). While serum GGT is not specific to sporidesmin toxicity, it is a quantitative measure of liver injury as serum GGT levels correlate well ($r^2 > 0.75$) with the severity of liver damage (Towers and Stratton, 1978).

The heritability of resistance to FE was exploited to establish resistant and susceptible selection lines in 1975 (a control line where no selection was exercised was added in 1982) (Morris et al., 1989). Initially randomly selected ewes were mated with rams which had been identified by progeny testing as resistant or susceptible to FE (Campbell et al., 1975). Progeny testing involved randomly chosen offspring from each ram; these were challenged at four months of age with sporidesmin and a liver damage score assigned upon post mortem examination six weeks later. Performance testing replaced progeny testing in 1982 (Towers et al., 1983; Morris et al., 1989), and rams were then selected on the basis of a performance test in which they were challenged with small doses of sporidesmin and the degree of liver injury was measured by examining serum GGT levels. An elevation of greater than 30% above pre-dose GGT concentration was taken as indicative of liver injury and of genetic susceptibility to sporidesmin. After 30 years of genetic selection, which equates to approximately 10 generations, the lines are divergent enough that the resistant line is estimated to be 11 times more resistant to sporidesmin than the susceptible line (Morris et al., 2004). The divergence of the resistant-line from the
control-line has been approximately 20% faster than the divergence of the susceptible-line from the control-line, and it is thought that this may be due to fertility issues in the susceptible line (Morris et al., 2004). The large differences in response to sporidesmin seen between the resistant and susceptible selection lines together with well documented pedigree information make these selection lines useful for studying the genetic basis of resistance or susceptibility to FE.

Sheep that have been bred for resistance to FE also show resistance to several other production diseases. In particular, sheep resistant to FE disease are also resistant to ryegrass staggers (Morris et al., 1995a). Ryegrass staggers is a neuromuscular condition caused by ingestion of the mycotoxin lolitrem-B. Interestingly, mice bred for resistance to endophyte (Neotyphodium coenophialum) infested fescue also show resistance to sporidesmin (Hohenboken et al., 2000). However, resistant sheep showed no tolerance to another mycotoxin called zearalenone (Smith et al., 1988), suggesting that zearalenone may have a different mechanism of toxicity from sporidesmin and lolitrem-B. It is of interest that sporidesmin, lolitrem-B, zearalenone and the ergot alkaloids thought to be responsible for fescue toxicosis show little structural similarity. Therefore the cross resistance between sporidesmin and the other toxins may indicate a common mechanism of metabolism, transport or tissue repair.

Resistance to FE has been shown to have a significant correlation with resistance to internal parasites (Morris et al., 1996; Morris et al., 2000) which is the most important production disease in New Zealand (Bunson, 1988). It is thought that this commonality may suggest that resistance to these two conditions involves some immunological components (Morris et al., 2004).

1.6.2 Breed specific differences (Ovis aries)
In addition to the Romney resistant and susceptible selection lines detailed above, there have been significant differences reported in the susceptibility of different sheep breeds to sporidesmin toxicity. A small study reported that Merinos are significantly more resistant to FE than both Romney and Border Leicester sheep (Smith et al., 1980). This is consistent with an in vitro study performed with sheep hepatic microsomes where microsomes derived from Merino sheep metabolised
sporidesmin more efficiently than Romney sheep microsomes (Fairclough et al., 1978). The correlation of this *in vitro* study with the *in vivo* toxicity studies suggests a role for the cytochrome P450 dependent monooxygenase enzymes in conferring resistance to FE.

East Friesian rams were found to have comparable resistance to the Romney resistant selection line rams (Morris et al., 2001). A preliminary experiment using very small numbers of animals suggests that the Finnish Landrace breed are more resistant than both Romneys and Texels (Morris et al., 2004). The difference in resistance between Finnish Landrace and Texels has been exploited to generate genetic crosses that are being utilised for a genome scan looking for areas of the genome that segregate with resistance to FE disease (S. H. Phua, pers. comm.) (refer 1.7.1).

### 1.6.3 Species specific differences

In addition to the variation found within sheep breeds with resistance to FE, there is also significant species variation (Table 1.2). Feral goats require a dose eight times higher than sheep to induce the same degree of liver pathology and histology (Smith and Embling, 1991). These differences could be due to altered activity of drug metabolising enzymes seen in goats as compared to sheep (Cook and Wilson, 1970; Szotakova et al., 2004).

Comparison of different toxicity studies (Table 1.2) shows that in general ruminants tend to be more susceptible to sporidesmin with a couple of exceptions, the first being goats (Smith and Embling, 1991) the second being horses, as to date horses have not been shown to be sensitive to sporidesmin (Morris et al., 2004). Fallow deer appear to be more sensitive than sheep (Mortimer and Smith, 1981) and of

<table>
<thead>
<tr>
<th>Resistant</th>
<th>Moderate</th>
<th>Susceptible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goat (<em>Capra hircus</em>)$^1$</td>
<td>Cattle (<em>Bos taurus</em>)$^2$</td>
<td>Fallow deer (<em>Dama dama L.</em>)$^2$</td>
</tr>
<tr>
<td>Horse (<em>Equus caballusi</em>)$^2$</td>
<td>Red deer (<em>Cervus elaphus</em>)$^2$</td>
<td>Sheep (<em>Ovis aries</em>)$^2$</td>
</tr>
<tr>
<td>Rat (<em>Rattus norvegicus</em>)$^2$</td>
<td>Rabbit (<em>Oryctolagus cuniculus</em>)$^2$</td>
<td>Alpaca (<em>Vicugna pacos</em>)$^3,4$</td>
</tr>
<tr>
<td>Mice (<em>Mus musculus</em>)$^2$</td>
<td>Guinea pig (<em>Cavia porcellus</em>)$^2$</td>
<td></td>
</tr>
</tbody>
</table>

$^1$(Smith and Embling, 1991), $^2$(Mortimer and Ronaldson, 1983), $^3$(Coulton et al., 1997), $^4$(Morris et al., 2004)
Chapter One: Introduction

Intermediate susceptibility are the experimental models including guinea pigs, rats and swine (Mortimer and Ronaldson, 1983). Mice are incredibly resistant to sporidesmin requiring a cumulative dose of 100x higher than that used to induce injury in sheep (Mortimer and Ronaldson, 1983). Additionally mice show a sex specific effect where females are more sensitive than males (Mortimer, 1970), which again supports a role for the sex specific P450 enzymes (Hodgson and Smart, 2001) in sporidesmin metabolism.

A recent report suggests that alpacas are also reasonably sensitive to sporidesmin (Coulton et al., 1997). Cattle and red deer seem to be of intermediate susceptibility (Morris et al., 2004). Marsupials were thought to be unaffected by sporidesmin exposure but a recent report has discussed the finding of FE in the eastern grey kangaroo (Macropus giganteus) (Hum, 2005).

1.7 RESEARCH INTO RESISTANCE TO FACIAL ECZEMA

Interestingly, a number of biochemical differences have been noted between resistant and susceptible animals. For example, a difference in the sedimentation rate and resistance to physical stress of erythrocytes was found between resistant and susceptible sheep (Upreti et al., 1991). Although the implications of the difference is not known, it is suggested that susceptible sheep have more fluid erythrocyte membranes (Upreti et al., 1991), which may impact on the function of the erythrocytes in these animals. This is particularly important as sporidesmin is known to interact with lipid bilayers (Upreti and Jain, 1993) and differences in membrane composition could alter the kinetics of this interaction.

However, to employ marker assisted selection DNA markers or genes that are associated with resistance to FE disease must be discovered. To do this two main approaches to identification of genes that are important in polygenic production traits have been used. The first being the whole genome scan, and the second the candidate gene approach.
1.7.1 Genome Scan
The goal of a genome scan is to find areas of the genome that display linkage
disequilibrium between a region on the chromosome and the trait in the population
being studied (Knott et al., 1996). These regions are often very broad and alternative
approaches must be used to narrow down the quantitative trait loci (QTL) region.

To date two whole genome scans have been undertaken with the intent of
discovering candidate genes involved in conferring resistance or susceptibility to FE
disease. The first was carried out using outcross half-sib pedigrees derived from the
resistant and susceptible Romney selection lines (refer section 1.6.1). Briefly, four
F1 Romney rams were produced from reciprocal crosses of resistant and susceptible
selection-line animals and mated with 130 to 170 Romney ewes to generate four
outcross pedigrees, with 124 to 167 progeny per family. The severity of liver
damage in sporidesmin-dosed progeny was assessed by measuring serum levels of
the liver-specific enzymes GGT and GDH three weeks post-dosing. The QTL
analysis using the outcross families was conducted via a primary and secondary
genome screens. Microsatellite markers that were reasonably evenly spaced
throughout the genome and that were heterozygous in at least one of the four sires
were used. In the primary screen inheritance of the sire alleles was determined for
22 of the most resistant and 22 of the most susceptible progeny (based on GGT data).
If a suggestive or significant QTL was detected all progeny (n = 124 to 167 per sire)
of the heterozygous sires were genotyped and analysed (Phua et al., 1999; Duncan et
al., 2007). Results from this analysis indicated an area of the genome that showed
association with the trait of interest.

The second genome scan is currently underway and is utilising Finnish Landrace and
Texel animals, which as previously described are extremely resistant and susceptible
respectively (S. H. Phua, pers. comm.)
1.7.2 Candidate Gene Approach

There are four classes of genes that have characteristics that would make them likely candidates for conferring resistance or susceptibility to sporidesmin.

1.7.2.1 Genes involved in phase I and phase II drug metabolism.

In phase I drug metabolism a polar group is introduced to the toxic molecule thus increasing the water solubility of the compound and facilitating phase II metabolism. A considerable amount of evidence suggests that sporidesmin is a target for phase I metabolism. Merino sheep have a significantly shorter sodium pentobarbital sleep time (pentobarbital is inactivated by the hepatic microsomal cytochrome P450 system) than Romney sheep. Correspondingly Merino sheep are significantly more resistant to sporidesmin than Romneys (Smith et al., 1980). These observations are consistent with the finding that Merino sheep have a significantly higher degree of sporidesmin metabolism by the microsomal fraction of a liver biopsy (Fairclough et al., 1978). Additionally, a highly significant correlation was found between sporidesmin metabolism rates by the microsomal fraction of a liver biopsy in 17 Romney rams and the average liver damage score of their progeny (Fairclough et al., 1978). Moreover, treating sheep with hexachlorobenzene which is known to induce several enzymes of phase I metabolism, primarily the P-450 isoforms (Linko et al., 1986), resulted in significant protection against sporidesmin in field trials (Mortimer et al., 1978). Furthermore, a mouse line that has been selected for resistance to both tall fescue toxicosis and sporidesmin (Hohenboken and Blodgett, 1997) had shorter sleep times following anaesthesia than the line that was genetically susceptible (Arthur et al., 2003). Overall these data are strongly suggestive of a role for the cytochrome P-450 isoforms not only in sporidesmin metabolism, but also as a key player in the mechanism of resistance to sporidesmin.

Phase II hepatic detoxification enzymes, which include the glutathione-S-transferases and uridine diphosphate glucuronosyl transferases, catalyse conjugation reactions with toxic molecules; these increase their aqueous solubility and facilitate their excretion in the urine. Again, in the mouse lines that were selected to be genetically resistant to tall fescue toxicosis, which also showed resistance to sporidesmin, the activities of both of these enzymes were higher in the resistant line than the
susceptible line (Hohenboken and Blodgett, 1997), indicating the possible involvement of phase II enzymes in sporidesmin metabolism.

Due to the large and pleiotropic nature of the enzymes involved in Phase I and II metabolism and a significant degree of functional redundancy, these gene families are difficult to investigate.

1.7.2.2 Antioxidant genes

It has been established that sporidesmin is capable of generating free radicals (Munday, 1982; Munday, 1984a; Munday, 1984b; Munday, 1987) (refer 1.3.4.1) and thus genes that encode anti-oxidant enzymes are obvious candidate genes for conferring resistance or susceptibility to FE disease. In particular the superoxide dismutases (SOD), catalase and glutathione peroxidases are of interest due to their role in detoxifying superoxide and hydrogen peroxide respectively.

Alleles of markers surrounding the catalase locus have been shown to segregate in the selection line animals (Phua et al., 1999). But the absence of a QTL in this chromosomal region indicated that the effect conferred by the catalase locus was likely to be recessive or minor (Phua et al., 1999). A similar study showed no involvement of CuZn-SOD, Mn-SOD, glutathione peroxidase and glutathione reductase in FE resistance (Phua et al., 1998).

A recent study examined the circulating plasma levels of SOD, glutathione peroxidase, catalase and glutathione reductase in the blood of Romney selection line sheep (Hohenboken et al., 2004) and found that SOD activity was lower and catalase activity was higher in the resistant than the susceptible line. Glutathione peroxidase activity was higher in the resistant compared to the susceptible line, whilst neither glutathione reductase or glutathione concentration differed significantly between lines. However, these differences were not large enough to be a reliable indicator of resistance to FE, therefore these genes may contribute to resistance but probably only play a minor role (Morris et al., 2004).
1.7.2.3 Genes for xenobiotic transport

The third class of genes are those that are involved in the efflux of xenobiotics, the so-called drug transporters, in particular the ABC transporters. The ABC transporter superfamily represents the largest family of transmembrane proteins and is a ubiquitous family of proteins found in all organisms from bacteria (Davidson and Chen, 2004) to humans (Dean and Allikmets, 1995). The transporters bind and hydrolyse ATP using the resulting energy to drive the transport of various molecules across all cell membranes (Higgins, 1992; Childs and Ling, 1994; Dean and Allikmets, 1995).

The ABC superfamily was first highlighted as a candidate gene family in 1994 when it was discovered that a *Saccharomyces cerevisiae* protein called PDR5 (STS1) was capable of conferring a multi-drug resistance phenotype to the yeast (Bissinger and Kuchler, 1994). Over-expression of PDR5 conferred resistance to sporidesmin, whereas a series of PDR5 deletion mutants caused the yeast to be hypersensitive to sporidesmin. At the time, the closest mammalian ortholog of PDR5 was ABCB1 (MDR1, P-glycoprotein) and this gene was subsequently examined for potential involvement in conferring resistance to FE (Longley, 1998). In that study a microsatellite marker (*OarMEM3*) and an RFLP (restriction fragment length polymorphisms) in the ABCB1 coding region were genotyped across relatively small numbers of resistant, susceptible and control selection-line sheep. No association between alleles of these markers was seen with resistance or susceptibility. In addition, there was no evidence for a QTL in this genome region in outcross pedigrees (refer 1.7.1) and it was concluded that that there was no genetic association between ABCB1 and resistance to FE (Longley, 1998).

However, using competitive RT-PCR, the same study reported that ABCB1 was more highly expressed in the livers of resistant but not susceptible selection line animals in response to sporidesmin exposure (Longley, 1998) (refer chapter 3). This observation suggested that resistant animals were able to induce expression of ABCB1 following sporidesmin exposure while susceptible animals could not. This was of interest as ABCB1 is a drug efflux transporter expressed in the liver, and increased expression of ABCB1 has been associated with resistance to chemotherapeutic compounds in humans (reviewed in Endicott and Ling, 1989).
1.7.2.4 Genes involved in metal absorption, transport and storage.

A fourth class of candidate genes are the genes that regulate transition metal bioavailability, genes that are involved in metal absorption, transport and storage. Of particular interest are the genes involved in copper and iron metabolism as these metals are predicted to participate in redox cycling of sporidesmin (refer 1.3.4.1). Ceruloplasmin, which is the copper storage protein, and transferrin, which is an iron transport protein, have been examined for association with resistance or susceptibility to FE. RFLPs in ceruloplasmin and transferrin did not segregate in the FE selection lines (Phua et al., 1998). Another study analysed the association of transferrin allozymes with resistance or susceptibility to FE and found no association (Morris et al., 1988).

1.8 OBJECTIVES OF THIS RESEARCH

The overall goal of this research project was to identify and investigate new genes associated with conferring resistance or susceptibility to FE disease in sheep. To do this several approaches were taken.

The first was to confirm the reported differential expression of ABCB1 (MDR1, P-glycoprotein) between resistant and susceptible sheep (Longley, 1998) using qRT-PCR. The aim of doing this was to subsequently determine what genetic factors were responsible for the differential expression (chapter three).

The second approach involved determining the closest mammalian ortholog to the yeast PDR5 protein using phylogenetic analyses of several whole genome sequences. This ortholog was then assessed as a possible candidate gene, by isolation and characterisation of the corresponding full length sheep cDNA sequence, expression analysis and association of SNP markers in this gene in the selection line animals (chapter four).

The third approach used microarray analyses to examine the differences in the global gene expression profile between resistant and susceptible sheep both pre- and post-exposure to sporidesmin. The aim of this approach was to identify novel genes or
biochemical pathways involved in the mechanism of resistance to FE disease (chapter five).

Finally, the last aim of this project was to characterise HepG2 cells as an in *vitro* model for examining the mechanism of zinc protection against sporidesmin toxicity. Any genes modulated by zinc could be identified using microarray expression profiling, and would be candidate genes for resistance to FE (chapter six).
Chapter Two

Materials and Methods.

2.1 APPROVALS AND PERMITS

Modification of *Escherichia coli* by the introduction of non-conjugative plasmids was carried out in the University of Otago Transitional Containment Facility (#2767) under Environmental Risk Management Authority (ERMA) approval GMD001138 (update GMD002509). In accordance with these approvals all work was carried out under PC1 (physical containment 1) conditions.

All tissue and blood samples collected for this study were collected in accordance with the 1987 Animals Protection (Codes of Ethical Conduct) Regulations. Approval for each experiment was granted by the AgResearch Invermay Animal Ethics Committee (AEC). The AEC approvals governing this work are; AECP484, AECP534 and AEC_inv596_03. Animal tag numbers and relevant details of experimentation are given in section 2.2.

Radioactivity ([α-³²P]dCTP) was used in accordance with safety requirements and under NRL site licences issued to Dr. Allan Crawford, Dr. Theresa Wilson (both AgResearch Molecular Biology Unit) and Prof. Warren Tate (Department of Biochemistry, University of Otago).
2.2 ANIMALS AND ANIMAL TISSUES

All sheep used in this study were Romney FE selection-line animals (refer to 1.6.1). Animal tag numbers and relevant experimental details such as age, gender and dose rates are given in Table 2.1 (resistant animals) and Table 2.2 (susceptible animals).

Animals R1 - R7 and S1 – S7 were sacrificed for a previous study (Longley, 1998), the animal codes used in the original study are provided in the last column of Tables 2.1 and 2.2 for cross-referencing purposes. These animals had needle biopsies taken from their livers prior to exposure to sporidesmin. Post-exposure samples were taken by necropsy, two weeks after sporidesmin exposure. Tissues from these experiments were limited, and none were available for this study. However, RNA samples from these animals which had been used in the previous study (Longley, 1998) and stored at -80°C for approximately 3 years were used in this study.

For tissues collected specifically for this study (animals R8 – R26 and S8 – S25) animals were housed at the AgResearch Invermay Agricultural Centre in Dunedin, New Zealand. This area is known to be free of FE and therefore the animals used for experimentation are naïve, having never been exposed to environmental sporidesmin. Sheep were pasture reared, and while they had no natural exposure to sporidesmin, exposure to other pathogens was not controlled.

A crude extract of sporidesmin (concentration was quantified using toxicity assays) was dissolved in ethanol and was diluted in water to the appropriate concentrations immediately before being administered to the animals orally by intra-ruminal intubation. Animals were sacrificed 24 h later using the captive bolt method. An experienced animal scientist removed the liver tissues and ensured that all samples were taken from the central region of the minor lobe (right lobe) of the liver. Tissue samples were dissected into slices approximately 1 cm x 0.2 cm x 0.2 cm and snap frozen in liquid nitrogen in Nunc Cryotubes (Nalgene Nunc). The samples were stored at -80°C. The time between captive bolt delivery and samples being frozen in liquid nitrogen was approximately 10 min.
Table 2.1: Details of resistant-line animals used in this study.

<table>
<thead>
<tr>
<th>New code</th>
<th>Animal Tag</th>
<th>N/D*</th>
<th>Birth Year</th>
<th>Age (months)</th>
<th>Gender</th>
<th>Dose rate (mg/kg)</th>
<th>Exposure Time</th>
<th>Year of Sacrifice</th>
<th>Other codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>O7200</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.35</td>
<td>2 wks</td>
<td>1998</td>
<td>O2 a/b</td>
</tr>
<tr>
<td>R2</td>
<td>O7231</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.35</td>
<td>2 wks</td>
<td>1998</td>
<td>O3 a/b</td>
</tr>
<tr>
<td>R3</td>
<td>O7385</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.35</td>
<td>2 wks</td>
<td>1998</td>
<td>O7 a/b</td>
</tr>
<tr>
<td>R4</td>
<td>O7281</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.35</td>
<td>2 wks</td>
<td>1998</td>
<td>O4 a/b</td>
</tr>
<tr>
<td>R5</td>
<td>O7141</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.35</td>
<td>2 wks</td>
<td>1998</td>
<td>O1 a/b</td>
</tr>
<tr>
<td>R6</td>
<td>O7304</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.35</td>
<td>2 wks</td>
<td>1998</td>
<td>O5 a/b</td>
</tr>
<tr>
<td>R7</td>
<td>O7330</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.35</td>
<td>2 wks</td>
<td>1998</td>
<td>O6 a/b</td>
</tr>
<tr>
<td>R8</td>
<td>O1107</td>
<td>N</td>
<td>2001</td>
<td>9</td>
<td>M</td>
<td>-</td>
<td>-</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>R9</td>
<td>O1104</td>
<td>N</td>
<td>2001</td>
<td>9</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>R10</td>
<td>O465</td>
<td>N</td>
<td>2000</td>
<td>2-3</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Dec 2000</td>
<td></td>
</tr>
<tr>
<td>R11</td>
<td>O0468</td>
<td>D</td>
<td>2000</td>
<td>21</td>
<td>M</td>
<td>0.1</td>
<td>24 h</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>R12</td>
<td>O0462</td>
<td>D</td>
<td>2000</td>
<td>21</td>
<td>F</td>
<td>0.1</td>
<td>24 h</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>R13</td>
<td>O1108</td>
<td>D</td>
<td>2001</td>
<td>60</td>
<td>M</td>
<td>0.1</td>
<td>24 h</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>R14</td>
<td>O1117</td>
<td>D</td>
<td>2001</td>
<td>9</td>
<td>F</td>
<td>0.1</td>
<td>24 h</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>R15</td>
<td>O5206</td>
<td>D</td>
<td>1995</td>
<td>60</td>
<td>F</td>
<td>0.4</td>
<td>24 h</td>
<td>Dec 2000</td>
<td></td>
</tr>
<tr>
<td>R16</td>
<td>O7274</td>
<td>D</td>
<td>1997</td>
<td>36</td>
<td>F</td>
<td>0.4</td>
<td>24 h</td>
<td>Dec 2000</td>
<td></td>
</tr>
<tr>
<td>R17</td>
<td>O2163</td>
<td>N</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R18</td>
<td>O2195</td>
<td>N</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R19</td>
<td>O2201</td>
<td>N</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R20</td>
<td>O2216</td>
<td>N</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R21</td>
<td>O2286</td>
<td>N</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R22</td>
<td>O2033</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R23</td>
<td>O2047</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R24</td>
<td>O2081</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R25</td>
<td>O2246</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>R26</td>
<td>O2277</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
</tbody>
</table>

* N = Naïve, D= sporidesmin dosed, ND = Samples were pre-sporidesmin exposure by needle biopsy, the animals were then exposed to sporidesmin and the livers recovered by necropsy for the post-exposure sample (Longley, 1999).

Table 2.2: Details of susceptible-line animals used in this study.

<table>
<thead>
<tr>
<th>New code</th>
<th>Animal Tag</th>
<th>N/D*</th>
<th>Birth Year</th>
<th>Age (months)</th>
<th>Gender</th>
<th>Dose rate (mg/kg)</th>
<th>Exposure Time</th>
<th>Year of Sacrifice</th>
<th>Other codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>P7051</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.15</td>
<td>2 wks</td>
<td>1998</td>
<td>P2 a/b</td>
</tr>
<tr>
<td>S2</td>
<td>P7153</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.15</td>
<td>2 wks</td>
<td>1998</td>
<td>P4 a/b</td>
</tr>
<tr>
<td>S3</td>
<td>P7267</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.15</td>
<td>2 wks</td>
<td>1998</td>
<td>P5 a/b</td>
</tr>
<tr>
<td>S4</td>
<td>P7354</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.15</td>
<td>2 wks</td>
<td>1998</td>
<td>P6 a/b</td>
</tr>
<tr>
<td>S5</td>
<td>P5003</td>
<td>ND</td>
<td>1995</td>
<td>&lt;12</td>
<td>M</td>
<td>-</td>
<td>-</td>
<td>1996</td>
<td>F7 a/b</td>
</tr>
<tr>
<td>S6</td>
<td>P7029</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.15</td>
<td>2 wks</td>
<td>1998</td>
<td>P1 a/b</td>
</tr>
<tr>
<td>S7</td>
<td>P7056</td>
<td>ND</td>
<td>1997</td>
<td>&lt;12</td>
<td>M</td>
<td>0.15</td>
<td>2 wks</td>
<td>1998</td>
<td>P3 a/b</td>
</tr>
<tr>
<td>S8</td>
<td>P463</td>
<td>N</td>
<td>2000</td>
<td>2-3</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Dec 2000</td>
<td></td>
</tr>
<tr>
<td>S9</td>
<td>P1105</td>
<td>N</td>
<td>2001</td>
<td>9</td>
<td>M</td>
<td>-</td>
<td>-</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>S10</td>
<td>P1110</td>
<td>N</td>
<td>2001</td>
<td>9</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>S11</td>
<td>P7015</td>
<td>D</td>
<td>1997</td>
<td>36</td>
<td>F</td>
<td>0.15</td>
<td>24 h</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>S12</td>
<td>P0469</td>
<td>D</td>
<td>2000</td>
<td>21</td>
<td>M</td>
<td>0.1</td>
<td>24 h</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>S13</td>
<td>P0467</td>
<td>D</td>
<td>2000</td>
<td>21</td>
<td>F</td>
<td>0.1</td>
<td>24 h</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>S14</td>
<td>P1106</td>
<td>D</td>
<td>2001</td>
<td>9</td>
<td>M</td>
<td>0.1</td>
<td>24 h</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>S15</td>
<td>P1111</td>
<td>D</td>
<td>2001</td>
<td>9</td>
<td>F</td>
<td>0.1</td>
<td>24 h</td>
<td>June 2002</td>
<td></td>
</tr>
<tr>
<td>S16</td>
<td>P2018</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S17</td>
<td>P2097</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S18</td>
<td>P2134</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S19</td>
<td>P2186</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S20</td>
<td>P2384</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>-</td>
<td>-</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S21</td>
<td>P2010</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S22</td>
<td>P2099</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S23</td>
<td>P2329</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S24</td>
<td>P2362</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
<tr>
<td>S25</td>
<td>P2369</td>
<td>D</td>
<td>2002</td>
<td>6</td>
<td>F</td>
<td>0.03</td>
<td>24 h</td>
<td>Mar 2003</td>
<td></td>
</tr>
</tbody>
</table>

* N = Naïve, D= sporidesmin dosed, ND = Samples were pre-sporidesmin exposure by needle biopsy, the animals were then exposed to sporidesmin and the livers recovered by necropsy for the post-exposure sample (Longley, 1999).
2.3 MATERIALS AND EQUIPMENT

All buffers, salts, acids and bases were AnalaR grade unless otherwise stated. All solutions were made up according to the methods outlined in Molecular Cloning (Sambrook et al., 1989) or Current Protocols in Molecular Biology (Ausubel et al., 2004).

All solutions were either autoclaved (15 psi, 121°C, 20 min) or filter sterilised. Solutions containing heat labile compounds such as glucose or sodium dodecyl sulphate (SDS) were filter sterilised.

The supplier for all chemicals and reagents used in this study is indicated in parentheses after the product. DNA restriction and modifying enzymes were obtained from New England Biolabs or Roche Applied Science.

In addition, several molecular biology kits were used during this study. The basic principle of each kit is briefly described in the appropriate sections, without giving an exhaustive protocol.

2.4 DNA METHODS

2.4.1 Genomic DNA isolation

DNA samples were extracted from whole blood using the high-salt method (Montgomery and Sise, 1990), resuspended in TE (10 mM Tris-HCl pH 7.2, 0.1 mM EDTA) and stored at -20°C. Alternatively, DNA was isolated from liver tissue that had been snap frozen in liquid nitrogen and stored at -80°C, using standard protocols (Strauss, 2004). Briefly, 200 mg of frozen tissue was cut into small pieces and incubated overnight at 50°C in 2.4 ml of digestion buffer (100 mM NaCl, 10 mM Tris-HCl pH 8.0, 25 mM EDTA pH 8.0, 0.5% SDS, 0.1 mg/ml proteinase K). Undigested tissue was removed by centrifugation for 1 min at 1700 × g. DNA was then phenol/ chloroform extracted (refer 2.4.6) and precipitated by adding ½ × volume of 7.5 M ammonium acetate and 2 × volume ice cold absolute ethanol,
followed by centrifugation at $1700 \times g$ for 20 min. The pellet was then rinsed with 70% ethanol, resuspended in 1 ml of TE and the DNA concentration was estimated spectrophotometrically (refer 2.4.10). Working stocks of DNA were adjusted to 100 ng/μl.

2.4.2 Oligonucleotide design and synthesis
Oligonucleotide primers were designed using Vector NTI Suite 7.1 (InforMax Inc.). Primers were designed to be between 18 – 25 nucleotides in length with a melting temperature (Tm) of 58 – 65°C, and GC content between 35 - 60%. The propensity of the primer pairs to form secondary structures such as dimers and hairpin loops that would interfere with the PCR reactions was also taken into consideration for primer design.

Primers were synthesised by the either the University of Otago primer synthesis facility (synthesised on an Applied Biosystems 380A DNA synthesiser), or were obtained commercially from Proligo (previously Geneset Oligonucleotides), or Applied Biosystems. The sequences of all oligonucleotides used in this study are given in Appendix A.

2.4.3 Polymerase chain reaction (PCR)
The polymerase chain reaction (PCR) is based on the ability of a thermostable DNA polymerase to amplify DNA, primed from oligonucleotides annealed to denatured single-stranded templates (Saiki et al., 1988). The procedure involves thermal denaturation of the DNA template, followed by annealing of two specific oligonucleotides to complementary sequences on opposite strands of the DNA, flanking the sequence to be amplified. The annealed primers are orientated with their 3' ends facing each other, such that DNA polymerase will extend across the region of the original DNA template between the primers. Each synthesised strand is complementary to one of the primers, and can serve as template in further cycles of annealing and extension. The denaturation, annealing and extension steps are repeated for 25–35 cycles resulting in exponential amplification of the DNA region of interest.
The template used for amplification reactions was either 1 μl of cDNA (reverse transcribed from 1 μg of total RNA in a total volume of 20 μl; refer 2.5.4), 100 ng of genomic DNA or 2-10 ng of plasmid DNA. PCR reactions were performed in a total volume of 20 μl with a final concentration of 1 x PCR buffer (Invitrogen), 0.2 mM dNTPs (Amersham Biosciences), 1.5 mM MgCl₂ (unless otherwise stated) (Invitrogen), 1 μM each forward and reverse primers and 1 U of Taq Polymerase (Invitrogen). Primer sequences and specific amplification conditions are given in Appendix A. PCR reactions were set-up on ice and immediately placed into an Eppendorf Mastercycler thermal cycler (Eppendorf). Routinely the reaction cycles were as follows:

1. 94°C, 2 min [initial denaturation]
2. 94°C, 30 sec [denaturation]
3. oligo annealing temperature, 30 sec
4. 72°C, 1 min [extension]
5. steps 2-4 above repeated, 30 – 35 cycles
6. 72°C, 3 min [final extension and renaturation]

The annealing temperature was typically set at 3°C below the theoretical melting temperature (Tm) of the oligonucleotides being used. The Tm was calculated using the formula \( Tm = 2 \times (A+T) + 4 \times (G+C) \), where A, T, G and C refer to the base composition of the oligonucleotide, where the oligonucleotide was shorter than 20 bp. For longer oligonucleotides the following formula was used: \( Tm = 64.9°C + 41°C \times (\text{number of G’s and C’s in the primer} - 16.4)/N \), where N is the length of the oligo. Pairs of PCR primers were routinely optimised by performing an annealing temperature gradient (usually 5°C either side of the 3°C below the hypothetical Tm) and were tested at four magnesium concentrations (1 mM, 1.5 mM, 2 mM and 3 mM).

The extension time used depended on expected length of PCR product (~1 min per kb).
2.4.3.1 5' and 3' RACE

5' and 3' RACE (rapid amplification of cDNA ends) were used to obtain 5' and 3' UTR (untranslated region) sequences. RACE was performed according to the manufacturer’s instructions (Ambion FirstChoice™-RLM-RACE). Briefly, 5' RACE involves calf intestinal phosphatase treatment to remove the 5' phosphate group from degraded mRNA, rRNA, tRNA and any contaminating DNA. The intact mRNA is then de-capped using tobacco acid pyrophosphatase and the de-capped mRNA is ligated to a 5' adapter sequence. 3' RACE differs slightly as the adapter sequence utilises the poly (A) tail of the mRNA. In both cases the mRNA is then reverse transcribed and cDNA is amplified using an oligonucleotide primer corresponding to the adapter sequence and a gene specific reverse primer. A second nested PCR reaction can be performed to increase both the specificity and yield of the product obtained.

Both the first and second (nested) PCR reactions were carried out in a 50 μl volume using 1.5 mM MgCl₂, an annealing temperature of 57°C and 35 PCR cycles. Additionally, in the second PCR, gene specific primers were used to confirm the specificity of the RACE product. The RACE PCR products were gel purified with the High Pure PCR product purification kit (Roche Applied Science), and cloned as detailed in 2.4.7. Positive clones were sequenced in both directions using universal primers T7 and SP6 with the Big-Dye Terminator Cycle Sequencing kit (Applied Biosystems) as detailed in 2.4.11.

2.4.4 Agarose gel electrophoresis

DNA was electrophoresed in 0.6 - 3.0% agarose gels in 1 × TAE buffer (0.4 M Tris-HCl, 50 mM sodium acetate, 10 mM EDTA, pH 7.8). Ethidium bromide (5 mg/ml stock solution stored at 4°C) was added to a final concentration of 0.5 μg/ml. Samples were mixed with loading dye (30% w/v glycerol, 0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol FF) prior to electrophoresis. For reference purposes a commercial size standard was also electrophoresed. Samples were electrophoresed at 80 - 100 V until appropriate separation had been achieved. DNA was visualised under UV light and documented either using the Alphalmager 2000 digital camera
Chapter Two: Materials and Methods

(Alpha Innotech Corporation) or the Gel Doc XR System with Quantity One Software (Bio-Rad).

2.4.5 Purification of PCR products
Linear DNA fragments (PCR products or cleaved DNA) were purified directly from PCR reactions, using the High Pure PCR product purification kit (Roche Applied Science) or gel-purified using the QIAquick Gel Extraction Kit (Qiagen).

The High Pure PCR product purification kit (Roche Applied Science) utilises the specific binding of DNA fragments greater than 100 bp to a glass fibre fleece in the presence of the chaotropic salt guanidine thiocyanate. The binding buffer facilitates efficient binding of PCR products to the glass fibre surface. Impurities such as primers, dNTPs and Taq polymerase do not bind to the glass fibre and are found in the flow through. Residual reaction components are washed off with an included wash buffer before the purified PCR product is eluted in the supplied TE elution buffer.

The QIAquick Gel Extraction Kit (Qiagen) uses a similar technology excepting that the DNA is bound to a silica-gel-membrane in the presence of a high salt buffer. Again, impurities are washed through the column and the purified DNA is eluted in TE or sterile water. Either kit yielded DNA pure enough to use directly in ligation or sequencing reactions.

In all cases the purification was carried out according to the manufacturers’ instructions and the purified products were quantitated by spectrophotometry (refer 2.4.10) or by gel electrophoresis and comparison with the Low DNA Mass™ Ladder (Invitrogen).

2.4.6 Phenol extraction and ethanol precipitation of DNA
Contaminating proteinaceous debris was removed from DNA solutions by phenol/chloroform extractions, followed by ethanol precipitation of the DNA. For nucleic acid extractions, an equal volume of buffered phenol and chloroform (1:1 ratio v/v) was mixed with the DNA sample, vortexed, and centrifuged for 4 min at 12000 × g. The aqueous phase was retained and the extraction repeated with an
equal volume of chloroform to remove traces of phenol. The DNA was then ethanol precipitated as follows: 1/10 volume of 3.5 M sodium acetate (pH 5.2) and 2 volumes of 100% ice-cold ethanol were added to the aqueous solution. The contents were mixed, and incubated at -20°C for a minimum of 15 min, before centrifugation at 12000 × g for 10 min to pellet the DNA. The DNA was washed in 500 μl of 70% ethanol, air-dried and resuspended in a suitable volume of MilliQ H2O.

2.4.7 Cloning of DNA fragments
2.4.7.1 Ligation and desalting
All cloning of PCR products utilised the pGEM T-easy vector (Promega), which is a TA overhang vector designed to easily clone PCR products. PCR products were “A-tailed” to maximise ligation efficiency as follows; 7 μl of purified PCR product (refer 2.4.5) was incubated at 70°C for 30 min with 1 × PCR buffer (Invitrogen), 1.5 mM MgCl₂ and 0.2 mM dATP (Roche Applied Science). Of the A-tailed PCR products, 3 μl was used in a ligation reaction with pGEM T-easy as per manufacturer’s instructions. Briefly 50 ng of pGEM T-easy vector was incubated with the appropriate amount of A-tailed PCR product, 3 Weiss units of T4 DNA ligase and 1 × rapid ligation buffer (30 mM Tris-HCl pH 7.8, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, 5% PEG). Ligations were routinely incubated overnight at 4°C to maximise the ligation efficiency.

Ligation mixes were desalted by dialysis prior to electroporation (Atrazhev and Elliott, 1996). Briefly, a 1.5 ml microcentrifuge tube was filled with 1.5 ml of molten 1% agarose (Invitrogen) with 100 mM glucose, and a 200 μl pipette tip was used to make a well. The ligation mixture was then added to the well and incubated on ice for 60-90 min, a proportion of this sample (usually 3-5 μl) was then used to transform competent E. coli cells either by electroporation (refer 2.4.7.4) or by heat shock transformation (refer 2.4.7.5).

2.4.7.2 Bacterial strains
The two E. coli strains routinely used in this study were DH5α (supE44 ΔlacU169Δ80 lacZΔ M15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1) and XL1-Blue (supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1lac F′[proAB lacF ΔlacZΔM15]
Chapter Two: Materials and Methods

*Tn10(tet^+)*]). Bacterial strains were maintained as permanent stocks in 15% glycerol in Luria-Bertani (LB) broth and stored at -80°C.

2.4.7.3 Broth, agar and antibiotics

LB agar plates were used throughout this study for growing bacterial strains from frozen stocks, general culturing of strains, and selection of transformants. Bacterial strains were routinely grown in LB broth (1.0% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.17 M NaCl, pH 7.0) and plated out on LB agar plates (1.0% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.17 M NaCl, pH 7.0, 1.5% (w/v) agar).

SOC medium (2.0% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 20 mM glucose) was used to aid recovery of bacterial transformants.

All stock antibiotic solutions were stored in aliquots at -20°C and those prepared in MilliQ H₂O were sterilised by filtration through 0.22 μm Millex-GV syringe driven unit (Millipore). Ampicillin was prepared as a 50 mg/ml stock solution in MilliQ H₂O and used at a working concentration of 50 μg/ml.

2.4.7.4 Electrocompetant cell preparation and electroporation

Electrocompetant *E. coli* cells were prepared using a method modified from Sambrook *et al.* (1989). The stock was streaked out on LB agar plates and grown overnight at 37°C, a single colony was then picked and used to inoculate 5 ml of LB which was grown overnight at 37°C with shaking (200 rpm). Typically 250 μl of an overnight culture of the strain to be made electrocompetant for transformation was used to inoculate 250 ml of LB broth and grown to an OD₆₀₀ₙₐₚ of 0.6. The cells were incubated on ice for 30 min and then pelleted by centrifugation at 3500 × g for 20 min and the bacterial pellet resuspended in 100 ml of sterile, ice-cold MilliQ H₂O. The cells were again pelleted by centrifugation at 3500 × g for 20 min and resuspended in 50 ml cold MilliQ H₂O. Following another centrifugation step, cells were resuspended in 20 ml of 10% v/v glycerol prior to another centrifugation step at 3500 × g for 20 min. Finally cells were resuspended in 2 ml of 10% v/v glycerol,
snap frozen on dry ice in 100 μl aliquots and stored at -80°C until required (Sambrook et al., 1989).

The DNA to be electroporated (0.1 – 1 μg) in a volume not exceeding 5 μl (refer 2.4.7.1) was added to a 100 μl aliquot of electrocompetant cells and incubated on ice for 1 min. The mixture was then transferred to a pre-chilled electroporation cuvette (Bio-Rad, 0.2 cm gap width). The cuvette was placed in the Gene Pulser II chamber (Bio-Rad) and an electroshock delivered (2.5 kV, 200 Ω, 25 μF). Pre-warmed SOC broth was added to the cuvette (900 μl) and the contents transferred to a sterile microcentrifuge tube and incubated with shaking at 37°C for 1 hour. This incubation allows phenotypic expression of the antibiotic resistance marker. Subsequently 100 μl of the transformation mix was plated onto appropriate selection plates. Cells to which no DNA had been added were treated in the same way to serve as a negative control.

2.4.7.5 Heat shock competent cell preparation and heat shock transformation

For routine transformation of E. coli strains with plasmid DNA or ligation mixtures the calcium chloride method was employed. An overnight culture of the strain to be made competent for transformation was used to inoculate 250 ml of LB broth and was grown to an OD₆₀₀ of approximately 0.4. The cells were incubated on ice for 30 min and then pelleted by centrifugation at 1800 × g for 10 min, and the bacterial pellet resuspended in 100 ml of ice-cold CaCl₂ solution (60 mM CaCl₂, 15% glycerol, 10 mM PIPES pH 8.0). The cells were spun again for 10 min at 1800 × g, the supernatant removed and resuspended in 100 ml of ice-cold CaCl₂ solution. After incubation on ice for 30 min cells were again harvested as described above, and resuspended in 10 ml ice-cold CaCl₂ solution. At this stage cells were distributed into 100 μl aliquots and used directly or stored at –80°C.

DNA (0.1–1 μg) in a volume not exceeding 10 μl, was added to 100 μl of competent cells and left on ice for 10 min. The tubes were then placed in a 42°C water bath for 45 sec; this heat-shock treatment allows uptake of DNA into the competent bacteria by an unknown mechanism (Mandel and Higa, 1970). Following heat shock 0.9 ml of SOC broth was added to the culture, which was then incubated with shaking at
37°C for 1 hour to allow expression of the plasmid-borne antibiotic resistant marker. Aliquots of 10 µl and 100 µl were plated onto appropriate selection plates. Cells to which no DNA had been added were treated in the same way and served as a negative control. Plates were incubated overnight at 37°C, and colonies were screened by α-complementation (refer 2.4.7.7) or PCR based screening (refer 2.4.7.8).

2.4.7.6 Determining the transformation efficiency of competent cells
The transformation efficiency of competent cells was assessed by transforming 0.1 ng of pBluescript II SK(-) supercoiled plasmid (Stratagene) into 100 µl of competent cells using the relevant transformation method (heat shock or electroporation). SOC broth (0.9 ml) was added to the culture, which was then incubated at 37°C for 1 hour to allow expression of the plasmid-borne antibiotic resistant marker. Cells were then plated at three dilutions (1 µl, 10 µl and 20 µl of cells diluted to 100 µl with SOC) on LB-ampicillin agar plates. The transformation efficiency was then calculated as the number of colonies formed per µg of DNA transformed. Competent cells were used if their transformation efficiencies exceeded $10^7$ CFU (colony forming units) for heat shock competent cells and $10^8$ CFU for electrocompetant cells.

2.4.7.7 Screening colonies by α-complementation
Both pBluescript II SK(-) (Stratagene) and pGEM T-easy vector (Promega) vectors contain a segment of DNA derived from the lac operon of E. coli that codes for the amino-terminal fragment of β-galactosidase. The synthesis of this fragment is induced by IPTG (isopropyl-β-D-thiogalactopyranoside) which is a gratuitous inducer of the lac promoter through inactivation of the LacI repressor. This fragment is capable of intra-allelic (α) complementation with a defective form of β-galactosidase encoded by the host bacteria. Bacteria containing the plasmid form blue colonies when grown on X-Gal (a chromogenic substrate for β-galactosidase) containing agar plates. However, insertion of foreign DNA into the multiple cloning site (MCS) of the vector disrupts the amino terminal fragment of β-galactosidase and therefore bacteria carrying recombinant plasmids give rise to white colonies.
X-Gal (5-bromo-4-chloro-3-indoyl-β-D-galactoside) was prepared as a 20 mg/ml stock solution in N, N-dimethyl formamide and stored in the dark at -20°C. X-Gal was used in agar plates at a final concentration of 20 μg/ml.

IPTG was prepared as a 100 mM stock solution in MilliQ H₂O, filter sterilised, and stored at -20°C. IPTG was used at a final concentration of 0.5 mM in agar plates to induce expression of genes under the control of the lac promoter.

**2.4.7.8 Screening colonies by PCR**

In cases where screening colonies by α-complementation was not appropriate or when the size of cloned DNA fragments required verification, colonies were screened by a PCR based method. Colonies were selected and re-suspended in 8 μl of LB. A portion of the resuspended colony (5 μl) was added to 8 μl of 0.5% Tween 20 and boiled for 30 seconds which served to break down the bacterial cell walls and denature DNA. Of this mixture 1 μl was then used in a standard PCR reaction (refer 2.4.3) with primers specific for the cloned DNA of interest or universal primers designed to the T7 or SP6 sequences (Appendix A). Once positive colonies were identified, the remaining 5 μl of resuspended colony was inoculated into 5 ml of LB and grown overnight as detailed below.

**2.4.8 Plasmid DNA isolation**

Cultures for plasmid DNA isolation were typically grown by inoculating single colonies into 5 ml of LB broth (with the appropriate antibiotic) in a 12 ml sterile culture-tube and incubating at 37°C overnight with shaking at 200 rpm.

The Wizard® Plus minipreps DNA purification system (Promega) was routinely used to extract plasmid DNA from 5 ml cultures according to the guidelines provided. The procedure is based on a modified alkaline lysis method where bacteria are lysed and proteins denatured (by SDS) in the presence of protease inhibitors. RNA was removed using RNase A, and DNA denatured by sodium hydroxide. The lysis mixture was neutralised with potassium acetate, causing denatured protein and chromosomal DNA to precipitate. The debris was pelleted by centrifugation. The
supernatant containing plasmid DNA was desalted using a silica membrane, and eluted in 50 μl of sterile water.

2.4.9 Restriction endonuclease digestion
Restriction enzyme digests were carried out according to the manufacturer’s recommendations and using the buffers provided. Typically 0.5 – 2.0 μg of plasmid DNA or purified PCR product were digested with 10 U of restriction enzyme in a 30 μl volume. Reactions were incubated at 37°C for 1 – 2 h unless otherwise recommended. DNA products were analysed using agarose gel electrophoresis.

2.4.10 Determination of nucleic acid concentration
The concentration of DNA and RNA samples was determined spectrophotometrically by measuring the absorbance at 260 nm. Samples were diluted (typically 1:200) in MilliQ H₂O before reading the OD₂₆₀nm in a quartz cuvette on a Jasco V-530 UV/visible spectrophotometer (Jasco Inc.). Alternatively 1.5 μl of solution was measured on the Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies).

The concentration of nucleic acid was then determined according to the following formulae:
An OD₂₆₀nm value of 1 corresponds to -
50 μg/ml of double-stranded DNA
40 μg/ml of single-stranded DNA or RNA.

The purity of DNA or RNA was assessed by measuring the A₂₆₀nm/A₂₈₀nm ratio. For clean DNA, uncontaminated by proteins or residual phenol the ratio of A₂₆₀nm to A₂₈₀nm is 1.8 while for RNA uncontaminated by proteins or residual phenol, the ratio of A₂₆₀nm to A₂₈₀nm is 2.0 (Sambrook et al., 1989). The A₂₆₀nm/A₂₃₀nm ratio was used as a further measure of purity, a ratios between 2 and 2.2 were standard. Lower ratios are indicative of contaminants including EDTA or phenol carried over from the purification process.
2.4.11 DNA Sequencing

DNA Sequencing was carried out initially by the Centre for Gene Research (Department of Microbiology, University of Otago, Dunedin), a commercial sequencing service, using an ABI 377 DNA Sequencer (Applied Biosystems).

Subsequently automated sequencing was carried out using an ABI PRISM® 3100 Genetic Analyzer in the AgResearch Molecular Biology Unit (Department of Biochemistry, University of Otago, Dunedin). Sequencing was carried out using Applied Biosystems BigDye® Terminator v3.1 fluorescent dye sequencing technology.

Fluorescent automated sequencing utilises a modification of the Sanger dideoxy chain-termination sequencing method (Sanger et al., 1977). It relies on the incorporation of fluorescently labelled dideoxynucleotides in cyclic rounds of primer annealing and extension. The dideoxynucleotides lack a 3' OH group that renders them incapable of forming a phosphodiester bond thus preventing chain elongation. The fragments are then separated electrophoretically and fluorescence detected using a scanning laser.

Reactions were carried out in a 10 μl total volume with a 1:4 dilution of BigDye® reaction mix (diluted in Sequencing Buffer) and 3.2 pmol of the relevant primer. Cycling was carried out as follows,

1. 96°C, 1 min [initial denaturation]
2. 96°C, 30 sec [initial denaturation]
3. 50°C, 15 sec [oligo annealing]
4. 60°C, 4 min [extension]
5. steps 2-4 above repeated, 25 cycles

Sequencing reactions were purified using a modified ethanol/ sodium acetate precipitation as follows: to the 10 μl reaction 24.5 μl MilliQ H₂O, 62.5 μl 95% ethanol and 3 μl 3 M NaOAc (pH 4.6) were added. Samples were vortexed to mix and incubated in the dark at room temperature for 30 min to 2 h. The DNA was then pelleted by centrifugation at 12000 × g for 25 min, the supernatant was discarded and
the pellet washed with 250 µl of 70% ethanol, before centrifuging at 12000 × g for another 5 min. Pellets were air dried at room temperature for 15 min in the dark and samples were re-suspended in 10 µl of HiDi formamide (Applied Biosystems). The sample was then transferred to a MicroAmp Optical 96-well plate (Applied Biosystems), denatured for 5 min at 95°C and sequenced on the ABI 3100 genetic analyser.

2.4.11.1 **Analysis of DNA sequences**

DNA sequence was edited using either ContigExpress within the Vector NTI advance v 9.1.0 (Invitrogen), Chromas (v 2.2.2) (Technelysium), or by editing the text output file based on sequence traces. DNA sequence was checked visually, and any errors detected were corrected and primer sequences removed. Regions of low sequence quality (high noise to signal ratio) were also removed.

Contig assembly, multiple alignments, pairwise alignments and translations were carried out using the appropriate tools in SeqLab (Wisconsin Package v 10.3, Accelrys Inc) or VectorNTI advance v 9.1.0 (Invitrogen).

Homology searches were performed using BLAST (Altschul et al., 1990) on the National Centre for Biotechnology Information (NCBI) BLAST web server (http://www.ncbi.nlm.nih.gov/BLAST/). Alternatively, the in house (AgResearch) BLAST servers were used to search the proprietary bovine and ovine expressed sequence tag (EST) libraries.

2.4.12 **SNP genotyping**

Single nucleotide polymorphism (SNP) markers were genotyped using primer extension with analyte detection on a MALDI-TOF mass spectrometer (Sequenom Inc) (Tang et al., 1999). Briefly, 2.5 ng of genomic DNA was amplified with 0.2 µM of primers A and B, which flank the SNP of interest. The amplification was carried out in a 10 µl volume using HotMaster Mix (Eppendorf) which contains a self-optimising magnesium concentration. Thermal cycling was carried out under the following conditions.
Chapter Two: Materials and Methods

1. 94°C, 1 min [initial denaturation]
2. 94°C, 30 sec [denaturation]
3. oligo annealing temperature, 30 sec
4. 65°C, 45 sec [extension]
5. steps 2-4 above repeated, 35 cycles
6. 65°C, 4 min [final extension and renaturation]

The amplified DNA was dephosphorylated using 0.3 U of shrimp alkaline phophatase at 37°C for 20 min, the shrimp alkaline phosphatase was then inactivated by incubation at 85°C for 10 min. This treatment dephosphorylates any residual amplification nucleotides, preventing their future incorporation and interference with the primer extension assay. The amplified PCR product was used as a template in a second, modified single-primer mini-sequencing reaction, whereby either single-base extension and chain termination or two to three base extensions occurs at the variant allele. Extension reactions contained 600 nM of extension primer, 50 μM d/ddNTP in Thermosequenase reaction buffer and 0.126 U Thermosequenase (Amersham Biosciences). Thermal cycling was then carried out as follows,

1. 94°C, 2 min [initial denaturation]
2. 94°C, 5 sec [denaturation]
3. 52°C, 5 sec [annealing]
4. 72°C, 5 sec [extension]
5. steps 2-4 above repeated, 45 cycles
6. 65°C, 4 min [final extension and renaturation]

The single primer mini-sequencing reactions were then desalted by addition of SpectroClean resin (Sequenom). The purified reactions were then spotted onto a chip (SpectroCHIPS, Sequenom) containing matrix pads. The matrix aids in desorption and ionization of the DNA. Chips were individually analyzed using the MALDI-TOF mass spectrometer. The resulting spectra were converted to genotype data using SpectroTYPER-RT software (Sequenom), which interprets the spectral output based on information for expected allele-specific oligonucleotide lengths generated during the assay design phase.
2.5 RNA METHODS

As far as practicable all solutions and plasticware used in the isolation and manipulation of RNA were treated with 0.1% diethyl pyrocarbonate (DEPC). DEPC inactivates RNases by covalent modification of -NH, -SH and -OH groups and is inactivated by autoclaving, where DEPC is hydrolysed to CO₂ and ethanol. Solutions for RNA work were treated with 0.1% DEPC and left to stand for 12 h at 37°C prior to autoclaving. Glassware that could not be DEPC treated were baked at 240°C for 4 h. All disposable plasticware was purchased RNAase free and used for RNA work only. All non-disposable plasticware was treated with a solution of 0.1 M NaOH plus 1 mM EDTA prior to being rinsed with DEPC treated water. Electrophoresis tanks dedicated for RNA analysis were cleaned with 0.5 % SDS, rinsed with distilled water, dried with ethanol and then filled with 3% H₂O₂ for 10 min. The tanks were then rinsed thoroughly with DEPC treated water.

Formamide for use in RNA solutions was deionised by treating with 2.5 % (w/v) Amberlite MB-1 resin on a magnetic stirrer for 1 hour. The resin was then removed by filtration through two sheets of Whatman No.1 filter paper. Deionised formamide was stored at 4°C in a dark bottle and used within 7 days.

2.5.1 RNA isolation

Liver tissue was snap frozen in liquid nitrogen and stored at -80°C (refer 2.2). RNA was extracted from frozen liver tissue using TRIzol reagent (Invitrogen). TRIzol is a mono-phasic solution of phenol and guanidine isothiocyanate and is a modification of the single-step RNA isolation method developed by Chomczynski and Sacchi (1987). During sample homogenization or lysis, TRIzol Reagent maintains the integrity of the RNA, while disrupting cells and cell components. Addition of chloroform followed by centrifugation separates the solution into an aqueous phase and an organic phase. RNA remains exclusively in the aqueous phase. After transfer of the aqueous phase, the RNA was recovered by precipitation with isopropanol.

Routinely, a tube of liver was removed from the -80°C freezer and kept on dry ice while 100 mg of liver tissue was chipped off the frozen sample. The remainder of
the sample was immediately returned to the -80°C freezer. The 100 mg of liver tissue was homogenised in 1 ml of TRIzol reagent using a Polytron mechanical homogeniser. Un-homogenised tissue was then removed from the mixture by centrifugation (12000 \times g for 10 min at 4°C). The homogenized samples were incubated for 5 min at RT to permit the complete dissociation of nucleoprotein complexes. Following this, 0.2 ml of chloroform was mixed with the sample and the sample was left to stand for 5 min at RT. Samples were then centrifuged (12000 \times g for 15 min at 4°C) and 500 µl of the upper (aqueous phase) was removed. The RNA was then precipitated from the aqueous phase by mixing with 0.5 ml of isopropanol, and incubating samples at room temperature for 10 min followed by centrifugation (12000 \times g for 10 min at 4°C). The RNA pellet was then washed with 1 ml of 75% ethanol, vortexed and centrifuged at 7500 \times g for 5 min at 4°C. The supernatant was discarded and RNA pellet dried at room temp for 10 min before re-suspending in 100 µl of DEPC treated water. In rare cases where the pellet would not dissolve entirely (presumably due to glycogen contamination) the RNA was heated to 55°C for a few minutes to aid re-suspension.

The quality of the isolated RNA was assessed by determining the \( \frac{A_{260 nm}}{A_{280 nm}} \) and \( \frac{A_{260 nm}}{A_{230 nm}} \) ratios (refer 2.4.10), and the integrity of the RNA was examined by formaldehyde-agarose gel electrophoresis (refer 2.5.3). Separation of the RNA on a denaturing (formaldehyde) gel with staining by ethidium bromide allowed the gel to be scanned (using the Personal Molecular Imager FX (Bio-Rad)) and the intensity of the 28S and 18S rRNA bands to be determined. For intact RNA the intensity of the 28S ribosomal RNA band should theoretically be approximately twice that of the 18S rRNA band, as the 28S rRNA is approximately twice the length of the 18S rRNA. Since the 28S rRNA is more labile than the 18S rRNA, equal intensities of the two bands generally indicates that some degradation of the RNA has occurred. Due to the nature of sample collection and the unavoidable delay in removing the liver from the deceased animal, RNA extracted from these tissues routinely showed some signs of degradation. Therefore in this study a 28S rRNA / 18S rRNA ratio of 1.6 or higher was considered acceptable.
2.5.1.1 Precipitation of RNA

In rare cases where RNA was not of sufficiently high concentration for further manipulations RNA was precipitated by the addition of 0.1 × volume 5 M NaAc, 5 μg glycogen (nuclease free) and 2.5 × volume 100% ethanol; the glycogen acts as a co-precipitant and aids RNA recovery. The solutions were incubated for 1 hour at -80°C followed by centrifugation at 4°C at 12000 × g for 20 min, the pellet was washed with 1 ml of 70% ethanol and re-centrifuged. The pellet was allowed to air dry prior to re-suspension in DEPC treated MilliQ H₂O.

2.5.2 DNase treatment of RNA

RNA was treated with recombinant DNase I (RNase free) (Roche Applied Science) according to the manufacturers instructions. Briefly, 2 μg of RNA was incubated in 10 mM sodium acetate and 0.5 mM MgSO₄ for 4 h at 37°C with 20 U of recombinant DNAase I (RNase free) and 2 U of RNasin RNase inhibitor (Roche Applied Science). The enzymes were then heat inactivated by heating for 15 min at 72°C. Following DNase treatment the RNA was precipitated as detailed in section 2.5.1.1 to avoid MgSO₄ carryover that may cause interference with the reverse-transcription reaction or subsequent PCR.

Effectiveness of the DNase treatment was assessed by PCR amplification of +/- RT controls with GAPDH. PCR was carried out as described in section 2.4.3 using 1 μl of the RT reaction and 1 μM each of the GAPDH qRT-PCR forward and reverse primers (Appendix A). PCR products were visualised by agarose gel electrophoresis in a 3% agarose gel.

2.5.3 Formaldehyde agarose gel electrophoresis

RNA was electrophoresed in 1-1.2 % formaldehyde agarose gels in 1 × MOPS buffer (20 mM MOPS (free acid), 5 mM sodium acetate, 1 mM EDTA, pH 7.0) with formaldehyde. Agarose was dissolved in 1 × MOPS and cooled prior to adding 0.22 M formaldehyde and 0.1 μg/ml of ethidium bromide (final concentration). Electrophoresis buffer was 1 × MOPS with 0.25 M formaldehyde. Samples were diluted in 5 × RNA loading buffer (0.0016% v/v saturated bromophenol blue
solution, 4 mM EDTA pH 8.0, 0.89 M formaldehyde, 20% v/v glycerol, 31% v/v deionised formamide and 4 × MOPS).

Formaldehyde gels were left to equilibrate in 1 × MOPS for 30 min prior to electrophoresis. Prior to loading the RNA sample, the sample was heated for 5 min at 65°C and loaded onto the equilibrated formaldehyde gel. Gels were electrophoresed at 5 V/cm. RNA was visualised under UV light using a UV transilluminator and documented either using the AlphaImager 2000 (Alpha Innotech Corporation) digital camera, the Gel Doc XR System with Quantity One Software (Bio-Rad) or Personal Molecular Imager FX (Bio-Rad) with Quantity One Software (Bio-Rad). For reference purposes 3 μg of the “0.5 - 10 Kb RNA ladder” (Invitrogen) was also electrophoresed.

2.5.4 cDNA synthesis

First strand cDNA synthesis was performed with 1 μg of total RNA using the Invitrogen SuperScript II reverse transcriptase system. SuperScript™ II Reverse Transcriptase is an engineered version of M-MLV (Moloney murine leukaemia virus) reverse transcriptase with reduced RNase H activity and increased thermal stability.

One μg of isolated RNA was mixed with 0.5 mmol of each dNTP, and 25 μg/ml (final concentration) of Oligo(dT)₁₂₋₁₈ (Invitrogen) or 12.5 μg/ml (final concentration) of random hexanucleotides (Roche Applied Science). Samples were heated to 65 °C for 5 min to denature the RNA, primer annealing was then done by quickly chilling the reaction on ice. Following incubation on ice the following were added; 1 × first strand buffer (Invitrogen), 20 mM dithiothreitol (DTT) and 20 U of RNasin® RNase inhibitor (Promega). Oligo(dT)₁₂₋₁₈ primed reactions were then heated to 42 °C for 2 min and random hexamer primed reactions heated to 25°C for 2 min prior to the addition of 200 U of SuperScript™ II reverse transcriptase (Invitrogen). Reverse transcriptions were carried out at 42°C for 50 min, prior to inactivation of the reverse transcriptase by heating to 70°C for 15 min.
Chapter Two: Materials and Methods

2.5.5 Northern blotting

The principle of Northern blotting is that a transcript within immobilised RNA can be quantified after hybridisation with a specific probe. The procedure involves (i) resolution by size of an RNA sample by denaturing gel electrophoresis, (ii) transfer of the resolved RNA to nylon membrane and (iii) hybridisation and detection of the RNA with radio-labelled probes.

For each gene of interest a separate Northern blot was made, blots were not stripped and re-probed.

2.5.5.1 Electrophoresis and electrotransfer

Briefly, 20 μg of RNA (and 7 μg of the “0.5 - 10 Kb RNA Ladder” (Invitrogen)) was electrophoresed on a 1.2% formaldehyde agarose gel (refer 2.5.3) either at 38 V for 16 h or at 120 V for 4 h, until the dye front was ~1 cm from the bottom of the gel. The gel was then washed in 1 x TAE and documented using the Personal Molecular Imager FX (Bio-Rad) with Quantity One Software (Bio-Rad) for quantitation of the 18S and 28S rRNA.

Gels were washed in 1 x TAE for 40 min and electroblotted onto Hybond H+ (Amersham Biosciences) using a TE series transphor electrophoresis unit (Hoefer Scientific Instruments) in 1 x TAE buffer (5 h at 20 V). RNA was fixed to the membrane using 0.05 M NaOH and the membrane rinsed in 2 x SSC prior to hybridization. The efficiency of the transfer was assessed by visualising the amount of RNA left on the gel following transfer. For sizing purposes the positions of the 18S and 28S rRNA bands and the RNA ladder were marked on the membrane.

2.5.5.2 Preparation of radioactively labelled DNA probes

Probe DNA was generated either by direct PCR amplification of the target followed by gel purification (refer 2.4.5) or by restriction endonuclease digestion of plasmid DNA containing the cloned DNA of interest (refer 2.4.9).

Labelling of probes for Northern blotting was performed using the RadPrime DNA labelling system (Invitrogen). Random primers (octamers) were annealed to the
Chapter Two: Materials and Methods

denatured DNA template and extended by Klenow fragment in the presence of \([\alpha-^{32}P]dCTP\) to produce high-specific activity DNA fragments.

Briefly, 25 ng of DNA was denatured by heating to 95°C for 5 min. Following the denaturation the sample was chilled on ice and 0.5 nmol of dATP, dGTP and dTTP were added together with 40% (v/v) random primers solution (50 mM Tris-HCl (pH 6.8), 5 mM MgCl2, 10 mM 2-mercaptoethanol, 60 µg/ml oligodeoxyribonucleotide primers (random octamers) with approximately 30 flCi \([\alpha-^{32}P]dCTP\). The samples were mixed thoroughly before adding 40 U Klenow fragment. The reactions were incubated at 37 °C for 15 min to 1 hour before purification using the High Pure PCR product purification kit (Roche Applied Science) to remove un-incorporated nucleotides. A 1 µl sample of labelled probe was spotted onto a piece of Whatman filter paper prior to purification and post-purification. The incorporation of radionucleotide into the probe was then estimated by Cerenkov counting and the probe was considered acceptable if the percentage incorporation exceeded 15% (usual incorporation ranged from 25 - 35%).

2.5.5.3 Prehybridisation and hybridisation
The membrane was pre-hybridised for 4 h at 42°C in a solution of 5 × SSC, 10 × Denhardt's, 0.05 M Tris-HCl pH 7.5, 0.1% saturated NaPPi, 0.85% SDS and 100 µg/ml of sheared denatured salmon sperm DNA (Roche Applied Science).

The radioactively labelled probe was denatured in a boiling water bath with the 100 µg/ml of sheared and denatured salmon sperm DNA (Roche Applied Science) for 5 min. The probe was then hybridised to the blot in a solution of 7 × SSC, 2 × Denhardt's, 0.1 M Tris-HCl pH 7.5, 0.2% saturated NaPPi, 1.7% SDS, 50% deionised formamide and for 18 h at 42°C.

Following hybridisation the membrane was washed for 20 min each in low stringency wash buffer (6 × SSC, 0.1% SDS, 50°C), medium stringency wash buffer (3 × SSC, 0.1% SDS, 60°C) and high stringency wash buffer (1 × SSC, 0.1% SDS, 60°C). The membrane was then exposed to Kodak BioMax XAR film (Eastman Kodak), with an intensifying screen at -80°C. If, following initial exposure to film
Chapter Two: Materials and Methods

the background levels were still high the membrane was subjected to an ultra-high
stringency wash (0.1 \times SSC, 0.1\% SDS, 65\degree C) and re-exposed to film.

2.5.5.4 Quantification and analysis of the autoradiographs.

As film is known to have a small linear range of detection for radio-active signals,
care was taken to ensure that the autoradiographs were not over-exposed. In most
cases multiple exposures of the same blot were taken and quantified, in an attempt to
ensure that exposures were in the linear range.

The autoradiographs were scanned using an ImageScanner™ II and LabScan™
(Amersham Biosciences), and quantified using ImageQuant TL™ (Amersham
Biosciences). The sizes of transcripts were estimated by comparison with the “0.5 -
10 Kb RNA size ladder” (Invitrogen).

Total absorbance units (AU) of each band was determined and normalised either to
the total AU for GAPDH in the same sample or to the intensity of the 18S rRNA
band obtained from the scanned ethidium bromide stained gel. This normalisation
attempts to correct for any variations in loading or transfer. Due to the narrow linear
range of the film used, it is important to note that the quantifications presented in this
thesis indicate possible trends in gene expression between samples rather than exact
numerical relationships.

2.5.6 TaqMan qRT-PCR

Quantitative RT-PCR (qRT-PCR) was used to quantify ABCB1 mRNA expression
in ovine liver RNA samples relative to GAPDH expression. Quantification was
performed using the TaqMan fluorogenic 5’ nuclease assay on an ABI PRISM 7700
Sequence Detection System (Applied Biosystems). The TaqMan assay has an added
layer of specificity which is useful in distinguishing between closely related gene
sequences, this specificity is provided by having a probe sequence that must also
bind to the target (in addition to the oligonucleotide primers).

2.5.6.1 TaqMan assay design

The TaqMan assays were designed using publicly available ovine mRNA sequences.
To ensure the amplification of cDNA rather than contaminating genomic DNA,
assays were designed to span an intron sequence. Intron positions were estimated using human genome sequence with the SPIDEY algorithm (Wheelan et al., 2001).

The primers and probe were designed using Primer Express version 1.5 (Applied Bioscience) following guidelines recommended by the manufacturer. Specifically the primers had a Tm of 50 - 60°C and a GC content of 30 - 80%. In general the primers were designed as close as possible to the probe binding region in order to minimise the PCR product size. The amplicon length was designed to be less than 200 bp and care was taken to regulate the GC content in the 3’ end of the primer to prevent GC clamping which is known to promote secondary structure formation. Probes were designed to have a Tm of 68 - 70°C, with no G on the 5’ end and long strings of G nucleotides in the probe were avoided.

2.5.6.2 TaqMan Assays
TaqMan assays were performed in a total volume of 25 μl using the relevant concentrations of forward and reverse primers, probe and a 1:2 dilution of TaqMan Universal PCR Master Mix (Applied Biosystems) with 0.5 μl of cDNA (refer 2.5.4). TaqMan Universal PCR Master Mix contains AmpliTaq Gold DNA Polymerase, AmpErase uracil-N-glycosylase, dNTPs with dUTP, ROX™ (a passive reference dye that is used to monitor the quality of the reaction and can be used to normalise the data if significant fluctuations occur), and optimized buffer components including MgCl₂ (final concentration is 3.5 mM). The TaqMan universal master mix contains dUTP rather than dTTP which allows the AmpErase uracil-N-glycosylase to remove any uracil incorporated into single- or double-stranded DNA (Longo et al., 1990) and prevent the re-amplification of carryover-PCR products.

Amplification was performed in a MicroAmp Optical 96-well reaction plate (Applied Biosystems). Thermal cycling was performed ABI Prism 7700 Sequence Detection System (Applied Biosystems) using the standard thermal cycling conditions.
Chapter Two: Materials and Methods

1. 50°C, 2 min [required for AmpErase UNG activity]
2. 95°C, 10 min [required for AmpliTaq Gold DNA polymerase activation and initial denaturation]
3. 95°C, 15 sec [denaturation]
4. 60°C, 1 min [annealing / extension]
5. steps 3-4 above repeated, 40 cycles

On all plates for quantification a relative standard curve of RNA was constructed using 5 × dilution series of cDNA from a specified animal (susceptible animal number 5, refer Table 2.2), allowing relative calculations of target concentrations. All reactions were performed in triplicate on a single plate.

2.5.6.3 TaqMan Assay Optimisation

The primer and probe concentrations were optimised for all TaqMan assays. This optimisation ensures that primers are not limiting in the amplification process, and also optimises the efficiency of the PCR reaction by compensating for differences in Tm between primers. The purpose of this procedure is to optimise detection levels by determining the lowest concentrations of primers and probe which give the maximum ΔRn (change in normalised fluorescence) at the minimum C_T (threshold cycle).

Primers were tested in combinations of 50, 300 and 900 nM and the probe was tested in 5 concentrations (50, 100, 150, 200 and 250 nM). Reactions were set up using a 5 × dilution of cDNA (refer 2.5.4) and a 1:2 dilution of TaqMan Universal PCR Master Mix. Thermal cycling was carried out as detailed in 2.5.6.2. Data was extracted and analysed as detailed in 2.5.6.4 using one-way ANOVA to assess statistical significance. The optimal primer and probe concentrations determined were used in all subsequent TaqMan assays.

Amplification efficiencies were calculated from the slope of the Log_{10} relative amount of RNA versus. C_T plot, using the following formula,

$$Ex = 10^{(-1/slope)} - 1$$
2.5.6.4 Analysis of Results

Results were analysed and data extracted using the sequence detection system software version 1.7 (Applied Biosystems). Data was normalised automatically to the passive reference dye (ROX) by dividing the emission intensity of the reporter dye (either TET or FAM) by the emission intensity of the Passive Reference (ROX) to obtain a ratio defined as the Rn (normalized reporter) for a given reaction. For each reaction the +Rn (the plateau of the reactions) and –Rn (the Rn at which there is a significant increase in Rn) were calculated and the threshold (C_T) determined to be the cycle at which the reaction was halfway between the initial and plateau phases in the exponential phase of amplification.

The data were exported into Excel (Microsoft) and the average and standard deviation C_T for each sample were calculated using the three replicates. The expression of the gene of interest was then expressed relative to the control gene expression (GAPDH) and the standard deviation estimated.

To determine if the two reactions could be multiplexed, and also if the comparative C_T method could be used for quantification a validation experiment was performed. For the validation experiment the ΔC_T was calculated by subtracting C_T for the gene of interest from the C_T for the control gene (GAPDH). If the two PCRs have an equal efficiency the slope of the line would be zero. In order to multiplex TaqMan assays or to use the comparative C_T method for quantitation the slope of the line must be less than 0.1 (as recommended by Applied Bioscience).

Statistical significance of differences among treatments was assessed with MINITAB (release 14, Minitab Inc.) using one way ANOVA with Fisher’s LSD post-hoc test. Graphs were plotted in SigmaPlot (version 9.0, SPSS Inc.), and curves (where appropriate) were fitted using linear regression. The standard method for calculating the standard deviation of a ratio was used when calculating ABCB1 expression normalised to GAPDH expression.

2.5.7 SYBR green and melt curve analysis

SYBR Green is a double stranded minor groove DNA binding dye that becomes incorporated into the PCR product during amplification. Following 40 cycles of
amplification the PCR product is then subjected to melt curve analysis. The temperature is sequentially increased until the fluorescence drops sharply; this drop in fluorescence is due to denaturation of the DNA strands and the subsequent release of the SYBR dye. The temperature at which this decrease occurs is indicative of the composition of the PCR product present and can indicate if there is one or more PCR products present in a particular sample. Amplification and melt curve analysis was carried out on the ABI Prism 7500 Sequence Detection System (Applied Biosystems).

Amplification with SYBR green followed by melt curve analysis was used to detect the presence of different amplicons in a single PCR reaction. The presence of different products in the PCR would be indicated by differences in the melting temperature of those different DNA species. This analysis was carried out using the Platinum SYBR Green qPCR supermix UDG (Invitrogen) according to the manufacturer's instructions. Briefly, PCR was carried out in 20 μl reactions with at least one sample amplified in duplicate. The reactions contained 1 × SYBR green qPCR supermix (Invitrogen), 0.1 × ROX (passive reference dye), 300 nM of each primer and 1 μl of cDNA (refer 2.5.4). Reactions were heated to 50°C for 2 min. then denatured at 95°C for 10 min.

Reactions were cycled as follows;

1. 50°C, 2 min [required for AmpErase UNG activity]
2. 95°C, 10 min [required for AmpliTaq Gold DNA polymerase activation and initial denaturation]
3. 95°C, 15 sec [denaturation]
4. oligo annealing temperature, 30 sec
5. 68°C, 30 sec [extension]
6. steps 3-5 above repeated, 40 cycles

Thermal cycling was carried out using the 7500HT sequence detection system (Applied Biosystems). A melt curve analysis was carried out at the end of each reaction.
2.5.8 Microarray methods

Global gene expression profiling was carried out with AgResearch in-house microarray slides consist of PCR-amplified bovine cDNA clones. The PCR amplified products were robotically spotted onto poly-lysine coated glass slides by D. Hyndman and Dr. T. Wilson (AgResearch MBU, Dunedin). The slides consisted of 1550 unique EST clones. The 1550 EST sequences on the array were clustered into a total of 1447 contigs, of which 1289 were represented only once on the array. Of the 1550 ESTs, 1058 had significant BLAST hits to the human RefSeq database, and these represented a total of 859 unique genes. Of the 1550 EST sequences, 1409 have been publicly released.

2.5.8.1 Reverse Transcription for microarray sample generation

Twenty μg of isolated total RNA was mixed with 2 μg of Oligo(dT)\textsubscript{12-18} (Invitrogen). Samples were heated to 70 °C for 10 min to denature the RNA, primer annealing was then done by quickly chilling the reaction on ice for 5 min. Following incubation on ice the following were added: 1 × first strand buffer (Invitrogen), 10 mM dithiothreitol (DTT), 20 U of Rnasin\textsuperscript{®} RNase inhibitor (Promega), 10 mM dNTPs (containing a 4:1 ratio of aminoallyl-dUTP (Sigma)) and 200 U of SuperScript\textsuperscript{TM} II reverse transcriptase (Invitrogen). Reverse transcriptions were allowed to continue at 42°C for 1 hour, at which point another 200 U of SuperScript\textsuperscript{TM} II reverse transcriptase (Invitrogen) was added and reactions allowed to continue for another hour.

The RNA was then hydrolysed and the reverse transcriptase inactivated by treating with 50% v/v 0.1 M NaOH at 70°C for 10 min. The reactions were cooled at room temperature and the pH neutralised by the addition of an equal volume of 0.1M HCl.

The reverse transcription reactions were cleaned up in a MicroCon YM-30 (Millipore) spin column, using two washes of 100 mM Na\textsubscript{2}CO\textsubscript{3} (pH 9) and the sample was concentrated down to a final volume of ~ 50 μl.
2.5.8.2 Dye labelling and clean-up

The purified cDNA with aminoallyl-dUTP incorporated was conjugated to monofunctional NHS-ester Cy3 and Cy5 dyes (Amersham) by combining and incubating for two h at room temperature in the dark.

Quenching of the un-incorporated fluorescent dyes was unnecessary as the two coupling reactions were performed and purified separately. The un-incorporated dye was removed using a QIA-Quick PCR purification kit (Qiagen) according to manufacturer’s instructions. Each dye was eluted in a total volume of 30 µl and combined to give a total volume of 60 µl. To this mixture 40 µl of ULTRAhyb buffer (Ambion) was added and loaded onto the pre-hybridised microarray slide underneath a glass lifter cover-slip (Erie Scientific).

2.5.8.3 Pre-hybridisation, hybridisation and slide washing

Slides were pre-hybridised for 20 min at 42°C in filtered (0.22 µm) pre-hybridisation buffer (5 × SSC, 0.1% SDS, 0.25% BSA). Pre-hybridisation was performed in a Falcon tube (Greiner) in a 42°C water bath.

The pre-hybridised slides were rinsed twice in deionised water, followed by a single wash in isopropanol and air dried. The slide was then covered with a Lifter cover-slip (Erie Scientific) and the 100 µl of denatured labelled cDNAs (heated to 95°C for 5 min) was loaded onto the pre-hybridised slide.

Slides were hybridised in a humidified chamber (CMT-Hybridisation Chambers, Corning) in a dark 42°C water bath for 18 h. Following hybridisation the slides were washed in covered Falcon tubes (Greiner). All wash solutions were made fresh and filtered through 0.22 µm filters. The washes were as follows: 7 min vigorous shaking in wash solution 1 (2 × SSC, 0.1% SDS) and 5 min each in solutions 2 (1 × SSC) and 3 (0.1 × SSC). Prior to scanning the slide was dried by centrifuging the slide inside a Falcon tube for 5 min at 50 × g.
2.5.8.4 Scanning and data acquisition
Hybridised slides were scanned using a ScanArray 5000 (Packard Biosciences). Both the power and PMT (photomultiplier tube gain) were altered to obtain similar peak heights for both the Cy3 and Cy5 channels. The dual images were then collected in TIFF format. The two TIFF images were imported into GenePix Pro (Axon Instruments) which then displays a visual composite of the array with the background fluorescence subtracted. Blocks are then overlaid over the image specifying the location of each EST, these were manually adjusted. Bad spots were then manually flagged if either the hybridisation was not uniform, the background level was high or there was spotting or streaking surrounding the EST. The clone tracking file was then imported and the correct localisation of ESTs and blocks was double checked. The slide was then analysed and the data exported as a GPR file. This file was then uploaded into the AgResearch microarray database.

2.5.8.5 Microarray statistics
The data for each slide was normalised using standardised procedures (Baird et al., 2004). This analysis was carried out by D. Baird (AgResearch, Lincoln) using GenStat release 6.1 implementing the average information restricted maximum likelihood algorithm. For each EST on the array, the ratio of Cy5 to Cy3 fluorescence was calculated and log2 transformed. The REML (restricted maximum likelihood) algorithms were used to adjust data for within slide bias such as dye bias, pin, row and carry over effects. The standardised residual of the normalised log ratio of the mean was then calculated. This is done by dividing the normalised log ratio of the mean of that EST by the standard deviation of the normalised log ratio of the mean for all good ESTs on the array. The corrected ratios were then standardised to the estimated standard errors and weighted across the duplicate ESTs on the slide. ESTs were excluded from further analysis if they had more than 11/22 bad spots or if their mean log intensity was less than 9. The remaining ESTs were sorted based on the modified T value (Smyth, 2004) for the log ratio of the mean. ESTs were counted as differentially expressed if the P value was less than or equal to 0.05 and if the EST exhibited a fold change of greater than 1.1 or less than 0.9.
2.5.8.6 Bioinformatics

The bovine EST libraries were annotated by the AgResearch bioinformatics group as follows. All bovine ESTs along with all publicly available bovine EST sequences were assembled into contigs using CAP3 (Huang and Madan, 1999) after an initial clustering step using BLAST. ESTs on the microarray were annotated by finding the human RefSeq (RefSeq release as at 11/4/2005) corresponding to the contig to which they belonged using BLASTN (Altschul et al., 1990). Each EST was annotated with the top human RefSeq hit. In cases where the EST matched more than one transcript variant of a gene then the top hit is listed.

All further bioinformatics analyses were carried out using the annotated human RefSeq genes. Firstly, possible biomarkers were identified by comparing the differentially expressed genes with the information contained in the plasma proteome database (http://www.plasmaproteomedatabase.org) (Ping et al., 2005). The aim of this analysis was to identify differentially expressed genes that generate proteins that are synthesised in the liver and secreted into serum and as such are possible biomarker targets.

The second approach utilised the gene ontology (GO) database (Harris et al., 2004b). This database provides biological classification of gene function through membership to functional categories that relate to certain biological processes, molecular functions or to cellular components. The aim of this analysis was to identify GO terms within a list of differentially expressed genes in the hope that this may aid in forming functional hypotheses about these differentially expressed gene sets. Gene Ontology terms significantly associated with the differentially expressed genes were found using the Expression Analysis Systematic Explorer (EASE). The list of differentially expressed genes was submitted to EASEonline (http://david.niaid.nih.gov/david/). The background list submitted included all human RefSeqs on the array. EASE calculates overrepresented functional gene categories compared to all the genes on the array (Hosack et al., 2003).
2.6 CELL CULTURE METHODS

2.6.1 Cells, media and consumables
Mammalian cell culture used HepG2 cells and HeLa cells, both of which were obtained as a frozen aliquot from the Cancer Genetics laboratory (Department of Biochemistry, University of Otago, Dunedin). HepG2 cells are a hepatocellular carcinoma cell line that retain many of the morphological features (Sormunen et al., 1993), drug metabolising activity (Knowles et al., 1980; Lu and Huang, 1994; Urani et al., 1998) and gene expression characteristics of primary hepatocytes (Harris et al., 2004a), making them useful for hepatotoxicity studies. HeLa cells are a cervical epithelial cell line and were used for comparison with published data concerning sporidesmin toxicity.

HepG2 and HeLa cells were cultured in α-MEM (Invitrogen) supplemented with 10% (v/v) fetal bovine serum (FBS) (Invitrogen) and 50 U/ml penicillin G and 50 μg/ml streptomycin sulphate.

Cultures were grown at 37°C in a humidified atmosphere of 95% air and 5% CO₂, and were maintained in the logarithmic growth phase by passage every 3 – 4 days (~80% confluence).

2.6.2 Passage of cells
Cells were harvested from several flasks by aspirating off the growth medium and washing once with pre-warmed phosphate-buffered saline (PBS) to remove any residual trypsin inhibitor present in the FBS. PBS containing 0.125% (w/v) trypsin was then added to each flask. The flasks were then incubated at 37°C for 5 – 10 min until the cells had detached from the substrate. Cells were then washed from the flasks with α-MEM + 10 % FBS and pelleted by centrifugation (100 × g for 5 min). The cell pellets were combined, washed in PBS and finally re-suspended in fresh α-MEM + 10 % FBS.
Maintenance cultures of HepG2 cells were generally passaged at a ratio of 1:4 or 1:6 but cultures that were used for experiments were counted using a hemocytometer and plated at a density of $1 \times 10^6$ cells/ml.

### 2.6.3 Treatment conditions

For clarity a time course of all treatments is given in Fig. 2.1.

#### 2.6.3.1 Sporidesmin treatment

Purified sporidesmin (refer Table 1.1), which was a mixture of approximately 94% sporidesmin A, 2.1% sporidesmin D, and 3 – 7% sporidesmin E was obtained from the AgResearch Toxinology Group (AgResearch Ruakura). Sporidesmin was dissolved in ethanol at a stock concentration of 5 mg/ml. In all experiments where cells were exposed to sporidesmin or other chemicals, appropriate solvent only controls were included.

To assess the cytotoxicity of sporidesmin, cells were cultured for 36 h post-passage before treatment with various concentrations of sporidesmin (0 – 10 µg/ml). Cells were harvested 16 h later (total of 52 h post-passage) and the percentages of non-viable cells were determined using the lactate dehydrogenase (LDH) assay (refer 2.6.4.1).

A time course study was also performed where cells were treated with 5 µg/ml of sporidesmin at 36 h post-passage, and cells were harvested every 4 h for 24 h (from 40 to 60 h post-passage) for determination of percent non-viable cells.

#### 2.6.3.2 Copper and zinc treatment

To assess zinc and copper cytotoxicity, the HepG2 cells were treated with either different zinc sulphate concentrations (0 – 500 µM) or different copper sulphate concentrations (0 – 500 µM) at 20 h post-passage (both compounds were dissolved in water). To measure the toxicity of these compounds together, cells were either treated with 200 µM zinc or copper together with different concentrations of the other metal for 16h. After the 16 h incubation with the metals the percentages of non-viable cells were measured.
Chapter Two: Materials and Methods

Figure 2.1: The timelines detail the experimental treatments of HepG2 cells. The top line (A) refers to experiments involving 16 h of zinc pre-treatment. The bottom line (B) refers to experiments involving 2 h of zinc pre-treatment. Note that the actinomycin D treatment was only performed in experiments involving 2 h of zinc pre-treatment as detailed in Section 2.6.3.3.

In experiments on zinc protection against sporidesmin toxicity, cells were incubated at either 20 h or 34 h post-passage with different concentrations of zinc sulphate (0 – 200 µM) and the incubation periods were for either 16 h or 2 h, respectively. At 36 h post-passage, the zinc-containing media were removed and the cells were washed with PBS before being treated with sporidesmin-containing media for a further 16 h. At 52 h post-passage, the percentages of non-viable cells were determined using the LDH assay.

2.6.3.3 Actinomycin D treatment
The effect of transcriptional inhibitor, actinomycin D, on zinc protection was determined by treating the cells at 33 h post-passage with 10 µg/ml of actinomycin D (dissolved in DMSO) (Fig. 2.1), then at 34 h post-passage with 200 µM zinc sulphate for a further 2 h (total of 36 h post-passage). The media containing zinc sulphate and actinomycin D were then removed and replaced with fresh medium containing sporidesmin (5 µg/ml) for a further 16 h of incubation. The percentages of non-viable cells were measured at 52 h post-passage.
2.6.4 Assays for cell death

2.6.4.1 Lactate dehydrogenase (LDH) assay

Lactate dehydrogenase (LDH) (EC 1.1.1.28) is a cytosolic enzyme that catalyses the conversion of pyruvate to lactate and oxidising NADH to NAD⁺.

\[
\text{Pyruvate} + \text{NADH} + H^+ \rightarrow \text{Lactate} + \text{NAD}^+
\]

The LDH leakage assay was used to measure the percentage of non-viable cells as a marker for cytotoxicity (reviewed in Fotakis and Timbrell, 2006). This assay measures the proportion of the cytosolic LDH released into the culture media by dead and dying cells as compared to the total LDH in the sample well.

In practice, the culture medium was removed and its LDH level measured. The adherent cells that remained in the sample well were washed briefly with PBS before lysing with 0.1% Triton X-100 in PBS; this lysate was then analysed for LDH activity. Routinely, 200 µl of assay buffer (0.2 mM NADH, 1.6 mM sodium pyruvate, 200 mM NaCl, 80 mM Tris-HCl pH 7.5) was aliquoted into each well of a 96-well plate, and 10 µl of either culture media or cell lysate was added. Data were collected using an ELx808iu ultramicroplate reader (Bio-Tek Instruments, Winooski, VT, USA). The absorbance at 340 nm was measured over 3 min and the initial rate of the oxidation of NADH to NAD⁺ was calculated. The percentage of non-viable cells was then calculated by determining the proportion of LDH in the media compared to the total LDH in a well (total media plus cell lysate) as follows:

\[
\text{Total LDH activity} = \text{LDH activity in culture medium} + \text{LDH activity in cell lysate}
\]

\[
\% \text{non-viable cells} = \left( \frac{\text{LDH activity in culture medium}}{\text{Total LDH activity}} \right) \times 100
\]

2.6.4.2 DNA laddering

Low molecular weight (LMW) DNA was extracted using a modification of the procedure described by Wyllie et al. (1980). Cells were lysed in 1 ml of lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5% Triton X-100) for 30 min on ice. The LMW DNA was then separated from intact chromatin by centrifugation at 15000 × g for 15 min at 4°C. The samples were then treated with RNase A (5 µg/ml)
for 1 hour at 37°C, followed by a 2 hour incubation at 50°C with proteinase K (500 μg/ml) and 0.5% SDS. The supernatants were extracted with phenol:chloroform (saturated with 10 mM Tris-HCl pH 8.0 and 1 mM EDTA) and DNA was precipitated with ethanol. The LMW DNA was then resolved on a 1.8% agarose gel which was documented with a Molecular Imager Fx gel scanner and Quantity One software (Bio-Rad).

2.7 COMPUTATIONAL METHODS AND STATISTICAL ANALYSES

2.7.1 General statistical analyses

2.7.1.1 Software
All statistical analyses were performed in either Microsoft Excel 2003 or with MINITAB® (release 14, Minitab Inc).

2.7.1.2 Detecting significant differences between sample groups
Statistical significance is derived from consideration of probability and allows conclusions to be drawn about groups of observations. In all cases the null hypothesis states that the observed difference is wholly accounted for by inherent variability. The alternate hypothesis states that the observed difference in group means is not due to inherent variability, but rather it is the result of experimental intervention.

Statistical significance was assessed using a two sample t-test with the Bonferroni correction for multiple testing (instead of $\alpha = 0.05 \alpha = 0.05/n$ where n is the number of tests performed), or one way ANOVA. Following a significant F-value obtained with ANOVA, a post-hoc test was used to refine the confidence intervals for all individual pairwise differences between a mean. In this study two post-hoc tests were used, Fisher’s LSD (least significant differences test) and Tukey’s HSD (honestly significant difference). The main difference between these tests is that Tukey provides a family error rate while Fisher provides an individual error rate. Therefore while the Fisher’s LSD is criticised for not adequately controlling for type I errors (concluding that there is a difference between sample groups when there isn’t), it is
very effective at detecting differences in means if it is applied only after the F-test in the ANOVA is significant at the 5% confidence interval (Brown, 2005). The Tukey (HSD) method, is designed to make all pair wise comparisons of means, while maintaining the error rate of the pre-established significance level. Because of this the calculated test statistic takes the number of possible comparisons into account, and as such results in a reasonably conservative estimate of significance.

2.7.2 Graphing
Graphs were plotted in SigmaPlot® (version 9.0, SPSS Inc.) and curves fitted using either sigmoidal or linear regression as indicated in the text. Image files were exported as either TIFF files or JPEG files.

2.7.3 Phylogenetic analyses
The aim of molecular phylogenetic analyses is to estimate the evolutionary relationships between organisms and between genes or proteins within an organism using DNA or protein sequence data. All phylogenetic algorithms are based on sequence homology and the assumption that the observed similarity is due to common ancestry (Baldauf, 2003).

2.7.3.1 Multiple sequence alignment.
All multiple sequence alignments in this study were constructed using CLUSTAL X version 1.81 (Thompson et al., 1997) using default parameters. CLUSTAL X is a progressive sequence alignment algorithm building an alignment up in a stepwise fashion based on a preliminary guide tree. The algorithm starts with the most similar sequences and progressively adds the more divergent ones (Baldauf, 2003).

Manual editing of the resulting alignments is required to minimise insertion / deletion events and to remove flanking unaligned sequences (Baldauf, 2003). Manually editing was carried out using Se-Al v2.0a11 (Rambaut, 1996).

The alignments were imported into PAUP* 4.0 (beta 10 version) (Swoford, 2002) for phylogenetic analysis as detailed below.
2.7.3.2 Neighbour joining analysis.
Many different methods are available for phylogenetic analyses (reviewed in Baldauf, 2003), one of the least computationally demanding is the neighbour-joining method. Neighbour-joining methods are based on distance-matrices. Simply, the evolutionary distance (which approximates the percentage sequence difference) is calculated for all pairwise comparisons in the multiple sequence alignment.

Aligned sequences were used to generate matrices of mean distances among proteins, and these matrices were used to generate a phylogenetic tree according to the neighbour-joining algorithm (Saitou and Nei, 1987). This tree was refined using the SPR branch-swapping technique under the minimum evolution criterion, implemented by PAUP*4.0b10 (Swofford, 2002). Bootstrapping (Felsenstein, 1985) was used to determine the relative support for the various branches of the tree (1000 replicates), and nodes with less than 50% support were collapsed to form polytomies. Phylogenies were visualised and manipulated using TreeView 1.6.2 (Page, 1996) and MacClade 3.0.4 (Maddison and Maddison, 1989).

2.7.3.3 Maximum parsimony analysis.
Unlike neighbour-joining, maximum parsimony is a computationally intensive analysis. It is a discrete data method where each column of the alignment is considered separately and a tree is constructed that best accommodates all of this information at each site.

Maximum parsimony analysis was carried out using PAUP*4.0b10 (Swofford, 2002) using heuristic searches and TBR branch-swapping. The analyses were done by stepwise random addition of taxa with 1,000 replications; confidence in the inferred topologies was estimated by bootstrapping (100 bootstrap pseudo-replicates, each with 10 random additions of sequences). Phylogenies were visualised and manipulated using TreeView 1.6.2 (Page, 1996) and MacClade 3.0.4 (Maddison and Maddison, 1989).

2.7.3.4 Bootstrapping
Bootstrapping is the most common way of determining levels of confidence in a constructed tree. Bootstrapping functions by taking random sub-samples from the
data set, and building trees from each of these. The frequencies with which the various parts of the constructed tree are reproduced in the random sub-samples indicate how reliable the tree nodes are. For 100% confidence the node has to be present in the constructed tree and in all of the random sub-samples (Baldauf, 2003).

2.7.4 Comparative mapping
Comparative mapping is a method by which genetic information can be extrapolated from information rich species such as humans or rodents, to information poor species such as ruminants. Comparative mapping utilises the extensive blocks of synteny conserved between animal species to estimate gene order in a particular chromosomal location (Burt, 2002).

The human genome map viewer was used to obtain human map positions for all of the genes of interest (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi). Comparative map positions for sheep were then assigned in two ways; the first was using the human and sheep synteny maps of (Maddox et al., 2003), the second was using an in house tool called Biomesh (McCulloh A, unpublished data). Biomesh uses the integrated bovine map to detect the nearest flanking sheep markers to a human locus.

A third approach utilised the Bovine Genome Assembly (v2.0) Browser v1.0 (also called gBrowse) (available at http://www.livestockgenomics.csiro.au/perl/gbrowse.cgi/bova2/). For genes annotated in the bovine genome the gene order in the surrounding region was compared with human to confirm synteny.

For genes that had not been annotated in the bovine genome the gene order in humans was used to extrapolate the location of the gene of interest in the bovine genome. This was done in conjunction with available supplementary information (Band et al., 2000; Itoh et al., 2005).

In both cases surrounding sheep microsatellite markers were obtained from gBrowse and the map positions of these microsatellite markers (sex averaged) were obtained from the best position maps (v 4.5) from http://rubens.its.unimelb.edu.au/%7Ejillm/jill.htm.
However, because not all sheep markers are annotated on gBrowse it was possible to further refine these putative map positions. This was done by determining whether the flanking markers indicated by gBrowse had any intervening ovine markers (http://rubens.its.unimelb.edu.au/%7Ejillm/jill.htm). For the intervening markers only those with NCBI sequences were analysed. Those markers with NCBI sequence were compared with the bovine genome using BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) to estimate the ovine marker positions relative to the bovine genome. In this way most of the genes were flanked by additional (un-annotated) ovine markers and the comparative map position could be refined.

2.7.5 Linkage mapping

Linkage mapping utilises recombination fractions within a pedigree to assess the genetic distance between two markers. Recombination fractions range from 0 where the loci are very close to each other to 0.5 where the loci are on different chromosomes (unlinked). In this study the pedigree used was the AgResearch International Mapping flock (IMF) which was developed for genetic mapping in sheep. The IMF consists of 9 three generation pedigrees comprising a total of 127 individuals (Crawford et al., 1995). Five breeds contributed to the pedigrees; Texel, Coopworth, Perendale, Romney and Merino. The different breed crosses were used to try and maximise the heterozygosity of the F1 generations.

The recombination fraction is the proportion of recombination events occurring between two loci. The lod (logarithm of the odds) score is the most common way of presenting the results of a linkage analysis (Morton, 1955). A maximum lod score of three is taken as evidence of linkage. This represents 1000:1 odds that the two loci are linked.

In order for linkage mapping to be successful an informative marker must be found within or close to the gene of interest. The most abundant markers and easiest to genotype are SNP markers which are common within introns. Thus the approach taken to discover SNPs was to sequence small introns in both directions in the IMF sires.
Chapter Two: Materials and Methods

To estimate intron size and position, the human mouse, and bovine genome sequences were extracted from the relevant databases using Golden Path (human and mouse) and using megaBLAST against the bovine genomic sequence (Zhang et al., 2000). These were then aligned with the ovine cDNA sequence using SPIDEY (Wheelan et al., 2001) set to adjust for divergent sequences. Primers were then designed to flank introns of a suitable size (< 800bp). The introns were then PCR amplified, purified, and sequenced from the four sires from the International Mapping Flock (IMF) (Crawford et al., 1995).

SNPs were genotyped as detailed in 2.4.12 and the genotypes were entered into the International Mapping Flock Database (AgResearch). Data were then extracted for CRI-MAP analysis (Lander and Kruglyak, 1995).
Chapter Three

Quantification of ABCB1 Gene Expression.

3.1 INTRODUCTION

Several classes of genes have been investigated as candidates for resistance to FE (refer to 1.7.2) but the ABC-transporter super-family is of particular interest. The potential of this gene family as a class of candidate genes was specifically highlighted when sporidesmin was shown to be a substrate for the yeast ABC transporter PDR5 (STS1) (Bissinger and Kuchler, 1994).

ABC transporters are known to transport a variety of both endogenous and exogenous compounds. Mutations or amplification of several of these transporters, including ABCB1, ABCC1 and ABCG2, can confer a multiple drug resistance phenotype. These proteins are known to contribute to resistance to chemotherapeutic

![Diagram of ABC transporters]

A. A ‘full’ ABC transporter consists of two sets of six transmembrane domains and two nucleotide binding domains (ABC cassettes, labelled in the diagram as ATP) encoded by a single polypeptide. In this diagram the ABC cassettes lie toward the C-terminus of the transmembrane segments in the ‘standard’ orientation.

B. A ‘half’ ABC transporter consists of one set of six transmembrane domains and one nucleotide binding domain. In this diagram the ABC cassette lies toward the N-terminus of the transmembrane segment, this is referred to as the ‘reverse’ orientation. Diagram adapted from Stefkova et al. (2004).
Chapter Three: Quantification of ABCB1 Gene Expression

compounds, such as doxorubicin, in humans (reviewed in Gottesman et al., 2002; Leslie et al., 2005).

3.1.1 ABC Transporters

The ABC transporter superfamily is the largest known family of transmembrane proteins. These transporters are found in all organisms from bacteria (reviewed in Davidson and Chen, 2004) through to higher mammals such as humans (reviewed in Dean et al., 2001) and sheep (Longley, 1998; Longley et al., 1999; Duncan et al., 2007). The ABC transporters bind and hydrolyse ATP to drive the transport of various molecules across cell and organelle membranes (reviewed in Higgins, 1992). The ABC transporters found in bacteria are primarily required for the import of essential nutrients. However in eukaryotes most of the ABC transporters are involved in the efflux of compounds from the cytoplasm. These compounds are transported to an intracellular compartment such as the endoplasmic reticulum or mitochondria, or outside of the cell.

3.1.1.1 Mammalian ABC Transporters

Generally ABC transporters can be classified as either full (Fig. 3.1A) or half-transporters (Fig. 3.1B). Full transporters have 12 transmembrane domains and two ABC cassettes, while half transporters have 6 transmembrane domains and a single ABC cassette. The diagrams in Fig. 3.1 also illustrate the two possible domain orientations observed in the ABC transporter super-family. The ‘standard’ orientation is illustrated in Fig. 3.1A where the transmembrane domains lie toward the N-terminus of the ABC cassettes. The so called ‘reverse’ orientation is observed in some ABC-transporters where the ABC cassette precedes the transmembrane domains and the transmembrane domains lie toward the C-terminus of the ABC cassette (Fig 3.1B). These domain orientations are thought to have arisen as a result of gene duplication or domain switching.

The mammalian ABC transporters are divided into 7 sub-families (A to G) based on sequence identity (reviewed in Dean et al., 2001). Interestingly, the ABCE and ABCF sub-families of mammalian ABC transporters lack transmembrane domains and are not thought to be active transporters (reviewed in Dean et al., 2001) although the mature proteins could associate with another ABC transporter and modulate the
function of this transporter. Alternatively, these proteins may require an interaction with another protein module containing transmembrane segments to form a mature protein, similar to that seen with some prokaryotic ABC transporters (Higgins, 1992). Although, this hypothesis is reasonably unlikely as several of the ABCE and F proteins are not membrane associated and have roles in transcription particularly in initiation of transcription (Tyzack et al., 2000; Dong et al., 2004).

3.1.1.2 The ABC cassette

The ABC cassette protein sequence is extremely well conserved both between ABC transporter family members within the same species, but more importantly the high levels of conservation extends to comparisons of ABC cassette sequences between very evolutionary divergent species (Sheps et al., 2004). Functionally, the ABC cassette is responsible for coupling ATP binding and hydrolysis to the transport of substrates. The ABC cassette is defined by several conserved protein motifs which include the Walker A and Walker B motifs (separated by 90 – 120 amino acids) that are common to all ATP binding proteins. The Walker A motif embodies a structure known as the phosphate-binding loop, which is a glycine rich loop followed by an uncapped \( \alpha \)-helix. This motif binds the ATP molecule through electrostatic interactions with the triphosphate moiety, while the Walker B motif is involved in binding magnesium ions which are required for ATP hydrolysis (Walker et al., 1982). In addition to these two regions, three other conserved motifs are present in the ABC cassette; the switch region, which contains a histidine loop and is thought to polarize the attaching water molecule for hydrolysis (Linton and Higgins, 1998), and the Q-motif which interacts with the \( \gamma \)-phosphate of ATP through a hydrogen bond (Diederichs et al., 2000). Finally, the motif that distinguishes the ABC transporters from other ATP binding proteins is the signature C motif (also called the linker peptide or Walker C motif) which is located just upstream of the Walker B site (Hyde et al., 1990).

Recent structural and biochemical studies have demonstrated that the ABC cassettes of ABC transporters dimerize in the presence of ATP (Moody et al., 2002; Smith et al., 2002). The signature C motif (LSGGQ) stabilises this dimer by hydrogen bonding via the hydroxyl group of the highly conserved serine and the amide of the invariant second glycine with the \( \gamma \)-phosphate of ATP (Smith et al., 2002).
3.1.2 PDR5 (Pleiotropic drug resistance protein 5, STS1, suppressor of sporidesmin toxicity 1)

PDR5 is a full yeast ABC transporter consisting of 12 transmembrane domains and 2 ABC cassettes in the reverse orientation (Fig. 3.2). PDR5 is known to transport a multitude of xenobiotics and has some overlapping substrate specificity with other yeast ABC transporters like YDR011W (SNQ2) (Rogers et al., 2001).

PDR5 was first isolated from S. cerevisiae. It was identified during a genetic screen of S. cerevisiae that had been transformed with a genomic library. Transformants were selected that were able to grow in the presence of 500 μg/ml sporidesmin (1.1 mM) (Bissinger and Kuchler, 1994).

It was subsequently shown that sporidesmin was a substrate for PDR5 mediated transport. Over-expression of PDR5 in S. cerevisiae conferred tolerance to sporidesmin and a series of PDR5 deletion mutants conferred a hypersensitive phenotype to sporidesmin and a structurally unrelated drug, cycloheximide (Bissinger and Kuchler, 1994).

3.1.3 ABCB1 (P-glycoprotein, MDR1)

The first studies that attempted to characterise a mammalian ortholog of the yeast PDR5 protein focussed on ABCB1 (P-glycoprotein, MDR1). ABCB1 was thought to be the likely ortholog to PDR5 based on sequence similarity and also based on its role as a major determinant for tumour resistance to chemotherapeutic compounds in
humans (reviewed in Endicott and Ling, 1989). Many of these compounds are small, electrophilic and lipid soluble molecules with which sporidesmin shares some biophysical similarity.

The ABCB1 gene was localised to ovine chromosome 4q15-20 (Longley et al., 1999) (syntenic to human 7q20-30). Two DNA markers in this region were genotyped across FE selection-line animals. The first was a microsatellite marker, OarMEM3, which was linked to ABCB1. This marker was genotyped in 65 resistant, 65 susceptible and 40 control-line animals. The second marker was an EcoRI RFLP in the coding region of ABCB1 and was genotyped in only 30 resistant and 20 susceptible selection-line animals. Alleles of OarMEM3 or the EcoRI RFLP did not segregate with resistance to FE in these animals. The OarMEM3 microsatellite was also genotyped across one FE outcross QTL pedigree (refer 1.7.1) and no evidence for association of this locus was found.

However, the same study showed that resistant animals exposed to sporidesmin had consistently higher expression of ABCB1 mRNA in liver tissue as determined by competitive RT-PCR (Longley, 1998). This difference in expression corresponded to an approximate 4-fold induction of ABCB1 expression in the livers of resistant animals exposed to sporidesmin as compared to naïve resistant animals and also to naïve and sporidesmin-dosed susceptible animals (Longley, 1998).

As DNA markers surrounding the ABCB1 locus did not segregate with resistance to FE the author suggested that this locus was weak or possibly recessive with respect to FE resistance (Longley et al., 1999). It was therefore possible that the 4-fold induction of ABCB1 gene expression observed could be due to a cis- or trans- acting factor affecting ABCB1 expression. Interestingly, the reported 4-fold induction would raise ABCB1 expression to levels consistent with that seen in a multi-drug resistance phenotype (Chaudhary and Roninson, 1993) and therefore could conceivably modulate resistance to sporidesmin.
3.1.4 ABCB gene family

All of the ABC family members share significant similarities in domain structure and organisation (refer 3.1.1.2 and also chapter 4) as well as a high degree of sequence similarity (particularly within the ABC cassette). This sequence similarity within the ABC cassette has been used to group the ABC transporter super-family sequences into sub-families, such that all members of a sub-family are more similar in sequence to each other than to any member of another sub-family (reviewed in Dean et al., 2001; Sheps et al., 2004).

Members of the ABCB subfamily include full transporters, such as ABCB1, ABCB4 (MDR3) and ABCB11 (SPGP or BSEP), and half-transporters, including ABCB2 and ABCB3 (TAP1 and TAP2). Other half-transporters within this sub-family are mitochondrial transporters ABCB6, ABCB7, ABCB8 and ABCB10, as well as ABCB9 which has a putative lysosomal localization. The full transporters of the subfamily are localized to the plasma membrane (in the apical membrane compartment), while the half-transporters are found in the membranes of various organelles. Therefore, the full transporters are likely to be of interest when considering cellular resistance to sporidesmin, as they are positioned on the cell membrane and would be capable of extruding the toxin from the cells.

3.1.5 Aims and rationale

Prior to further investigation of potential cis- or trans- regulators of ABCB1 from animals resistant and susceptible to FE, it was imperative to confirm the differential expression of ABCB1 using a method unrelated to the original competitive RT-PCR experiments (Longley, 1998). For this purpose TaqMan qRT-PCR was chosen as it is a very specific assay relying on binding of primers and a probe to the target sequence, negating possible cross reactions with other sub-family members.
3.2 RESULTS

3.2.1 TaqMan Assay Design

The TaqMan assay (Fig. 3.3) utilises an oligonucleotide probe with a reporter fluorescent dye on the 5' end and a quencher dye on the 3' end. While the probe is intact, the proximity of the quencher dye reduces the fluorescence emitted by the reporter dye by Förster resonance energy transfer. During PCR amplification using gene specific primers the annealed probe is cleaved by the 5' nuclease activity of Taq DNA polymerase as the primer is extended (Fig. 3.3). The cleavage of the probe separates the reporter dye from the quencher dye, resulting in an increase in fluorescence intensity proportional to the amount of amplicon produced. Due the requirement for both primers and probe to anneal to the template, the TaqMan assay is very specific. However, assay design is critical and there must be an area of the mRNA that has a sequence amenable to the design of discriminating primers and probe sequences.

3.2.1.1 ABCB1

The ABCB sub-family of ABC transporters are known to have a common evolutionary relationship (Fig. 3.4). These proteins share significant sequence similarity and it was therefore important to design a TaqMan qRT-PCR assay that was specific to ABCB1. Phylogenetic analysis carried out as part of this study using the nucleotide sequences of the ABCB family (Fig. 3.4) shows that the ABCB family can be divided into two clades, although this division does not have strong bootstrap

Figure 3.3: Diagram illustrating TaqMan chemistry (Fluorogenic 5' Nuclease Assay). Diagram courtesy of R Harrison (Applied Biosystems).
Chapter Three: Quantification of ABCB1 Gene Expression

Figure 3.4: Neighbour-joining tree for the human ABCB sub-family. mRNA sequences were extracted from Genbank, the accession numbers are ABCB1 (NM_000927), ABCB2 (NM_000593), ABCB3 (NM_000544), ABCB4 (NM_00443), ABCB5 (NM_178559), ABCB6 (NM_005689), ABCB7 (NM_004299), ABCB8 (NM_007188), ABCB9 (NM_019625), ABCB10 (NM_012089), ABCB11 (NM_003742) and ABCC1 (NM_004996). In the case of the full transporters (ABCB1, ABCB4, ABCB11 and ABCC1) the mRNA was split into two halves a C-terminal (CT) and N-terminal (NT) half. Each of these halves contained a transmembrane domain and an ABC cassette. The sequences were aligned using ClustalW and the phylogenetic tree generated using PAUP*4.0b10 and a neighbour joining algorithm with 100 bootstrap replicates (refer 2.7.3.2). The human ABCC1 sequence was included as an outgroup and used to root the tree. The ABCB family can be subdivided into two clades (A and B) with moderate bootstrap support.

support. However, the resolution within these clades, particularly clade A, is excellent. In particular, there is very strong bootstrap support showing that the ABCB1 N-terminal half of the mRNA is most closely related to the N-terminal half of the ABCB4 mRNA, with some similarity to the N-terminal half of ABCB11. The same relationship is observed for the C-terminal domains of these mRNA sequences excepting that the half transporter, ABCB5, is also located in this group.
The degree of sequence similarity between the ABCB sub-family members, in particular ABCB1 and ABCB4, must be taken into consideration when designing qRT-PCR assays for a specific transporter.

To determine the level of nucleotide conservation across the ABCB1 and ABCB4 genes, the human ABCB1 and ABCB4 mRNA sequences were extracted from NCBI and aligned using CLUSTALW. Comparison of human ABCB4 and ABCB1 shows an overall identity of 75%. However, there are two main areas of similarity seen, 81% in domain I and 80% in domain II (Fig. 3.5), with a highly variable leader and linker region. Using RPS-BLAST (Reverse Position Specific) (Marchler-Bauer and Bryant, 2004) it was determined that each of these conserved domains contain a transmembrane domain (consisting of 6 putative transmembrane helices) and an ABC cassette.

Comparison of human (NM_000927) and sheep (NM_001009790) ABCB1 yields an identity of 87%, and partial sequence of ABCB4 (117 bp from the 3’ region of the coding sequence, unpublished data) from sheep (Longley, 1998) shows an identity of 92% to human ABCB4 (NM_000443).

Pairwise comparisons of the coding domain sequences of human ABCB1 with ABCB4, ABCB5 and ABCB11 confirm the phylogenetic analysis (Fig. 3.4) that ABCB1 shares most sequence identity to ABCB4. Global alignment of ABCB1 and ABCB4 (NM_000927) yields an identity of 76%, ABCB1 and ABCB5 an identity of

![Figure 3.5: Sequence similarity of the human ABCB1 (NM_000927) and ABCB4 (NM_000443) coding domain sequences. Areas of homology were initially detected using BLAST 2 (Tatusova and Madden, 1999), the areas were then refined by aligning the two sequences using AlignX (Invitrogen). In the diagram grey bars correspond to regions of sequence identity. There are two unique regions, labelled the leader and linker regions, respectively.](image-url)
Chapter Three: Quantification of ABCB1 Gene Expression

41%, ABCB1 and ABCB11 (NM_003742) an identity of 59%. Note that only the coding domains were used for this analysis as the untranslated regions are usually very divergent due to the lack of functional constraint.

TaqMan qRT-PCR assays were designed using Primer Express (Applied Biosystems) as detailed in 2.5.6.1. Both the ‘leader’ and ‘linker’ DNA sequences identified as being divergent from the other ABCB sub-family members (Fig. 3.5) were evaluated for assay design using Primer Express. The linker region was not suitable for assay design and therefore the leader region was used to design the TaqMan qRT-PCR assay. Primers were designed to flank intron 2 of the ABCB1 gene, this intron is 10 kb in humans. The sequence flanking the intron was amenable to designing the probe such that the probe binding site straddles the two intron / exon boundaries. This is desirable in TaqMan qRT-PCR assay design as it removes the complication of amplification from contaminating genomic DNA.

Primers were tested (Fig. 3.6) and were shown to produce a single specific amplicon of the expected size (~90 bp) in the cDNA samples but not in minus reverse transcriptase controls or the genomic samples.

3.2.1.2 GAPDH (Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12))

Gene expression studies usually normalise expression of their gene of interest to a constitutively expressed endogenous control such as GAPDH. This normalisation
negates some of the experimental variation, such as differences in RT efficiency, or RNA loading. GAPDH is commonly used as an endogenous control as it is a key enzyme in the glycolytic pathway and is therefore thought to be constitutively expressed. Additionally, GAPDH has been used as an endogenous control in more recent studies examining the expression of ABC transporters in humans (Langmann et al., 2003).

DNA sequences annotated as encoding ovine GAPDH were extracted from the NCBI database. Three sequences were obtained (accession numbers AF022183, U94889

<table>
<thead>
<tr>
<th>DNA sequence</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAR_AF022183</td>
<td>(1)</td>
</tr>
<tr>
<td>OAR_AF035421</td>
<td>(325)</td>
</tr>
<tr>
<td>Consensus</td>
<td>(325)</td>
</tr>
<tr>
<td>OAR_AF022183</td>
<td>(51)</td>
</tr>
<tr>
<td>OAR_AF035421</td>
<td>(375)</td>
</tr>
<tr>
<td>Consensus</td>
<td>(375)</td>
</tr>
<tr>
<td>OAR_AF022183</td>
<td>(101)</td>
</tr>
<tr>
<td>OAR_AF035421</td>
<td>(475)</td>
</tr>
<tr>
<td>Consensus</td>
<td>(475)</td>
</tr>
</tbody>
</table>

Figure 3.7: Multiple sequence alignment of publicly available ovine GAPDH sequences. Sequence labels are NCBI accession numbers preceded by the abbreviation OAR to indicate the sequence is from *O. aries*. Shaded areas of the sequence indicate identity between the sequences. The positions of the forward, reverse primers (indicated in blue and by the blue arrows) and probe sequence (in red and double underlined) used in qRT-PCR are indicated. The position of human intron 7 (position 483) is indicated by the arrow head.
and AF035421). The sequences were aligned using CLUSTAL X version 1.81 (Thompson et al., 1997) using default parameters (Fig. 3.7). The positions of human introns relative to the ovine GAPDH sequence were determined by comparison with the only human genomic DNA sequence with homology to the GAPDH mRNA sequence available at the time which was AC002389. This contig has been subsequently found to map to human chromosome 19 and actually encodes GAPDHS (a spermatogenic version of GAPDH).

Subsequent comparison of the intron / exon structures for both GAPDH and GAPDHS carried out using SPIDEY (Wheelan et al., 2001) and the corresponding human genome sequences suggests high conservation of the intron / exon structure between these two genes. Although the first intron position is not conserved between GAPDHS and GAPDH (due to an N-terminal extension of GAPDHS), the remainder of the intron positions are conserved with two exceptions; introns 8 and 9 which are found in GAPDHS and not in GAPDH.

The TaqMan qRT-PCR assay was designed surrounding intron 7 of GAPDH as detailed in 2.5.6.1. In human intron 7 in GAPDHS is 314 bp and in GAPDH the intron 7 is 194 bp. The subsequent release of the bovine genome has allowed the confirmation of the intron / exon structure in GAPDH between human and cattle (in cattle intron 7 is 180 bp in length).

The intron / exon structure was then superimposed on the sequence alignment (Fig. 3.7) and the TaqMan qRT-PCR primers and probe were designed using Primer Express (Applied Biosystems). The positions of the primers and probe are indicated in Fig. 3.7 and yield a product size of 140 bp. Although ideally the binding site of the probe would overlie the intron / exon boundary, in this case the sequence surrounding the boundary was unacceptable for probe design.

The TaqMan qRT-PCR primers were tested to confirm their specificity (Fig. 3.8A). A single PCR product was amplified of the expected size from cDNA. However, the same product was obtained from sheep genomic DNA and indicating that the primers are able to amplify a pseudo-gene in the sheep genome or that the position or size of intron 7 is not conserved in sheep. The primers were also able to amplify directly
from RNA (-RT controls) indicating that the RNA preparations contained a significant quantity of genomic DNA. This genomic DNA contamination is problematic for the quantification of GAPDH expression. Therefore it was decided to DNAse treat the RNA samples prior to reverse transcription.

The RNA was treated with DNAsel (RNAse free) prior to reverse transcription as detailed in 2.5.2. The success of the DNAse treatment / reverse transcription was assessed by PCR amplification of the samples (including –RT controls) with the

---

**Figure 3.8: Testing of the GAPDH qRT-PCR primers.** Primers were tested using cDNA from two animals, S5 and R2 (Table 2.1 and 2.2) (A) Isolated RNA (without DNAse treatment) was reverse transcribed (including the appropriate minus RT controls), genomic DNA from *O. aries* (OAR) and *B. taurus* (BTA) are included for comparison. A no template PCR control (indicated) was run on each gel. B) RNA samples were treated with DNAse prior to reverse transcription. PCR products were resolved in a 3% agarose gel. The size of marker bands in bp are indicated on the left hand side of the gel.
qRT-PCR GAPDH primers. A typical result is seen in Fig. 3.8B where the expected product size is obtained for the reverse-transcribed DNAse treated samples, but no corresponding amplification product could be detected in the –RT controls. This confirms the absence of genomic DNA contamination in the samples. All RNA samples used for the quantification of ABCB1 gene expression were treated with DNAse in this way.

3.2.2 TaqMan Assay Optimisation
The purpose of optimising primer concentrations is to determine the minimum primer concentration giving the maximum ΔRn, which maximises the sensitivity of the TaqMan assays (as detailed in 2.5.6.3).

Primer optimisation was carried out by testing the nine different combinations of three primer concentrations (50, 300 and 900 nM). At the nine different primer concentration combinations tested (Fig. 3.9A) there was no significant difference in efficiency of GAPDH amplification (overall ANOVA $P = 0.071$) (Fig. 3.9A, light bars) and therefore 300 nM of each primer was used in further experiments. However, for ABCB1 there was a significant difference (Fig. 3.9A, dark bars) with an overall ANOVA of $P = 0.001$. Using Tukey’s post-hoc test, the amplification efficiency at 300/900 was significantly higher than 300/300. Therefore in all further experiments 300 nM of forward primer and 900 nM of reverse primer were used.

The purpose of optimising probe concentrations is to determine the minimum probe concentration that gives the lowest $C_T$ for each probe target thus improving the sensitivity of the TaqMan assay. Five different probe concentrations (50, 100, 150, 200 and 250 nM) were tested for both GAPDH and ABCB1 (Fig. 3.9B). As can be seen in Fig. 3.9B no significant difference was observed when comparing the $C_T$ for any of these concentrations with the exception of 50 nM for which the $C_T$ was significantly higher for both ABCB1 and GAPDH ($P = 0.008$, $P = 0.008$, respectively) than the remaining concentrations. Therefore 150 nM was chosen as the optimal probe concentration for both ABCB1 and GAPDH.
Chapter Three: Quantification of ABCB1 Gene Expression

A. Primer optimisation, both sets of primers were tested at three concentrations (in 9 combinations), no significant difference in the ΔRn was seen for GAPDH, however 300/900 nM resulted in a significantly higher ΔRn for ABCB1 (* P<0.05). B) probe optimisation, probes were tested in five different concentrations and in the case of both GAPDH and ABCB1 100, 150, 200 and 250 nM performed equally well, but better than 50 nM. ** P<0.01 C) relative efficiency experiment, both ABCB1 and GAPDH were tested across a 2 × dilution series of RNA (from animal S5), the ΔCt was calculated by subtracting the Ct for GAPDH from ABCB1 and a linear regression fitted, the slope of this regression is 0.3049. In all cases the data presented are the mean ± standard deviation.

Therefore in all further experiments the following primer and probe concentrations were used:

GAPDH: 300 nM of both forward and reverse primers, and 150 nM of probe.
ABCB1: 300 nM of forward primer, 900 nM of reverse primer and 150 nM of probe.
Chapter Three: Quantification of ABCBJ Gene Expression

The validation experiment is designed to assess the relative efficiencies of each of the PCR reactions for GAPDH and ABCBJ and was performed in order to determine if quantification of ABCBJ could be achieved using the comparative C_T method as opposed to the relative standard curve method.

The PCR efficiencies were assessed by amplification of a dilution series of calibrator cDNA, in this case using a 2 \times dilution series of cDNA obtained from resistant animal R2 (ranging from undiluted to a 72 times dilution). This calibrator cDNA was then analysed for both GAPDH and ABCBJ expression. The data were extracted in the standard way using SDS software and the ΔC_T was calculated by subtracting the C_T obtained for GAPDH from that obtained for ABCBJ (the standard deviation was calculated by assuming the covariance of the two values was 0 (i.e. independently obtained).

A linear regression curve was fitted in sigma plot and the corresponding curve had a slope of 0.3 (with an R^2 value of 0.894). According to manufacturer’s guidelines, for the comparative C_T method to be used the slope of this line must not exceed 0.1. Therefore the relative efficiencies of the ABCBJ and GAPDH amplification reactions were too divergent to use comparative C_T method. Therefore all quantifications were performed and analysed using the relative standard curve method.

3.2.3 Quantification of ABCBJ using the TaqMan Assay

3.2.3.1 Standard Curve Generation

As the comparative C_T method could not be used for quantification of ABCBJ with GAPDH as the endogenous control due to significant differences in amplification efficiencies, the relative standard curve method was used instead. Therefore on each plate of samples a 5 \times dilution series of a single susceptible animal (S5) cDNA was used to generate a relative standard curve.

Typical standard curves for ABCBJ and GAPDH are shown in Fig. 3.10C. Examples of the amplification curves are shown in Figs. 3.10A (GAPDH) and 3.10B.
Chapter Three: Quantification of ABCB1 Gene Expression

Figure 3.10: qRT-PCR standard curves for GAPDH and ABCB1. A) and B) Sequence detection software output of amplification curves of GAPDH and ABCB1 respectively. ΔCt values were obtained by determining the cycle at which the amplification was approximately half of the linear amplification phase. This value was then plotted against the log10 of the relative amount of RNA (C). In both cases the graphs are linear over a range of RNA concentrations (125 fold dilution).

(ABCB1). The Cₜ values obtained from these amplification curves were then plotted against the log₁₀ of the relative amount of RNA (relative to the RNA input in the original RT reaction). The amplification efficiencies for ABCB1 and GAPDH were determined from the slope of the standard curve (according to the formula given in 2.5.6.3) and were 92.3% and 61.2%, respectively.

The relative amount of RNA was set at a value of 50 (assuming 1 µg of total RNA is used in a 20 µl RT reaction then the approximate ‘RNA’ concentration in the undiluted cDNA would be 50 ng/µl) for the undiluted sample. Linear regression was then used to fit a curve and in all cases an R² value of greater than 0.94 was obtained.
Chapter Three: Quantification of ABCB1 Gene Expression

The equations generated by the linear regression were then used to determine the relative amounts of target (ABCB1 and GAPDH) in the samples.

3.2.3.2 Sample Analysis

The C₇₅s obtained for all samples (in triplicate) were compared with the relative standard curve for the plate that they were analysed on to obtain the relative amounts

![Graph A](image1)

![Graph B](image2)

Figure 3.11: Graphs show the relative levels of target mRNAs (as determined using the standard curve method of quantification) for each of the animals tested for both ABCB1 (A) and GAPDH (B). Resistant animals are on the left (animals R1 – R7), susceptible animals are on the right (animals S1 – S4 and S6 – S7). Dark bars the expression pre-exposure to sporidesmin and the light bars represent the exposure post-exposure to sporidesmin. Data is shown as the mean ± SD of triplicate samples.
of the target mRNA. For every sample a -RT control was analysed simultaneously with the cDNA sample. In all cases the value obtained for the -RT controls was less than 1% of the relevant measurement.

The target (ABCB1 and GAPDH) mRNA levels were obtained for each animal. As can be seen from Fig. 3.11 there is a substantial degree of variation seen in the expression of both ABCB1 (Fig. 3.11A) and GAPDH (Fig. 3.11B). Variation can originate from many sources including experimental variation and inter-animal variation; however the most likely source for some of this variation is the amount of RNA used in the reverse transcription reaction. If this was a significant contributor then the expression of the target genes would be correlated with the concentration of RNA (as determined by the UV absorbance values, refer 2.4.10). Pearson's correlation coefficients were calculated for both ABCB1 and GAPDH expression with the relative amount of RNA both pre- and post- DNase treatment.

In general the expression of GAPDH does not co-vary with the RNA concentration used in the initial RT reaction. The Pearson correlation co-efficient for GAPDH expression and RNA concentration pre- and post- DNase treatment were -0.051 ($P = 0.804$) and 0.321 ($P = 0.110$), respectively. However, the Pearson correlation co-efficients for ABCB1 expression and RNA concentration pre- and post- DNase treatment were -0.249 ($P = 0.220$) and 0.436 ($P = 0.026$) respectively. This suggests that the variation in ABCB1 expression is due (at least in part) to variation in the amount of RNA used in the initial reverse transcription reaction.

More interestingly, there is also a correlation between the $C_T$ obtained for ABCB1 and the $C_T$ obtained for GAPDH with a Pearson correlation co-efficient of 0.830 ($P = 0.001$). This suggests that for the most part variation seen in ABCB1 expression is co-ordinated with variation in GAPDH expression. This suggests that GAPDH is an adequate endogenous control that would account for some of the experimental variation observed and allow us to detect true differences in target gene expression between resistant and susceptible animals.
Chapter Three: Quantification of ABCB1 Gene Expression

Figure 3.12: Expression of ABCB1 normalised to GAPDH for individual animals. Data is expressed relative to R1 which has been set to an arbitrary value of 1.0. Resistant animals are on the left (animals R1 – R7), susceptible animals are on the right (animals S1 – S4 and S6 – S7). Dark bars the expression pre-exposure to sporidesmin and the light bars represent the exposure post-exposure to sporidesmin. Data is the mean ± SD.

3.2.3.3 Normalised Data.

The expression data for ABCB1 was normalised to the expression of the constitutively expressed GAPDH gene. Additionally the pre-dose ABCB1 expression level of animal R1 has been arbitrarily assigned a value of 1 and the remaining animals are expressed relative to that (Fig. 3.12).

Figure 3.12 summarises the ABCB1 expression of individual resistant and susceptible sheep both pre- and post-exposure to sporidesmin. As the RNA samples used in this study were from the same animals used in the initial study (Longley, 1998) a direct comparison of the qRT-PCR data could be made with the competitive RT-PCR data. While some of the data for the individual animals are consistent with the trend observed in the original study (Longley, 1998) only one of the animals examined shows a statistically significant increase in ABCB1 expression post-exposure to sporidesmin, this is animal R6 (P = 0.04). However, R6 was not analysed in the original study (Longley, 1998). The remaining animals in this study
Chapter Three: Quantification of ABCB1 Gene Expression

1.4 Pre-dose
Post-dose

Figure 3.13: Expression of ABCB1 normalised to GAPDH for pooled resistant and pooled susceptible animals. Data is shown as the mean ± SEM.

do not show any significant difference in ABCB1 expression when normalised to GAPDH and do not confirm the results of Longley (1998).

When the data from individual animals is pooled and resistant and susceptible animals are compared both pre- and post- exposure to sporidesmin (Fig. 3.13) it appears that ABCB1 expression in susceptible animals is more responsive to sporidesmin dosing (increase in expression ratio from 0.82 to 1.10) although this is not statistically significant ($P = 0.595$). This suggests that ABCB1 is not differentially expressed as a result of either genetic background or as a result of exposure to the sporidesmin toxin, and contradicts the findings of Longley (1998).

3.2.4 Characterisation of the Competitive RT-PCR product

The original study reported by Longley (1998) showed that ABCB1 was induced approximately 4-fold in resistant animals following exposure to sporidesmin. This could not be replicated using a different technique (qRT-PCR). Therefore attempts were made to characterise the PCR product that was used to quantitate ABCB1
Figure 3.14: Comparison of oligonucleotide primers used for competitive RT-PCR The oligonucleotide primers used in the Longley (1998) study for quantification of ABCB1 gene expression is compared with the cDNA sequences of both human and ovine ABCB1 and ABCB4.

Comparison of the competitive RT-PCR oligonucleotide primer sequences with the human and ovine ABCB1 and ABCB4 cDNA sequences (Fig. 3.14) indicates that the primers may also bind and amplify ABCB4 in addition to ABCB1. The forward primer (Con8) is 100% identical to both human ABCB1 and ABCB4, while there is one base pair mismatch with ovine ABCB1 (no ovine sequence is available for ABCB4). The reverse primer (Con7) also has a one base pair mismatch for both ovine ABCB1 and ABCB4, while a different single base pair mismatch is observed for both human ABCB1 and ABCB4. The expected ovine ABCB1 amplicon would be 452 bp in length and the ovine ABCB4 amplicon would be of an indeterminate size. In human the expected ABCB1 amplicon would be 452 bp in length and the ABCB4 amplicon would be 443 bp in length. The size difference would be virtually indistinguishable in the original competitive RT-PCR experiments (Longley, 1998).

In humans, these two amplicons share a sequence similarity of 84% across the length of the amplicon.

This comparison indicates that the competitive RT-PCR primers (Con7 and Con8) may in fact amplify both ABCB1 and ABCB4 with very similar sized amplicons. It was therefore of interest to determine if ABCB4 or another gene amplified using these primers could be responsible for the 4-fold induction of ABCB1 expression reported in the original study (Longley, 1998).
The presence of a second amplicon was also suggested by the fact that repeated attempts to sequence the gel purified competitive RT-PCR product directly failed to yield any ‘clean’ results.

3.2.4.1 Restriction Endonuclease Digestion

Sequence analysis of the ovine ABCB1 cDNA sequence (NM_001009790) and a 465 bp EST corresponding to ovine ABCB4 (160903CS1802485600001) indicated a HaeIII restriction site present in the ABCB1 sequence that was not present in the ABCB4 sequence.

This differential HaeIII restriction site was used to determine if the competitive RT-PCR primers used by Longley (1998) to quantitate ABCB1 expression were also amplifying ABCB4. To do this eight cDNA samples, 2 each of naïve resistant (R8, R20), naïve susceptible (S4, S17), sporidesmin dosed resistant (R11, R23) and

![Figure 3.15: Agarose gel showing restricted competitive RT-PCR product. Animal numbers are indicated above the gel (R = resistant, S = susceptible; for details refer to Table 2.1 and Table 2.2). NTC = no template PCR control. Undigested (-) and HaeIII digested (+) PCR products generated by the competitive RT-PCR primers con7 and con8. M = size marker (size of marker bands in bp are indicated on the left hand side of the gel).](image-url)
sporidesmin dosed susceptible (S13, S21) were PCR amplified with the competitive RT-PCR primers (Con7 and Con8) under the conditions described by Longley (1998). The PCR product was then digested with HaeIII (as detailed in 2.4.9) and the restricted PCR products were analysed on a 3% agarose gel (Fig. 3.15).

Bioinformatic analysis of the available ovine ABCB1 and ABCB4 sequences confirms the largest restriction fragment for the ABCB1 product would be 202 bp while the largest fragment for the ABCB4 would be at least 241 bp. Comparison with size markers indicates that the undigested product is running at ~441 bp (consistent with the size expected of 452 bp) and that the restricted fragments are ~164, 200, 243 and 274 bp. The total size of the restriction enzyme digested fragments is ~881 bp which is nearly exactly twice that of the undigested product (~441 bp). This analysis confirms that there are at least two amplicons generated by the competitive RT-PCR primers (Con7 and Con8) and that these two amplicons are likely to be ABCB1 and ABCB4.

Two animals, S13 and R11, showed a different restriction pattern to the other animals with the disappearance of the 274 bp fragment. This suggests that these animals have a SNP (single nucleotide polymorphism) in this sequence that generates an additional HaeIII restriction site. This is reasonably likely as HaeIII is a relatively frequent base cutter with a 4 bp recognition sequence (5'-G C C-3'). However, despite the disappearance of the 274 bp fragment, no additional fragments could be detected on the agarose gel. By chance the SNP in these animals could generate a HaeIII site that cleaves the 274 bp fragment into sizes consistent with the remaining three bands, or that it generates a fragment or fragments of less than 100 bp which would be virtually undetectable on the gel system used.

3.2.4.2 Melt Curve Analysis
A mixture of amplicons present in a PCR reaction may be detected using PCR amplification in the presence of the minor groove binding dye SYBR green followed by melt-curve analysis. Any difference in sequence between the two amplicons will result in a difference in the melting temperature (Tm) which could be detected using melt curve analysis.
Amplification with SYBR green followed by melt curve analysis was used in this study to detect the presence of multiple amplicons in a single PCR reaction (as detailed in 2.5.7). The only difference between the conditions used in the competitive RT-PCR experiment and this experiment was the magnesium concentration, the competitive RT-PCR required 1.5 mM MgCl₂ while 3 mM MgCl₂ is incorporated into the qRT-PCR kit. A test amplification detected no extra bands when 3 mM MgCl₂ was used (data not shown).

Thermodynamic analysis of the two possible human ABCB1 and ABCB4 amplicon sequences reveals that the ABCB1 amplicon will have a Tm of approximately 79.3 °C and ABCB4 will have a Tm of 78.8 °C. The corresponding ovine ABCB1 sequence has a predicted Tm of 81.5 °C. Seven cDNA samples were analysed by quantitative RT-PCR with melt curve analysis.

Seven cDNA samples (R8, R20, R23, R11, S17 and S21) were subjected to qRT-PCR in the presence of SYBR green DNA binding dye on the ABI Prism 7500 Sequence Detection System using the Invitrogen Platinum SYBR Green qPCR

![Figure 3.16](image)

**Figure 3.16:** SYBR Green Melt curve analysis for competitive RT-PCR products (showing derivative data). The light grey line with a Tm of 78.4 °C represents the no template control (NTC). The remaining samples have melting temperatures between 86.5 °C and 87.1 °C.
supermix (refer 2.5.7) and con7 / con8 as primers (Fig. 3.16). One sample, R20, was analysed in duplicate to determine the reproducibility of the Tm measurements. This sample when measured in duplicate gave a Tm of 86.5 °C and 86.8 °C suggesting an experimental error of at least ±0.3 °C.

The seven cDNA samples analysed by this method gave Tms ranging from 86.5 – 87.1°C (Fig. 3.16) which suggests that the sequence difference between these amplicons is not sufficient to yield a detectable difference in the Tms of the amplicons. In fact the sequences are so similar as to give a Tm difference of less than 0.6 °C.

Interestingly, one of the samples that gave a Tm of 87.1 °C was R11 which also gave an anomalous result in the restriction digestion analysis (Fig. 3.15). Note that S13 (the other sample that gave anomalous restriction digest results) was not analysed by melt curve analysis. The anomalous restriction digest and slightly higher Tm could indicate a substantial deviation in DNA sequence as compared to the other animals analysed.

Taken together the results indicate that the original competitive RT-PCR primers were amplifying two products and based on the restriction digestion, these products are likely to be ABCB1 and ABCB4. This is supported by the relatively similar melting temperatures for these products, which suggests a high level of sequence similarity.

3.2.5 Northern Blot using competitive RT-PCR product as a probe.

The differential expression of ABCB1 previously observed (Longley, 1998) could not be replicated using an alternative technique (qRT-PCR), and data suggested that at least two amplification products were generated in the original study.

As was hypothesised that the competitive RT-PCR primers were not specifically amplifying ABCB1 (refer 3.2.4) Northern hybridisation was carried out using the competitive RT-PCR product as a probe. This PCR product is likely to contain
Figure 3.17: Northern analysis of the competitive RT-PCR product. (A) Northern Blot showing the two hybridisation bands obtained using the competitive RT-PCR product as a probe. Below the Northern is the ethidium bromide stained gel, the 18S rRNA band intensity was used as a loading control. Samples for which there is a cross beneath the agarose gel were excluded from the analysis in (B) due to abnormal hybridisation. (B) Combined data showing the expression of Con7/8 hybridised gene products relative to 18S rRNA for resistant and susceptible animals, pre- and post-exposure to sporidesmin. Data are shown as the mean ± SD. Animals S10, R13, R14 and S15 (as indicated in (A)) have been excluded due to abnormal hybridisation. * = P < 0.05

multiple amplicons and therefore a probe generated with this product should be capable of detecting multiple transcripts on the Northern blot. It was hoped that using this relatively non-specific probe the transcript that was described to be induced 4-fold in the original study could be identified (Longley, 1998).
Unfortunately, the RNA that was used in the competitive RT-PCR experiments, and also here for the TaqMan qRT-PCR experiments was limiting and not enough sample was available for Northern blotting. Therefore RNA samples were prepared from tissues harvested for this study (refer Tables 2.1 and 2.2). This unfortunately means that the results are not directly comparable with the original study, particularly as the post-exposure liver samples were harvested 24 h post-exposure as compared with 2 weeks as in the original study.

The ABCB sub-family is unique in that it contains four full transporters (ABCB1, ABCB4, ABCB5 and ABCB11) and seven half transporters (ABCB2, ABCB3, ABCB6, ABCB7, ABCB8, ABCB9 and ABCB10) (Dean et al., 2001). As these sub-families are defined based on sequence similarity it is likely that the probe would be able to cross hybridise to other members of the sub-family and would facilitate detection of a differentially expressed transcript.

Northern analysis was performed as detailed in 2.5.5, the autoradiographs were scanned and quantified as detailed in 2.5.5.4 and the size of transcripts estimated by comparison with the 0.5 - 10 kb RNA Ladder (Invitrogen) size ladder. The scanned Northern blot is shown in Fig. 3.17A. Band A was determined to be 3.95 kb and band B 1.77 kb. The larger band is consistent in size with both ABCB1 (4.9 kb in human, 4.4 kb in sheep) and ABCB4 (3.8 kb in human). However, the smaller band (band B) is consistent in size with a half transporter rather than a full transporter. Band b could also represent unannotated alternative transcripts of either ABCB1 or ABCB4.

Quantification of the Northern blot was achieved by determining the total absorbance units of each band from the scanned Northern blot, and normalising this data to the 18S rRNA band from the scanned ethidium bromide stained gel. This normalisation attempts to correct for any variations in RNA loading of the Northern blot as the level of 18S rRNA is assumed to be constitutive and it can therefore act as an endogenous control.
Chapter Three: Quantification of ABCB1 Gene Expression

For intact RNA the intensity of the ethidium bromide stained 28S rRNA band should be approximately twice that of the 18S rRNA band. The ethidium bromide stained blot in Fig. 3.17 clearly shows the 18S band with a higher intensity than the 28S band. On this particular Northern blot this phenomenon is partially an artefact due to migration of the ethidium bromide dye. Nevertheless it is clear that some degradation of the RNA sample has occurred, and this may distort the quantification. However, much of the variation seen in Figure 3.17B may be attributable to irregular hybridisation or RNA degradation. In particular animals when S10, R13, R14 and S15 were excluded from the analysis (as they all show dark hybridisation smears), then the variation was significantly reduced (Fig. 3.17C). In this case there is still no significant difference in the expression of band A ($P = 0.054$), but band B is differentially expressed as determined by one way ANOVA ($P = 0.002$). Using a Tukey’s HSD post-hoc test significant differences were observed between the four groups. Naïve susceptible animals and dosed resistant animals express band B more highly than naïve resistant animals. This is broadly consistent with the findings in the original study (Longley, 1998), however, the fold difference in expression reported here is ~1.2 fold rather than the 4 fold previously reported (Longley, 1998).
3.3 DISCUSSION

The objective of the work presented in this chapter was to verify the differential expression of ABCB1 reported by Longley (1998) prior to exploring the biochemical mechanism of gene induction. In this study ABCB1 expression was quantified using a qRT-PCR assay with TaqMan chemistry. The results from the TaqMan assay failed to verify the induction of ABCB1 expression. There were several possible explanations for this.

One of the biggest confounding factors with this experiment was the design of the TaqMan qRT-PCR assay. Due to the high degree of sequence similarity between ABCB1 and ABCB4, the TaqMan qRT-PCR assay design was limited to two areas of the ABCB1 cDNA sequence, namely the leader and the linker regions. The linker region was not suitable for assay design and as such the assay used was designed to the extreme 5' region of the coding sequence. Because the total transcript size is 4.4 kb and the reverse transcription reaction was primed with oligo(dT), the reverse transcriptase reaction would have to be optimal in order to facilitate accurate quantification in this region. Consistent with this hypothesis is the very high C\textsubscript{T} values obtained for ABCB1 (~36 cycles) which is near to the limit of detection (40 cycles) and may have made true differences in ABCB1 expression difficult to detect. The co-variance of ABCB1 expression and RNA concentration post-DNase treatment is also indicative of poor amplification due to low reverse transcriptase efficiency. Unfortunately, the endogenous control, GAPDH, would not be able account for variation in reverse transcriptase efficiency as GAPDH has a relatively short mRNA transcript (~ 1.2 kb) and therefore the reverse transcription of this transcript should be more efficient than a longer transcript such as ABCB1.

Comparison of the 3' UTR sequence of the ABCB1 and ABCB4 genes shows that they are more sequence divergent. So in hindsight using the 3' UTR sequence may have yielded a more specific sequence for assay design while avoiding some of the issues with quantifying the 5' region of the gene. However, this was not done as the UTR sequences have relaxed functional constraint and generally contain many polymorphisms. These polymorphisms could cause amplification biases due to primer mismatches.
Another confounding factor in this study was the use of GAPDH as the endogenous control gene to normalise ABCB1 expression. GAPDH primers were designed to the only ovine GAPDH sequence that was available at the time, but subsequently sequence has become available which indicates that there are at least 2 polymorphisms in the GAPDH F primer binding site. This is consistent with the very poor amplification efficiency obtained for GAPDH (~61%) and this would have a significant impact on the quantification of GAPDH and its utility as an endogenous control. Additionally, the use of GAPDH necessitated DNAse treatment of the RNA samples prior to quantification, as amplicons generated from genomic DNA and cDNA could not be distinguished (suggesting the absence of an intron in the ovine GAPDH that is present in the human GAPDH gene). The DNAse treatment also avoided the problem of pseudo-gene amplification which is known to be a problem when using GAPDH as an endogenous control, as mammalian genomes contain a number of GAPDH processed pseudo-genes (Harper et al., 2003). Comparison of the human GAPDH RefSeq (NM_002046) with the bovine genome sequence (Btau_2.0) using BLAT (Kent, 2002) reveals at least 7 significant hits, of which 4 appear to be processed pseudogenes.

However, the DNAse treatment caused a ~10-fold decrease in RNA concentration as assessed by UV absorbance and while there was not enough RNA remaining to examine it using agarose gel electrophoresis, it is unlikely that the genomic contamination was sufficient to account for a ~10-fold decrease in the nucleic acid concentrations measured following DNAse treatment. It is therefore likely that some of the RNA was degraded by the DNAse (despite being certified as RNAse free). This degradation is unlikely to have occurred in a predictable or reproducible way and may have contributed to the variation seen in the experiments presented here. The measurements of RNA concentration could be problematic as due to limited availability of sample no duplicates measurements were made so there is no available estimate of error associated with these measurements.

It is known that GAPDH expression is not uniform across all tissues and cell types. It is therefore possible that the expression of GAPDH is varying along with ABCB1 and is masking the true differences in expression. This is supported by the finding
that ABCB1 expression and GAPDH expression are highly correlated (Pearson's correlation coefficient = 0.830, P = 0.001). It is usually recommended that the experiment be performed with a minimum of 2 housekeeping genes, and in this case it would have been useful to use β-actin which was the endogenous control used in the original study (Longley, 1998).

Additionally, GAPDH is a medium abundance transcript and ideally copy numbers of the individual housekeeping gene should be in a similar range to that of the target gene, facilitating comparative quantification. So a rarer transcript such as hypoxanthine guanine phosphoribosyltransferase or porphobilinogen deaminase may be more appropriate for quantification of ABCB1 expression by qRT-PCR. Ideally, when examining disease states as compared to normal tissues where the metabolism of the tissues could be radically different, several endogenous controls would be tested and validated (Kim and Kim, 2003; Szabo et al., 2004; Herrera et al., 2005); however, for practical reasons such as cost and time to develop these assays in sheep, this was not done.

While it is possible that variation in GAPDH expression may have absorbed some of the differences reported by Longley (1998), it is unlikely that a four fold difference in expression would be masked by differences in GAPDH expression.

Despite these confounding factors, it is likely that the quantification assay used in the original study (competitive RT-PCR) (Longley, 1998) was not specific for ABCB1 and that another gene is responsible for the differential expression observed.

Bioinformatic analysis suggested that the sequence conservation in the region targeted by competitive RT-PCR in the original study is high, in particular between ABCB1 and ABCB4. As the TaqMan qRT-PCR assay was designed to amplify ABCB1 specifically, it was thought that the discrepancies between the qRT-PCR experiment and the competitive RT-PCR experiment could be due to the amplification of a different family member by the competitive RT-PCR primers. To test this hypothesis PCR was carried out using these competitive RT-PCR primers (con7 and con8). Sequencing data suggested that there were more than one PCR product obtained using the competitive RT-PCR primers, by the inability to obtain
Chapter Three: Quantification of ABCB1 Gene Expression

'clean' sequence using direct sequencing. Further bioinformatic analysis showed that the oligonucleotide primers used in the original experiment were likely to bind and amplify ABCB4 as well as ABCB1. An attempt was made to clone and sequence the competitive RT-PCR amplicons and preliminary evidence was obtained to show the presence of a second (minor) amplicon in addition to ABCB1. However, the identity of that second amplicon could not be confirmed (data not shown).

Restriction analysis with the enzyme HaeIII confirmed the presence of at least two amplicons within the PCR products. To determine the number, similarity and possibly the relative abundance of the different PCR products qRT-PCR was carried out in the presence of SYBR green. Some variation was seen in the melting curves, and this variation was consistent with differences in the restriction digests. However, based on the melt curve analyses the two amplicons have a very similar sequence (as evidenced by the similar Tms obtained).

A final attempt was made to confirm the differential expression observed by Longley (1998) by using Northern analysis with the radio-labelled competitive RT-PCR product as a probe. The competitive RT-PCR product that was hypothesised to contain at least two amplicons was deliberately used as a probe for the Northern analysis. Due to the sequence similarity between the ABCB mRNA sequences it is likely that a Northern probe in this region will cross hybridise with other ABCB family members, and it was hoped that the transcript responsible for the 4-fold levels of induction observed in the original experiments (Longley, 1998) would be detectable using such a heterogeneous probe.

Two bands were observed on the Northern blot. The first, band A, was 3.95 kb in size and is broadly consistent with a full sized transporter. The second, band B, was 1.77 kb in size and broadly consistent with a half transporter, suggesting that the competitive RT-PCR probe has a sequence similar enough to cross hybridise with a sequence other than ABCB1 and ABCB4 (as these are both full transporters). The sizes obtained for the hybridised bands are not consistent with that expected, i.e. the size of the full length transcript of ovine ABCB1 is 4.4 kb (Longley, 1998; Longley et al., 1999) and in human the sizes of the full ABCB transporters range from 2.9 kb to 5.8 kb and the half transporters from 2.4 kb to 3.9 kb. There is also evidence of
partial degradation of some of the RNA samples as seen by smearing of the RNA on the Northern blot, which can complicate both the sizing and quantification of the hybridisation bands.

Using Northern hybridisation, band B was found to be differentially expressed ~1.2 fold more highly in naïve susceptible and sporidesmin dosed resistant animals as compared with naïve resistant animals. Additionally sporidesmin dosed resistant animals were shown to express band B 1.2-fold more highly than sporidesmin dosed susceptible animals. This pattern of differential expression is only broadly consistent with that previously reported (Longley, 1998) as no differential expression between naïve resistant and susceptible animals was reported and we only observed a 1.2-fold induction of band B, versus a 4-fold induction previously reported. Therefore this observation does not confirm previously reported data.

Additionally, the size of band B is 1.77 kb which is slightly smaller than annotated for other ABCB family members (2.5 – 5.8 kb) and is also inconsistent with any alternative transcripts for ABCB1 reported in the ECgene database (Kim et al., 2005a; Kim et al., 2005b) or in EMBL-EBI alternative splicing database project (Stamm et al., 2006). Therefore it is difficult to speculate what this differential expression may mean both in terms of the data previously presented concerning ABCB1 expression (Longley, 1998) but also in terms of which gene is being differentially expressed, and whether a 1.2 fold induction is sufficient to mediate a significant biological response.

It is important to note that this experiment is not conclusive as it utilised a cross hybridising radio-labelled probe to detect gene expression. It is therefore possible that the true differentially expressed transcript is expressed at a low level and the Northern signal is being ‘swamped’ by other transcripts. A more specific method for detecting ABCB1 and ABCB4 could be used, such as a qRT-PCR assay designed to the 3'UTR regions or alternatively using radio-labelled oligonucleotides to probe the expression of these transcripts on a Northern blot. Oligonucleotide probes tend to be shorter and can be designed to be more specific than the probes used in this study.
Chapter Three: Quantification of ABCB1 Gene Expression

Taken together these data indicate that ABCB1 is not differentially expressed in livers of resistant and susceptible selection line sheep, contradicting previous work (Longley, 1998). While it is possible that ABCB4 was being amplified in addition to ABCB1 more work must be done to confirm this.
Chapter Four

SECTION A: Identification of the Closest Mammalian Ortholog of the *S. cerevisiae* ABC transporter PDR5.

4.1 INTRODUCTION

4.1.1 ABC Transporters

The ABC (ATP-binding cassette) transporter super-family has been described in 3.1.1. The proteins within this family all have two common features: the first is the ABC cassette which consists of several conserved protein motifs (refer 3.1.1.2) and the second is a series of transmembrane segments (usually 6 transmembrane segments for every ABC cassette). The ABC cassette protein sequence is particularly well conserved across species and is commonly used for the classification of new ABC transporters as well as for phylogenetic analysis of this super-family (Sheps *et al.*, 2004).

As previously discussed (refer 3.1.2), a putative role for ABC transporters in conferring resistance to FE was first indicated by work done in *S. cerevisiae* (Bissinger and Kuchler, 1994). In that study a yeast ABC transporter, PDR5, was found to be capable of conferring resistance to sporidesmin in *S. cerevisiae*. Subsequently, a putative ovine ortholog of PDR5, ABCB1, was investigated as a candidate gene for conferring resistance to FE in sheep but no genetic evidence was found (Longley, 1998, also chapter three).
4.1.1.1 PDR5 (Pliotropic drug resistance protein 5, STS1, suppressor of sporidesmin toxicity I)

The drug efflux profile of PDR5 is very pleiotropic and covers structurally dissimilar antifungal and anti-tumour drugs, hormones, and ionophores (Bissinger and Kuchler, 1994; Kolaczkowski et al., 1996). The diversity of compounds transported by PDR5 is characteristic of the ABC super-family as substrate preference appears to be defined by size and hydrophobicity (Golin et al., 2003). Recent analysis of the human ABCC1 protein indicated that charged or polar amino acids in the transmembrane helices close to the cytoplasmic interface affect the substrate specificity of these efflux pumps (Haimeur et al., 2004; Koike et al., 2004; Zhang et al., 2004a). There is, however, no non-empirical way to determine if a compound is a substrate for a particular ABC transporter.

PDR5 (Fig. 4.1) is a full ABC transporter with 12 α-helical transmembrane domains and two ABC domains. PDR5 also has an uncommon domain organisation in that the ABC domain is located toward the N-terminus of the protein (ABC-TM6)2 in the reverse orientation to the transmembrane domains. This is contrasted with the more common domain organisation where the ABC domain lies toward the C-terminus of the protein (TM6-ABC)2. As such PDR5 falls into a smaller class of ABC transporters called the 'reverse' transporters.

![Figure 4.1: Schematic of the S. cerevisiae ABC transporter PDR5, illustrating the location of ABC cassettes (NBD1, NBD2) and membrane topology (picture adapted from Egner et al. (1998)).](image-url)
The ABC superfamily is very well studied in humans and rodents, and several members have been implicated in conferring resistance to chemotherapeutic drugs (reviewed in Leslie et al., 2005). Mining of the human genome has found 48 ABC transporter genes, of which only 16 have a known function or substrate and 14 are associated with disease (Borst and Elferink, 2002). Of the 48 human ABC transporters, 38 are conserved across all vertebrate genomes. Among the vertebrate species examined, rodents have been found to have the highest number of species specific ABC transporters, with seven transporters not found in hominid species (Dean and Annilo, 2005).

There is no conclusive evidence for the mechanism controlling the coupling of ATP hydrolysis to substrate transport but it is known that two ABC domains interact to form an interface (Locher, 2004). The supposition that an interface is formed by the ABC domains is indicated by the fact that the half ABC transporters (with only one ABC domain) must form higher oligomers such as dimers or tetramers for functionality (Xu et al., 2004). The importance of the ABC domain for the function of the transporter accounts for its exceptionally high sequence conservation and favours the use of this sequence in phylogenetic analyses.

Interestingly, despite the evolutionary divergence between yeast and higher mammals there is some evidence for functional conservation of ABC transporters. For example: the human ABCB7 and ABCB10 proteins and the yeast ATM1, MDL1 and MDL2 proteins (Sheps et al., 2004) all localise to the mitochondria and are involved in iron metabolism (Dean et al., 2001; Lill and Kispal, 2001; Chloupkova et al., 2003). This functional constraint is observed despite the gene family having undergone significant levels of evolution including gene duplication.

4.1.2 Functional inference from comparative genomics and phylogenetics

The aim of this section of work was to identify a mammalian ortholog of PDR5. Several lines of enquiry taken together can point to orthology, but uncertainties are common as inferring function from sequence identity is problematic and complex. Sequence homology can provide evidence in the prediction of molecular function (Ponting, 2001). However, the greatest confidence in functional assignment should be placed in orthologous pairs (genes which have arisen from speciation events) as
opposed to paralogs (a homolog that performs different but related functions within the same organism) (Ponting, 2001). Orthology and paralogy can only be accurately assigned in genomes that have been fully sequenced (Ponting, 2001) and can be complicated by large scale gene duplications, particularly occurring early in the vertebrate lineage (Holland, 1999). Defining orthology is further complicated by multi-domain proteins where domain architecture must be considered.

One of the most convincing methods for predicting functional homology is the consideration of structural similarity and protein fold conservation (Eisenstein et al., 2000; Shapiro and Harris, 2000). As structural data are limited for many species and proteins, in particular transmembrane proteins, any functional assignment based on sequence similarity should be interpreted with caution. In addition all available information including known function, structure, domain organisation, secondary structure (including membrane topology), substrate specificity and sub-cellular localisation should be taken into account. In this way a phylogenetic relationship can be used to generate a biological hypothesis for further testing (Ponting, 2001).

4.1.3 Comparative Mapping

Despite the excellent coverage of the ovine genome linkage map (Maddox et al., 2001) the information it contains is a long way behind that of fully sequenced genomes such as human and mouse. Comparative mapping provides a method for exploiting the ‘information-rich’ genome maps of mouse and human to ‘information-poor’ genomes of farmed ruminants such as sheep (Burt, 2002).

Comparative mapping takes advantage of the extensive conservation of synteny and gene order between two species (Echard et al., 1994). This phenomenon can be extended beyond comparison of closely related species and can give important information about evolutionary relationships (O’Brien et al., 1993). The size of conserved syntenic blocks varies in accordance with the evolutionary distance between the animal species being compared. These syntenic blocks can then be used to localise loci of interest in an unmapped species provided enough ‘anchoring’ or surrounding markers have been mapped.
For comparative mapping only type I markers (gene associated markers) are useful as type II (anonymous highly polymorphic markers such as microsatellites) are thought to be poorly conserved between species. Comparative mapping in sheep has been further facilitated by the recent release of the bovine genome.

4.1.4 Aims and Rationale

Despite the evolutionary distance between yeast and higher mammals there is evidence supporting the conservation of substrate specificity for some ABC transporters (Dean et al., 2001; Lill and Kispal, 2001; Chloupkova et al., 2003). Given that sporidesmin is known to be a substrate for the yeast PDR5 protein (Bissinger and Kuchler, 1994) the aim of this study was to identify a mammalian ortholog of PDR5 (section A) and to assess the involvement of this ortholog in conferring resistance to sporidesmin in sheep (section B).
4.2 RESULTS

4.2.1 Identifying the mammalian ortholog of the yeast PDR5 protein.

4.2.1.1 Retrieval of the ABC domain sequences from the species of interest

To accurately detect possible mammalian orthologs of the yeast PDR5 protein a complete phylogenetic analysis using all of the ABC transporters from five fully-sequenced organisms was undertaken. The five organisms selected for analysis were S. cerevisiae, Mus musculus, Homo sapiens, Caenorhabditis elegans and Drosophila melanogaster. These latter two organisms were included to bridge the evolutionary gap between S. cerevisiae and mammals facilitating the accurate grouping of yeast and human ABC sequences.

The ABC-transporter protein sequences were extracted from the organism specific RefSeq protein databases with PSI-BLAST (Altschul et al., 1997) using the PDR5 reference sequence as the query (NP_014796). PSI-BLAST is a position specific blast algorithm capable of detecting distant homologs. The PSI-BLAST algorithm was iterated until no further significant homologies were found. Using this algorithm a total of 232 proteins were found; 62 homologs from C. elegans, 51 from D. melanogaster, 45 from H. sapiens, 44 from M. musculus and 30 from S. cerevisiae (Table 4.1).

The 232 ABC-transporter sequences were extracted in FASTA format and imported into the Wisconsin Package v 10.3 (Accelrys Inc, San Diego, CA). The HmmerSearch algorithm (Eddy, 1998) was then used to isolate the ABC domain sequences using the ABC transporter Pfam motif (PF00005). This algorithm isolates only the ~180 amino acid ABC cassette, which is the most highly conserved region of the protein and facilitates comparisons across very divergent species. Lastly, the

| Table 4.1: The number of sequences and domains that were extracted from each species using PDR5 (NA_014796) as a query sequence. |
|-----------------|---------------|---------------|-----------------|-----------------|---------------|
|                 | S. cerevisiae | C. elegans    | D. melanogaster | M. musculus     | H. sapiens    |
| RefSeqs         | 30            | 62            | 51              | 44              | 45            |
| Domains         | 53            | 98            | 81              | 68              | 76            |
| Full Transporters| 23            | 36            | 30              | 24              | 31            |
| Half Transporters| 7             | 26            | 21              | 20              | 14            |
domain architecture of these proteins was confirmed by comparison with the Pfam database (Bateman et al., 2004).

Using HmmerSearch a total of 376 ABC cassette protein sequences were extracted from the 232 protein sequences; 98 homologs from C. elegans, 81 from D. melanogaster, 76 from H. sapiens, 68 from M. musculus and 53 from S. cerevisiae (Table 4.1). All phylogenetic analyses were done using these ABC cassette protein sequences only.

The final step in the preparation of the dataset was to remove sequences that did not group with an established sub-family of mammalian transporters (Dean et al., 2001). This step was important as the focus of this analysis was to examine the yeast ABC-transporter sequences in the context of the mammalian sub-families. Inclusion of outlier sequences would not only complicate the interpretation of the phylogeny but also compound problems associated with long-branch attraction (Bergsten, 2005).

To identify outlier sequences the human and mouse domain sequences were aligned using CLUSTAL X version 1.81 (Thompson et al., 1997) to create a phylogenetic framework. The C. elegans, D. melanogaster and S. cerevisiae sequences were then individually aligned against this framework to identify outlier sequences that did not group with a defined mammalian sub-family. Using this approach 24 of the 376 domain sequences did not group with a mammalian sub-family. These 24 sequences were excluded from further analyses.

4.2.1.2 Generation of the Neighbour-Joining Phylogenetic tree
The resulting set of 352 domain sequences was then aligned using CLUSTAL X version 1.81 (Thompson et al., 1997). The resulting alignments were manually edited using Se-Al v2.0a11 (Rambaut, 1996) to remove flanking un-aligned sequences. The alignments were imported into PAUP* 4.0 (beta 10 version) (Swofford, 2002) for phylogenetic analysis.

Due to the large number of sequences in this initial dataset neighbour-joining analysis was used as a first pass approach. Neighbour joining analysis is very useful for large datasets as it is a distance based method for estimating the level of
dissimilarity between two sequences and is much less computationally intensive than character based phylogenetic methods such as maximum parsimony (Brinkman and Leipe, 2001).

Neighbour-joining phylogeny was constructed as detailed in 2.7.3.2.

4.2.1.3 Tree topology and orthologous groups.
A schematic version of the neighbour-joining tree is shown in Fig. 4.2 and the individual branches are shown in Fig. 4.3 (sub-families A through G). While the exact architecture of the tree is intricate, there were eight major groupings of sequences. Six of these groupings corresponded directly with six of the mammalian ABC sub-families (A, B, D, E, F and G). The remaining mammalian sub-family, ABCC, was split into 2 clades with strong bootstrap support (91%). The first clade contained the ABCC N-terminal domain sequences and the second contained the C-terminal sequences. This split was consistent with published phylogenies for H. sapiens, D. melanogaster (Dean et al., 2001) and C. elegans (Sheps et al., 2004).

In general, the neighbour-joining tree shown here agrees with those previously published. For instance the ABC (C-terminal) sequences were more closely related to the ABCB sub-family of proteins than the ABCC (N-terminal). This relationship had also been previously observed (Dean et al., 2001; Sheps et al., 2004) and may indicate ancient gene duplication and subsequent divergence. However, there are slight differences in the topology of the tree reported here. Specifically, the ABCA and ABCG sub-families were found to be more similar to each other than to any of the other sub-families, similar to that observed when just the human sequences were examined (Dean et al., 2001). This differs from previously published phylogeny for S. cerevisiae and C. elegans (Sheps et al., 2004) where the ABCA and ABCG sub-families formed monophyletic sister groups.
Figure 4.2: Simplified neighbour-joining phylogenetic tree of ABC cassette domain sequences from *H. sapiens*, *M. Musculus*, *D. melanogaster*, *C. elegans* and *S. cerevisiae*. 352 ABC domain sequences were aligned using Clustal X, and a neighbour joining tree constructed in PAUP*4.0b10. Bootstrapping was used to determine the relative support for various branches of the tree (100 replicates) and nodes with less than 50% support were collapsed to form polytomies.
Chapter Four (Section A): Identification of an Ortholog of the S. cerevisiae ABC transporter PDR5

[Diagram of phylogenetic trees for ABCA, ABCB, and ABCC proteins]
Figure 4.3: (including facing page). Detailed neighbour-joining tree of ABC domain sequences, generated using 352 ABC cassette sequences from H. sapiens, M. Musculus, D. melanogaster, C. elegans and S. cerevisiae. The overall tree topology is shown in Fig. 4.2, and the individual branches are shown here. The genome of origin for each protein is indicated by prefixes before each gene name, according to the following scheme HSA, H. sapiens; MMU, M. Musculus; DME, D. melanogaster, CEL, C. elegans; SCE, S. cerevisiae. The sequence name is then given followed by the NCBI accession number and the position of the ABC domain within the full transporter sequence. The two PDR5 ABC domain sequences are shaded in blue.
All of the eight clades (Fig. 4.3) contained representatives from all of the 5 species sampled, except the ABCE clade. The ABCE clade is unique as it contained ABC domain sequences from all species except *D. melanogaster*. It is known that the ABC family has undergone large scale duplication and divergence events since the split of the *D. melanogaster* lineage (Dean *et al.*, 2001), providing a possible explanation for this observation.

Surprisingly, very few orthologous groups were found in this phylogeny. With the exception of the rodent specific transporters, all of the human and mouse sequences formed orthologous pairs. However, only 11 orthologous groups contained a non-mammalian species. Of the 11 orthologous groups six were found within the ABCF sub-family. Each of these six orthologous groups contained a representative sequence from *C. elegans*, two of the six groups contained a sequence from *S. cerevisiae*, and the remaining four a sequence from *D. melanogaster*.

### 4.2.1.4 The ABCG sub-family.

Based on the phylogenetic analyses presented in Fig. 4.2 and Fig. 4.3, both the N- and C-terminal ABC domains of the yeast PDR5 protein group with the mammalian ABCG family. The sequences that make up this ABCG sub-family are listed in Table 4.2.

In addition to the two yeast PDR5 ABC domain sequences, a further 10 yeast proteins are represented in this cluster. Interestingly, all 49 domain sequences in this cluster correspond with half transporters with the exception of nine of the ten proteins from *S. cerevisiae*. These nine *S. cerevisiae* transporters are all full transporters containing two ABC cassettes.

The 49 ABC cassettes that had been determined to be part of the ABCG sub-family (Fig. 4.2) were realigned as previously detailed and were analysed using both neighbour-joining analysis and the more rigorous method of maximum parsimony (refer 2.7.3.3). Neighbour-joining analysis (Fig. 4.4) and maximum parsimony analysis (Fig. 4.5) was carried out as detailed in 2.7.3. Bootstrapping (Felsenstein, 1985) was used to determine the relative support for the various branches of the tree.
(1000 replicates), and nodes with less than 50% support were collapsed to form polytomies.

Phylogenies were rooted using the ABCA1 sequence. ABCA1 was selected as an out-group based on the results presented in Fig. 4.2 where the ABCA sub-family was found to be a sister clade to the ABCG sub-family.

### Table 4.2: Sequences that grouped with the mammalian ABCG sub-family. Name, sequence accession numbers and location of ABC domain (as determined by HmmSearch (Eddy, 1998)) for sequences from *S. cerevisiae, C. elegans, D. melanogaster, M. musculus* and *H. sapiens* that group together according to neighbour joining analysis (Fig. 4.3).

<table>
<thead>
<tr>
<th>Name</th>
<th>Accession</th>
<th>ABC Dom</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H. sapiens</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCG1</td>
<td>NP_058198</td>
<td>111-293</td>
</tr>
<tr>
<td>ABCG2</td>
<td>NP_004818</td>
<td>77-262</td>
</tr>
<tr>
<td>ABCG4</td>
<td>NP_071452</td>
<td>95-277</td>
</tr>
<tr>
<td>ABCG5</td>
<td>NP_071881</td>
<td>79-269</td>
</tr>
<tr>
<td>ABCG8</td>
<td>NP_071882</td>
<td>98-289</td>
</tr>
<tr>
<td><strong>M. Musculus</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCG1</td>
<td>NP_033723</td>
<td>379-560</td>
</tr>
<tr>
<td>ABCG2</td>
<td>NP_036050</td>
<td>466-643</td>
</tr>
<tr>
<td>ABCG3</td>
<td>NP_084515</td>
<td>372-538</td>
</tr>
<tr>
<td>ABCG4</td>
<td>NP_620405</td>
<td>111-293</td>
</tr>
<tr>
<td>ABCG5</td>
<td>NP_114090</td>
<td>706-912</td>
</tr>
<tr>
<td>ABCG8</td>
<td>NP_080456</td>
<td>99-290</td>
</tr>
<tr>
<td><strong>D. melanogaster</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG11069</td>
<td>NP_651307</td>
<td>47-227</td>
</tr>
<tr>
<td>CG17632</td>
<td>NP_523824</td>
<td>59-237</td>
</tr>
<tr>
<td>CG17646</td>
<td>NP_608618</td>
<td>141-327</td>
</tr>
<tr>
<td>CG2759</td>
<td>NP_476787</td>
<td>123-317</td>
</tr>
<tr>
<td>CG3164</td>
<td>NP_722605</td>
<td>78-262</td>
</tr>
<tr>
<td>CG3327</td>
<td>NP_524648</td>
<td>215-403</td>
</tr>
<tr>
<td>CG4314</td>
<td>NP_524108</td>
<td>101-292</td>
</tr>
<tr>
<td>CG5853</td>
<td>NP_788019</td>
<td>136-320</td>
</tr>
<tr>
<td>CG9663</td>
<td>NP_608760</td>
<td>180-364</td>
</tr>
<tr>
<td>CG9664</td>
<td>NP_722889</td>
<td>53-237</td>
</tr>
<tr>
<td><strong>C. elegans</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D571</td>
<td>NP_494495</td>
<td>86-277</td>
</tr>
<tr>
<td>3F238</td>
<td>NP_497825</td>
<td>67-253</td>
</tr>
<tr>
<td>3H174</td>
<td>NP_498322</td>
<td>20-213</td>
</tr>
<tr>
<td>3H232</td>
<td>NP_498332</td>
<td>92-284</td>
</tr>
<tr>
<td>3H561</td>
<td>NP_498425</td>
<td>56-245</td>
</tr>
<tr>
<td>3M311</td>
<td>NP_499442</td>
<td>1-180</td>
</tr>
<tr>
<td>3N478</td>
<td>NP_499616</td>
<td>132-334</td>
</tr>
<tr>
<td>4M237</td>
<td>NP_502164</td>
<td>61-250</td>
</tr>
<tr>
<td>4N103</td>
<td>NP_502352</td>
<td>129-320</td>
</tr>
<tr>
<td><strong>S. cerevisiae</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADP1</td>
<td>NP_009937</td>
<td>416-607</td>
</tr>
<tr>
<td>AUS1P</td>
<td>NP_014654</td>
<td>59-249</td>
</tr>
<tr>
<td>AUS1P</td>
<td>NP_014654</td>
<td>775-954</td>
</tr>
<tr>
<td>PDR10</td>
<td>NP_014973</td>
<td>206-406</td>
</tr>
<tr>
<td>PDR10</td>
<td>NP_014973</td>
<td>952-1141</td>
</tr>
<tr>
<td>PDR11</td>
<td>NP_012252</td>
<td>59-247</td>
</tr>
<tr>
<td>PDR11</td>
<td>NP_012252</td>
<td>775-954</td>
</tr>
<tr>
<td>PDR12</td>
<td>NP_015267</td>
<td>174-373</td>
</tr>
<tr>
<td>PDR12</td>
<td>NP_015267</td>
<td>871-1060</td>
</tr>
<tr>
<td>PDR15</td>
<td>NP_010694</td>
<td>196-386</td>
</tr>
<tr>
<td>PDR15</td>
<td>NP_010694</td>
<td>913-1102</td>
</tr>
<tr>
<td>PDR5</td>
<td>NP_014796</td>
<td>186-386</td>
</tr>
<tr>
<td>PDR5</td>
<td>NP_014796</td>
<td>898-1087</td>
</tr>
<tr>
<td>SNQ2</td>
<td>NP_10294</td>
<td>186-386</td>
</tr>
<tr>
<td>SNQ2</td>
<td>NP_10294</td>
<td>882-1070</td>
</tr>
<tr>
<td>YNR070W</td>
<td>NP_014468</td>
<td>56-257</td>
</tr>
<tr>
<td>YNR070W</td>
<td>NP_014468</td>
<td>759-946</td>
</tr>
<tr>
<td>YO1075C</td>
<td>NP_014567</td>
<td>55-263</td>
</tr>
<tr>
<td>YO1075C</td>
<td>NP_014567</td>
<td>720-916</td>
</tr>
</tbody>
</table>
Chapter Four (Section A): Identification of an Ortholog of the *S. cerevisiae* ABC transporter PDR5

Figure 4.4: Neighbour-joining analysis of the ABCG sub-family of ABC domains. The relationship between the ATP-binding domains of 47 ABC sequences from *H. sapiens* (HSA) (5 sequences), *M. musculus* (MMU) (6 sequences), *D. melanogaster* (DME) (15 sequences), *C. elegans* (CEL) (11 sequences) and *S. cerevisiae* (SCE) (10 sequences) is depicted. Alignment was generated with CLUSTALW, edited manually using SeqAI and subjected to neighbour-joining analysis using PAUP*. Bootstrap values (in percentage) out of 1000 replications are indicated. The two PDR5 ABC domain sequences are highlighted in blue.
Figure 4.5: Maximum parsimony analysis of the ABCG sub-family of ABC domains. The relationship between the ATP-binding domains of 47 ABC sequences from *H. sapiens* (HSA) (5 sequences), *M. musculus* (MMU) (6 sequences), *D. melanogaster* (DME) (15 sequences), *C. elegans* (CEL) (11 sequences) and *S. cerevisiae* (SCE) (10 sequences) is depicted. Alignment was generated with CLUSTALW, edited manually using SeqAI and subjected to maximum parsimony using PAUP*. Bootstrap values (in percentage) out of 100 replications are indicated. The two PDR5 ABC domain sequences are highlighted in blue.
Overall the neighbour-joining tree topology was consistent with that previously obtained (Fig. 4.3) although there are some subtle differences. Specifically, several minor changes in the bootstrap support for particular branches were detected while the overall relationships were upheld. Additionally, DME_CG4314 has gone from being a paraphyletic sister-group to an out-group within the first cluster. HSA_ABCG5 and MMU_ABCG5 were initially reported to be an orthologous pair, but in Fig. 4.4 these two sequences have been joined by DME_CG11069 to form an orthologous group with a moderate (55%) bootstrap value. Lastly, in the cluster containing the ABCG1 and ABCG4 domain sequences, DME_CG17646 and DME_CG5853 have been included as a sister group to the mammalian ABCG1 proteins.

The neighbour-joining tree (Fig. 4.4) and the maximum parsimony tree (Fig. 4.5) also have very similar topology with some subtle differences, specifically *C. elegans* 3H232 is most closely related to 3M237 in the neighbour-joining analysis but using maximum parsimony it is most closely related to 3F238 with good bootstrap values (86%). This is also observed with *S. cerevisiae* PDR15 which using neighbour-joining analysis is reported to be most closely related to *S. cerevisiae* PDR5. However using maximum parsimony analysis it is reported to be most closely related to *S. cerevisiae* PDR15 with PDR5 as an sister group. Also the relationship between *D. melanogaster* CG17646 and CG5853 has been demoted to sister groups and the *D. melanogaster* CG3164 is indicated as the progenitor of the mammalian ABCG1/G4 group in neighbour-joining analysis, but this has been demoted to sister group to three other DME domain sequences (CG17646, CG5853, CG9663). All other features are identical between the two phylogenetic analyses.

Neither the maximum parsimony analysis nor the neighbour-joining analysis could further resolve the ABCG sub-family phylogeny into orthologous groups. Therefore while PDR5 and the ABCG mammalian sub-family appear to share a common origin, the individual relationships could not be delineated further, limiting the use of phylogenetic analyses for selecting a candidate gene from the ABCG sub-family for further studies.
4.2.2 Defining the ABCG family in sheep

Based on the phylogenetic data presented (Fig. 4.3, 4.4 and 4.5) and published data (Sheps et al., 2004) the yeast PDR5 protein was most closely related to the mammalian ABCG sub-family of proteins. The ABCG sub-family contains six half transporters (ABCG1, 2, 3, 4, 5, 8), one of which, ABCG3, appears to be rodent specific. It was of interest to determine the evolutionary origin of the ABCG3 gene in rodents to determine the likelihood of ruminants possessing an ABCG3 homolog.

As a first pass approach the mouse ABCG3 cDNA sequence (NM_030239) was used to query the publicly available and in-house ovine EST databases using BLAST (Altschul et al., 1997). Significant hits ($P < 1 \times 10^{-20}$) were obtained for all mammalian ABCG genes with the exception of ABCG3 and ABCG5, suggesting that sheep do not possess orthologs of these genes.

While no ortholog of ovine ABCG5 could be detected, ABCG5 does have a bovine ortholog (DQ001133) so it is reasonable to assume that there is in fact an ovine ortholog.

The second, more in depth, approach was attempted to determine the origin of the ABCG3 gene in rodents. Based on sequence identity and conservation of intron position (data not shown) ABCG3 is known to be most closely related to the ABCG2 gene and is thought to have arisen as a recent duplication event (Mickley et al., 2001). To assess the possible origin of the ABCG3 gene the conservation of gene order around the ABCG2 locus was examined, utilising the fully sequenced genomes of five organisms. The five organisms selected were *H. sapiens* (build 35.1) which was

![Figure 4.6 Simplified phylogeny of eutherian mammals. This diagram was modified from (Strachan and Read, 1996). Species in bold font are species that were utilised for this analysis. Numbers given at node branches show estimated divergence times in millions of years.](image)
used as a framework sequence, *M. musculus* (build 34.1), *R. norvegicus* (build 3.1), *C. familiaris* (build 1.1), *Bos taurus* (build 3). These five species were chosen as they represent species from four of the major groupings of eutherian mammals (Fig. 4.6).

A 20 Mb region surrounding the 68.6 kb ABCG2 locus was chosen for analysis and the chromosomal position for all annotated loci were compared between the five genomes of the five species. In order to expand the number of shared loci, the human cDNA sequence was used as a query to blast against the genome of interest using cross species megaBLAST (Zhang *et al.*, 2000). This was particularly useful where the annotation for the genome of interest was incomplete.

The starting position of the gene was then obtained from significant BLAST scores ($E < 1 \times 10^{-20}$), and synteny was assessed by comparison with the human gene order. In cases where there were more than one significant hit, loci were not included in the analysis as these loci are possibly part of large well conserved gene families, or could possibly represent pseudo-genes.

In humans there are 102 annotated loci in this 20 Mb area, however, only 56 of them were useful for comparative mapping purposes. The remaining 46 either did not have homologues in the other four species examined, did not have a single significant BLAST hit, were localised to an unmapped contig, or represented a known pseudo-gene. A stylised comparative map of this region based on these 56 loci is presented in Fig. 4.7.

Based on this diagram, there has been significant genome re-organisation in this region since the divergence of rodents from other eutherian mammals. There are at least two chromosomal breakpoints in this 20 Mb region (Fig. 4.7) and several inversions (data not shown). Both of these chromosomal breakpoints are found in identical positions in *M. musculus* and *R. norvegicus* indicating that this chromosomal rearrangement has been stable since the divergence of the rodent lineage.
The chromosomal location of the ABCG2 locus relative to the breakpoint is indicated by the arrows on Fig. 4.7 while the location of the ABCG3 locus is indicated by an asterisk for the rodent species. The ABCG2 locus and ABCG3 locus (in rodents) are found on either side of one of these breakpoints. The proximity of these loci to the breakpoint in synteny implies that the ABCG3 locus has evolved following chromosomal breakage and rearrangement.

Surprisingly a breakpoint in a similar region had been found in cattle using radiation hybrid mapping (Itoh et al., 2005). The analysis presented here confirmed the presence of a breakpoint in synteny in this region. However, closer analysis showed that the breakpoint is not located in exactly the same position as in the rodent lineage. In cattle the synteny is broken at least two loci proximal to the ABCG2 locus (~ 500 kb, based on the human map). This provides evidence that the region surrounding the ABCG2 locus may be fairly fragile and is a hot spot for

![Stylised syntenic map of the 20 Mb region surrounding the ABCG2 locus.](image-url)

**Figure 4.7:** Stylised syntenic map of the 20 Mb region surrounding the ABCG2 locus. The positions of loci in other species are expressed relative to the human map position. The position of the ABCG2 gene is indicated in the human map. The corresponding map positions for each species are indicated in Mbp. The locations of the ABCG2 loci in each species are indicated (♀) and the ABCG3 locus is indicated by an asterisk (*) for *R. norvegicus* and *M. musculus.*
chromosomal rearrangement (Pevzner and Tesler, 2003). More importantly there is no evidence for an ABCG3 locus or an ABCG2 pseudo-gene in the bovine genome assembly. Therefore the ABCG sub-family in ruminants is likely to be restricted to five members: ABCG1, 2, 4, 5 and 8.

4.2.3 Collation of information concerning the ABCG sub-family

Inferring function of a protein based on phylogenetic data alone can be erroneous (refer 4.1.2) and when considering function of putative orthologues it is advisable to consider information from a variety of sources, these include not only sequence homology, but also known function, substrate specificity, and domain organisation (Ponting, 2001). Additionally, when using this information to identify possible candidate genes the examination of tissue expression profiles and phenotypes associated with known mutations can yield useful information.

As the phylogenetic analyses (Fig. 4.4, Fig. 4.5) indicated that the PDR5 protein is equally related to all 5 members of the ABCG sub-family, additional information was required to select a candidate gene from this family for further analysis. Information concerning known function, substrates, expression and pathologies was collated from published literature and is summarised in Table 4.3.

Analysis of published expression profiles reveals that all five members of the ABCG sub-family are fairly ubiquitously expressed and all were expressed in the liver which was the tissue of interest in this study. The functions and known pathologies revealed that four out of five of the ABCG sub-family members (namely ABCG1, ABCG4, ABCG5 and ABCG8) were primarily involved in cholesterol transport and metabolism. The final member of this family, ABCG2, is functionally very interesting and a very good candidate gene, as it is known to function as a xenobiotic transporter in humans (Table 4.3).
Table 4.3: Collation of information about the mammalian ABCG proteins.

<table>
<thead>
<tr>
<th>Expression Pattern</th>
<th>Known Functions</th>
<th>Pathologies caused by mutations in this gene</th>
</tr>
</thead>
</table>
| **ABCG1 (ABC8, White)**
  **Human** – lung, placenta, liver, kidney, skeletal muscle (Oldfield et al., 2002), macrophages (Klucken et al., 2000)
  **Rat** – eye, liver, lung, spleen (Oldfield et al., 2002)
  **Mouse** – adrenals, thymus (Savary et al., 1996), lung, brain, spleen (Savary et al., 1996; Croop et al., 1997) and macrophages (Su et al., 2002) | Macrophage cholesterol transport (Klucken et al., 2000). Over-expression in mice causes increased biliary cholesterol excretion, suggesting a role for this protein in cholesterol homeostasis in the liver (Ito, 2003) | Over-expressed in macrophages from patients with Tangier disease compared to control macrophages (Lorkowski et al., 2001). It is expressed in foamy macrophages within the atherosclerotic plaque (Wang et al., 2004a). May play a role in cholesterol homeostasis as hepatic expression of ABCG1 can be modulated by dietary cholesterol intake in rodents (Ito, 2003). Additionally the expression in macrophages is modulated by cholesterol loading (Klucken et al., 2000; Laffitte et al., 2001; Schmitz et al., 2001). |
| **ABCG2 (ABCP, BCRP1, MXR1)**
  **Human** – placenta (Allikmets et al., 1998), vessels and normal breast epithelium (Faneyte et al., 2002), placentantrophoblasts, epithelium of intestine and colon, liver canalicular membranes, ducts and lobules of the breast (Maliepaard et al., 2001)
  **Rat** – Liver, kidney, heart, brain, spleen, lung, skeletal muscle (Honscha et al., 2000)
  **Mouse** – placenta (Allikmets et al., 1998) | Xenobiotic transporter that appears to play a major role in the multidrug resistance phenotype of the MCF-7 breast cancer cell line. When ABCG2 is over-expressed the transfected cells become resistant to mitoxantrone, daunorubicin and doxorubicin. They also display diminished intracellular accumulation of daunorubicin, and manifest an ATP-dependent increase in the efflux of rhodamine 123. | Mice lacking ABCG2 displayed a unique environmentally induced photosensitivity, that was due to increased levels of pheophorbide a in the blood (Jonker et al., 2002). |
| **ABCG3 (putative rodent specific transporter)**
  **Mouse** – Expressed highly in spleen, thymus and lung. Expressed at moderate levels in other tissues, including liver (Mickley et al., 2001). | ABCG3 has no known function. The ABC cassette has degenerated it is therefore thought to be incapable of ATP binding and hydrolysis. It has been proposed that it may act as a functional regulator of ABCG2 by binding and forming non-functional heterodimers. However, ABCG2 and ABCG3 have only partially overlapping tissue expression profiles (Mickley et al., 2001). | |

129
ABCG4 (White 2)

**Human** – Brain (cerebellum, cerebral cortex, medulla, occipital pole, frontal lobe, temporal lobe, putamen) (Annilo et al., 2001), eye (Oldfield et al., 2002), heart, thymus, kidney, liver, small intestine (Annilo et al., 2001).

**Mouse** – Spleen, (Annilo et al., 2001; Yoshikawa et al., 2002), eye, bone marrow (Yoshikawa et al., 2002).

**Rat** – eye, brain (Oldfield et al., 2002).

ABCG5 (White 3, Sterolin 1)

**Human** – liver, small intestine, colon (Berge et al., 2000)

**Mouse** – liver (Berge et al., 2000), hepatocytes (uniform distribution across the hepatic lobule) (Repa et al., 2002a), duodenum, jejunum (Berge et al., 2000) mostly in enterocytes lining the villi (Repa et al., 2002b), ileum (Berge et al., 2000).

ABCG8 (White 4, Sterolin 2)

ABCG5 and ABCG8 are co-ordinately expressed and appear to be functionally equivalent.
4.2.4 *In silico* mapping (comparative mapping).

The final piece of evidence considered when selecting a candidate gene from the five members of the ABCG sub-family was to determine if there was any evidence for genetic association with resistance or susceptibility to FE. To do this *in silico* mapping was used to determine the probable map positions for the five ABCG sub-family members and to compare them to quantitative trait loci (QTL) identified in the genome scan experiment (S. H. Phua, *pers. comm.*).

The first step in the *in silico* mapping of the ovine ABCG family was to determine the bovine map positions for the gene sequences of interest. This was done using the Bovine Genome Assembly (v2.0) Browser (also called gBrowse) (available at http://www.livestockgenomics.csiro.au/perl/gbrowse.cgi/bova2/). For genes annotated in the bovine genome the gene order in the surrounding region was compared with human to confirm synteny. But for genes that had not been annotated in the bovine genome the gene order in humans was used to extrapolate the location of the gene of interest in the bovine genome. This was done in conjunction with the information provided by Itoh *et al.* (2005) and Band (2000).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Human Map (Mb)</th>
<th>Mouse Map (Mb)</th>
<th>Cattle Map (Mb)</th>
<th>Sheep Map (Mb)</th>
<th>Flanking Markers in Sheep</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCG1</td>
<td>21</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>BM3205 – DIK5034</td>
</tr>
<tr>
<td></td>
<td>42.49 - 42.59</td>
<td>28.87-28.93</td>
<td>92.66 – 92.74</td>
<td>295.2 – 305.8</td>
<td></td>
</tr>
<tr>
<td>ABCG2</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>OarAE101 - OarHH55</td>
</tr>
<tr>
<td></td>
<td>89.37-89.44</td>
<td>58.81-58.85</td>
<td>33.55 – 33.61</td>
<td>50.2</td>
<td></td>
</tr>
<tr>
<td>ABCG4</td>
<td>11</td>
<td>9</td>
<td>15</td>
<td>15</td>
<td>MNB85A – JAB1</td>
</tr>
<tr>
<td></td>
<td>118.52-118.54</td>
<td>44.26-44.27</td>
<td>14.25 -14.35</td>
<td>38 – 45.9</td>
<td></td>
</tr>
<tr>
<td>ABCG5</td>
<td>2</td>
<td>17</td>
<td>11</td>
<td>11</td>
<td>BM304 – BMS1953</td>
</tr>
<tr>
<td></td>
<td>43.95-43.98</td>
<td>82.49-82.51</td>
<td>13.46 – 15.79</td>
<td>128.9 – 134.8</td>
<td></td>
</tr>
<tr>
<td>ABCG8</td>
<td>2</td>
<td>17</td>
<td>11</td>
<td>3</td>
<td>BM304 – BMS1953</td>
</tr>
<tr>
<td></td>
<td>43.95-43.98</td>
<td>82.51-82.53</td>
<td>13.46 – 15.79</td>
<td>128.9 – 134.8</td>
<td></td>
</tr>
</tbody>
</table>
In both cases sheep microsatellite markers flanking the loci of interest were obtained from gBrowse and the map positions of these markers (sex averaged) were obtained from the best position maps (v 4.5) from http://rubens.its.unimelb.edu.au/~7Ejillm/jill.htm (Table 4.4).

Only subsets of sheep microsatellite markers were annotated on gBrowse. Therefore by manually placing ovine markers on the bovine genome it was often possible to further refine these putative map positions. The manual assignment of ovine microsatellite markers to the bovine genome was done by determining whether the ovine markers flanking the gene of interest (as indicated by gBrowse) had any intervening ovine markers, as assessed by comparison with the current ovine linkage map (http://rubens.its.unimelb.edu.au/~7Ejillm/jill.htm). Those ovine markers with NCBI sequence were compared with the bovine genome using BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) to estimate the ovine marker positions relative to the bovine genome. The best human, mouse, cattle and putative sheep map positions are given in Table 4.4.

These probable map positions of the ABCG family members were compared with quantitative trait loci (QTL) identified in whole genome screens (S. H. Phua, pers. comm.). This comparison was performed to determine if any of the ABCG genes fell in QTL regions (Table 4.5). ABCG5 and ABCG8 were localised to a chromosomal position with significant evidence for a FE QTL, while ABCG2 is found in a chromosomal region with weak evidence for a FE QTL.

Table 4.5: Comparison of map positions with existing QTL data. Map positions obtained for the ovine ABCG sub-family with pre-existing genome scan data (S. H. Phua, unpublished data), ns = not significant.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sheep chromosome and position (cM)</th>
<th>Flanking Markers in Sheep</th>
<th>Evidence for QTL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCG1</td>
<td>1: 295.2 – 305.8</td>
<td>BM3205 &amp; DIK5034</td>
<td>ns</td>
</tr>
<tr>
<td>ABCG2</td>
<td>6: 50.2</td>
<td>OarAE101 &amp; OarHH55</td>
<td>Weak evidence</td>
</tr>
<tr>
<td>ABCG4</td>
<td>15: 38 - 45.9</td>
<td>MNB85A &amp; JAB1</td>
<td>ns</td>
</tr>
<tr>
<td>ABCG5</td>
<td>3: 128.9 – 134.8</td>
<td>BM304 &amp; BMS1953</td>
<td>Significant</td>
</tr>
<tr>
<td>ABCG8</td>
<td>3: 128.9 – 134.8</td>
<td>BM304 &amp; BMS1953</td>
<td>Significant</td>
</tr>
</tbody>
</table>
SECTION B: Evaluation of ABCG2 as a candidate gene in facial eczema disease.

4.3 INTRODUCTION
As shown in section A of this chapter, the yeast PDR5 protein is most closely related at the protein level to the mammalian ABCG sub-family. In most mammalian species the ABCG sub-family contains five members (ABCG1, ABCG2, ABCG4, ABCG5 and ABCG8), four of which (ABCG1, ABCG4, ABCG5 and ABCG8) are known to function in lipid homeostasis, specifically cholesterol homeostasis (reviewed in Schmitz et al., 2001). The final member of this family, ABCG2, functions as a xenobiotic transporter and is therefore a putative functional candidate for resistance to FE. Additionally, mice deficient in ABCG2 display sensitivity to diet-dependent phototoxicity (Jonker et al., 2002), a phenotype similar to FE-afflicted animals. ABCG2 has also been shown to transport phylloerythrin which is the major photosensitising agent in FE (Robey et al., 2006). ABCG2 is also a positional candidate for facial eczema disease as comparative mapping localised it to ovine chromosome 6 in a region containing weak evidence for a QTL. Therefore the focus of this section of work was to evaluate ABCG2 as a positional candidate gene for FE.

4.3.1 ABCG2
ABCG2 is a reverse half ABC-transporter, consisting of six transmembrane domains localised towards the C-terminus of the protein, while the ATP-binding domain is localised towards the N-terminus (refer to 3.1.1.1). ABCG2 was discovered simultaneously by three research groups; it was cloned from a breast cancer cell line (MCF7) selected for drug resistance (Doyle et al., 1998), from human placenta (Allikmets et al., 1998) and from mitoxantrone-resistant human colon carcinoma cell lines (Miyake et al., 1999).
Chapter Four (Section B): Evaluation of ABCG2 as a candidate gene for FE

4.3.1.1 Tissue distribution
The ABCG2 transcript is fairly ubiquitously expressed and is abundant in the placental trophoblast, the epithelium of the small intestine and the liver canalicular membrane (Doyle and Ross, 2003). ABCG2 has also been shown to be expressed in the gall bladder (Suzuki et al., 2003) as well as the ducts and lobules of the breast (Jonker et al., 2005). It is highly induced during lactation and has been shown to concentrate drugs and xenobiotics into milk in several animal species, including cattle (Jonker et al., 2005). In polarised cells such as the intestinal epithelial cells the ABCG2 protein is localised to the apical membrane, consistent with its role as a drug efflux pump (Maliepaard et al., 2001; reviewed in Krishnamurthy and Schuetz, 2006).

The expression of ABCG2 is known to be regulated by physiological levels of estrogen and a functional estrogen response element has been found in the ABCG2 promoter (Ee et al., 2004). Although the response is cell line specific (reviewed in Krishnamurthy and Schuetz, 2006), estrogen and several estradiol compounds have been shown to down-regulate ABCG2 expression through an estrogen receptor α-interaction pathway (Imai et al., 2005). Consistent with this, much higher levels of ABCG2 expression are seen in the livers of male mice, partly through estrogen mediated inhibition and also through testosterone mediated induction (Tanaka et al., 2005). ABCG2 expression has also been shown to be responsive to tissue hypoxia in stem cells via the transcription factor HIF-1 (Krishnamurthy et al., 2004).

4.3.1.2 Function and substrate specificity
The ABCG2 protein is known to function as a cellular efflux pump for hydrophobic compounds, and it can also transport large anionic compounds such as drug conjugates (Sarkadi et al., 2006). The full complement of endogenous substrates is not well defined, but it includes steroid hormones, in particular 17-β estradiol (Pavek et al., 2005) as well as glucuronide and sulphate conjugates of estrogen (Imai et al., 2003). Recently ABCG2 has been shown to transport phylloerythrin, a photoreactive breakdown product of chlorophyll (Robey et al., 2006).

134
ABCG2 is involved in both phase 0 and phase III drug metabolism as it is able to modulate the absorption of xenobiotics in the intestine (phase 0) and also facilitates the elimination of phase II metabolites from cells (phase III) (Sarkadi et al., 2006; Wakabayashi et al., 2006). ABCG2 is over-expressed in several forms of cancer and is thought to contribute to the multi-drug resistance phenotype observed in some humans (reviewed in Sarkadi et al., 2006).

The tertiary structure of ABCG2 has not been resolved and there is some debate over the nature of the active transporter. ABCG2 had been thought to function as a homodimer (Kage et al., 2002; Litman et al., 2002; Mitomo et al., 2003) which is stabilised by disulfide bonds involving a single conserved cysteine residue at position 608 in the human protein. However, recent evidence suggests that ABCG2 may form a tetrameric structure composing 4 homodimeric ABCG2 complexes (McDevitt et al., 2006). More recently, higher oligomeric structures of ABCG2 have been found. Specifically a dodecameric ABCG2 oligomer has been characterised and the oligomerisation domain mapped to the large intracellular loop between transmembrane domains 5 and 6 (Xu et al., 2007).

Very little is known about the substrate recognition and transport properties of ABCG2, although for other ABC transporters (specifically ABCB1 and ABCC1) the transmembrane domains have been found to be responsible for the recognition of substrates (reviewed in Sarkadi et al., 2006).

The activity of the ABCG2 protein can be inhibited by several compounds. These compounds are mostly polyphenols and include some naturally occurring flavonoids (Zhang et al., 2004b). The inhibition by common dietary constituents raises the possibility of dietary regulation of ABCG2 activity, particularly in the intestine.

4.3.1.3 Mutations and polymorphisms
There are at least 80 naturally occurring variants of the ABCG2 gene found in human populations and a number of other variants found in drug selected cell lines (reviewed in Wakabayashi et al., 2006). A few of these mutations are known to affect the transport activity of the ABCG2 protein (reviewed in Wakabayashi et al., 2006). For instance, the Q141K mutation affects the transport efficiency of the
protein (Mizuarai et al., 2004) as well altering the localisation of the mature protein from the cell membrane to an intracellular location. Similar results have also been found for the S441N mutation (Tamura et al., 2006).

At present only a few mutations are known to affect the substrate specificity of the ABCG2 protein, and most of these have only been found in drug-selected cell lines and not in human populations. The best characterised mutation is the R482X mutation which was discovered in two drug selected cell lines (Honjo et al., 2001). In these cell lines the polar amino acid arginine is replaced with either the small non-polar amino acid glycine, or the somewhat larger polar amino acid threonine. Both of these substitutions radically alter the substrate specificity of the ABCG2 protein (reviewed in Krishnamurthy and Schuetz, 2006). No corresponding polymorphisms have yet to be detected in human populations (Honjo et al., 2002).

4.3.2 Linkage Mapping
In the absence of a fully assembled genome sequence, linkage mapping is the most accurate method of determining the map position of a locus of interest. Linkage maps are assembled from co-segregating groups of markers. Within these groups the proportion of recombinants (crossover events) detected between linked markers is used as a measure of the distance between them. Therefore the chance of crossover is proportional to the distance between the two markers. This distance is represented by the recombination fraction, $\theta$. Recombination fractions range from $\theta = 0$ for two loci very close to each other to $\theta = 0.5$ for loci far apart or on different chromosomes. Therefore loci are said to be linked when $\theta < 0.5$. The unit of recombination is the centimorgan (cM) which corresponds to an approximate recombination frequency of 1% and a physical distance of $\sim 1$ Mb. Because of the occurrence of multiple crossovers, $\theta$ is not an additive measure of distance and it is transformed by a map function into a map distance. The method used to do this in the present study is Kosambi map units. This function translates $\theta = 0.27$ into 30 cM (Ott, 1991).

4.3.2.1 The ovine linkage map
For linkage mapping to be successful large numbers of informative genetic markers are required in addition to large pedigrees in which the inheritance of each marker can be studied (reviewed in Montgomery and Crawford, 1997). The most recent
version of the sheep linkage map (third generation linkage map) (Maddox et al., 2001) contains 1,062 loci (941 anonymous loci and 121 known genes). The map spans 3,400 cM (sex-averaged) for the autosomes and 132 cM (female) on the X chromosome. Additional assignments for ~250 genes have been made using somatic cell hybrids (Saidi-Mehtar et al., 1979; Burkin et al., 1991) and fluorescence in situ hybridization.

This linkage map is a compilation of genotype data generated by 15 laboratories using the AgResearch International Mapping Flock (IMF) pedigrees. The IMF pedigrees were developed for genetic mapping in sheep and consist of 9 three generation pedigrees comprising a total of 127 individuals (Crawford et al., 1995).

4.3.3 Association Studies

The objective of association studies is to find consistent differences in allele frequencies at the loci of interest between two groups of phenotypically extreme animals. Association studies are often more useful than large pedigree linkage studies (QTL analyses) when analysing complex disease traits like FE. This is because complex disease traits can involve many genes each contributing a small amount to the overall phenotype. Also, the modes of inheritance are more complex and standard QTL experiments have limited power to detect these interacting, small or minor effect genes (Risch and Merikangas, 1996). Association studies require large numbers of unrelated samples, and these can often be easier to collect than samples from defined pedigrees.

The ability to detect single nucleotide polymorphisms (SNPs) in sheep has improved with the availability of the bovine genome sequence. SNPs are relatively easy to find and are valuable type I makers. The use of biallelic SNP makers in association studies is facilitated by high levels of linkage disequilibrium (LD) in ruminants. These high levels of linkage disequilibrium mean that even small numbers of SNPs can be used to detect association across a relatively wide genomic region. LD is fairly conservative in humans (40 – 60 kb) (Reich et al., 2001) but in sheep LD has been shown to extend for tens of centimorgans (McRae et al., 2002).
4.4 RESULTS

4.4.1 Ovine ABCG2 cDNA sequence
At the time of this study no ruminant sequence for ABCG2 was available. Therefore it was imperative to determine the full-length cDNA sequence to facilitate further experimental investigation.

Oligonucleotide primers were designed as detailed in section 2.4.2 using the human RefSeq (NM_004827) as a template sequence (refer to Appendix A for primer sequences). PCR amplification was carried out as detailed in 2.4.3 using liver cDNA as the template (refer 2.5.4). PCR products were purified and sequenced as detailed in 2.4.5 and 2.4.11. The entire ABCG2 coding domain was amplified and sequenced in four overlapping fragments. These fragments were sequenced from four individual selection-line animals: one individual each from dosed resistant, dosed susceptible, naïve resistant, and naïve susceptible groups. The 5' and 3' untranslated-region (UTR) sequences were obtained by 5' and 3' RACE as detailed in 2.4.3.1.

The full-length transcript obtained in this study was 2191 bp excluding the poly(A) tail. The sequence and selected features are presented in Fig. 4.8 (NCBI accession: DQ886530). The transcript had a single open reading frame (ORF) of 1977 bp and a predicted peptide product of 655 amino acids. The 5'-UTR was between 147 and 156 bp in length, and the 3'-UTR was 66 bp. The uncertainty in determining the size of the 5'-UTR is due to the presence of two methionine start codons that are 6 bp apart. Neither of the two putative start codons have a strong Kozak sequence (Kozak, 1986): GXXATGt for the first methionine and AXXATGt for the second. Analysis using the ATGpr algorithm (Salamov et al., 1998) suggests that the first methionine is the favoured initiation codon with a reliability score of 0.35 as compared to 0.14 for the second. However, the first methionine is not found in other species except cattle (Cohen-Zinder et al., 2005). Tandem initiation codons have been found to increase translational efficiency in vitro (Wakiyama et al., 1993) and may indicate differences in the regulation of ABCG2 between ruminants and other mammalian species.
Chapter Four (Section B): Evaluation of ABCG2 as a candidate gene for FE

-147 ttggtgtaa catcaaggtt ctagaggttg aggggttaacc cagagactgc atgaaactct

-86 atgctgcttc tctctgtgctt accctatgga tctctctctat tgggtgttctc ctaggttgta

-26 tcttttataa tacaccgaa ggccggaatg TGCAAAATGT CCTCCCAATAG TTACGAGGTT

33 M FK MS SNS SYE V

34 TCCATTCCAA TGTCAAATAAA ACCCCACGCC ATTCAGAGAA CAACCTCTAA GGACCTGCAG

94 S IFM SK PK NNP ET TSK DLQ

153 AGTTGACTG AAGGACTGTT GTTAAGTGT TTACAACATCT GCTATCGAGT AAAAAATGAAAG

TLEGAVL SFHNGLI PHSICK V

213 ACTGGCTTTC TACTGTCG AGAAACAATT GAAAGAAAAT TACTGACAGA TATCAATGGA TGFLLCR KTIIEKI LANG

273 GTCATGAAC GTCGCTCAAA TGCATTCTCA GGACCACAG GCTGAGGCAA ATCTCGTGGT

333 VMKPGNLNAI GPTGGGK RRSSL

393 TAGATATCT TACGGCTGAG GAGAGACCA CAAGGATTAT TTTGATCAAT

513 L DIII AAR KDP HGL SGGV L

394 GGAGCACCCT GACCTGCCTA TTTTAAATGT AACCTAGGTT ATGTTGCTCA AGATGATGTC

453 GA PR FAPAN FKNCGSY VVQDDV

454 GTGATGGGGA CCTGGACTG GAGAGAAC ATCGTGGTCT CATGAGGCAA ATCTCGTGGT

513 TTMNYYEKNERINKVIQELG

573 CTGGATTAAG TGCGACATTG CAAGGTTGGA ACICAGTTTA TCGTGTTGT GTCTGGGCA

633 LDKVADSKVGTQFIRGWSGG

634 GAAAGAAAAGA GAACAGTAGT TGCATGGAG CCTATTACTGT ATCCATCCAT CTTGTTCCCTG

693 ERKRDTSIAMELITDSPILFL

693 GATGAGCCCA CAACCTGGTTT AGATCCAGAC ACAGCAAAATG CTGCCTTTTG GCTTTGGAG

753 DEPTGSGSSTNANAVLLKL

754 AGGATGTCTA AACAGGAGCG ACAAATCATC TTTCCACTCC ATCAGCCCTG TTATCCACAT

813 RSMKSQGRTIIPFSIHQPRYSI

814 TCCAGTATTG TGGTAGACCT CACCTGGTTG GCTCAGAGAA GACTAAATTG CCATGGGCTC

873 FKLFDSLTLASGRMLFHPG

873 GCTCAGGAGG CTTGGGTTTA CTTGAGAGC AATGGTTTCC ACTGTGAGCC ATGATAAAGC

AQLGYFEDINCHCPEYNN
Chapter Four (Section B): Evaluation of ABCG2 as a candidate gene for FE

874
CCTGCAGACT TCCTCCTGGA CATCATTAAT GGAGACTCTT CTGCTTGCGT GTAATATAGA
PADF FLD IINGDSSAVV LNR

934
GAGACTCAG ATGATGAGAC TAAGAGAGCC GAGAGCTCTT CCAAAAAGCA TACCTCACCT
EDS DDEA KETE EPSK NDKTS L

994
ATAGAAAAAT TAGCCTGGTT TTAGTGCAC TCTCCTTCT TCAAGGAAAC AAAAGTGGAA
IEKL AGFY VNSSS FKFETKEV

1054
TTAGATATAAT TCTCAGGAGG GCGGAGGAGG AAGAGCTTCTT CATCCTATAA GGGATCTCCT
LDKF S GRRR KKLSSYKEIT

1114
ATGCCCACT CCTTCTGTCA TGACCTCAAA TGAGTTCCCA AGCCTTCACT CAAAATTTA
YATS FQHQLWISRFSKKNL

1174
CTGGGTAATC CCCAGGCTCC TAGTACGTAG TTAAATGGTA CAGCTCTTCT GGGACTGTTT
LGNP QAIAQLIITVLFGL

1234
ATAGGGCCA TTTCTCTGTA TCTAAAAAAT CATCCTCTAG GAATCCAGGA CAGAGCCGGG
IGAPIFYDLKNDPSGIQRAG

1294
GTGCCTCTCT TCTGACGAC CAAACAGTCTT TGTCAGCCGT TGGCTGGTGT VLFLLTTNQCFS3VSAVELL

1354
GTTGGGAA GAGAGCTTCT TAGATGATGA AAATACTAGTG AGATGTACAT VEEKKLFHIEYISGYRVSS

1414
TACCTCTTTG GAAACTGTTT ATCTGATTATA CTCCCCCATGA GGATGTCCTT AAGTATTATA
YFFGKLSSDLPLMLPSII

1474
TTACCTGCTA TAAATACCTT CTGTTAGGA CTGAAAACGA AGGTGAGGCC TCTCTCACCC
FTCI TYFLLLGLKPFVEAFFI

1534
ATGATGTTTA CCGCTAGATG GTGAGCTTATT TCACTGCTGT CATGTCAGCT GGCCATAGCA
MMFTLMMVAYSSSMAALAI

1594
GCGGTCAGA GTGGGCTGTC CAGTACCACT CGACTCTCTA CCATCTCTT CTGTTAGTAT
AGQS VVSIATTLLMTISFVFM

1654
ATGATATTT CAGGGCTTGT GTTAAACTCT AAACCATTGG GGGTGGAGTT GTGCTGGAGG
MIFSGLLVNLKTIAGAWLSW

1714
CAGTACTTGA GCATCCCTCG ATACGGCTAT GGGCTTGGTG AGCATATATGA ATTTTTGGGA
QYLSIPRYGVAALQHNEFLG

1774
CAAAACTTCT CCACAGGACT CAAGTACACA GCACAAACATA CCTGATGCTA TGCAATATCT
QNFCPGLNVTRANNTCSYAIN

1834
ACTGGTGGAAG AATTCGTCAC AACAGGGCC ATGATATCTT CACTTCGGG CCGTGGGAG
TGEEFLTNQGIDISPWGLWK
Chapter Four (Section B): Evaluation of ABCG2 as a candidate gene for FE

Figure 4.8: Nucleotide and predicted amino acid sequence of the sheep ABCG2 cDNA. Nucleotides are numbered positively in the 5' to 3' orientation beginning with residue 1 of the first putative ATG initiation codon, and ending with the last residue of the poly (A) tail. Nucleotides preceding the first in-frame initiation codon are numbered negatively. The first (l) and second (t) in-frame initiation codons are indicated. The Walker A motif corresponding to position 79–86 in the amino acid sequence (GXXGXGKS) is underlined, the two walker B motifs (XXXD where X is hydrophobic) corresponding to positions 205–209 and 291–295 in the amino acid sequence are double underlined and the signature motif is highlighted in grey (amino acids 186–192). The termination codon is indicated by an asterisk and the putative polyadenylation signal (non-canonical) is indicated in red.

Examination of the ovine ABCG2 protein sequence confirms the presence of a full ABC domain: it contains the Walker A and Walker B motifs, both of which are involved in ATP binding (Walker et al., 1982), and the signature motif (Hyde et al., 1990). In the ovine ABCG2 sequence the nucleotide binding domain (NBD) is localised towards the N-terminus of the protein and the transmembrane domains are localised towards the C-terminal in the so-called ‘reverse’ orientation. This ‘reverse’ domain organisation is characteristic of the ABCG family of mammalian ABC transporters (Dean et al., 2001).

Prediction of secondary structure (Fig. 4.9) using TopPred (v. 2) (von Heijne, 1992; Claros and von Heijne, 1994) infers that the ovine ABCG2 protein has six transmembrane domains (TM), consistent with the human and rodent ABCG2 proteins. There is also a long extracellular loop between TM5 and TM6, which is also found in other mammalian ABC proteins, and contains three conserved cysteine residues as indicated in Fig. 4.9. One of these, Cys606, is thought to be involved in dimerisation of the ABCG2 protein, while the other two Cys595 and Cys611 are thought to form an intra-molecular disulphide bond which is important for N-glycosylation and correct localisation of the ABCG2 protein to the plasma membrane. In total there are 11 cysteine residues present in the ovine protein and
Figure 4.9: The hydropathy plot of the deduced ovine ABCG2 protein sequence. The hydropathy profile was calculated with the Kyte and Doolittle hydropathy algorithm (Kyte and Doolittle, 1982) implemented using TopPred2 (http://bioweb.pasteur.fr/seganal/interfaces/toppred.html) using the default parameters. The positions of conserved cysteine residues are indicated by the pink asterisks.

these are all in conserved positions with respect to the human ABCG2 protein sequence.

The 3' UTR sequence was 66 bp in length and the putative polyadenylation signal highlighted in Fig. 4.8 was a non-canonical but common polyadenylation sequence (AUUAAA). It has been suggested that such non-canonical sequences may be processed less efficiently than the canonical signal thus providing an extra level of regulation (Beaudoin et al., 2000). The non-canonical polyadenylation signal is observed for ABCG2 in several mammalian species including human, rat, mouse and cattle. However, humans also have a second canonical polyadenylation site ~1800 bp downstream which is not annotated in other species. The D. melanogaster white protein (NM_057439) has a canonical polyadenylation signal.

4.4.1.1 Sequence Homology

The ovine ABCG2 cDNA sequence shares 88% identity at the nucleotide level with the human ABCG2 sequence (NM_004827) and less than 50% identity with the other four members of the ABCG sub-family. The same pattern is reflected in the
comparison of protein sequences where the highest identity and similarity, 85% and 92.1% respectively, are shared with the human ABCG2 protein. This strongly supports that the sequence isolated in this study is indeed the ovine ortholog of ABCG2.

Comparison of the ovine ABCG2 protein sequence with sequences from seven other species revealed high levels of sequence identity (>75%) for the four mammalian species (mouse, rat, human and pig) and moderate levels (30% - 63%) for the three non-mammalian species. Specifically, the ovine ABCG2 sequence shared 85% identity and 92.1% similarity with the human ABCG2 protein (NP_004818), 79.6% identity and 91.5% similarity with the mouse ABCG2 protein (NP_036050), 79.1% identity and 90.7% similarity with the rat ABCG2 protein (NP_852046), and 30.6% identity and 53.9% similarity with the *D. melanogaster* CG2759-PA (white) protein (NP_476787). The ovine sequence was also compared with the yeast PDR5 protein (Z75061) and nucleotide sequences (CAA99359). To do this the PDR5 protein sequence was split into two, an N-terminal and a C-terminal region, each containing an ABC domain and 6 transmembrane domains. The ovine ABCG2 sequence shares similar levels of identity to each of the two halves of the yeast PDR5 protein as the two halves to themselves.

Taken together, the conserved motif structure, domain organisation and sequence similarity suggest that the cDNA sequence obtained in this study is in fact the ovine ortholog of ABCG2.

### 4.4.2 ABCG2 sequence variation

The complete ABCG2 coding region was sequenced in two resistant and two susceptible selection-line sheep. Sequence analysis revealed two synonymous nucleotide polymorphisms in one of the susceptible animals. The first single nucleotide polymorphism (SNP) was an A/G transition at position 554 of the cDNA sequence (DbSNP ss65824075:G>A, AgR AS-000346). The second SNP was also a transition at position 973 (DbSNP ss65824134:C>T, AgR AS-000347).

At the time of this study the only mutation that was known to change the substrate specificity of the ABCG2 protein was the R482X mutation (refer to 4.3.1.3).
order to screen for this mutation in resistant and susceptible sheep, direct sequencing of exons 9 – 13 (427 bp, encompassing this R482X codon) was undertaken using cDNA from sixteen FE selection-line animals. While this is a relatively small number of animals, the FE selection lines represent genetically isolated populations and it is expected that mutations that contribute significantly to resistance or susceptibility to FE would be fixed in the population. Analysis of the DNA sequences obtained detected no polymorphisms between animals in this region.

4.4.3 Linkage mapping of the ABCG2 locus.
ABCG2 was selected both as a functional and positional candidate for FE, but its chromosomal location had been inferred by comparative mapping. Therefore linkage mapping was used to confirm the location of the ABCG2 locus. Linkage mapping requires a marker that is informative in the family of interest, in this case the IMF. The marker must be in or close to the gene of interest, and the ABCG2 intron sequences were seen as a source for SNP discovery.

The sizes and positions of introns within the ovine ABCG2 gene were estimated by comparison with the human, mouse and bovine genome sequences. The relevant genome sequences were obtained from Golden Path (human and mouse) and the bovine genome sequence was obtained using megaBLAST (Zhang et al., 2000). These genome sequences were then aligned with the ovine cDNA sequence using SPIDEY (Wheelan et al., 2001) set to adjust for divergent sequences. Three introns were selected that were small enough to be PCR amplified and sequenced (<1000 bp): these were introns 3, 4 and 14.

Primers were then designed to flank these introns (Appendix A). All three introns were PCR amplified (refer 2.4.3), purified (refer 2.4.5) and sequenced (refer 2.4.11) from all four sires of the IMF. Only intron 4 resulted in a single sequence and contained several SNPs.

The intron 4 sequence obtained was 812 bp in length (NCBI accession: DQ886529, Fig. 4.10), and was bounded by the classical GT donor AG acceptor splice boundaries (Fig. 4.10). Six SNPs and a single insertion/deletion were discovered within this intron.
Figure 4.10: DNA sequence of ovine ABCG2 intron 4. The canonical splice sites are boxed (GT donor, AG acceptor). Six SNPs were discovered (highlighted in blue). The SNPs are represented by their corresponding IUPAC code (k = T/G, r = G/A, y = C/T). A single base pair insertion/deletion was also discovered (highlighted in red). Arrow indicates the location of the SNP used for linkage mapping and association studies (DbSNP ss65824074:G>A, AgR AS000343).

The SNP at position 671 (DbSNP ss65824074:G>A, AgR AS000343) was selected as the marker for linkage mapping as three out of the four IMF sires were heterozygous (maximising the number of informative meioses). The SNP genotyping assay was designed by K. Patterson (AgResearch) and the genotyping carried out by H. Henry (AgResearch). The genotyping of this SNP was carried out
Table 4.6: Two point linkage analysis of the SNP marker ss65824074:G>A. Markers with associated LOD scores and θ values are tabulated.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Position (cM)</th>
<th>Recombination (θ)</th>
<th>LOD Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGF</td>
<td>22.8</td>
<td>0.22</td>
<td>3.4</td>
</tr>
<tr>
<td>MCM53</td>
<td>26.9</td>
<td>0.19</td>
<td>4.5</td>
</tr>
<tr>
<td>JL1A</td>
<td>32.9</td>
<td>0.09</td>
<td>4.7</td>
</tr>
<tr>
<td>CENPE</td>
<td>39.5</td>
<td>0</td>
<td>4.8</td>
</tr>
<tr>
<td>MCMA14</td>
<td>42.2</td>
<td>0.02</td>
<td>12</td>
</tr>
<tr>
<td>LSCV43</td>
<td>44.9</td>
<td>0.01</td>
<td>16.1</td>
</tr>
<tr>
<td>AE101</td>
<td>46.8</td>
<td>0</td>
<td>15.7</td>
</tr>
<tr>
<td>HH55</td>
<td>51.2</td>
<td>0</td>
<td>10.8</td>
</tr>
<tr>
<td>BM143</td>
<td>54.6</td>
<td>0.03</td>
<td>8.8</td>
</tr>
<tr>
<td>JMP36</td>
<td>57.9</td>
<td>0.06</td>
<td>6.6</td>
</tr>
</tbody>
</table>

using primer extension with analyte detection on a MALDI-TOF mass spectrometer (Sequenom Inc.) (Tang et al., 1999) as detailed in section 2.4.12.

The IMF flock (127 individuals) was genotyped with this SNP marker (ss65824074:G>A, AS-000343). The resulting mass spectra were automatically scored using SpectroTYPER (Sequenom), manually checked, then subjected to multipoint linkage analysis using CRI-MAP (Lander and Kruglyak, 1995). CRI-MAP analysis was carried out by T. Van Stijn and A. Beattie (AgResearch). The

![Figure 4.11: Localisation of the ABCG2 locus in the sex-averaged map of ovine chromosome 6 using the SNP marker (ss65824074:G>A, AS000343).](image)
Chapter Four (Section B): Evaluation of ABCG2 as a candidate gene for FE

Figure 4.12: QTL analysis of ovine chromosome 6. A region detected on OAR6, associated with FE in outcross pedigrees. Selective genotyping was used in the primary genome screen (left), where only animals with extremely resistant or susceptible phenotypes were genotyped. In the secondary screen (right), all progeny of informative sires were genotyped and analysed (refer 1.7.1). The vertical axis plots the F-value for the allele-within-sire term (1 numerator d.f. for each of the four families). The three horizontal lines indicate the genome-wide suggestive levels as determined by permutation analysis for the three FE traits analysed (logeGDH, logeGGT and logeGGT21). The positions of microsatellite markers used in the genome scan are indicated by the arrows below each graph: A = OarCP125, B = BM9058, C = MCM53, D = BM1329, E = OarAE101, F = BM143, G = BM4621, H = BM4311, I = OarJMP8 and J = BL1038. Figure reproduced from Duncan et al. (2007).

results from two-point linkage analysis are shown in Table 4.6.

The ss65824074:G>A polymorphism was linked to markers on OAR6, (Table 4.6) with a maximum two-point lod score of 16.1 with LSCV43 on the framework map of Maddox et al. (2001) and a recombination fraction of 0.01.

Multipoint mapping placed ABCG2 to a position 2 cM distal to OarAE101 (Fig. 4.11) on the same framework map (Maddox et al., 2001). This map position was consistent with that predicted from comparative mapping between human, cattle and sheep. This ABCG2 locus was mapped between AE101 (marker E) and BM143 (marker F) and falls within the peak of a minor FE QTL on OAR6 identified in a genome-scan experiment (Fig. 4.12) (Duncan et al., 2007).

4.4.4 Association studies of markers in FE selection lines

Association studies were undertaken to assess the involvement of ABCG2 in conferring resistance to FE. This approach has been used to evaluate ABCB1 (Longley, 1998) and catalase (Phua et al., 1999) as candidate genes for resistance to FE.
Chapter Four (Section B): Evaluation of ABCG2 as a candidate gene for FE

The three SNP markers discovered above (ss65824074:G>A, ss65824075:G>A, ss65824134:C>T) were genotyped in 66 resistant, 66 susceptible and 44 control line animals.

The distribution of allele frequencies of the SNP markers were obtained for each line and the allele frequencies were compared for each individual SNP by calculating chi-square statistics. The significance level for this statistic was found by Dr. K Dodds (AgResearch) using the peddrift simulation method (Dodds and McEwan, 1997), which accounts for genetic drift, founder and sampling effects by incorporating the actual pedigree data into the analysis.

There were no significant differences in the allele frequencies observed for the two SNPs in the coding region of ABCG2 (ss65824075:G>A and ss65824134:C>T). However, the allele frequencies distribution for the intronic SNP marker (ss65824074:G>A) were significantly different in the susceptible line compared with the resistant and control lines (P=0.044) (Table 4.7, Fig. 4.13).

Pre-existing genotype data for three closely neighbouring microsatellite markers (JL36, CSAP14E, OarAE101) also showed no significant differences in their allele

<table>
<thead>
<tr>
<th>Alleles of SNP Markers</th>
<th>Number of alleles</th>
<th>Peddrift analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Resistant line</td>
<td>Susceptible line</td>
</tr>
<tr>
<td>ss65824074:G&gt;A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>67 (64 animals)</td>
<td>22 (66 animals)</td>
</tr>
<tr>
<td>G</td>
<td>61 (64 animals)</td>
<td>110 (66 animals)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ss65824075:G&gt;A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>32 (59 animals)</td>
<td>21 (64 animals)</td>
</tr>
<tr>
<td>G</td>
<td>86 (59 animals)</td>
<td>107 (64 animals)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ss65824134:C&gt;T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>97 (60 animals)</td>
<td>122 (66 animals)</td>
</tr>
<tr>
<td>T</td>
<td>23 (60 animals)</td>
<td>10 (66 animals)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.7: Allele distributions of ABCG2 SNP markers in FE selection lines.
Table 4.8: Haplotype distribution for the three ABCG2 SNP markers (ss65824074:G>A, ss65824075:G>A and ss65824134:C>T) in FE selection lines.

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>Control line</th>
<th>Resistant line</th>
<th>Susceptible line</th>
<th>Peddrift analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAC*</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>AGC</td>
<td>42</td>
<td>58</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>GAC*</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>GAT</td>
<td>6</td>
<td>22</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>GGC</td>
<td>13</td>
<td>24</td>
<td>87</td>
<td>$P = 0.031$</td>
</tr>
</tbody>
</table>

* The AAC and GAC haplotypes were collapsed for peddrift analysis due to the small frequencies of each.

In addition to analysing the three SNPs individually, haplotypes were estimated from pedigree information by B. Auvray (AgResearch) using SNPhap (http://www-gene.cimr.cam.ac.uk/clayton/software/snphap.txt). Haplotype data provide more information about the combination of alleles observed in an individual by combining the genotypes for the biallelic SNP markers (reviewed in Romero et al., 2002). The haplotypes were treated as a single multi-allelic maker (as no recombination between the three SNPs was detected) and the haplotype frequencies were compared by calculation of chi-square statistics, and the significance level for this statistic was found using the peddrift simulation method (Dodds and McEwan, 1997). Overall there is a significant association of the ABCG2 haplotype ($P = 0.031$) with resistance to FE (Table 4.8). This association confirms ABCG2 is a candidate gene for resistance to FE.

Figure 4.13: Association studies of markers in FE selection lines. The markers (including the relative spacing) and their corresponding significance levels as determined by peddrift analysis are indicated.
4.4.5 Northern analysis of ABCG2

Increased expression of ABCG2 is known to confer resistance to xenobiotic compounds. This increase in expression can be modulated by a cis-acting regulatory region or by gene duplication (reviewed in Doyle and Ross, 2003). Therefore in light of the positive association of the ABCG2 locus with resistance to FE, Northern hybridisation was carried out to determine if there was any differential expression of ABCG2 in livers of resistant versus susceptible sheep under naïve and sporidesmin-dosed conditions.

The Northern blotting procedure was carried out as detailed in section 2.5.5. The ABCG2 probe was a 750 bp PCR product generated using primers OAR_F and OAR_R. The GAPDH probe was derived from cloned and sequence verified GAPDH cDNA. The probes were labelled as detailed in section 2.5.5.2. The sizes of the transcripts were determined by comparison with 0.24 - 9.5 Kb RNA Ladder (Invitrogen). The autoradiograph was scanned and analysed as detailed in 2.5.5.4. Total absorbance units of each band was determined and normalised to the total AU for GAPDH as an internal control, to correct for any variations in loading or transfer. Data was compared using one-way ANOVA with Fisher’s LSD post-hoc test.

Northern analysis detected three hybridisation bands. The most abundant band was band B which was ~ 2.1 kb size (Fig. 4.14). Since the size of the full-length transcript sequenced in this study was 2191 bp it is likely that the most abundant 2.2 kb band (band B) is the full length ABCG2 transcript.

Figure 4.14: (facing page) Northern analysis of ovine ABCG2 (A) Northern hybridisation showing expression of the ABCG2 mRNA in the livers of resistant and susceptible selection-line animals, both pre- and post-exposure to sporidesmin. The gender of the individual animals is indicated below the blot (M = male, F = female). Band B is the predominant transcript with a size consistent with the full length cDNA sequence obtained in this study (2.2 kb). The two remaining products labelled “A” and “C” are possible alternative transcripts and are sized 1.5 kb and 4 kb, respectively. (B) Northern blots were quantified by densitometry and expression normalised to control (GAPDH) expression. Expression levels are the mean ± SD for the four groups of animals: undosed resistant (n=3), undosed susceptible (n=3), dosed resistant (n=6) and dosed susceptible (n=5). Significance was assessed using one-way ANOVA. (C) Data for resistant and susceptible selection-line animals was pooled. Expression levels are the mean ± SD for the two groups of animals: undosed (n=6) and sporidesmin dosed (n=11). Significance was assessed by least squares (ANOVA) methods, * = P = 0.0165, *** P= 0.0009.
Chapter Four (Section B): Evaluation of ABCG2 as a candidate gene for FE

A.

<table>
<thead>
<tr>
<th></th>
<th>Naive Resistant</th>
<th>Naive Susceptible</th>
<th>Sporidesmin dosed Resistant</th>
<th>Sporidesmin dosed Susceptible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band A</td>
<td>R8 R9 R10</td>
<td>S8 S9 S10</td>
<td>R11 R12 R13 R14 R15 R16</td>
<td>S11 S12 S13 S14 S15</td>
</tr>
<tr>
<td>(4.0 kb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.1 kb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.5 kb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAPDH</td>
<td>M F F F</td>
<td>F M F M F M</td>
<td>F F M M</td>
<td>M F M F</td>
</tr>
<tr>
<td>(control)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B.

![Graph showing expression ratios for bands A, B, and C.]

C.

![Graph showing expression ratios for ABCG2 transcript.]

151
The other two bands were much less abundant and were approximately 4 kb (band A) and 1.5 kb (band B) in size. Comparison of these sizes with those for alternative transcripts annotated for the human ABCG2 gene in the ECgene database (Kim et al., 2005a; Kim et al., 2005b) showed that the size of the smaller product (band C) would be consistent with the size of an annotated alternative transcript of ABCG2 in humans. In some PCR reactions a smaller PCR product consistent with the size of the small putative alternative transcript was obtained. However, the significance of these alternative transcripts is not known. The third alternative transcript (band A) is much larger (~4000 bp): this is consistent with the size of one of the known human alternative ABCG2 transcripts which has a large 3' UTR of approximately 2 kb.

There is a high degree of variation between animals in ABCG2 expression in livers of resistant and susceptible sheep, and this variation is reflected in the relatively large standard deviations when the data were grouped. The expression ratios were log-transformed to stabilise the variance and then analysed by least squares (ANOVA) methods. The initial model included the main effects of sporidesmin exposure and genetic line along with their interaction. This analysis shows that there is no contribution of resistance or susceptibility to the expression of ABCG2 (Fig. 4.14B). However, exposure of animals to sporidesmin results in an increase in ABCG2 expression (Fig. 4.14C). Specifically, the most abundant transcript (band B) is differentially expressed being induced 1.6 fold in animals exposed to sporidesmin ($P = 0.0165$). The smaller, possibly alternatively spliced transcript, transcript C is induced 2.1 fold ($P = 0.0009$). Transcript A is not differentially expressed ($P = 0.3164$).

Further, while estrogen has recently been shown to modulate ABCG2 expression in cell culture (Imai et al., 2005), no sex specific effect was seen in ABCG2 expression in these sheep ($P = 0.147$). Although sample sizes were low (female $n= 11$, male $n = 6$).
4.3 DISCUSSION

Sporidesmin is known to be a substrate for the yeast PDR5 protein, as over-expression of PDR5 in *S. cerevisiae* conferred resistance to sporidesmin and deletion mutations resulted in sensitivity (Bissinger and Kuchler, 1994). The objective of this study was to identify a mammalian ortholog of the yeast PDR5 protein and assess the role of the ovine ortholog in FE.

Phylogenetic analyses were used to identify a mammalian ortholog of the yeast PDR5 protein. The ABC-domain sequences were extracted from the fully-sequenced genomes of five organisms, namely *S. cerevisiae*, *C. elegans*, *D. melanogaster*, *M. musculus* and *H. sapiens*. *C. elegans* and *D. melanogaster* were chosen because they bridge the evolutionary gap between *S. cerevisiae* and the mammalian species of interest, thus limiting spurious phylogenies.

A total of 376 domain sequences from these five organisms were included in the preliminary neighbour-joining phylogeny. Of these protein domains, 24 did not group with one of the six established mammalian sub-families. These sequences probably belong to the hypothetical H sub-family of transporters (Roth *et al.*, 2003) and were excluded from further analyses. The remaining domain sequences were re-aligned and subjected to neighbour-joining analysis, the resulting phylogenetic tree was broadly similar to published phylogenies (Dean *et al.*, 2001; Sheps *et al.*, 2004).

As function can be most confidently predicted from orthologous groups or pairs, the neighbour-joining phylogeny can make several functional predictions. For instance, the *S. cerevisiae* YDR091C protein forms an orthologous group with the human ABCE1 protein and the *C. elegans* 3O249 sequence. YDR091C is an iron sulphur protein that is required for translation initiation (Dong *et al.*, 2004), the human ABCE1 protein has recently been shown to also have a critical role in translation initiation (Chen *et al.*, 2006). It is therefore reasonable to suggest that the *C. elegans* 3O249 protein may also have a play a role in this process.

However, the phylogenetic analyses were not able to identify an orthologous group containing the yeast PDR5 protein. Although the phylogeny did indicate that both
the N- and C-terminal ABC domain sequences of the yeast PDR5 protein (and several other yeast transporters) were most closely related to the mammalian ABCG sub-family.

In general, the sequence identity between the two ATP-binding domains of a single full-transporter is greater than that between domains from two different transporters. This phenomenon is thought to reflect functional constraints or be a consequence of concerted evolution (Higgins, 1992). Unexpectedly the phylogeny places the two PDR5 ABC-domains in separate sub-clusters, indicating that the N- and C-terminal ABC domains share greater sequence similarity with other yeast ABC domains than they do to each other. This may indicate functional differences between the two domains or may be a result of the evolution of the PDR5 protein via domain switching.

To elucidate the relationship between the ABCG sub-family and the yeast PDR5 protein further, the 49 domains within this cluster were re-aligned and subjected to neighbour-joining and maximum parsimony analysis. A similar tree topology was obtained and no orthologous pairs or groups could be identified in this analysis. Therefore, the conclusion from the phylogenetic analysis was that the yeast PDR5 protein is most closely related to the mammalian ABCG sub-family of transporters, and the exact evolutionary relationships between these proteins could not be further delineated using this analysis. The use of phylogenetics to ascertain orthology in such diverse species as yeast and higher mammals is problematic, as it assumes that segmental gene duplication and gene conversion are not prevalent in any of the species. This is an issue when analysing sequence data from S. cerevisiae, which is known to have undergone whole genome duplication. This duplication has facilitated rapid sequence and functional divergence in S. cerevisiae proteins (Wolfe and Shields, 1997). It is therefore likely that the ABCG family actually represent paralogs of the yeast PDR5 protein rather than true orthologs.

As a candidate gene could not be confidently identified from the phylogeny alone, information concerning function, expression and polymorphisms of the mammalian ABCG sub-family were considered. In human the ABCG sub-family contains five members, while rodents have an additional transporter, ABCG3. Examination of the
chromosomal organisation surrounding the ABCG2 and ABCG3 loci in rodents indicates that the ABCG3 locus may have originated in rodents as a result of chromosomal rearrangement and repair mechanisms. Interestingly, the ABCG3 locus appears to be under positive selection and is not co-ordinately expressed with ABCG2 (Mickley et al., 2001) indicating that it may have diverged functionally from the ABCG2 protein. The rhesus macaque (Macaca mulatta) has an annotated ABCG3 locus, but based on sequence similarity and synteny in this chromosomal region, the ABCG3 locus is likely to be mis-annotated and should actually be annotated as ABCG2. No evidence for a ruminant ortholog of ABCG3 could be detected and it was concluded that the ruminant ABCG family is likely to consist of five members.

Of the five members of the ABCG sub-family, ABCG2 was selected for further investigation as it is known to function as a xenobiotic transporter (Doyle et al., 1998; Litman et al., 2000; Robey et al., 2001; Kim et al., 2002). Additionally, mice deficient in ABCG2 were sensitive to diet dependent photo-toxicity (Jonker et al., 2002), a phenotype similar to FE afflicted animals. Recently, ABCG2 has been shown to transport phylloerythrin which is the major photosensitising agent in FE (Robey et al., 2006). Therefore ABCG2 could conceivably contribute to resistance by preventing the photosensitisation and secondary damage caused by the disease. Moreover, the ABCG2 gene was shown by comparative mapping to be localised to sheep chromosome 6, to within a region with weak evidence for a FE QTL. This map position was subsequently confirmed by linkage mapping in the AgResearch international mapping flock (IMF). Thus ABCG2 was considered both as a functional and positional candidate for resistance to FE. Although, it should be noted that both ABCG5 and ABCG8 are likely to map to sheep chromosome 3 in an area of the genome with strong evidence for a QTL, making them theoretically better candidates for FE. However, both ABCG5 and ABCG8 proteins function in cholesterol homeostasis, specifically cholesterol efflux from the liver into the bile and they have never been shown to transport xenobiotics or be associated with drug resistance in any species (Kusuhara and Sugiyama, 2007). In addition, both ABCG5 and ABCG8 have been subsequently examined as candidate genes for FE, with no evidence found for association of ABCG8 and work is continuing on ABCG5 (S. H. Phua, pers. comm.).
As no ruminant sequence was available for ABCG2, it was amplified and sequenced in four animals. 5’ and 3’ RACE methods were used to obtain complete cDNA sequence. Both the nucleotide sequences and the predicted protein sequences of the ovine ABCG2 gene share high sequence identity with their mammalian counterparts. Additionally the domain and motif organization resemble that of the human and mouse ABCG2 proteins. Further support that the sequence obtained is the true ovine homolog to human ABCG2, and a possible paralog to the yeast PDR5, is their sharing of the unique ‘reverse’ orientation of the nucleotide binding domain to the transmembrane domains: this orientation is not found in any other group of human ABC transporters.

As mutations in the ABCG2 protein sequence are known to modulate the function of this protein, cDNA sequencing was used to identify changes in the amino acid sequence of ABCG2. The coding region of ABCG2 was sequenced in two resistant and two susceptible selection line sheep, and the region encompassing a known mutational hotspot was sequenced in sixteen animals. Overall, only two synonymous nucleotide changes in exons 6 and 9 were detected, both occurring in one of the susceptible sheep. These data suggest that there are no fixed differences in the primary structure of the ABCG2 protein between these resistant and susceptible animals. However, it has been recently discovered that synonymous polymorphisms in ABC-transporters, specifically ABCB1, may actually affect the activity of the active protein. It is hypothesised that this effect may be mediated by rare codon usage, which could affect co-translational protein folding resulting in altered protein function (Kimchi-Sarfaty et al., 2007). It is therefore possible that the synonymous polymorphisms found in this study may contribute to altered ABCG2 activity and disease resistance. However, this is unlikely as no differences in the allele frequencies for these polymorphisms were seen between resistant and susceptible animals (Table 4.7).

The two synonymous SNPs discovered in the ABCG2 coding sequence together with the SNP used for linkage mapping were genotyped in a number of resistant, susceptible and control selection-line sheep. The allele frequencies for each marker were compared between the resistant, susceptible and control selection lines using
Chi-squared statistics. The significance level for these statistics was determined using peddrift (Dodds and McEwan, 1997). A significant difference in the allele frequency distribution of the intron 4 SNP (ss65824074:G>A) was detected, with the G allele being present at a significantly higher frequency in the susceptible line. Estimation of haplotypes using SNPHap (Clayton, 2002) also revealed a significant difference in the distribution of haplotype frequencies, similar to the analysis of the single SNPs.

While the association detected between the intronic SNP marker and susceptibility to FE is statistically significant, it is not as highly significant as commercial markers for other ruminant traits (Van Eenennaam et al., 2007). Nor is it as highly associated as other genes already found to be involved in FE (Phua et al., 1999). It would therefore be of interest to investigate this locus further either by generating and testing more SNPs or microsatellite markers in this region. Alternatively, the association of these SNPs with FE should be tested in an independently selected flock of sheep.

As the association studies indicated a weak but significant association of the ABCG2 locus with FE, ABCG2 was further investigated by examining the expression profile of ABCG2 transcripts in FE resistant and susceptible animals. Three hybridisation bands were detected, the most abundant of which is consistent in size with the full length transcript obtained in this study (~2.1 kb). The other two hybridisation bands are possible alternative transcripts (~4 kb and ~1.5 kb) broadly consistent with those annotated in the ECgene database (Kim et al., 2005a; Kim et al., 2005b) but the functions of these transcripts are unknown.

Quantification of the Northern blot must be interpreted with caution as it is complicated by the limited linear range of the film and the apparent degradation of some samples. Therefore the differences in expression should be taken as an indicator of trends, rather than as a strict numerical representation of gene expression. Analysis of the blot shows that there is no difference in levels of transcription for the putative 2.1 kb ABCG2 transcript between the resistant and susceptible selection-line sheep. Interestingly, there was a 1.6-fold increase in the expression of the main transcript (band B) and a 2.1 fold increase in the expression of the minor transcript (band C) was detected in animals exposed to sporidesmin as
compared to naïve animals. The induction of these transcripts following exposure to sporidesmin indicates that sporidesmin induces a physiological process in the livers of these sheep that is shared between both resistant and susceptible sheep.

The expression of ABCG2 in stem cells is known to be modulated by the transcription factor as HIF-1 (hypoxia-inducible factor 1) (Krishnamurthy et al., 2004) and HIF-1 is also expressed in liver, it is therefore possible that the observed increase in ABCG2 transcription may be via HIF-1 activation in response to a cytotoxic hypoxic state induced by sporidesmin. Moreover, sporidesmin mediated induction of ABCG2 may enable it to act as a phase III defence mechanism by accelerating removal of the xenobiotics via enhanced excretion into the bile (reviewed by Krishnamurthy and Schuetz, 2006).

It is known that ABCG2 in mice can not only modulate the hepato-biliary excretion of substrates but can also modulate intestinal absorption (Staud and Pavek, 2005). Thus it is therefore possible that expression in tissues other than the liver, such as the intestine, may be of interest. It is possible that a polymorphism in the promoter of ABCG2 may alter the transcriptional profile in different tissues due to tissue specific expression of transcription factors.

Further, mutations leading to cryptic splicing could also not be excluded, although abnormal transcripts were not consistently detected in selection-line animals. It is possible that the critical cryptic splice variant is present at a low level and is therefore not detectable by a relatively insensitive technique such as Northern hybridisation. Such a cryptic splice variant may introduce a pre-mature stop codon thus causing the transcript to be rapidly degraded by non-sense mediated decay (Stamm et al., 2005). It is also possible that a small change in the expression of a cryptic splice variant may mediate a larger physiological effect due to alteration of protein turn-over or by affecting mRNA stability (Sureau et al., 2001).

Besides levels of transcription and translation, expression of ABCG2 also depends on proper localisation of the protein to the cell surface, and in polarised cells to the right membrane compartment. Therefore any perturbation in the post-translational processing of the ABCG2 protein may affect its activity (Sarkadi et al., 2006).
Finally, it cannot be ruled out that an unidentified locus in linkage disequilibrium with the ABCG2 locus is the causative gene and the target of selection in the genetic lines. Linkage disequilibrium is reported to extend up to tens of centimorgans in sheep (McRae *et al.*, 2002) and the QTL region detected in this study covers at least 40 cM. The corresponding syntenic region in humans is relatively gene dense, with 11 annotated genes in the megabase of DNA sequence surrounding the ABCG2 locus.
Chapter Five

cDNA Microarrays as a Tool for Candidate Gene Identification.

5.1 INTRODUCTION

Microarray technology has become an increasingly useful and important tool in biological research. Microarrays facilitate the simultaneous quantification of expression levels for thousands of genes in a biological sample. Identifying patterns of gene expression in the sample and grouping of genes into functional classes can provide insight into the biological function and relevance of these genes.

5.1.1 cDNA microarray technology

Microarray technology was developed in the mid-1990s (Schena et al., 1995) and continues to be refined and adapted. The technology has been used extensively in biological research, and this is in part due to significant technological advances such as the use of glass as a nonporous solid support and the development of fluorescence-hybridization detection (Schena et al., 1995; Lockhart et al., 1996; Schena et al., 1996).

Nucleic acid microarrays use either short oligonucleotides (15 – 25 nucleotides), long oligonucleotides (50 – 120 nucleotides), or PCR-amplified cDNAs (100 – 3000 bp) as array elements. The PCR-amplified cDNAs produce strong signals and high specificity (DeRisi et al., 1996). Additionally, the cDNA elements are readily obtained from cDNA / EST (expressed sequence tag) libraries and are typically used for organisms for which only a limited part of the whole genome information is available.

The goal of a microarray experiment is to extract meaningful biological information from the expression data. This information can include patterns of relative gene expression, participation in biochemical pathways and “genetic signatures” (Lettieri, 2006) of the disease state or altered biological state.
5.1.2 Microarrays and complex disease traits

It is commonly thought that individual genes associated with complex diseases are unlikely to fully explain disease pathogenesis or variation within these pathogenic mechanisms (Shai, 2006). The underlying basis of a complex disease is likely to be dictated by the abnormal expression of tens to hundreds of genes (Friedmann, 1992). Therefore gene expression profiling can be a powerful analytical tool for dissecting complex diseases (Shai, 2006). Microarray analysis facilitates the discovery of new genes as well as metabolic and regulatory pathways involved in complex diseases and may identify new treatment targets (Lyons, 2002).

Microarray technology does have several limitations, including some technical aspects such as probe sensitivity and data condensation (Yuen et al., 2002). It is also imperative to have a sufficient number of samples to ensure adequate statistical power. But perhaps the major limitation, especially when working with an information-poor species such as sheep, is data interpretation. It can be difficult to put the hundreds of differentially expressed genes into a meaningful biological context and to generate biological hypotheses from the data. Microarray data can be used to complement positional cloning strategies and to identify novel candidates for disease susceptibility loci (Niculescu et al., 2000). Microarray analysis has been used extensively for studying complex disease traits in humans (Shai, 2006) including various forms of cancer such as bowel (reviewed in Alvarado et al., 2006) and breast cancers (reviewed in Modlich et al., 2006), Alzheimer’s disease (reviewed in Blalock et al., 2005), Parkinson’s disease (reviewed in Miller and Federoff, 2006) and obesity (reviewed in Baranova et al., 2005).

This technology has also been successfully applied to complex diseases in livestock such as mastitis resistance in cattle (Burton et al., 2001) and more recently to parasite resistance in sheep (Diez-Tascon et al., 2005; Keane et al., 2006; Keane et al., 2007). The integration of microarray expression data with genotype data has been used to identify a genetic variation in DQA1 that is associated with parasite resistance (Keane et al., 2007). These studies demonstrate the utility of this approach for investigating complex disease traits.
5.1.3 Toxicogenomics

Toxicogenomics aims to study the complex interaction between an organism’s genome, their sensitivity to toxins, and disease processes. The field of toxicogenomics has progressed rapidly since the development of microarray technology (Afshari et al., 1999). Gene expression studies are a useful and sensitive indicator of exposure to a toxin as well as disease state and cellular metabolism. Microarrays are a unique way of characterizing how cells and organisms adapt to changes in the external environment (Lettieri, 2006). The measurement of gene expression following exposure to a chemical or toxin can provide information about the mechanism of toxicity by establishing a signature of gene expression changes both in vitro (Burczynski et al., 2000; Waring et al., 2001) and in vivo (Hamadeh et al., 2002).

Toxicogenomics is now a rapidly evolving field of study and numerous publications have evaluated the potential uses of microarray technology (Simmons and Portier, 2002; Tennant, 2002; Ulrich and Friend, 2002) as well as illustrating the practical use of gene expression profiling in toxicology (Bartosiewicz et al., 2001; Bulera et al., 2001; Waring et al., 2001; Hamadeh et al., 2002). The plethora of experimental data has shown that patterns of gene expression relating to biological pathways are robust enough to allow insight into mechanisms of toxicity, and gene expression data can provide meaningful information about the organ systems affected by the toxicity (reviewed in Pennie et al., 2004). This latter finding is particularly important as gene expression profiling has been used to show that the specific genes repressed or induced upon exposure to a toxic stress vary depending on the cell type as well as the type of toxicants to which the cells were exposed (Troester et al., 2004).

5.1.4 Biomarkers

Biomarkers are biochemical features that can be used to track disease progression or treatment. Microarray technology has facilitated biomarker identification as they provide a rapid, efficient, and systematic approach for biomarker discovery (Allgayer et al., 1997; Brien et al., 1998). Good biomarkers must be highly accurate and reliable predictors for either disease diagnosis or prognosis (Xiong et al., 2001). Alternatively they may identify putative therapeutic targets or aid in understanding the basic biology of the disorder (Chow et al., 2001).
Using microarray technology to identify possible biomarkers is particularly relevant when examining the gene expression of liver tissue. This is because the liver is the primary site of synthesis for many serum proteins and specific serum biomarkers are known to reflect the status of the entire liver (Regev et al., 2002; Ratziu et al., 2005).

While the focus in FE research has been to discover and develop genetic markers for resistance to FE (refer section 1.7), biomarkers are considered a viable alternative to genetic markers for facilitating marker assisted selection in FE. A recent study examined serum levels of several antioxidant enzymes including catalase, superoxide dismutase, glutathione peroxidase and glutathione reductase as possible biomarkers (Hohenboken et al., 2004). Significant differences were detected between the selection-lines for superoxide dismutase (lower in resistant animals), catalase (higher in resistant animals) and glutathione peroxidase (higher in resistant animals). However, the differences between the genetic lines were not profound enough to facilitate the use of these proteins as biomarkers (Hohenboken et al., 2004).

In addition to discovering biomarkers for resistance and susceptibility, there may also be some utility for biomarkers that can indicate sporidesmin exposure or disease progression. Such biomarkers could be utilised for routine surveillance of flocks in high-risk areas, or be used to determine how severely the animals have been afflicted. Currently two serum biomarkers are used to track disease progression, including gamma-glutamyltransferase (GGT) and glutamate dehydrogenase (GDH). Both of these proteins are expressed in many tissues, but elevated levels in the serum are indicative of liver injury. However, these markers are only elevated in the late stages of liver disease and also are not specific for sporidesmin toxicity (reviewed in Reichling and Kaplan, 1988). Therefore it may be possible, using expression profiling, to discover more sensitive or reliable biomarkers.
5.1.5 Aims and Rationale.

The aim of this study was to use microarray technology to compare the hepatic expression profiles of resistant and susceptible sheep, both pre- and post-exposure to sporidesmin. It was hoped that this analysis would reveal new candidate genes and candidate biochemical pathways related to resistance to FE disease.

Also, as the liver is the site of biosynthesis of many serum proteins, it was hoped that expression profiling of liver tissue might identify putative serum biomarkers. These biomarkers would be useful for predicting resistance to FE or for detecting sporidesmin exposure or for tracking the progression of the disease.
Chapter Five: cDNA Microarrays as a Tool for Candidate Gene Identification

5.2 RESULTS

5.2.1 Description of the Microarray Slide

In accordance with MIAME (Minimum Information about a Microarray Experiment) guidelines (Brazma et al., 2001) a full description of the array, including EST clones, position on the array, NCBI accession numbers (where available), contig membership, and best human RefSeq hit (as obtained by BLAST (Altschul et al., 1990)) are included as an Excel file (Array description file_MIAME.xls) on the attached CD (Appendix B). The MIAME standards were developed at the European Bioinformatics institute (EBI) and describe the minimum information required to ensure that microarray data can be easily interpreted and that results derived from its analysis can be independently verified.

The microarray work presented here was done early in the AgResearch microarray program (in 2001 / 2002) and the array used in this study was synthesised from an in-house bovine liver EST library (cloned in the pBK-CMV phagemid vector (Stratagene)) which contained a total of 6650 ESTs (manufactured by AgResearch in conjunction with Genesis Corporation). Of these ESTs 1550 were selected for inclusion on the microarray. These ESTs included two ABC transporters (ABCE1 and ABCC6). In general, functional categories that were over-represented on the array included genes involved in metabolism (including fatty acid metabolism and amino acid metabolism), biosynthetic pathways (including cellular biosynthesis and protein biosynthesis) and also complement activation and the immune system.

The EST sequences were PCR amplified using universal primers T7 and T3. The PCR products were printed onto Poly-L-Lysine coated glass slides (Sigma-Aldrich) at the University of Otago Genomics Facility.

The 1550 EST sequences on the array were clustered into a total of 1447 contigs, of which 1289 were represented only once on the array. Of the 1550 EST sequences 1409 have been publicly released and are available for download from the NCBI and 1058 had significant BLAST hits to the human RefSeq database which represented a total of 859 unique genes represented on the microarray.
5.2.2 Tissue isolation and RNA extraction

Liver tissue samples for microarray analysis were obtained from four selection-line animals; two from the resistant line (R16 and R10) and two from the susceptible line (S11 and S8) (refer to Tables 2.1 and 2.2). All four animals were female, but differed quite significantly in age (R16 = 3 years, R10 = 2-3 months, S16 = 6 months, S11 = 3 years). Sporidesmin-dosed animals were exposed to sporidesmin orally (resistant animal: 0.4 mg/kg and susceptible animal 0.15 mg/kg). Animals were sacrificed 24 hours after sporidesmin exposure and the livers snap frozen in liquid nitrogen.

5.2.3 Experimental design

Experimentally, a modified factorial design was used with 22 microarray hybridisations incorporating four direct comparisons; naïve susceptible vs. naïve resistant, naïve susceptible vs. sporidesmin-dosed susceptible, naïve resistant vs. sporidesmin-dosed resistant and sporidesmin-dosed resistant vs. sporidesmin-dosed susceptible (Fig. 5.1). Each comparison consisted of a minimum of five individual

![Figure 5.1: Schematic of the microarray experimental design. Each of the four experimental groups (sporidesmin-dosed resistant, sporidesmin-dosed susceptible, naïve resistant and naïve susceptible) were represented by a single animal. The comparisons indicated by the black arrows represent experimental comparisons while the blue and red arrows both indicate statistical comparisons. The animal codes are given in parentheses (refer to section 2.2 for details) and are provided to facilitate cross-referencing with the Northern analysis.](image-url)
Chapter Five: cDNA Microarrays as a Tool for Candidate Gene Identification

slides and incorporated dye swaps to minimise the bias related to dye incorporations.

A further two statistical comparisons were made by combining data from naïve and dosed animals (Fig. 5.1 comparison A) to determine which genes were modulated by sporidesmin exposure, and the second by combining data from resistant and susceptible animals (Fig. 5.1 comparison B) to determine which genes were differentially expressed due to the genetic line of the animals irrespective of exposure to sporidesmin.

5.2.4 Microarray hybridisations and post-processing

The RNA was isolated as detailed in section 2.5.1. The RNA was reverse transcribed, labelled and the microarrays hybridised as detailed in 2.5.8. The hybridised slides were washed and scanned using a ScanArray 5000 (Packard Biosciences), and the dual images were collected in TIFF format. The combination and processing of images were performed using the GenePix Pro software (Axon Instruments). An example of a scanned and processed slide is shown in Fig. 5.2.

5.2.4.1 Quality control

Slides that showed high background, widespread irregular spot morphology, unexpectedly low signal or comet tails were discarded. Within a slide bad spots were flagged and were excluded from the analysis. For statistical analysis data was

![Figure 5.2: Pseudo-coloured image of a hybridised microarray slide. The slide consists of 1550 ESTS arrayed as tandem duplicates. As much as possible the order of the ESTs was randomised, although this was limited by 96 well plate formats and the robotics capability at the time (Hyndman, pers. comm.).](image-url)
exported as a GPR file which can be manipulated in Excel. The GPR files were checked to ensure the correct EST assignments had been made prior to statistical analysis.

5.2.5 Statistical analysis
A very small number of animals \((n = 4)\) were used for the microarray hybridisations. The small number of animals meant that some of the statistical comparisons had one degree of freedom. To combat this, the normalised data from all 22 slides was combined and a number of average statistics calculated. In particular, the expression data for each individual animal were calculated as the difference between the expression data for an individual animal and the mean of all four animals. Therefore genes are either over or under-represented in a particular animal as compared to the average expression across all four animals. This calculation uses all the hybridisation data to estimate values for each animal so that the effect of a single animal is not just a mean of the slides that that animal is represented on, effectively increasing the degrees of freedom in the statistical analyses (Baird, pers. comm.).

Statistical analysis was performed by David Baird (AgResearch, Lincoln) using restricted maximum likelihood (REML) algorithms (Baird et al., 2004) implemented by GenStat (refer to section 2.5.8.5).

Briefly, the ratio of Cy5 to Cy3 fluorescence was calculated and \(\log_2\) transformed. The data were then adjusted for various parameters and normalised. The ESTs were sorted based on the modified \(T\) value (Smyth, 2004) for the log ratio of the mean. ESTs were counted as differentially expressed where the \(P\) value from the associated \(T\)-test was less than 0.05 and where the EST exhibited a fold change of greater than 1.1 or less than 0.9.

Statistical comparison of the four individual animals used in the microarray experiments yielded a total of 159 differentially expressed ESTs (Tables 5.2 – 5.5, summarised in Table 5.1). These 159 ESTs correspond with 145 unique differentially expressed ESTs, 14 were found to be differentially expressed in two of the four animals. Of the 145 unique ESTs reported as differentially expressed 116
had corresponding human RefSeq sequences and the remainder were classified as ‘unknown’. The 116 ESTs corresponded with 106 unique RefSeqs.

On the whole, relatively low numbers of differentially expressed genes were reported for the four individual animals as compared to the mean of these animals (summarised in Table 5.1). However, the one notable exception was the naïve resistant animal (Table 5.4), where a total of 97 ESTs were differentially expressed (37 down-regulated and 60 up-regulated). Within the set of 97 differentially expressed EST sequences five RefSeq sequences were represented by more than one EST. In all cases the differential expression reported from these EST sequences was consistent.

When the data from the two naïve animals (R10 and S8) were combined and compared with the combined data obtained from sporidesmin-dosed animals (R16 and S11) 15 ESTs were identified as being differentially expressed (Table 5.6). These 15 ESTs represented genes that were differentially expressed in response to sporidesmin exposure but independent of the genetic background of the animals (independent of resistance or susceptibility). These 15 ESTs correspond with 11 known RefSeq genes.

The final comparison involved combining the data from resistant animals (R10 and R16) and comparing it with susceptible animals (S8 and S11) in order to determine gene expression that was influenced by the genetics of the animals but not by their exposure to sporidesmin. This analysis revealed only three ESTs, two down-regulated in resistant animals as compared to susceptible and one up-regulated (Table 5.7).
Table 5.2: ESTs differentially expressed in sporidesmin-dosed resistant sheep R16.

<table>
<thead>
<tr>
<th>Bovine EST sequence</th>
<th>Human Refseq:</th>
<th>Gene Description</th>
<th>Gene Symbol</th>
<th>Cytogenetic band</th>
<th>Contig BLAST</th>
<th>Fold Change</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLIB001013HT</td>
<td>NM_002946</td>
<td>replication protein A2, 32kDa</td>
<td>RPA2</td>
<td>1p35</td>
<td>0</td>
<td>0.70</td>
<td>1.86E-02</td>
</tr>
<tr>
<td>BLIB004547HT</td>
<td>NM_013237</td>
<td>Homo sapiens px19-like protein (PX19), mRNA LanC lantibiotic synthetase component C-like 1 (bacterial)</td>
<td>PX19</td>
<td>5q35.3</td>
<td>0</td>
<td>0.84</td>
<td>1.20E-02</td>
</tr>
<tr>
<td>BLIB0099080HT</td>
<td>NM_006055</td>
<td></td>
<td>LANCL1</td>
<td>2q33-q35</td>
<td>2E-40</td>
<td>0.85</td>
<td>2.48E-02</td>
</tr>
<tr>
<td>BLIB001343HT</td>
<td>NM_00102510</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.86</td>
<td>6.25E-03</td>
</tr>
<tr>
<td>BLIB003710HT</td>
<td>5</td>
<td>casin kinase 1, alpha 1</td>
<td>CSNK1A1</td>
<td>5q32</td>
<td>0</td>
<td>0.87</td>
<td>2.79E-03</td>
</tr>
<tr>
<td>BLIB001129HT</td>
<td>NM_006684</td>
<td>complement factor H-related 4</td>
<td>CFHR4</td>
<td>1p32</td>
<td>6E-15</td>
<td>0.87</td>
<td>8.10E-03</td>
</tr>
<tr>
<td>BLIB004786HT</td>
<td>NM_002310</td>
<td>leukemia inhibitory factor receptor</td>
<td>LIFR</td>
<td>5p13-p12</td>
<td>3E-17</td>
<td>0.88</td>
<td>3.58E-04</td>
</tr>
<tr>
<td>BLIB999003HT</td>
<td>NM_020978</td>
<td>amylase, alpha 2B; pancreatic G protein-coupled receptor, family C, group 5, member B</td>
<td>AMY2B</td>
<td>1p21</td>
<td>4E-41</td>
<td>0.90</td>
<td>2.25E-03</td>
</tr>
<tr>
<td>BLIB999041HT</td>
<td>NM_016235</td>
<td></td>
<td>GPRC5B</td>
<td>16p12</td>
<td>2E-70</td>
<td>0.90</td>
<td>3.96E-04</td>
</tr>
<tr>
<td>BLIB006550HT</td>
<td>NM_001710</td>
<td>complement factor B</td>
<td>CFB</td>
<td>6p21.3</td>
<td>0</td>
<td>1.10</td>
<td>5.23E-02</td>
</tr>
<tr>
<td>BLIB005245HT</td>
<td>NM_002276</td>
<td>keratin 19</td>
<td>KRT19</td>
<td>17q21.2</td>
<td>0</td>
<td>1.10</td>
<td>1.69E-02</td>
</tr>
<tr>
<td>BLIB005543HT</td>
<td>NM_018231</td>
<td>protein kinase, cAMP-dependent, regulatory, type II, alpha</td>
<td>PRKAR2A</td>
<td>3p21.3-p21.2</td>
<td>6E-11</td>
<td>1.11</td>
<td>5.34E-02</td>
</tr>
<tr>
<td>BLIB004837HT</td>
<td>NM_004157</td>
<td>complement component 5</td>
<td>C5</td>
<td>9q33-q34</td>
<td>0</td>
<td>1.11</td>
<td>1.67E-03</td>
</tr>
<tr>
<td>BLIB001633HT</td>
<td>NM_001735</td>
<td>coagulation factor X</td>
<td>F10</td>
<td>13q34</td>
<td>0</td>
<td>1.12</td>
<td>1.68E-03</td>
</tr>
<tr>
<td>BLIB006027HT</td>
<td>NM_022461</td>
<td>5-azacytidine induced 2</td>
<td>AZI2</td>
<td>3p24.1</td>
<td>0</td>
<td>1.12</td>
<td>3.28E-03</td>
</tr>
<tr>
<td>BLIB006220HT</td>
<td>NM_006533</td>
<td>melanoma inhibitory activity</td>
<td>MIA</td>
<td>7q22-13.33</td>
<td>0</td>
<td>1.13</td>
<td>1.88E-02</td>
</tr>
<tr>
<td>BLIB008561HT</td>
<td>NM_013397</td>
<td>chromosome 6 open reading frame 49</td>
<td>C6orf49</td>
<td>6p21.31</td>
<td>1E-117</td>
<td>1.16</td>
<td>3.56E-02</td>
</tr>
<tr>
<td>BLIB002361HT</td>
<td>NM_020213</td>
<td>poly (ADP-ribose) polymerase family, member 6</td>
<td>PARP6</td>
<td>15q23</td>
<td>4E-27</td>
<td>1.19</td>
<td>2.88E-02</td>
</tr>
<tr>
<td>Bovine EST sequence</td>
<td>Human Refseq:</td>
<td>Gene Description</td>
<td>Gene Symbol</td>
<td>Cytogenetic band</td>
<td>Contig BLAST</td>
<td>Fold Change</td>
<td>P value</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------</td>
<td>------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>BLIB006344HT</td>
<td></td>
<td>leukocyte cell-derived chemotaxin 2</td>
<td>LECT2</td>
<td>5q31.1-q32</td>
<td>1E-105</td>
<td>0.85</td>
<td>1.46E-04</td>
</tr>
<tr>
<td>BLIB001262HT</td>
<td>NM_002302</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.88</td>
<td>1.13E-06</td>
</tr>
<tr>
<td>BLIB004438HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.88</td>
<td>7.87E-03</td>
</tr>
<tr>
<td>BLIB003925HT</td>
<td>NM_015459</td>
<td>Homo sapiens DKFZP564J0863 protein</td>
<td>DKFZP564J</td>
<td>0863</td>
<td></td>
<td>0.89</td>
<td>6.43E-03</td>
</tr>
<tr>
<td>BLIB005223HT</td>
<td></td>
<td>arginase, liver</td>
<td>ARG1</td>
<td>6q23</td>
<td></td>
<td>0.90</td>
<td>8.61E-04</td>
</tr>
<tr>
<td>BLIB003472HT</td>
<td>NM_000045</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.90</td>
<td>1.34E-02</td>
</tr>
<tr>
<td>BLIB001901HT</td>
<td>NM_015331</td>
<td>nicastrin</td>
<td>NCSTN</td>
<td>1q22-q23</td>
<td></td>
<td>1.10</td>
<td>4.20E-04</td>
</tr>
<tr>
<td>BLIB006341HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.11</td>
<td>2.47E-05</td>
</tr>
<tr>
<td>BLIB005698HT</td>
<td>NM_016255</td>
<td>family with sequence similarity 8, member A1</td>
<td>FAM8A1</td>
<td>6p22-p23</td>
<td>1E-29</td>
<td>1.11</td>
<td>2.25E-03</td>
</tr>
<tr>
<td>BLIB007850HT</td>
<td>NM_014899</td>
<td>Rho-related BTB domain containing 3</td>
<td>RHOBTB3</td>
<td>5q15</td>
<td></td>
<td>1.12</td>
<td>2.75E-03</td>
</tr>
<tr>
<td>BLIB008563HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.13</td>
<td>1.89E-02</td>
</tr>
<tr>
<td>BLIB001902HT</td>
<td>NM_080417</td>
<td>septin 4</td>
<td>38231</td>
<td>17q22-q23</td>
<td></td>
<td>1.15</td>
<td>4.82E-04</td>
</tr>
<tr>
<td>BLIB001616HT</td>
<td>NM_199002</td>
<td>Rho guanine nucleotide exchange factor (GEF) 1</td>
<td>ARHGEF1</td>
<td>19q13.13</td>
<td></td>
<td>1.21</td>
<td>4.83E-06</td>
</tr>
</tbody>
</table>
Table 5.4: ESTs differentially expressed in naïve resistant sheep R10.

<table>
<thead>
<tr>
<th>Bovine EST sequence</th>
<th>Human Refseq:</th>
<th>Gene Description</th>
<th>Gene Symbol</th>
<th>Cytogenetic band</th>
<th>Contig BLAST</th>
<th>Fold Change</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLIB005485HT</td>
<td>NM_018834</td>
<td>matrin 3</td>
<td>MATR3</td>
<td>5q31.2</td>
<td>0</td>
<td>0.79</td>
<td>1.34E-02</td>
</tr>
<tr>
<td>BLIB005486HT</td>
<td>NM_013237</td>
<td>Homo sapiens px19-like protein (PX19), mRNA</td>
<td>PX19</td>
<td>5q35.3</td>
<td>0</td>
<td>0.80</td>
<td>5.13E-04</td>
</tr>
<tr>
<td>BLIB006122HT</td>
<td>NM_013237</td>
<td>matrin 3</td>
<td>MATR3</td>
<td>5q31.2</td>
<td>0</td>
<td>0.81</td>
<td>2.17E-08</td>
</tr>
<tr>
<td>BLIB005223HT</td>
<td>NM_013237</td>
<td>matrin 3</td>
<td>MATR3</td>
<td>5q31.2</td>
<td>0</td>
<td>0.81</td>
<td>9.59E-07</td>
</tr>
<tr>
<td>BLIB005010HT</td>
<td>NM_013237</td>
<td>matrin 3</td>
<td>MATR3</td>
<td>5q31.2</td>
<td>0</td>
<td>0.84</td>
<td>5.29E-06</td>
</tr>
<tr>
<td>BLIB006244HT</td>
<td>NM_000667</td>
<td>alcohol dehydrogenase 1A (class I), alpha polypeptide</td>
<td>ADH1A</td>
<td>4q21-q23</td>
<td>3E-87</td>
<td>0.84</td>
<td>4.50E-07</td>
</tr>
<tr>
<td>BLIB003129HT</td>
<td>NM_018834</td>
<td>matrin 3</td>
<td>MATR3</td>
<td>5q31.2</td>
<td>0</td>
<td>0.85</td>
<td>1.64E-02</td>
</tr>
<tr>
<td>BLIB002951HT</td>
<td>NM_006708</td>
<td>glyoxalase I</td>
<td>GLO1</td>
<td>6p21.3-p21.1</td>
<td>0</td>
<td>0.86</td>
<td>5.17E-04</td>
</tr>
<tr>
<td>BLIB001262HT</td>
<td>NM_002302</td>
<td>leukocyte cell-derived chemotaxin 2</td>
<td>LECT2</td>
<td>5q31.1-q32</td>
<td>1E-105</td>
<td>0.86</td>
<td>2.61E-07</td>
</tr>
<tr>
<td>BLIB003935HT</td>
<td>NM_015459</td>
<td>Homo sapiens DKFZP564J0863 protein ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)</td>
<td>DKFZP564J0</td>
<td>11q13.1</td>
<td>0</td>
<td>0.87</td>
<td>1.96E-03</td>
</tr>
<tr>
<td>BLIB004918HT</td>
<td>NM_003348</td>
<td>mannosidase, alpha, class 1A, member 2</td>
<td>ST3GAL4</td>
<td>11q23-q24</td>
<td>0</td>
<td>0.88</td>
<td>7.88E-06</td>
</tr>
<tr>
<td>BLIB007494HT</td>
<td>NM_006278</td>
<td>ST3 beta-galactoside alpha-2,3-sialyltransferase 4</td>
<td>ST3GAL4</td>
<td>11q23-q24</td>
<td>0</td>
<td>0.88</td>
<td>4.15E-03</td>
</tr>
<tr>
<td>BLIB0990022HT</td>
<td>NM_006699</td>
<td>peptidylprolyl isomerase A (cyclophilin A)</td>
<td>MAN1A2</td>
<td>1p13</td>
<td>5E-39</td>
<td>0.88</td>
<td>2.89E-05</td>
</tr>
<tr>
<td>BLIB006170HT</td>
<td>NM_021130</td>
<td>peptidylprolyl isomerase A (cyclophilin A)</td>
<td>PPIA</td>
<td>7p13-p11.2</td>
<td>0</td>
<td>0.89</td>
<td>3.75E-03</td>
</tr>
<tr>
<td>BLIB004795HT</td>
<td>NM_005410</td>
<td>selenoprotein P, plasma, 1</td>
<td>SEPP1</td>
<td>5q31</td>
<td>0</td>
<td>0.89</td>
<td>2.75E-03</td>
</tr>
<tr>
<td>BLIB001263HT</td>
<td>NM_002225</td>
<td>selenoprotein P, plasma, 1</td>
<td>SEPP1</td>
<td>5q31</td>
<td>0</td>
<td>0.89</td>
<td>2.75E-03</td>
</tr>
<tr>
<td>BLIB004216HT</td>
<td>NM_182603</td>
<td>selenoprotein P, plasma, 1</td>
<td>SEPP1</td>
<td>5q31</td>
<td>0</td>
<td>0.89</td>
<td>2.75E-03</td>
</tr>
<tr>
<td>BLIB003789HT</td>
<td>NM_003407</td>
<td>selenoprotein P, plasma, 1</td>
<td>SEPP1</td>
<td>5q31</td>
<td>0</td>
<td>0.89</td>
<td>2.75E-03</td>
</tr>
<tr>
<td>BLIB001648HT</td>
<td>NM_017768</td>
<td>leucine rich repeat containing 40</td>
<td>LRRRC40</td>
<td>1p31.1</td>
<td>0</td>
<td>0.89</td>
<td>1.07E-02</td>
</tr>
<tr>
<td>BLIB001572HT</td>
<td>NM_000186</td>
<td>complement factor H</td>
<td>CFH</td>
<td>1q32</td>
<td>3E-11</td>
<td>0.89</td>
<td>4.10E-04</td>
</tr>
<tr>
<td>BLIB006228HT</td>
<td>NM_006759</td>
<td>UDP-glucose pyrophosphorylase 2</td>
<td>UGP2</td>
<td>2p14-p13</td>
<td>0</td>
<td>0.89</td>
<td>2.02E-03</td>
</tr>
<tr>
<td>BLIB004214HT</td>
<td>NM_198947</td>
<td>family with sequence similarity 111, member B</td>
<td>FAM111B</td>
<td>11q12.1</td>
<td>4E-51</td>
<td>0.90</td>
<td>2.56E-04</td>
</tr>
<tr>
<td>BLIB001728HT</td>
<td>NM_003143</td>
<td>single-stranded DNA binding protein 1</td>
<td>SSBP1</td>
<td>7q34</td>
<td>0</td>
<td>0.90</td>
<td>2.06E-02</td>
</tr>
<tr>
<td>BLIB003472HT</td>
<td>NM_000045</td>
<td>arginase, liver</td>
<td>ARG1</td>
<td>6q23</td>
<td>0</td>
<td>0.90</td>
<td>1.25E-02</td>
</tr>
<tr>
<td>BLIB009137HT</td>
<td>NM_00101084</td>
<td>Homo sapiens hypothetical protein LOC123876</td>
<td>LOC123876</td>
<td>16p12.3</td>
<td>0</td>
<td>0.90</td>
<td>3.78E-04</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene Name</td>
<td>Chromosome</td>
<td>Gene Id</td>
<td>Description</td>
<td>Final</td>
<td>ID</td>
<td>q-value</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------</td>
<td>----</td>
<td>----------</td>
</tr>
<tr>
<td>BLIB008139HT</td>
<td>NDRG family member 3</td>
<td>20q11.21</td>
<td>NM_032013</td>
<td>NDRG3</td>
<td>0.90</td>
<td>1.78E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB007045HT</td>
<td>apolipoprotein A-I</td>
<td>11q23-q24</td>
<td>NM_000039</td>
<td>APOA1</td>
<td>0.90</td>
<td>1.55E-02</td>
<td></td>
</tr>
<tr>
<td>BLIB008210HT</td>
<td>fibrinogen alpha chain</td>
<td>4q28</td>
<td>NM_000058</td>
<td>FGA</td>
<td>0.90</td>
<td>1.12E-04</td>
<td></td>
</tr>
<tr>
<td>BLIB006528HT</td>
<td>toulided-like kinase 1</td>
<td>2q31.1</td>
<td>NM_012290</td>
<td>TLK1</td>
<td>0.90</td>
<td>2.97E-02</td>
<td></td>
</tr>
<tr>
<td>BLIB006674HT</td>
<td>interferon, alpha-inducible protein (clone IFI-6-16)</td>
<td>1p35</td>
<td>NM_000203</td>
<td>G1P3</td>
<td>0.90</td>
<td>6.07E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB003918HT</td>
<td>ATG12 autophagy related 12 homolog (S. cerevisiae)</td>
<td>5q21-q22</td>
<td>NM_004707</td>
<td>ATG12</td>
<td>0.90</td>
<td>8.31E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB008220HT</td>
<td>ceruloplasmin (ferroxidase)</td>
<td>3q23-q25</td>
<td>NM_000096</td>
<td>CP</td>
<td>0.90</td>
<td>2.80E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB001690HT</td>
<td>fumarylacetoacetate hydrolase domain containing 2A</td>
<td>2p24.3-p11.2</td>
<td>NM_016044</td>
<td>FAHD2A</td>
<td>0.90</td>
<td>3.55E-02</td>
<td></td>
</tr>
<tr>
<td>BLIB003923HT</td>
<td>arginase, liver</td>
<td>6q23</td>
<td>NM_000045</td>
<td>ARG1</td>
<td>0.90</td>
<td>4.47E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB006315HT</td>
<td>chromosome 3 open reading frame 10</td>
<td>3p25.3</td>
<td>NM_018462</td>
<td>C3orf10</td>
<td>0.90</td>
<td>4.25E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB004114HT</td>
<td>glycine-N-acyltransferase</td>
<td>11q12.1</td>
<td>NM_201648</td>
<td>GLYAT</td>
<td>0.90</td>
<td>1.19E-02</td>
<td></td>
</tr>
<tr>
<td>BLIB004045HT</td>
<td>vitronectin</td>
<td>17q11</td>
<td>NM_000638</td>
<td>VTN</td>
<td>0.90</td>
<td>1.34E-02</td>
<td></td>
</tr>
<tr>
<td>BLIB003334HT</td>
<td>glutamine and serine rich 1</td>
<td>11p13</td>
<td>NM_024774</td>
<td>QSER1</td>
<td>0.90</td>
<td>6.87E-04</td>
<td></td>
</tr>
<tr>
<td>BLIB006611HT</td>
<td>cytochrome c oxidase subunit VIId polypeptide 2 (liver)</td>
<td>6q12</td>
<td>NM_001865</td>
<td>COX7A2</td>
<td>0.90</td>
<td>2.27E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB006210HT</td>
<td>ribosomal protein L9</td>
<td>4p13</td>
<td>NM_000661</td>
<td>RPL9</td>
<td>0.90</td>
<td>1.90E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB005805HT</td>
<td>guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1</td>
<td>5q35.3</td>
<td>NM_006098</td>
<td>GNB2L1</td>
<td>0.90</td>
<td>1.90E-03</td>
<td></td>
</tr>
<tr>
<td>BLIB008662HT</td>
<td>ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit (oligomycin sensitivity conferring protein)</td>
<td>21q22.1-q22.2</td>
<td>NM_001697</td>
<td>ATP5O</td>
<td>0.90</td>
<td>1.18E-02</td>
<td></td>
</tr>
<tr>
<td>BLIB008549HT</td>
<td>ribosomal protein L18</td>
<td>19q13</td>
<td>NM_000979</td>
<td>RPL18</td>
<td>0.90</td>
<td>4.78E-02</td>
<td></td>
</tr>
<tr>
<td>BLIB009127HT</td>
<td>neurolysin (metallopeptidase M3 family)</td>
<td>5q12.3</td>
<td>NM_020726</td>
<td>NLN</td>
<td>0.90</td>
<td>1.95E-04</td>
<td></td>
</tr>
<tr>
<td>BLIB006745HT</td>
<td>eukaryotic translation initiation factor 2, subunit 2 beta, 38kDa</td>
<td>20pter-q12</td>
<td>NM_003908</td>
<td>EIF2S2</td>
<td>0.90</td>
<td>1.97E-04</td>
<td></td>
</tr>
<tr>
<td>BLIB008034HT</td>
<td>complement factor B</td>
<td>6p21.3</td>
<td>NM_001710</td>
<td>CFB</td>
<td>0.90</td>
<td>3.19E-04</td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
<td>Chromosome</td>
<td>Location</td>
<td>P-value</td>
<td>FDR</td>
<td>Log2 Fold Change</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------</td>
<td>----------</td>
<td>---------</td>
<td>-----------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>BLIB001129HT</td>
<td>complement factor H-related 4</td>
<td>1q32</td>
<td>6E-15</td>
<td>1.11</td>
<td>3.27E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB003296HT</td>
<td>ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit (oligomycin sensitivity conferring protein)</td>
<td>21q22.1-q22.21q22.</td>
<td>11</td>
<td>7E-86</td>
<td>1.11</td>
<td>4.95E-02</td>
<td></td>
</tr>
<tr>
<td>BLIB006341HT</td>
<td>ATP synthase, H+ transporting, mitochondrial complex, 0 subunit (oligomycin sensitivity conferring protein)</td>
<td>21q22.1-q22.21q22.</td>
<td>11</td>
<td>7E-86</td>
<td>1.11</td>
<td>2.24E-05</td>
<td></td>
</tr>
<tr>
<td>BLIB002790HT</td>
<td>hemopexin</td>
<td>11p15.5-p15.4</td>
<td>9E-45</td>
<td>1.11</td>
<td>1.59E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB006563HT</td>
<td>alpha-1-microglobulin/bikunin precursor</td>
<td>9q32-q33</td>
<td>0</td>
<td>1.11</td>
<td>8.80E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB004706HT</td>
<td>inter-alpha (globulin) inhibitor H3</td>
<td>3p21.2-p21.1</td>
<td>0</td>
<td>1.11</td>
<td>1.35E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB002280HT</td>
<td>ribosomal protein L41</td>
<td>12q13</td>
<td>4E-45</td>
<td>1.11</td>
<td>4.13E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB001745HT</td>
<td>mannose-6-phosphate receptor binding protein 1</td>
<td>19p13.3</td>
<td>1E-107</td>
<td>1.11</td>
<td>2.32E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB005829HT</td>
<td>glycine-N-acyltransferase</td>
<td>11q12.1</td>
<td>3E-25</td>
<td>1.11</td>
<td>8.18E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB001949HT</td>
<td>apolipoprotein C-IV</td>
<td>19q13.2</td>
<td>4E-14</td>
<td>1.12</td>
<td>2.60E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB002801HT</td>
<td>apolipoprotein C-IV</td>
<td>19q13.2</td>
<td>4E-14</td>
<td>1.12</td>
<td>3.44E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB002495HT</td>
<td>amylo-1, 6-glucosidase, 4-alpha-glucanotransferase (glycogen debranching enzyme, glycogen storage disease type III)</td>
<td>1p21</td>
<td>4E-81</td>
<td>1.12</td>
<td>1.07E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB005761HT</td>
<td>ribosomal protein L35a</td>
<td>3q29-pter</td>
<td>0</td>
<td>1.12</td>
<td>1.27E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB004933HT</td>
<td>sulfotransferase family, cytosolic, 2A, dehydroepiandrosterone (DHEA)-preferring, member 1</td>
<td>19q13.3</td>
<td>8E-24</td>
<td>1.13</td>
<td>8.13E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB007835HT</td>
<td>dehydroepiandrosterone (DHEA)-preferring, member 1</td>
<td>3q29-pter</td>
<td>0</td>
<td>1.13</td>
<td>7.91E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB006642HT</td>
<td>ribosomal protein L24</td>
<td>3q12</td>
<td>0</td>
<td>1.13</td>
<td>1.64E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB001621HT</td>
<td>solute carrier family 29 (nucleoside transporters), member 1</td>
<td>6p21.1-p21.2</td>
<td>0</td>
<td>1.13</td>
<td>1.40E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB007603HT</td>
<td>solute carrier family 29 (nucleoside transporters), member 1</td>
<td>16q13</td>
<td>5E-52</td>
<td>1.14</td>
<td>1.95E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB004754HT</td>
<td>metallothionein 1A (functional)</td>
<td>11p13</td>
<td>0</td>
<td>1.14</td>
<td>6.34E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB007945HT</td>
<td>catalase</td>
<td>11p13</td>
<td>0</td>
<td>1.14</td>
<td>5.98E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB001303HT</td>
<td>catalase</td>
<td>11p13</td>
<td>0</td>
<td>1.14</td>
<td>8.91E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB006091HT</td>
<td>Homo sapiens PQQ repeat and WD repeat domain containing (PQWD), mRNA</td>
<td>19q13.11</td>
<td>0</td>
<td>1.14</td>
<td>1.68E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB007850HT</td>
<td>Rho-related BTB domain containing 3</td>
<td>5q15</td>
<td>0</td>
<td>1.15</td>
<td>4.04E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB003683HT</td>
<td>ribosomal protein L17</td>
<td>18q21</td>
<td>0</td>
<td>1.16</td>
<td>1.35E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB002292HT</td>
<td>XM_371617</td>
<td>cullin-associated and neddylation-dissociated 2 (putative)</td>
<td>CAND2</td>
<td>3p25.1</td>
<td>0</td>
<td>1.16</td>
<td>4.14E-05</td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td>----------------------------------------------------------</td>
<td>-------</td>
<td>--------</td>
<td>----</td>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>BLIB001058HT</td>
<td>NM_001333</td>
<td>cathepsin L2</td>
<td>CTSL2</td>
<td>9q22.2</td>
<td>0</td>
<td>1.16</td>
<td>1.71E-03</td>
</tr>
<tr>
<td>BLIB005686HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.16</td>
</tr>
<tr>
<td>BLIB008114HT</td>
<td>NM_005143</td>
<td>haptoglobin</td>
<td>HP</td>
<td>16q22.1</td>
<td>5E-68</td>
<td>1.16</td>
<td>8.43E-04</td>
</tr>
<tr>
<td>BLIB006101HT</td>
<td>NM_001622</td>
<td>alpha-2-HS-glycoprotein</td>
<td>AHS</td>
<td>3q27</td>
<td>5E-87</td>
<td>1.16</td>
<td>5.95E-04</td>
</tr>
<tr>
<td>BLIB004946HT</td>
<td>NM_004546</td>
<td>NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 2, 8kDa</td>
<td>NDUF2</td>
<td>7q34</td>
<td>6E-71</td>
<td>1.16</td>
<td>2.69E-04</td>
</tr>
<tr>
<td>BLIB004056HT</td>
<td>NM_002021</td>
<td>flavin containing monooxygenase 1</td>
<td>FMO1</td>
<td>1q23-q25</td>
<td>0</td>
<td>1.17</td>
<td>2.33E-05</td>
</tr>
<tr>
<td>BLIB002886HT</td>
<td>NM_001482</td>
<td>glycine amidinotransferase (L-arginine:glycine amidinotransferase)</td>
<td>GATM</td>
<td>15q21.1</td>
<td>0</td>
<td>1.17</td>
<td>8.51E-05</td>
</tr>
<tr>
<td>BLIB001526HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIB001029HT</td>
<td>NM_001353</td>
<td>aldo-keto reductase family 1, member C1 (dihydrodiol dehydrogenase 1; 20-alpha (3-alpha)-hydroxysteroid dehydrogenase)</td>
<td>AKR1C1</td>
<td>10p15-p14</td>
<td>0</td>
<td>1.18</td>
<td>3.47E-05</td>
</tr>
<tr>
<td>BLIB009208HT</td>
<td>NM_000371</td>
<td>transthyretin (prealbumin, amyloidosis type I)</td>
<td>TTR</td>
<td>18q12.1</td>
<td>1E-11</td>
<td>1.20</td>
<td>8.77E-06</td>
</tr>
<tr>
<td>BLIB004514HT</td>
<td>NM_001021</td>
<td>ribosomal protein S17</td>
<td>RPS17</td>
<td>15q</td>
<td>3E-83</td>
<td>1.21</td>
<td>1.43E-05</td>
</tr>
<tr>
<td>BLIB0999052HT</td>
<td>NM_001011</td>
<td>ribosomal protein S7</td>
<td>RPS7</td>
<td>2p25</td>
<td>0</td>
<td>1.21</td>
<td>2.37E-03</td>
</tr>
<tr>
<td>BLIB002867HT</td>
<td>NM_000998</td>
<td>ribosomal protein L37a</td>
<td>RPL37A</td>
<td>2q35</td>
<td>3E-89</td>
<td>1.21</td>
<td>5.95E-05</td>
</tr>
<tr>
<td>BLIB008661HT</td>
<td>NM_001752</td>
<td>catalase</td>
<td>CAT</td>
<td>11p13</td>
<td>0</td>
<td>1.22</td>
<td>3.31E-06</td>
</tr>
<tr>
<td>BLIB008245HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.22</td>
</tr>
<tr>
<td>BLIB004925HT</td>
<td>NM_002410</td>
<td>ATPase type 13A1</td>
<td>ATP13A1</td>
<td>19p13.11</td>
<td>0</td>
<td>1.23</td>
<td>1.61E-04</td>
</tr>
<tr>
<td>BLIB004934HT</td>
<td>NM_001752</td>
<td>catalase</td>
<td>CAT</td>
<td>11p13</td>
<td>0</td>
<td>1.23</td>
<td>4.71E-06</td>
</tr>
<tr>
<td>BLIB002777HT</td>
<td>NM_017883</td>
<td>WD repeat domain 13</td>
<td>WDR13</td>
<td>Xp11.23</td>
<td>0</td>
<td>1.23</td>
<td>1.07E-05</td>
</tr>
<tr>
<td>BLIB001616HT</td>
<td>NM_199002</td>
<td>Rho guanine nucleotide exchange factor (GEF) 1</td>
<td>ARHGEF1</td>
<td>19q13.13</td>
<td>0</td>
<td>1.33</td>
<td>5.10E-08</td>
</tr>
<tr>
<td>BLIB004730HT</td>
<td>NM_001622</td>
<td>alpha-2-HS-glycoprotein</td>
<td>AHS</td>
<td>3q27</td>
<td>5E-87</td>
<td>1.38</td>
<td>7.25E-08</td>
</tr>
<tr>
<td>BLIB009138HT</td>
<td>NM_005950</td>
<td>metallothionein 1G</td>
<td>MT1G</td>
<td>16q13</td>
<td>3E-51</td>
<td>1.40</td>
<td>4.97E-06</td>
</tr>
</tbody>
</table>
**Table 5.5: ESTs differentially expressed in naive susceptible sheep S8.**

<table>
<thead>
<tr>
<th>Bovine EST sequence</th>
<th>Human Refseq:</th>
<th>Gene Description</th>
<th>Gene Symbol</th>
<th>Cytogenetic band</th>
<th>Contig BLAST</th>
<th>Fold Change</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLIB009902HT</td>
<td>NM_015697</td>
<td>coenzyme Q2 homolog, prenyltransferase (yeast)</td>
<td>COQ2</td>
<td>4q21.23</td>
<td>0</td>
<td>0.79</td>
<td>5.78E-03</td>
</tr>
<tr>
<td>BLIB001902HT</td>
<td>NM_080417</td>
<td>septin 4 (NMD candidate)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.06E-04</td>
</tr>
<tr>
<td>BLIB003129HT</td>
<td>NM_018834</td>
<td>matrix 3</td>
<td>MATR3</td>
<td>5q31.2</td>
<td>0</td>
<td>0.88</td>
<td>5.42E-02</td>
</tr>
<tr>
<td>BLIB008581HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.89</td>
<td>3.87E-05</td>
</tr>
<tr>
<td>BLIB007045HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.89</td>
<td>1.05E-02</td>
</tr>
<tr>
<td>BLIB001913HT</td>
<td>NM_001167</td>
<td>baculoviral IAP repeat-containing 4</td>
<td>BIRC4</td>
<td>Xq25</td>
<td>0</td>
<td>0.90</td>
<td>2.43E-02</td>
</tr>
<tr>
<td>BLIB001812HT</td>
<td>XM_497815</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.90</td>
<td>1.06E-03</td>
</tr>
<tr>
<td>BLIB006304HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.90</td>
<td>9.38E-03</td>
</tr>
<tr>
<td>BLIB003174HT</td>
<td>NM_003824</td>
<td>Fas (TNFRSF6)-associated via death domain</td>
<td>FADD</td>
<td>11q13.3</td>
<td>2E-07</td>
<td>0.90</td>
<td>6.01E-03</td>
</tr>
<tr>
<td>BLIB003789HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.90</td>
<td>1.00E-02</td>
</tr>
<tr>
<td>BLIB004046HT</td>
<td>XM_496537</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.90</td>
<td>8.15E-03</td>
</tr>
<tr>
<td>BLIB001598HT</td>
<td>NM_021871</td>
<td>fibrinogen alpha chain</td>
<td>FGA</td>
<td>4q28</td>
<td>0</td>
<td>0.90</td>
<td>7.43E-03</td>
</tr>
<tr>
<td>BLIB001507HT</td>
<td>NM_152912</td>
<td>Homo sapiens mitochondrial translational initiation factor 3 (MTIF3), mRNA</td>
<td>MTIF3</td>
<td>13q12.2</td>
<td>1E-23</td>
<td>1.10</td>
<td>2.71E-02</td>
</tr>
<tr>
<td>BLIB001015HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.10</td>
<td>2.01E-02</td>
</tr>
<tr>
<td>BLIB004026HT</td>
<td>NM_005721</td>
<td>ARP3 actin-related protein 3 homolog (yeast)</td>
<td>ACTR3</td>
<td>2q14.1</td>
<td>0</td>
<td>1.10</td>
<td>1.54E-03</td>
</tr>
<tr>
<td>BLIB004585HT</td>
<td>XM_379786</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.10</td>
<td>3.28E-02</td>
</tr>
<tr>
<td>BLIB005566HT</td>
<td>NM_017460</td>
<td>cytochrome P450, family 3, subfamily A, polypeptide 4</td>
<td>CYP3A4</td>
<td>7q21.1</td>
<td>0</td>
<td>1.10</td>
<td>2.38E-04</td>
</tr>
<tr>
<td>BLIB004210HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.10</td>
<td>6.43E-04</td>
</tr>
<tr>
<td>BLIB001365HT</td>
<td>NM_033331</td>
<td>CDC14 cell division cycle 14 homolog B (S. cerevisiae)</td>
<td>CDC14B</td>
<td>9q22.33</td>
<td>1E-14</td>
<td>1.11</td>
<td>4.60E-03</td>
</tr>
<tr>
<td>BLIB001508HT</td>
<td>NM_001621</td>
<td>aryl hydrocarbon receptor</td>
<td>AHR</td>
<td>7p15</td>
<td>0</td>
<td>1.12</td>
<td>1.89E-03</td>
</tr>
<tr>
<td>BLIB009138HT</td>
<td>NM_005950</td>
<td>metallothionein 1G</td>
<td>MT1G</td>
<td>16q13</td>
<td>3E-51</td>
<td>1.12</td>
<td>3.02E-02</td>
</tr>
<tr>
<td>BLIB004496HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.13</td>
<td>4.16E-06</td>
</tr>
<tr>
<td>BLIB002886HT</td>
<td>NM_001482</td>
<td>glycin amidotransferase (L-arginine:glycine amidotransferase)</td>
<td>GATM</td>
<td>15q21.1</td>
<td>0</td>
<td>1.13</td>
<td>1.16E-03</td>
</tr>
<tr>
<td>BLIB007777HT</td>
<td>NM_000186</td>
<td>complement factor H</td>
<td>CFH</td>
<td>1q32</td>
<td>5E-51</td>
<td>1.14</td>
<td>2.34E-05</td>
</tr>
<tr>
<td>BLIB009269HT</td>
<td>NM_000596</td>
<td>insulin-like growth factor binding protein 1</td>
<td>IGFBP1</td>
<td>7p13-p12</td>
<td>6E-46</td>
<td>1.14</td>
<td>1.41E-03</td>
</tr>
<tr>
<td>BLIB003438HT</td>
<td>NM_080911</td>
<td>uracil-DNA glycosylase</td>
<td>UNG</td>
<td>12q23-q24.1</td>
<td>1E-110</td>
<td>1.32</td>
<td>1.07E-02</td>
</tr>
<tr>
<td>Bovine EST sequence</td>
<td>Human Refseq:</td>
<td>Gene Description</td>
<td>Gene Symbol</td>
<td>Cyto genetic band</td>
<td>Contig</td>
<td>Fold</td>
<td>P value</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------</td>
<td>------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>--------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>BLIB001013HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.82</td>
<td>4.69E-02</td>
</tr>
<tr>
<td>BLIB005223HT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.88</td>
<td>8.84E-06</td>
</tr>
<tr>
<td>BLIB004142HT</td>
<td>NM_005256</td>
<td>growth arrest-specific 2</td>
<td>GAS2</td>
<td>11p14.3-</td>
<td></td>
<td>0.89</td>
<td>4.15E-02</td>
</tr>
<tr>
<td>BLIB005486HT</td>
<td>NM_013237</td>
<td>Homo sapiens px19-like protein (PX19), mRNA</td>
<td>PX19</td>
<td>p15.2</td>
<td>0</td>
<td>0.90</td>
<td>8.50E-03</td>
</tr>
<tr>
<td>BLIB006122HT</td>
<td></td>
<td>Homo sapiens DKFZP564J0863 protein (DKFZP564J0863), mRNA</td>
<td>DKFZP564J0</td>
<td>5q35.3</td>
<td>0</td>
<td>0.90</td>
<td>2.82E-06</td>
</tr>
<tr>
<td>BLIB003935HT</td>
<td>NM_001752</td>
<td>catalase</td>
<td>CAT</td>
<td>11q13.1</td>
<td>0</td>
<td>0.90</td>
<td>1.70E-03</td>
</tr>
<tr>
<td>BLIB004934HT</td>
<td>NM_020410</td>
<td>ATPase type 13A1</td>
<td>ATP13A1</td>
<td>19p13.11</td>
<td>0</td>
<td>1.10</td>
<td>4.94E-04</td>
</tr>
<tr>
<td>BLIB004925HT</td>
<td>NM_001752</td>
<td>proteasome (prosome, macropain) 26S subunit, ATPase, 3</td>
<td>PSMC3</td>
<td>11p12-p13</td>
<td>2 x 10^-66</td>
<td>1.11</td>
<td>4.83E-02</td>
</tr>
<tr>
<td>BLIB001013HT</td>
<td>NM_002804</td>
<td>ribosomal protein S17</td>
<td>RPS17</td>
<td>15q</td>
<td>3 x 10^-83</td>
<td>1.11</td>
<td>1.74E-04</td>
</tr>
<tr>
<td>BLIB008661HT</td>
<td>NM_001752</td>
<td>catalase</td>
<td>CAT</td>
<td>11p13</td>
<td>0</td>
<td>1.11</td>
<td>6.23E-05</td>
</tr>
<tr>
<td>BLIB002361HT</td>
<td>NM_020213</td>
<td>poly (ADP-ribose) polymerase family, member 6</td>
<td>PARP6</td>
<td>15q23</td>
<td>4 x 10^-27</td>
<td>1.11</td>
<td>5.44E-02</td>
</tr>
<tr>
<td>BLIB009138HT</td>
<td>NM_005950</td>
<td>metallothionein 1G</td>
<td>MT1G</td>
<td>16q13</td>
<td>3 x 10^-51</td>
<td>1.16</td>
<td>8.36E-04</td>
</tr>
<tr>
<td>BLIB001616HT</td>
<td>NM_199002</td>
<td>Rho guanine nucleotide exchange factor (GEF) 1</td>
<td>ARHGEF1</td>
<td>19q13.13</td>
<td>0</td>
<td>1.16</td>
<td>1.79E-06</td>
</tr>
<tr>
<td>BLIB004730HT</td>
<td>NM_001622</td>
<td>alpha-2-HS-glycoprotein</td>
<td>AHSG</td>
<td>3q27</td>
<td>5 x 10^-87</td>
<td>1.17</td>
<td>5.18E-06</td>
</tr>
</tbody>
</table>
5.2.6 Bioinformatics

5.2.6.1 Unknown EST sequences

Across the six microarray comparisons a total of 26 ESTs were reported to be differentially expressed that could not be matched to a human RefSeq. Of these 26 sequences 10 subsequently failed “SeqClean” analysis, due to their length, low complexity or degree of vector contamination (A. McCulloch, pers. comm.). The remaining 16 unknown sequences had all been publicly released, and were compared against the RefSeq database (all organisms) using BLAST (Altschul et al., 1990) and against the BTAU2.0 genome assembly using BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start).

Based on the BLAST and BLAT information nine of the 16 EST sequences were assigned to RefSeqs from species other than human, and all of the 16 EST sequences resulted in significant BLAT hits to the bovine genome. However, 8 of the 16 ESTs corresponded to either intronic or intergenic DNA sequence. The fact that so many of the ‘unknown’ EST sequences actually corresponded to intronic or intergenic DNA sequence suggests that there is significant contamination of the EST libraries by genomic DNA.

5.2.6.2 Identification of possible biomarkers

The aim of this analysis was to identify putative biomarkers from the reported differentially expressed genes (refer to 5.1.4). Ideally, a biomarker would be detectable in blood as it is relatively easy and inexpensive to obtain with minimum harm to the animal.

There are many resources that can be used to define possible biomarkers, two of which were used in this study. The first approach used Gene Ontology (GO) classifications (Harris et al., 2004b). The microarray used in this experiment had a total of 1550 EST sequences of which 134 ESTs were associated with the gene ontology term ‘extracellular space’. The GO term ‘extracellular space’ includes many secreted proteins. These 134 EST sequences correspond with 76 unique protein sequences. Of these 134 EST sequences 16 were reported as differentially expressed in the microarray experiment. These 16 EST sequences correspond with
12 unique protein sequences. However, although these 12 proteins are classified as being associated with the extracellular space there is no way of knowing whether these proteins are actually detectable in the blood, so a second approach to classification of possible biomarker targets was undertaken.

The second approach utilised the plasma proteome database (http://www.plasmaproteomedatabase.org). This database is produced by the Plasma Proteome Project which is a consortium of 18 laboratories that have attempted to obtain a comprehensive inventory of all of the proteins found in human plasma. The initial database included 3020 proteins and that had been cross-referenced with the expression of these proteins in other tissues (Ping et al., 2005). For this study the sub-database of interest was the liver, as it identifies proteins that are synthesised in the liver and secreted into plasma, and as such are possible biomarker targets.

Potential biomarkers were identified using the Plasma Proteome Database by comparing the list of differentially expressed genes obtained (Tables 5.2 – 5.7) with information from this database. A total of 23 putative biomarker targets were identified. These 23 putative biomarkers included 11 of the original 12 identified using GO classification and the details of these 11 proteins are given in Table 5.8.

Of the 11 possible biomarker targets identified, three are particularly promising as they were identified by more than one EST and in more than one animal. These three markers were α-2-HS-glycoprotein, catalase and fibrinogen-alpha chain.

Alpha-2-HS-glycoprotein was reported to be up-regulated in the naïve resistant animal, and is also reported to be differentially expressed as a result of exposure to sporidesmin (irrespective of genetics). α-2-HS-glycoprotein has been implicated in many biological processes including modulating the immune response and it is annotated as an acute phase protein (reviewed in Arnaud and Kalabay, 2002).
Table 5.8: Putative biomarkers for either resistance to FE or exposure to sporidesmin. Genes that were found to be differentially expressed by microarray analysis were compared with the plasma proteome database and the gene ontology database in order to identify genes whose differential expression in the liver may translate to differential protein levels in the serum of animals. The 11 proteins listed were identified in both the plasma proteome database and the gene ontology databases.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene Name</th>
<th>RefSeq</th>
<th>EST name</th>
<th>Differentially expressed in</th>
<th>Expression ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHSG</td>
<td>alpha-2-HS-glycoprotein</td>
<td>NM_001622</td>
<td>BLIB006101HT</td>
<td>naïve resistant</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BLIB004730HT</td>
<td>naïve resistant</td>
<td>1.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BLIB004730HT</td>
<td>sporidesmin exposure resistant</td>
<td>1.17</td>
</tr>
<tr>
<td>C5</td>
<td>complement component 5</td>
<td>NM_001735</td>
<td>BLIB001633HT</td>
<td>sporidesmin-dosed resistant</td>
<td>1.11</td>
</tr>
<tr>
<td>CAT</td>
<td>catalase</td>
<td>NM_001752</td>
<td>BLIB007945HT</td>
<td>naïve resistant</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BLIB008661HT</td>
<td>naïve resistant</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BLIB004934HT</td>
<td>naïve resistant</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BLIB004934HT</td>
<td>sporidesmin exposure resistant</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BLIB008661HT</td>
<td>sporidesmin exposure resistant</td>
<td>1.11</td>
</tr>
<tr>
<td>CP</td>
<td>ceruloplasmin (ferroxidase)</td>
<td>NM_000096</td>
<td>BLIB008220HT</td>
<td>naïve resistant</td>
<td>0.90</td>
</tr>
<tr>
<td>F10</td>
<td>coagulation factor X</td>
<td>NM_000504</td>
<td>BLIB006138HT</td>
<td>sporidesmin-dosed resistant</td>
<td>1.12</td>
</tr>
<tr>
<td>FGA</td>
<td>fibrinogen alpha chain</td>
<td>NM_000508</td>
<td>BLIB008512HT</td>
<td>naïve resistant</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BLIB001598HT</td>
<td>naïve resistant</td>
<td>0.90</td>
</tr>
<tr>
<td>HPX</td>
<td>hemopexin</td>
<td>NM_000613</td>
<td>BLIB002790HT</td>
<td>naïve susceptible</td>
<td>1.11</td>
</tr>
<tr>
<td>IGFBP1</td>
<td>insulin-like growth factor binding protein 1</td>
<td>NM_000596</td>
<td>BLIB009269HT</td>
<td>naïve resistant</td>
<td>1.14</td>
</tr>
<tr>
<td>LECT2</td>
<td>leukocyte cell-derived chemotaxin 2</td>
<td>NM_002302</td>
<td>BLIB001262HT</td>
<td>sporidesmin-dosed susceptible</td>
<td>0.88</td>
</tr>
<tr>
<td>TTR</td>
<td>transthyretin (prealbumin, amyloidosis type I)</td>
<td>NM_000371</td>
<td>BLIB009208HT</td>
<td>naïve resistant</td>
<td>1.20</td>
</tr>
<tr>
<td>VTN</td>
<td>vitronectin</td>
<td>NM_000638</td>
<td>BLIB004045HT</td>
<td>naïve resistant</td>
<td>1.10</td>
</tr>
</tbody>
</table>

Fibrinogen-alpha (FGA) is also an excellent candidate for a biomarker as its expression is restricted to the liver, so that changes in the liver expression of FGA should accurately reflect changes in the serum concentrations of the FGA protein. FGA has a dual function, yielding monomers that polymerize into fibrin and acting...
as a cofactor in platelet aggregation. FGA was reported to be under-represented in both naïve resistant and naïve susceptible animals, making it a putative biomarker for sporidesmin exposure or disease progression.

The final protein of interest is catalase; catalase is responsible for the catabolism of hydrogen peroxide to water and the generation of hydrogen peroxide by sporidesmin is implicated in the pathogenesis of FE (refer 1.3.4.1). However, while catalase is readily detectable in the serum of sheep (Hohenboken et al., 2004) it is ubiquitously expressed. Therefore, serum catalase will increase in response to generalised tissue damage and is therefore not specific for sporidesmin. Additionally, small changes in the hepatic expression of catalase are unlikely to be detected in the serum.

Both catalase and fibrinogen genes were selected for Northern analysis to confirm their differential expression.

5.2.6.3 Gene Ontology.

The Gene Ontology (GO) database (Harris et al., 2004b) classifies differentially expressed genes on the basis of their function. Genes are assigned to functional categories that relate to certain biological processes, molecular functions, or to cellular components. The GO categories are constantly updated, with new terms and relationships being to reflect current knowledge in biology. Identification of GO terms within a list of differentially expressed genes can be used to form functional hypotheses about these gene sets, by highlighting functional categories or biological processes of interest.

The GO terms significantly associated with the differentially expressed genes were found using EASEonline (Dennis et al., 2003). EASEonline compares the distribution of functional categories within the set of differentially expressed genes relative to the distribution of functional categories of all the genes contained on the microarray. Observed and expected category population values are compared and the statistical enrichment or depletion of a category is quantified by using the Fisher exact test.
Chapter Five: cDNA Microarrays as a Tool for Candidate Gene Identification

Only three of the six comparisons from the microarray experiment resulted in statistically significant associations with particular GO terms. These are presented in two tables; the GO terms associated with genes differentially expressed in resistant animals (Table 5.9) and GO terms associated with genes differentially expressed in susceptible animals (Table 5.10).

**GO terms associated with genes differentially expressed in resistant animals.**

Three GO categories were obtained for genes that were under-represented in the sporidesmin dosed resistant animal, and three categories that were over-represented in this animal (Table 5.9).

The three GO categories that were identified for genes that were under-represented in this animal were transmembrane receptor activity, receptor activity and transducer activity and each category contained the same three genes. These genes were the leukaemia inhibitory factor receptor (LIFR) the LanC lantibiotic synthetase component C-like 1 (LANCL1) and the G protein coupled receptor, family C group 5 member B (GPRC5B) proteins. LIFR is a polyfunctional cytokine whose expression is reported to be up-regulated in alcoholic liver injury (Zhang et al., 2006) and the authors propose that it may function to ameliorate alcohol induced liver disease. LIFR has also found to be up-regulated following liver resection and heptectomy (Omori et al., 1996). LANCL1 is related to bacterial LanC proteins which are involved in the biosynthesis of antimicrobial peptides. LANCL1 is weakly associated with cell membranes and is thought to play a role in protein modification (Bauer et al., 2000). Interestingly, GPRC5B is not known to be expressed in human liver tissue (Brauner-Osborne and Krogsgaard-Larsen, 2000) therefore the expression detected on the microarray could be due to cross hybridisation with a pseudogene or close family member.

Three GO categories were also identified for genes that were over-represented in sporidesmin dosed resistant animals; these were enzyme regulator activity, extracelluar, and organismal physiological process. These three categories contained two proteins in common; the complement component 5 (C5) and coagulation factor X (F10). Both of these proteins are key players in the
Table 5.9: GO terms significantly associated with the differentially expressed genes in resistant sheep

<table>
<thead>
<tr>
<th>GO Term</th>
<th>No. of RefSeqs</th>
<th>Fisher exact score associated with term</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GO Terms associated with genes repressed in sporidesmin-dosed resistant sheep (R16)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Molecular function</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>transmembrane receptor activity</td>
<td>3</td>
<td>$6.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>receptor activity</td>
<td>3</td>
<td>$2.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>signal transducer activity</td>
<td>3</td>
<td>$1.9 \times 10^{-2}$</td>
</tr>
<tr>
<td><strong>GO Terms associated with genes more highly expressed in sporidesmin-dosed resistant sheep (R16)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Biological Process</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>organisinal physiological process</td>
<td>3</td>
<td>$4.0 \times 10^{-2}$</td>
</tr>
<tr>
<td><em>Molecular function</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>enzyme regulator activity</td>
<td>3</td>
<td>$7.5 \times 10^{4}$</td>
</tr>
<tr>
<td><em>Cellular component</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>extracellular</td>
<td>3</td>
<td>$2.7 \times 10^{-2}$</td>
</tr>
<tr>
<td><strong>GO Terms associated with genes repressed in naïve resistant sheep (R10)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Biological Process</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>response to stimulus</td>
<td>10</td>
<td>$7.6 \times 10^{-4}$</td>
</tr>
<tr>
<td>response to external stimulus</td>
<td>8</td>
<td>$4.2 \times 10^{-3}$</td>
</tr>
<tr>
<td>response to biotic stimulus</td>
<td>6</td>
<td>$1.4 \times 10^{-2}$</td>
</tr>
<tr>
<td>organisinal physiological process</td>
<td>7</td>
<td>$1.9 \times 10^{-2}$</td>
</tr>
<tr>
<td>immune response</td>
<td>5</td>
<td>$2.2 \times 10^{-2}$</td>
</tr>
<tr>
<td>defence response</td>
<td>5</td>
<td>$2.8 \times 10^{-2}$</td>
</tr>
<tr>
<td><em>Molecular function</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>catalytic activity</td>
<td>13</td>
<td>$5.7 \times 10^{3}$</td>
</tr>
<tr>
<td>transition metal ion binding</td>
<td>3</td>
<td>$2.4 \times 10^{-2}$</td>
</tr>
<tr>
<td><em>Cellular component</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>extracellular space</td>
<td>4</td>
<td>$1.7 \times 10^{-2}$</td>
</tr>
<tr>
<td><strong>GO Terms associated with genes more highly expressed in naïve resistant sheep (R10)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Biological Process</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>electron transport</td>
<td>5</td>
<td>$1.9 \times 10^{-2}$</td>
</tr>
<tr>
<td><em>Molecular function</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ion transporter activity</td>
<td>5</td>
<td>$5.2 \times 10^{-3}$</td>
</tr>
<tr>
<td>transporter activity</td>
<td>10</td>
<td>$2.2 \times 10^{-2}$</td>
</tr>
<tr>
<td>cation transporter activity</td>
<td>4</td>
<td>$1.6 \times 10^{-2}$</td>
</tr>
<tr>
<td>hydrogen ion transporter activity</td>
<td>3</td>
<td>$2.9 \times 10^{-2}$</td>
</tr>
<tr>
<td>primary active transporter activity</td>
<td>3</td>
<td>$2.9 \times 10^{-2}$</td>
</tr>
<tr>
<td>monovalent inorganic cation transporter activity</td>
<td>3</td>
<td>$2.9 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

immune response. C5 is an anaphylatoxin molecule that induces release of inflammatory mediators from mast cells and phagocytes, and serves to amplify the inflammatory response. It also stimulates oxidative metabolism and the production
of ROS in neutrophils (Chenoweth and Hugli, 1978). This is potentially relevant to FE as hepatic infiltration of neutrophils is an early response to tissue injury, cellular stress or systemic inflammation (reviewed in Jaeschke and Hasegawa, 2006). F10 production has been shown to modulate neutrophil chemoattractant activity in liver injury (Yamaguchi et al., 2000). Several of the other proteins identified in these categories have also been related to liver injury or disease, for example: melanoma inhibitory activity protein (MIA2) is expressed exclusively in the liver (Bosserhoff et al., 2003) and is found to be up-regulated in fibrotic liver injury (Hellerbrand et al., 2005).

A number of categories were found to be associated with genes over and under represented in the naïve resistant animal (Table 5.9). GO categories of interest under-represented in this animal include ‘responses to stimulus’ and also ‘immune and defence responses’ (which was a sub-set of the ‘response to stimulus’ GO category). The response to stimulus GO category showed an under-representation of genes such as complement factor H, which is a key regulator of innate immunity and protects the host cell from inappropriate complement activation (Zipfel et al., 1999). Also under-represented was interferon α inducible protein (IFI-6-16) which has been reported to be up-regulated in the early stages of liver fibrosis (Bieche et al., 2005). Leukocyte derived chemotaxin 2 (LECT2) is also under-represented. LECT2 is usually up-regulated in response to recovery from liver failure (Sato et al., 2004a; Sato et al., 2004b) and is thought to regulate the homeostasis of NK (natural killer) T-cells in the liver (Ovejero et al., 2004). Finally, SEPP1 was reported as under-represented, SEPP1 is implicated in protection against liver necrosis as an antioxidant (reviewed in Burk et al., 2003).

The naïve resistant animal also down regulate genes involved in catalytic activity and transition metal ion binding. Of particular interest is ceruloplasmin which is involved in copper transport and is thought to provide a defence against oxidative stress by scavenging superoxide radicals (Goldstein et al., 1979). Ceruloplasmin is of interest as copper is thought to be important in the cellular toxicity of sporidesmin (refer to 1.3.4.1). Ceruloplasmin has previously been examined as a candidate gene for FE and markers within the gene showed no segregation between resistant and susceptible selection lines (Phua et al., 1998).
GO categories that were overrepresented in naïve resistant animals include ‘electron transport’, which consists of genes such as catalase, cytochrome c oxidase subunit VIIa, flavin containing monooxygenase 1 and NADH dehydrogenase (NDUFB2). Both NDUFB2 and cytochrome c oxidase are integral components of the mitochondrial respiratory chain (MRC). Cytochrome c oxidase is the terminal component of the MRC and NDUFB2 is a component of complex I of the MRC. Up-regulation of these genes could indicate that naïve resistant animals have inherently different ATP requirements, or that they have enhanced MRC activity for other reasons. Up-regulation of the MRC has previously been shown to confer a degree of protection with respect to liver failure (Navarro et al., 2005) and deregulation of the MRC is known to hasten liver failure (reviewed in Morris, 1999) possibly due to the production of ROS.

Finally, two oxidoreductase enzymes were identified in this functional category. Firstly catalase which detoxifies hydrogen peroxide to water and oxygen and secondly Flavin containing monooxygenase 1 (FMO1). FMO1 is known to oxygenate xenobiotics that contain a soft nucleophile such as nitrogen or sulphur. FMO1 is not known to be markedly regulated by the environment and most variation in expression is genetic (Krueger and Williams, 2005). It is thought that FMO1 may control hydrogen peroxide levels as well as genes regulated by hydrogen peroxide, sulphydral / disulphide ratios and general redox state (Khomenko et al., 2004). In general the FMOs have very broad substrate specificity and oxygenation of target compounds makes the substrate less toxic (reviewed in Krueger and Williams, 2005).

Six groups of transporter proteins were also identified in the naïve resistant animal (Table 5.9). Five of these groups are subsets of the 6th group, ‘transporter activity’. The ten genes in this category include very diverse transporters and substrate specificities. For example, transthyretin is implicated in thyroid hormone transport, hemopexin in heme transport and complement factor H – related 4 which is involved in lipid transport. Several of these transporter proteins do share functional similarities, either as antioxidants or as members of the MRC.
GO terms associated with genes differentially expressed in susceptible animals.

Go terms identified as under-represented in the naïve susceptible animal are primarily associated with apoptosis (Table 5.10). The three that are under-represented in this category have somewhat opposing functions. In particular, BIRC4 (baculoviral IAP repeat-containing protein 4) acts as a potent inhibitor of apoptosis via the inhibition of caspase activation (Salvesen and Abrams, 2004; Eckelman et al., 2006). TNFRSF6 (tumor necrosis factor receptor superfamily, member 6) is known to function as one of the key receptors in the initiation of apoptosis but is also known to have proliferative activity in certain cell types (Hyer et al., 2006). Comparatively little is known about Septin 4 (SEPT4) except that it is a GTP binding protein and it is involved in active membrane movement such as cytokinesis and vesicle trafficking (Garcia et al., 2006). SEPT4 has also been shown to be pro-apoptotic (Ihara et al., 2003) similar to its close relative ARTS (Larisch, 2004).

The naïve susceptible animal shows a significant over-representation of genes associated with the GO category ‘response to stimulus’. This category includes genes such as CYP3A4, a protein that is highly induced by xenobiotic exposure and is involved in the metabolism of many drugs and xenobiotics (reviewed in Luo et al., 2004). Also induced is the aryl hydrocarbon receptor (AHR) which plays a key role in the cellular response to xenobiotic exposure by modulating the expression of several key CYP enzymes (although not CYP3A4) (Ramadoss et al., 2005). The

Table 5.10: GO terms significantly associated with the differentially expressed genes in susceptible sheep

<table>
<thead>
<tr>
<th>GO Term</th>
<th>No. of RefSeqs</th>
<th>Fisher exact score</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO Terms associated with genes repressed in naïve susceptible sheep (S8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biological Process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>regulation of apoptosis</td>
<td>3</td>
<td>1.1 × 10^-4</td>
</tr>
<tr>
<td>apoptosis</td>
<td>3</td>
<td>3.8 × 10^-4</td>
</tr>
<tr>
<td>programmed cell death</td>
<td>3</td>
<td>3.8 × 10^-4</td>
</tr>
<tr>
<td>cell death</td>
<td>3</td>
<td>8.2 × 10^-4</td>
</tr>
<tr>
<td>death</td>
<td>3</td>
<td>8.2 × 10^-4</td>
</tr>
<tr>
<td>GO Terms associated with genes more highly expressed in naïve susceptible sheep (S8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biological Process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>response to stimulus</td>
<td>4</td>
<td>4.7 × 10^-2</td>
</tr>
</tbody>
</table>
over-representation of these genes in the naïve susceptible animal is perhaps
surprising if the animal is truly naïve.

5.2.7 Re-sequencing of EST clones
The bovine EST libraries used to construct the microarray slides used in this study
are known to contain both positional and sequencing errors (Wilson, pers. comm.).
The exact error rate for the liver EST library was unknown so a number of EST
clones that had been reported to be differentially expressed were re-sequenced to
verify their identity.

Nineteen clones were re-transformed into XL1-Blue *E. coli* using heat shock
transformation. Mini-prep isolation of the DNA plasmids was used to isolate the
plasmids for direct sequencing using the universal primers T7 and SP6 (refer to
2.4.11).

The results of the re-sequencing are summarised in Table 5.11. The sequences were
compared with the human RefSeq database and also compared back to the bovine
EST databases to ensure the correct identity of the EST. Of these 19 ESTs, three
contained two plasmids. These were independently isolated and sequenced. One of
these three samples contained two plasmids of different sizes, but both contained the
same sequence insert. The two others contained different inserts that showed
BLAST matches to another EST sequence. Of the remaining 16 samples that
contained a single insert 6 contained EST sequences that did not match with their
assigned EST resulting in a final false identity rate of 37.5%. However, for one of
these sequences the top human RefSeq match was the same as the initial EST
(GSTA2), and therefore may not have matched to the expected EST due to non-
overlapping sequence with the EST of interest. Overall the false identity rate from
this subset of ESTs is estimated to be 31.25%.
Table 5.11: Table of EST clones that were re-sequenced to assess the accuracy of the EST library used to generate the microarray slides

<table>
<thead>
<tr>
<th>EST</th>
<th>Clone</th>
<th>RefSeq</th>
<th>ID</th>
<th>Sequence (bp)</th>
<th>Top blast hit</th>
<th>Top hit ID</th>
<th>E value</th>
<th>Top Bovine Hit</th>
<th>E-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLIB001262HT</td>
<td>NM_002302</td>
<td>LECT2</td>
<td>666</td>
<td>NM_002302</td>
<td>LECT2</td>
<td>7×10^-6</td>
<td>BLIB001262HT</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BLIB001541HT</td>
<td></td>
<td></td>
<td></td>
<td>689</td>
<td>NM_000509</td>
<td>FGA</td>
<td>1×10^-57</td>
<td>BLIB004537HT</td>
<td>0</td>
</tr>
<tr>
<td>BLIB001698HT</td>
<td>NM_005313</td>
<td>GRP58</td>
<td>515</td>
<td>NM_138927</td>
<td>SON</td>
<td>2×10^-134</td>
<td>na</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>BLIB001698HT</td>
<td>NM_005313</td>
<td>GRP58</td>
<td>536</td>
<td>NM_002032</td>
<td>FTH1</td>
<td>2×10^-180</td>
<td>BLIB999060HT</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BLIB001888HT</td>
<td></td>
<td></td>
<td></td>
<td>930</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>BLIB001888HT</td>
<td>1×10^-165</td>
</tr>
<tr>
<td>BLIB002703HT</td>
<td>NM_021871</td>
<td>FGA</td>
<td>696</td>
<td>NM_021871</td>
<td>FGA</td>
<td>na</td>
<td>BLIB002703HT</td>
<td>5×10^-53</td>
<td></td>
</tr>
<tr>
<td>BLIB002911HT</td>
<td>NM_000509</td>
<td>FGA</td>
<td>564</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>BLIB002911HT</td>
<td>1×10^-147</td>
<td></td>
</tr>
<tr>
<td>BLIB003342HT</td>
<td>NM_001013624</td>
<td>LOC509053</td>
<td>568</td>
<td>NM_001001349</td>
<td>KBRAS</td>
<td>3×10^-25</td>
<td>BLIB005856HT</td>
<td>9×10^-16</td>
<td></td>
</tr>
<tr>
<td>BLIB003544HT</td>
<td></td>
<td></td>
<td></td>
<td>568</td>
<td>NM_000509</td>
<td>FGA</td>
<td>3×10^-124</td>
<td>BLIB004486HT</td>
<td>0</td>
</tr>
<tr>
<td>BLIB003400HT</td>
<td>NM_000567</td>
<td>LOC527553</td>
<td>561</td>
<td>NM_000567</td>
<td>CRP</td>
<td>8×10^-30</td>
<td>BLIB004280HT</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BLIB003809HT</td>
<td>NM_145740</td>
<td>GSTA2</td>
<td>565</td>
<td>NM_145740</td>
<td>GSTA2</td>
<td>2×10^-141</td>
<td>BLIB001041HT</td>
<td>1×10^-123</td>
<td></td>
</tr>
<tr>
<td>BLIB004280HT</td>
<td>NM_018834</td>
<td>MATR3</td>
<td>545</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>BLIB005485HT</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BLIB005485HT</td>
<td>NM_018834</td>
<td>MATR3</td>
<td>745</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>BLIB005485HT</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BLIB006697HT</td>
<td>NM_002872</td>
<td>RAC2</td>
<td>585</td>
<td>NM_002872</td>
<td>RAC2</td>
<td>9×10^-164</td>
<td>BLIB006026HT</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BLIB007891HT</td>
<td>NM_001430</td>
<td>EPAS1</td>
<td>480</td>
<td>NM_016208</td>
<td>VPS28</td>
<td>1×10^-96</td>
<td>BLIB007491HT</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BLIB008950HT</td>
<td>XM_378876</td>
<td>LOC149351</td>
<td>630</td>
<td>NM_001430</td>
<td>EPAS1</td>
<td>3×10^-149</td>
<td>BLIB008500HT</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>BLIB008661HT</td>
<td>NM_001752</td>
<td>CAT</td>
<td>617</td>
<td>NM_001752</td>
<td>CAT</td>
<td>0</td>
<td>BLIB008661HT</td>
<td>1×10^-149</td>
<td></td>
</tr>
<tr>
<td>BLIB009138HT</td>
<td>NM_005950</td>
<td>MT1G</td>
<td>640</td>
<td>NM_005950</td>
<td>MT1G</td>
<td>9×10^-60</td>
<td>BLIB009138HT</td>
<td>1×10^-115</td>
<td></td>
</tr>
</tbody>
</table>
5.2.8 Confirmation of differential expression by Northern hybridisation

Five of the 159 differentially expressed ESTs were selected for confirmation by Northern analysis. These included the highest and lowest known differentially expressed genes, metallothionein 2 (MT2) and matrin 3 (MATR3), respectively. Three other genes were selected for confirmation based on known function. These genes were catalase (CAT), leukocyte derived chemotaxin 2 (LECT2) and ras-related C3 botulinum toxin substrate 2 (RAC2).

CAT was chosen as an association between catalase alleles and resistance to facial eczema disease is already known (Phua et al., 1999). LECT2 was chosen as it was reported as differentially expressed in four out of the six microarray comparisons. Additionally, it is a good functional candidate as its expression is known to be a marker for hepatic injury and regeneration and it is readily detectable in serum (Sato et al., 2004a; Sato et al., 2004b). Finally RAC2 was selected as Rho-like small GTPases are thought to mediate endogenous reactive oxygen species (ROS) production via NADPH oxidases. This endogenous ROS cascade can be induced in response to cell damage or inflammation.

A further set of three genes were selected for confirmation. These genes were significantly differentially expressed, but fell within initial cut-off of 0.9 – 1.1 fold expression. These three genes were examined to determine the appropriateness of these cut-off values. The genes selected in this category were, C-reactive protein (CRP), fibrinogen-α (FGA), glutathione-S-transferase alpha-2 (GSTA2).

Finally, two additional genes were selected for confirmation, these genes were not reported to be differentially expressed, but were functionally very interesting. These genes were endothelial PAS domain protein 1 (EPAS1) and I-k-B-interactive ras-like protein 2 (KBRAS2). EPAS1 is a transcription factor involved in the induction of oxygen regulated genes and so is functionally interesting as sporidesmin pathology is thought to be mediated (at least in part) by the generation of oxygen radicals. The second gene, KBRAS2, is an important component of regulation of the NFκB signalling cascade. KBRAS2 interacts with IκB-α and IκB-β and slows the rate of
degradation, enhancing signalling through the NFκB cascade. This is important as NFκB is a key transcription factor in the stress response.

Northern blots were generated using RNA extracted from the livers of the four animals used on the microarray and an additional 13 animals. Ten individual Northern blots were generated and each blot probed with one of the ten genes selected for confirmation. In each case the autoradiographs were quantified by densitometry and the expression normalised to ethidium bromide stained 18S rRNA levels. Statistical comparisons between the four treatment groups (naïve resistant, naïve susceptible, dosed resistant and dosed susceptible) were carried out and the data are presented graphically below each blot. Statistical comparisons were also carried out to detect differences in expression as a result of selection line (i.e. irrespective of exposure to sporidesmin) and also differences in expression as a result of exposure to sporidesmin (i.e. irrespective of genetic line). The data from these latter two analyses is also provided graphically, but only where a statistically significant difference was detected. For clarity, a summary of the Northern and microarray expression data for these genes is provided in Table 5.13.

5.2.8.1 Catalase (CAT)
Catalase is a peroxisomal enzyme that catalyses the detoxification of hydrogen peroxide (H₂O₂) to water (H₂O) and molecular oxygen (O₂).

Northern blotting using the EST BLIB008661HT as a probe revealed three hybridisation bands on the corresponding autoradiograph (Fig. 5.3A). These three bands were sized by comparison with the RNA ladder and were 4.2 kb (transcript A), 2.2 kb (transcript B) and 1.4 kb (transcript C) respectively. Transcript B is closest in size to the expected full-length transcript of 2.3 kb (based on the human RefSeq sequence NM_001752) and is the key transcript in this analysis. Mining of the EC gene database (Kim et al., 2005a; Kim et al., 2005b) revealed several transcripts consistent with the smaller transcript C (~1.4 kb), which may implicate transcript C as an alternative transcript of catalase. There was no human transcript corresponding with the larger transcript A (4.2 kb) in the EC gene database.
Chapter Five: cDNA Microarrays as a Tool for Candidate Gene Identification

A. Table showing the expression levels of transcripts A, B, and C in Naive and Sporidesmin-dosed groups.

<table>
<thead>
<tr>
<th></th>
<th>Naive</th>
<th>Naive</th>
<th>Sporidesmin-dosed</th>
<th>Sporidesmin-dosed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Resistant</td>
<td>Susceptible</td>
<td>Resistant</td>
<td>Susceptible</td>
</tr>
<tr>
<td>A (4.2 kb)</td>
<td>R8</td>
<td>R9</td>
<td>R10</td>
<td>S8</td>
</tr>
<tr>
<td>CAT B (2.2 kb)</td>
<td>R11</td>
<td>R12</td>
<td>R13</td>
<td>R14</td>
</tr>
<tr>
<td>C (1.4 kb)</td>
<td>18S rRNA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Bar graphs showing the expression ratios of different transcripts.

i. Transcript A (4.2 kb)

ii. Transcript B (2.2 kb)

iii. Transcript C (1.4 kb)

C. Comparison of expression ratios between Naive and Sporidesmin-dosed groups.

D. Comparison of expression ratios between Resistant and Susceptible groups.

192
Quantification of the Northern blot revealed that when the four treatment groups are considered individually all three transcripts were differentially expressed to varying degrees (Fig. 5.3B). The transcript that is consistent with the size of the human RefSeq transcript (transcript B) was found to be expressed 1.6 fold higher in the naïve susceptible animals as compared with the naïve resistant animals. Transcript B was also expressed 2.1 fold higher in the naïve susceptible animal than the dosed resistant animal (overall ANOVA \( P = 0.006 \)).

The larger transcript (transcript A) was differentially expressed between naïve susceptible animals and dosed resistant animals, naïve susceptible animals expressed transcript A 1.8 fold higher than dosed resistant animals, and 2 fold higher than dosed susceptible animals (overall ANOVA \( P = 0.008 \)).

The putative alternative transcript (transcript C) was also differentially expressed: naïve susceptible animals expressed transcript C 1.6 fold higher levels than naïve resistant animals. This transcript was also differentially expressed between naïve susceptible animals and dosed resistant animals by 1.5 fold (overall ANOVA \( P = 0.014 \)).

When genetic background of the animals is excluded and the expression of catalase is examined based on sporidesmin exposure (Fig. 5.3C), transcript A is found to be expressed 1.7 fold higher in naïve animals than sporidesmin dosed animals \( (P = 0.002) \). No differences were observed for either transcripts B or C.

However, when sporidesmin exposure is excluded and the expression of catalase is examined based on selection-line (Fig. 5.3D), both transcripts B and C were found to be expressed significantly higher in susceptible animals as compared to resistant

Figure 5.3: (opposite page) Northern hybridisation of CAT mRNA in ovine liver. (A) Scanned Northern blot showing expression of the catalase mRNA in livers of resistant and susceptible selection-line animals, both naïve and sporidesmin dosed. M = marker lane. Underneath the scanned Northern blot is the ethidium bromide stained 18S rRNA band used for normalisation. (B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean± SD. The expression levels of each of the three hybridisation bands for resistant and susceptible animals’ pre- and post- exposure to sporidesmin. Statistical significance was assessed by one-way ANOVA followed by Tukey’s post-hoc test \( (* P<0.05) \). (C) Northern expression data were grouped according to sporidesmin exposure (D) Northern expression data were grouped according to genetic line (resistant or susceptible). For C and D statistical significance was assessed using two sample T-tests \( *, P<0.05, ** P<0.01 \)
animals. Transcript B was expressed 1.5 fold higher in susceptible animals \( (P = 0.015) \) and transcript C was expressed 1.4 fold higher in susceptible animals \( (P = 0.007) \). The higher expression of catalase in susceptible animals irrespective of sporidesmin exposure may indicate that susceptible animals have higher oxidative loads than resistant animals requiring more catalase expression.

5.2.8.2 Leukocyte derived chemotaxin 2 (LECT2)
Leukocyte derived chemotaxin 2 (LECT2) is expressed almost exclusively in the liver and has neutrophil chemotactic activity. It is also able to promote chondrocyte proliferation. It is a putative candidate gene as its expression is known to be a

Figure 5.4: Northern hybridisation of LECT2 mRNA in ovine liver A) Scanned Northern blot showing expression of the LECT2 mRNA in livers of resistant and susceptible selection-line animals, both naive and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. For this analysis the expression data from S10, R13 and R14 were excluded as the RNA showed evidence of degradation. Significance was assessed by one way ANOVA followed by Tukey’s post-hoc test \( (* P < 0.05) \).
marker for hepatic injury and regeneration in humans and it is also easily detectable in serum (Sato et al., 2004a; Sato et al., 2004b) making it a possible biomarker (Table 5.8). Increased expression of LECT2 is also prognostic for recovery from liver failure.

Northern blotting using the EST BLIB001262HT as a probe yielded only one hybridisation band and it was 0.7 kb in size (Fig. 5.4) which is slightly smaller than the expected size of 1.1 kb based on the human RefSeq sequence (NM_002302), but is consistent with the size of the bovine LECT2 transcript of 0.72 kb (NM_174380).

Quantification of LECT2 expression showed a significant induction of 1.7 fold ($P = 0.007$) for resistant animals following exposure to sporidesmin. No significant differences were observed for susceptible animals. LECT2 expression in the liver is thought to be dependent on wnt / β-catenin signalling pathways (Ovejero et al., 2004) which are critical for heptocyte proliferation (Gonzalez, 2006).

5.2.8.3 Matrin 3 (MATR3)

MATR3 (matrin 3) is thought to play a role in transcription. It was selected for confirmation by Northern blotting as it was the most significantly down-regulated gene from the microarray experiments, and this down-regulation in naïve resistant sheep was reported by two independent EST sequences on the same microarray.

Northern blotting of sheep liver RNA using BLIB005485HT as a probe revealed a single hybridisation band of 1.8 kb (Fig. 5.5A). In humans two alternative transcripts of the MATR3 gene exist and both transcripts are 4 kb in size (NM199189, NM_08834). The hybridisation band obtained in this study was significantly smaller than the annotated 4 kb transcripts, and was also smaller than the known coding domain region of MATR3 (2.5 kb). A BLAST search of the human RefSeq database yielded two sequences similar to MATR3; XM_937971 and XM_4963791 both of which are likely to be pseduogenes and if processed both would yield products of 1.2 kb, which is 0.6 kb smaller than the 1.8 kb hybridisation band obtained. Additionally, neither of the rat pseudogenes were processed or are consistent in size with the transcript obtained here (Matsushima et al., 1998). This suggests that the BLIB005485HT probe may be detecting a gene other than MATR3.
A.

<table>
<thead>
<tr>
<th>Naïve Resistant</th>
<th>Naïve Susceptible</th>
<th>Sporidesmin dosed Resistant</th>
<th>Sporidesmin dosed Susceptible</th>
</tr>
</thead>
<tbody>
<tr>
<td>R8 R9 R10 S8 S9 S10</td>
<td>R11 R12 R13 R14 R15 R16 S11 S12 S13 S14 S15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MATR3 (1.8 kb)

18S rRNA

B.

Figure 5.5: Northern hybridisation of MATR3 mRNA in ovine liver A) Scanned Northern blot showing expression of the matrin 3 mRNA in livers of resistant and susceptible selection-line animals, both naïve and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. Significance was assessed by one way ANOVA followed by Tukey’s post-hoc test. C) data were grouped according to selection-line, the expression levels are mean ± SD for the two groups of animals. Significance was assessed using two sample T-tests *** P < 0.001.

Re-examination of the BLIB005485HT sequence using blastx (to compare the EST sequence to protein databases) indicates that it may be detecting cytochrome c
oxidase subunit I. This is a key component of the mitochondrial respiratory chain, and in cattle the transcript size is 1.5 kb, broadly consistent with the 1.8 kb hybridisation band. More work would be required to determine what transcript the BLIB005485HT EST sequence is detecting.

Quantification of the Northern blot revealed significant differences in the intensity of the hybridisation band; in particular naïve susceptible animals expressed this gene 1.3 fold higher than sporidesmin dosed resistant animals, and sporidesmin dosed susceptible animals expressed this gene 1.3 fold higher than sporidesmin dosed resistant animals (overall ANOVA $P = 0.006$).

Sporidesmin dose was then excluded as a variable and the expression of this gene was examined between selection-lines (Fig. 5.5C). Susceptible animals were found to express this gene 1.3 fold higher than susceptible animals ($P < 0.001$).

5.2.8.4 Metallothionein II (MT2)
The fourth gene examined was metallothionein II (MT2) which is a ubiquitously expressed gene that is part of a highly conserved gene family. Members of this family have very high cysteine content and function as transporters and cellular stores of heavy metals, primarily zinc and copper.

Metallothionein II (MT2) was selected for confirmation as it is the most significant differentially expressed gene detected on the microarray and is reported to be over-represented in naïve animals. MT2 is also functionally interesting as MT2 and other metallothionein proteins function as transporters and storage molecules for transition metals, in particular zinc and copper (Coyle et al., 2002). As zinc is known to act prophylactically to protect sheep from sporidesmin exposure (refer 1.5) MT2 is a good candidate gene for FE.
Figure 5.6: Northern hybridisation of MT2 mRNA in ovine liver. A) Scanned Northern blot showing expression of the MT2 mRNA in livers of resistant and susceptible selection-line animals, blot naïve and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. Significance was assessed using one-way ANOVA followed by Tukey’s post-hoc test. C) data were grouped according to selection-line. D) data were grouped according to whether animals had been exposed to sporidesmin or not. For both C and D, expression levels are mean ± SD for the two groups of animals. Significance was assessed using two sample T-tests * P < 0.05, ** P < 0.01.
Northern blotting of sheep liver RNA using BLIB009138HT as a probe revealed two hybridisation bands (Fig. 5.6A). The smaller hybridisation band (transcript B) was 0.7 kb and the larger band (transcript A) was 2.2 kb.

The metallothionein gene family members are highly homologous and the gene family is very pleiotropic between species. There is thought to be nine members of the metallothionein gene family in sheep (Peterson et al., 1988) of which six have been characterised. This six members include MT1a, MT1b, MT1c and MTII which are thought to be linked on ovine chromosome 14 (Peterson et al., 1988). All four isoforms have been detected in the sheep liver, but MT1a and MTII are expressed most highly (Peterson and Mercer, 1988). The remaining two characterised ovine metallothionein genes are MT3, which is expressed specifically in the brain (Chung et al., 2002), and MT4. This gene family has undergone significant divergence within the eutherian mammals as the human metallothionein family has 22 members including four processed pseudogenes, while mice only have 4. The majority of the 22 family members in humans fall in two clusters, one on chromosome 1 the other chromosome 16 suggesting that they have arisen relatively recently by gene duplication and divergence.

The localisation of four sheep metallothionein genes (MT1a, MT1b, MT1c and MTII) to two overlapping contigs (Peterson et al., 1988) on chromosome 14 suggests that these genes have also evolved due to gene duplication. Because of these recent gene duplications the members of the metallothionein gene family in both sheep and humans are fairly well conserved (> 60%) and the size of these transcripts varies between 0.3 kb and 0.6 kb, broadly consistent with the size of the smaller hybridisation band (transcript B). However, comparing the sequence of BLIB009138HT to the ovine sequences reveals that it has the highest identity (91%) to the annotated ovine MT1a (X04626) although it also shares 86% identity to the annotated ovine MTII. Because of the homology and consistency among transcript sizes it can not be determined with any certainty which metallothionein transcripts are being detected on the Northern blot.

The larger hybridisation band (transcript A) was 2.2 kb and was not consistent with any annotated alternative transcript and may represent cross hybridisation to another
gene family member. Three human metallothionein genes are larger than 0.6 kb, two of these are annotated pseudogenes, MT1JP (NG_005506) which is 1.8 kb and MT1L (NR_001447) which is 1.1 kb. The third is MTL5 (metallothionein like protein 5) which has a transcript variant ranging in size from 2.5 kb and 3.6 kb, broadly consistent with the larger size transcript. However, MT2 shares only ~50% identity and hybridisation kinetics suggest it would be unlikely for the Northern probe to bind. Therefore, due to this massive gene expansion in the human lineage it is impossible to accurately speculate what the identity of this larger hybridisation band may be as it could be a gene unique to the ruminant lineage.

Quantification of the Northern blot revealed only a very minor difference in expression for transcript A (Fig. 5.6Bi): sporidesmin-dosed susceptible animals expressing the 2.2 kb transcript 1.3 fold higher than sporidesmin-dosed resistant animals (overall ANOVA $P = 0.025$).

The major differences in expression were observed with the smaller transcript B (Fig. 5.6Bii). Transcript B was expressed 1.4 fold higher in naïve resistant, and sporidesmin-dosed susceptible animals as compared to sporidesmin-dosed resistant animals, and 1.6 fold higher in naïve susceptible animals (overall ANOVA $P = 0.003$). This implies that sporidesmin-dosed resistant animals repress the expression of transcript B by 30% following exposure to sporidesmin, while susceptible animals maintain relatively high levels of metallothionein gene expression irrespective of exposure to sporidesmin.

This effect is significant enough that when exposure to sporidesmin is excluded from the model and the expression of metallothionein is considered based on selection lines (Fig 5.6C) susceptible animals express 1.3-fold higher levels of metallothionein (transcript B) than do resistant animals ($P = 0.010$).

When the genetic line of the animals is excluded from the statistical model and the expression of MT is examined as a function of the genetic background of the animals (Fig 5.6D) it is clear that naïve animals express 1.3-fold higher levels of transcript B than animals exposed to sporidesmin ($P = 0.027$).
The differences detected relating to selection line (Fig 5.6C) and sporidesmin-exposure (Fig 5.6D) probably reflect the magnitude of the repression of metallothionein seen in sporidesmin-dosed resistant animals (Fig. 5.6Bii).

The data suggests that animals resistant to sporidesmin repress their expression of MT2 following exposure to sporidesmin. Given the role of MT2 in zinc metabolism, and the role of zinc in protecting against FE, this raises interesting questions regarding the functionality of this differential expression. Transcription of MT2 is enhanced by MTF1 (metal responsive element binding transcription factor 1) which in turn is negatively regulated by NFI (nuclear factor I) type proteins.

As metallothionein is a possible functional candidate for resistance to FE comparative mapping was undertaken for the metallothionein family, and the transcription factors MTF1 and NFI. The comparative map positions were then compared with existing QTL data (S. H. Phua, pers. comm.). None of the genes investigated fell into regions of the sheep genome known to be associated with resistance to FE.

5.2.8.5 Ras-related C3 botulinum toxin substrate 2 (RAC2)

RAC2 is specific to haematopoietic cell lineages and the observed expression in the liver could be due to the presence of Kupffer cells or infiltrating neutrophils. However, RAC2 is closely related to RAC1 (92 % identical) and it is possible that both the microarray and Northern blot analysis were detecting the RAC1 transcript which is fairly ubiquitously expressed. Both RAC1 and RAC2 are plasma membrane-associated small GTPases which cycle between an active GTP-bound and inactive GDP-bound state. In active state the RAC molecules bind to a variety of effector proteins to regulate cellular responses, such as secretory processes, phagocytosis of apoptotic cells and epithelial cell polarization.
A. 

<table>
<thead>
<tr>
<th></th>
<th>Naive Resistant</th>
<th>Naive Susceptible</th>
<th>Sporidesmin dosed Resistant</th>
<th>Sporidesmin dosed Susceptible</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R8</td>
<td>R9</td>
<td>R10</td>
<td>S8</td>
</tr>
<tr>
<td></td>
<td>R10</td>
<td>R11</td>
<td>R12</td>
<td>R13</td>
</tr>
<tr>
<td></td>
<td>R14</td>
<td>R15</td>
<td>R16</td>
<td>S11</td>
</tr>
<tr>
<td></td>
<td>R17</td>
<td>S12</td>
<td>S13</td>
<td>S14</td>
</tr>
<tr>
<td></td>
<td>S15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAC2</td>
<td>(2.1 kb)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>(1.7 kb)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>(1.2 kb)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

18S rRNA

B.

i. 

Transcript A (2.1 kb)

Expression Ratio

![Graph showing expression ratio for Transcript A (2.1 kb)]

ii. 

Transcript B (1.7 kb)

Expression Ratio

![Graph showing expression ratio for Transcript B (1.7 kb)]

iii. 

Transcript C (1.2 kb)

Expression Ratio

![Graph showing expression ratio for Transcript C (1.2 kb)]

C.

Transcript A (2.1 kb)

Transcript B (1.7 kb)

Transcript C (1.2 kb)

Expression Ratio

![Graph showing expression ratio for all transcripts]
Northern blotting of sheep liver RNA using BLIB006026HT as a probe revealed three hybridisation bands (Fig 5.7A), the sizes of these bands were determined to be 2.1 kb (transcript A), 1.7 kb (transcript B) and 1.2 kb (transcript C). The human RefSeq sequences are 1.5 kb in length for RAC2 (NM_002872) and 2.4 kb in length for the three transcript variants of RAC1 (NM_006908, NM_018890, NM_198829). By comparison with the sizes expected from the human RefSeqs the hybridisation transcripts A and B could represent the ovine orthologs of the RAC1 and RAC2 genes, respectively, or alternative transcripts of either of these genes. Interestingly the *B. taurus* RAC1 sequence is reported to be 1 kb in length which is not only a lot smaller than the human homolog but is also consistent with the size of transcript C.

Quantification of the Northern blot revealed no differences in expression of transcript A or B (Fig. 5.7Bi,ii). Differential expression was observed for the smaller transcript, transcript C (Fig. 5.7Biii), where the levels of this transcript in sporidesmin-dosed susceptible animals were 1.2-fold higher than the levels in either naïve resistant or sporidesmin-dosed resistant animals.

When sporidesmin exposure is excluded as a variable and the expression of RAC2 is considered based on selection line (Fig 5.7C) a significant difference is observed for both transcripts A and C, where transcript A is expressed 1.1 fold higher in susceptible animals than resistant animals (*F* = 0.029) and transcript C is expressed 1.1 fold higher in susceptible animals (*P* < 0.001).

Little is known about the transcriptional regulation of either RAC1 or RAC2 so the significance of this differential expression is unclear. However, if the enhanced expression of RAC1 in the livers of susceptible sheep (Fig. 5.7C) leads to increased activation of RAC1 this would hypothetically have a detrimental effect. The main physiological consequences of RAC1 activation involve cytoskeletal remodelling.

---

**Figure 5.7:** (opposite page) Northern hybridisation of RAC2 mRNA in ovine liver. A) Scanned Northern blot showing expression of the RAC2 mRNA in livers of resistant and susceptible selection-line animals, both naïve and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. Significance was assessed by one way ANOVA followed by Tukey's post-hoc test. C) data werea grouped according to selection-line. Expression levels are mean ± SD for the two groups of animals. Significance was assessed using two sample T-tests *P* < 0.05, ***P* < 0.001.
Chapter Five: cDNA Microarrays as a Tool for Candidate Gene Identification

and ROS production (Brown et al., 2006). Cytoskeletal remodelling has been indicated as one of the major morphological effects of sporidesmin in vitro (Jordan and Pedersen, 1986). Additionally, activation of RAC1 in hepatic stellate cells is known to increase the severity of liver disease (Choi et al., 2006) and inactivation of RAC1 is known to protect against ischemic/reperfusion injury in rats (Harada et al., 2003). Therefore it is possible that a small increase in expression of RAC1, such as that observed for transcript C, may mediate a significant physiological response.

5.2.8.6 C-reactive protein (CRP)

C-reactive protein (CRP) was chosen for confirmation as the microarrays predicted that CRP was differentially expressed between naïve and sporidesmin-dosed animals independent of the genetic line ($P = 0.02$). However, the expression ratio was reported to be 0.96 which is within the 0.9-1.1 fold cut-off selected for this experiment. Therefore CRP was one of the genes selected to examine the appropriateness of these cut-offs. CRP is also functionally interesting as it is an acute phase protein and has several functions associated with host defence: it promotes agglutination, bacterial capsular swelling, phagocytosis and complement fixation through its calcium-dependent binding to phosphorylcholine (reviewed in Mortensen, 2001). CRP would also be a good biomarker as it is almost exclusively expressed in the liver and changes in liver expression should be accurately reflected in the serum.

Northern blotting of sheep RNA using BLIB003400HT as a probe revealed a single hybridisation band (Fig. 5.8A). The size of this hybridisation band was inferred to be 3.2 kb. This is significantly bigger than the 2 kb expected based on the human RefSeq sequence (NM_000567). However, the 3' UTR is quite variable in size and in B. taurus the predicted CRP mRNA is 3.9 kb in size (XM_605946).

Quantification of the Northern blot (Fig. 5.8B) revealed a high degree of between-animal variation in the hepatic expression of CRP. Therefore while it appears that resistant animals had higher expression following exposure to sporidesmin, this difference was not statistically significant.
**Chapter Five: cDNA Microarrays as a Tool for Candidate Gene Identification**

Figure 5.8 Northern hybridisation of CRP mRNA in ovine liver. A) Scanned Northern blot showing expression of the C-reactive protein (CRP) mRNA in livers of resistant and susceptible selection-line animals, both naïve and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. Significance was assessed by one way ANOVA.

5.2.8.7 **Fibrinogen Alpha (FGA)**

FGA was chosen for confirmation as the microarray analyses indicated that FGA was differentially expressed in naïve resistant animals \( (P=0.014) \). However, the expression was within the 0.9 – 1.1 fold cut off for this experiment (0.95) and FGA was one of the genes selected for determining the appropriateness of these cut-offs.

FGA is expressed at very high levels in the liver. As an acute phase protein the plasma concentration of FGA greatly increases during the response to tissue injury, infection or other inflammatory stimuli thereby making FGA a potentially good biomarker for FE.
A.

<table>
<thead>
<tr>
<th></th>
<th>Naïve</th>
<th>Naïve</th>
<th>Sporidesmin dosed</th>
<th>Sporidesmin dosed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resistant</td>
<td></td>
<td></td>
<td>Resistant</td>
<td>Susceptible</td>
</tr>
<tr>
<td>R8</td>
<td>R9</td>
<td>R10</td>
<td>S8</td>
<td>S9</td>
</tr>
<tr>
<td>S10</td>
<td></td>
<td></td>
<td>M</td>
<td></td>
</tr>
</tbody>
</table>

FGG (1.6 kb)

B.

i. (Transcript A (1.6 kb))

ii. (Transcript B (1.2 kb))

C.

Figure 5.9: Northern hybridisation of FGA mRNA in ovine liver. A) Scanned Northern blot showing expression of the FGA mRNA in livers of resistant and susceptible selection-line animals, both naïve and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. Significance was assessed by one way ANOVA followed by Tukey’s post-hoc test. C) data were grouped according to whether animals had been exposed to sporidesmin or not. Expression levels are mean ± SD for the two groups of animals. Significance was assessed using a two sample T-test * P < 0.05.
Northern blotting of sheep liver RNA using BLIB002911HT as a probe revealed two hybridisation bands (Fig. 5.9A). These two hybridisation bands were sized at 1.6 kb (transcript A) and 1.2 kb (transcript B). The size of transcript A is broadly consistent with the 1.7 kb size of the human RefSeq (NM_000509). There is at least one alternative transcript of the human FGA, but it is 2.2 kb (Kim et al., 2005a; Kim et al., 2005b) and is larger than either of the two hybridisation bands obtained for sheep.

Quantification of the Northern blot revealed no differential expression between individual groups of animals for either transcript A or B (Fig. 5.9Bi, ii).

However, when the information concerning the genetic-line of the animals is excluded and the expression of FGA is examined between naïve and sporidesmin-dosed animals (Fig. 5.9C), sporidesmin-dosed animals have higher levels of both transcripts A and B. Transcript A is expressed 1.4 fold higher \((P = 0.048)\) and transcript B 1.5 fold higher \((P = 0.044)\) in sporidesmin-dosed animals.

The higher expression of FGA in sporidesmin-dosed animals may be indicative of an acute phase response induced by sporidesmin, or may indicate hepatocellular regeneration as levels of FGA have been found to peak approximately 12 hours following hepatic injury (Fulop et al., 2001).

5.2.8.8 Glutathione S-transferase, alpha-2 (GSTA2)

GSTA2 was reported to be expressed more highly in naïve susceptible animals \((1.06 \text{ fold}, P = 0.030)\) and also differentially expressed between resistant and susceptible sheep irrespective of exposure to sporidesmin \((1.04 \text{ fold}, P = 0.040)\) by microarray analysis. However, the expression in each case fell within the 0.9 - 1.1 fold cut-off threshold used in this study. GSTA2 is functionally interesting as it is involved in xenobiotic metabolism as a phase II enzyme which functions to conjugate reduced glutathione to a wide number of exogenous and endogenous hydrophobic electrophiles, possibly including sporidesmin (refer 1.7.2.1).
Chapter Five: cDNA Microarrays as a Tool for Candidate Gene Identification

A.

<table>
<thead>
<tr>
<th></th>
<th>Naive Resistant</th>
<th>Naive Susceptible</th>
<th>Sporidesmin dosed Resistant</th>
<th>Sporidesmin dosed Susceptible</th>
</tr>
</thead>
<tbody>
<tr>
<td>R8</td>
<td>R9</td>
<td>S8</td>
<td>R11</td>
<td>R12</td>
</tr>
<tr>
<td>R10</td>
<td>S9</td>
<td>S10</td>
<td>R13</td>
<td>R14</td>
</tr>
<tr>
<td>R11</td>
<td>R12</td>
<td>R13</td>
<td>R15</td>
<td>R16</td>
</tr>
<tr>
<td>S11</td>
<td>S12</td>
<td>S13</td>
<td>S14</td>
<td>S15</td>
</tr>
</tbody>
</table>

GSTA2 (0.8 kb)

18S rRNA

B.

![Northern blot image]

Figure 5.10: Northern hybridisation of GSTA2 mRNA in ovine liver. A) Scanned Northern blot showing expression of the GSTA2 mRNA in livers of resistant and susceptible selection-line animals, both naïve and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. Significance was assessed by one way ANOVA followed by Tukey’s post-hoc test.

Northern blotting of sheep liver RNA using BLIB004280HT as a probe revealed a single hybridisation band (Fig. 5.10A). This band was sized at 0.8 kb by comparison with the RNA ladder, which is slightly smaller than the expected size of 1 kb based on the size of the human RefSeq (NM_000846).

Quantification of the Northern blot revealed no differences in the expression of GSTA2. However, there was a relatively large degree of between-animal variation which may mask the small changes in expression predicted by the microarray experiments.
5.2.8.9 Endothelial pas domain protein 1 (EPAS1)

EPAS1 was one of the two genes selected for analysis by Northern blot that were not reported to be differentially expressed based on the microarray analysis. EPAS1 was selected as it is functionally interesting: it is fairly ubiquitously expressed and functions as a transcription factor involved in the induction of oxygen regulated genes.

Northern hybridisation of sheep liver RNA using BLIB007891HT as a probe revealed a single hybridisation band of 5.0 kb (Fig. 5.11A). This size is consistent with the expected size of 5.2 kb based on the human RefSeq sequence (NM_001430).

A.

![Blot showing EPAS1 and 18S rRNA](image)

B.

![Graph showing expression ratio](image)

**Figure 5.11:** Northern hybridisation of EPAS1 mRNA in ovine liver. A) Scanned Northern blot showing expression of the EPAS1 mRNA in livers of resistant and susceptible selection-line animals, both naive and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. Significance was assessed by one way ANOVA followed by Tukey’s post-hoc test.
Quantification of the Northern blot and normalisation to the 18S rRNA levels revealed no differences in expression of EPAS1 (Fig. 5.11B). Overall the expression of EPAS1 was extremely low and quantification of the Northern blot is likely to be inaccurate.

5.2.8.101F-κ-B-interacting ras-like protein 2 (KBRAS2, NKIRAS2)

KBRAS2 was not reported to be differentially expressed by the microarray experiment and it was selected for Northern analysis as it is functionally very interesting and is involved in modulating the expression of a great number of other genes. It is ubiquitously expressed and interacts with IkB to slow its rate of degradation. IkB are part of a family of proteins that bind to the transcription factor NFκB via ankyrin repeats and sequester it in the cytoplasm of the cell. Phosphorylation of IkB by intracellular signalling cascades cause ubiquitation of the

![Northern blot](image)

**Figure 5.12:** Northern hybridisation of KBRAS2 mRNA in ovine liver A) Scanned Northern blot showing expression of the KBRAS mRNA in livers of resistant and susceptible selection-line animals, both naïve and sporidesmin dosed. B) The Northern blot was quantified by densitometry and expression normalised to 18S rRNA levels. Expression levels are mean ± SD for the two groups of animals. Significance was assessed by one way ANOVA followed by Tukey’s post-hoc test.
The IκB molecule which targets it for degradation by the proteosome. This allows the NFκB molecule to translocate to the nucleus where it can act as a pleiotropic regulator of many aspects of cellular activity, in particular stress, injury and especially in pathways of the immune response. Based on the GO categories identified (Table 5.9, 5.10) NFκB is likely to play a significant role in modulating genes in these categories as differential expression of KBRAS2 could modulate the NFκB response.

Northern analysis of sheep liver RNA using BLIB003342HT as a probe revealed a single hybridisation band of 2.2 kb, which is smaller than the expected size of human RefSeq of 3 kb (NM_001001349). However, examination of the sequence of BLIB003342HT reveals that it is actually most closely related to a hypothetical protein (NM_001013624) which appears to be a zinc finger transcription factor. The expected size of this hypothetical gene is 2 kb in human and it may be this gene product that is being detected on the Northern hybridisation.

Quantification of the Northern blot revealed no differences in gene expression although quantification was made difficult by the extremely low levels of expression observed on the Northern hybridisation.

5.2.9 Evaluation of the accuracy of the microarray data
Northern hybridisation was carried out not only to verify the predicted differential expression in a larger set of animals than used for the microarray but more fundamentally to verify the accuracy of the microarray experiment in predicting differential expression.

To assess this accuracy Pearson correlation coefficients were calculated comparing the Northern hybridisation data for the four individual animals used in the microarray experiment with the microarray expression data (Table 5.12).

The only significant correlation obtained was between the larger transcript of metallothionein (transcript A) and the microarray data \( (P = 0.016) \). In all other cases
the expression data obtained from the microarray does not correlate with the data obtained by Northern analysis for the four microarray animals. However, previous studies have indicated that correlations of around 0.7 are expected for comparison of microarray data with that obtained by another method (qRT-PCR) (Czechowski et al., 2004). If a correlation of greater than 0.7 is accepted then GSTA2, LECT2, MT2 (transcript A and transcript B) are all correlated with the microarray data. Further to this, CAT and FGA are expressed on the Northern blot in the same general trend predicted by the microarray experiment (Table 5.13). However, as a strict mathematical correlation does not exist it must be concluded that the microarray data is not a good predictor of the true expression of these individual genes based on Northern blot analysis.

Table 5.12: Pearson's correlation coefficients for microarray data as compared with Northern expression data.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Transcript</th>
<th>Pearson's correlation coefficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT</td>
<td>A</td>
<td>0.312</td>
<td>0.688</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.108</td>
<td>0.892</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>-0.075</td>
<td>0.925</td>
</tr>
<tr>
<td>CRP</td>
<td></td>
<td>-0.478</td>
<td>0.522</td>
</tr>
<tr>
<td>EPAS1</td>
<td></td>
<td>0.397</td>
<td>0.603</td>
</tr>
<tr>
<td>FGA</td>
<td>A</td>
<td>0.538</td>
<td>0.462</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.190</td>
<td>0.810</td>
</tr>
<tr>
<td>GSTA2</td>
<td></td>
<td>0.923</td>
<td>0.077</td>
</tr>
<tr>
<td>KBRAS2</td>
<td></td>
<td>0.038</td>
<td>0.962</td>
</tr>
<tr>
<td>LECT2</td>
<td></td>
<td>0.823</td>
<td>0.177</td>
</tr>
<tr>
<td>MATR3</td>
<td></td>
<td>0.315</td>
<td>0.684</td>
</tr>
<tr>
<td>MT2</td>
<td>A</td>
<td>0.984</td>
<td>0.016*</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.863</td>
<td>0.137</td>
</tr>
<tr>
<td>RAC2</td>
<td>A</td>
<td>0.022</td>
<td>0.978</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>-0.563</td>
<td>0.497</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.119</td>
<td>0.881</td>
</tr>
</tbody>
</table>
Table 5.13: Summary of the trends of differential expression across the microarray and Northern hybridisation data.

<table>
<thead>
<tr>
<th></th>
<th>Microarray expression data</th>
<th>Northern expression data</th>
<th>Average Northern expression data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(for four animals)</td>
<td>(for the four microarray animals)</td>
<td>(data averaged across all animals used in the Northern analysis)</td>
</tr>
<tr>
<td>Naïve</td>
<td>Dosed</td>
<td>Naïve</td>
<td>Dosed</td>
</tr>
<tr>
<td>Resistant (R10)</td>
<td>Susceptible (S8)</td>
<td>Susceptible (R16)</td>
<td>Susceptible (S11)</td>
</tr>
<tr>
<td>CAT</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>CRP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPAS1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGA</td>
<td>↓*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTA2</td>
<td></td>
<td></td>
<td>↑</td>
</tr>
<tr>
<td>KBRAS2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LECT2</td>
<td>↓*</td>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>MATR3</td>
<td>↓</td>
<td></td>
<td>↑</td>
</tr>
<tr>
<td>MT2</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>RAC2</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
</tbody>
</table>

*"Sporidesmin exposure" refers to genes that are differentially expressed in response to sporidesmin exposure, irrespective of the genetic line of the animals (R10 and S8 vs. R16 and S11)

"Genetic line" refers to genes that were differentially expressed as a result of the genetic-line of the animals, irrespective of their exposure to sporidesmin (R10 and R16 vs. S8 and S11)

* denotes genes that were identified as significantly differentially expressed but fell within the 0.9-1.1 fold differential expression selected as the cut-off for this experiment.
5.3 DISCUSSION

Gene expression profiling via microarray technology has become an increasingly affordable and powerful tool to interrogate complex biological questions. Microarray technology has been applied to understanding individual susceptibility to complex disease traits in many species, including ruminants (Diez-Tascon et al., 2005; Keane et al., 2006) as well as to determine toxin exposure and mechanisms of toxicity (reviewed in Pennie et al., 2004).

FE is a complex disease and is likely to involve many genes each contributing a small effect. Therefore the primary objective of this study was to determine the potential for using microarray technology to examine the gene expression profiles of liver tissue obtained from sheep resistant and susceptible to FE. To further elucidate differences between resistant and susceptible animals the expression profiles of sheep pre- and post-exposure to sporidesmin were also examined. It was hoped that examination of these expression profiles might highlight not only mechanisms of genetic resistance, but also could provide information concerning the toxicity of sporidesmin; such results may ultimately lead the discovery of biomarkers or genetic markers. These markers could be used for selection of resistance or to detect sub-clinical FE.

The microarrays utilised in this study were generated from an in-house bovine EST library and cDNA probes were synthesised from ovine liver RNA. Cross-species hybridisation depends on a number of factors including the length of the probe, the percentage of similarity between the orthologous sequences and the type and distribution of base mismatches (Rinaudo and Gerin, 2004). Hybridisation of sheep cDNA against bovine microarrays was expected to produce specific hybridisation under the conditions used in this experiment because sheep and cattle sequences are highly similar in the coding region (~96%) (McEwan and Crawford, 2004). It has been previously demonstrated that hybridisation conditions, such as those used in this experiment, allow cross-species hybridisation between species with sequence homology of 96% (Adjaye et al., 2004; Rinaudo and Gerin, 2004; Wang et al., 2004b) and bovine EST arrays have been previously used with ovine RNA (Diez-Tascon et al., 2005). One implication of this is that, under the slightly relaxed
stringency used, it would be difficult to differentiate between members of multigene families that show high levels of sequence similarity.

Using fluorescence-labelled sheep liver cDNA as a probe on the bovine arrays, relatively low numbers of genes differentially expressed between FE resistant and susceptible animals were identified: a total of 123 for naïve animals, and 35 for sporidesmin exposed animals. The small number of differentially expressed genes identified is at least in part due to the size of the array (1550 probe sequences). Analysis of the microarray data was complicated by the small number of animals used in the analysis (n = 4) this meant that the expression of each gene from a particular animal was compared to an average expression for that gene across all four animals to obtain differentially expressed genes. This approach is further complicated by the differing ages of the animals used. Both the naïve animals were 2 – 3 months of age, but both the sporidesmin-dosed animals were 36 months of age (Table 2.1 and 2.2). These differences in ages are likely to have profound effects on gene expression in the liver. For example the expression of sex hormones will be markedly different in the naïve animals (as they have yet to reach sexual maturity) as compared to the sporidesmin-dosed animals. But also age related differences in the immune system (McClure et al., 2005), and even hepatic drug detoxification system (Mori et al., 2007) of other animals are well characterised. Therefore the comparison of gene expression across animals with such different ages is likely to have skewed the reported differential expression quite significantly.

The genes that were identified as differentially expressed were relatively diverse in function, and EASEonline (Dennis et al., 2003) was able to identify GO categories of interest (Table 5.9, 5.10). In particular, several genes that were identified as down-regulated in naïve resistant sheep were involved in ion transport and transition metal binding. Transition metals such as zinc and copper are known to modulate the toxicity of sporidesmin in vitro (Munday, 1984a; Munday, 1985; Duncan et al., 2005) and in vivo (Smith et al., 1977). Genes involved in metal absorption, transport and storage have been considered as good candidates in FE for some time (Morris et al., 1988; Phua et al., 1998). GO terms associated with the mitochondrial respiratory chain were also identified in naïve resistant animals and may indicate that these animals have different ATP requirements or dysfunction of the mitochondria. It has
previously been found that regulation of the respiratory chain is critical for the recovery from liver damage (Morris, 1999; Navarro et al., 2005) and this differential expression may implicate the respiratory chain or mitochondrial function in resistance to FE. Further, naïve susceptible sheep were found to have higher expression of genes involved in "response to stimulus", suggesting that even in the absence of sporidesmin the liver of these animals is under stress. There is also a possibility that this animal was not truly naïve, but this is unlikely as all the animals used in this experiment were grazed communally in an FE-free area.

In addition to the Gene Ontology analysis a number of differentially expressed genes were identified as putative biomarkers whether for genetic resistance or as a marker for sporidesmin exposure (Table 5.8). The most promising of these for detecting sporidesmin exposure was fibrinogen-alpha (FGA). FGA was reported to be under-represented in both naïve resistant and naïve susceptible animals. The expression of FGA was confirmed by Northern analysis (Fig. 5.9) and was found to be expressed 1.4 fold higher in sporidesmin-dosed animals as compared to naïve animals ($P = 0.048$). If this difference proves to be robust and is detectable in the serum of sheep it may have some utility as a biomarker for surveillance for sporidesmin exposure. However, the confidence interval for this difference is very small and it is difficult to speculate whether such a modest difference in hepatic expression of FGA would be detectable in the serum of the animals.

At the time of this study microarray technology was a relatively new tool and the bovine microarray slides used in this study had not been used before. Because of this two quality control experiments were carried out. Initially, a number of EST clones were sequence verified. A false identity rate of 31.25 % was identified, which in the first instance is quite high but consistent with that obtained for other cDNA microarray data sets (reviewed in Kothapalli et al., 2002).

The second quality control experiment involved confirming the differential expression of a subset of the differentially expressed genes using Northern analysis. The confirmation of differential expression was a critical step as microarrays are prone to several artefacts, including data condensation (Yuen et al., 2002). Additionally, as only a small number of animals were available for the microarray
experiment, it was imperative to confirm the differential expression in a larger group of animals. Lastly, an arbitrary cut-off threshold for the levels of differential expression was used (0.9 – 1.1 fold) and it was crucial to determine the appropriateness of these cut-offs.

Ten genes were selected for confirmation by Northern analysis. These genes were selected because they represented both the highest and lowest differentially expressed genes on the microarray (MT2 and MATR3). Three further genes were selected that were differentially expressed and were functionally interesting, in particular LECT2, CAT and RAC2 all of which are proposed to have a role in response to cell stress. These genes were grouped in one of the following three categories: significantly differentially expressed \((P < 0.05)\) and outside the thresholds used in this study \((0.9 – 1.1 \text{ fold}, \text{five genes})\), three genes were significantly differentially expressed \((P < 0.05, \text{three genes})\) but within the threshold used in this study and two genes were not differentially expressed \((P > 0.05, \text{two genes})\).

Northern analysis confirmed the differential expression of four of the six genes (FGA, LECT2, MATR3, MT2). Generally, there was no statistical correlation between the expression of any of these genes and the microarray data. The lack of correlation between the microarray data and Northern analysis may be caused by the well known phenomenon of microarray data to be condensed (Yuen et al., 2002). Condensation of microarray data causes the ratios of mRNA expression to be consistently under-represented thus masking true variation in the data. The problem of condensation was exacerbated by the small number of animals used in the study.

The quality of some of the Northern data could also be contributing to the lack of statistical correlation observed. In particular, the lack of linearity between the film and exposure to the radioactive signal was not specifically taken into account. Also, the expression of all ten genes was normalised to the ethidium bromide stained 18S rRNA. This normalisation strategy is not ideal as it fails to account for unequal transfer of the RNA from the gel to the blot and the 18S rRNA tends to be present at very high levels and may therefore mask true differences in gene expression. Additionally, ethidium bromide staining is semi-quantitative but previous studies have found it sufficiently sensitive for quantification of gene expression (Adjaye et
al., 2004). Also, two RNA samples (R13 and R14) are of dubious quality based on the intensity of the 18S and 28S rRNA bands. In addition to these complications many of the genes examined by Northern hybridisation exhibited a large degree of between-animal variation. As the number of animals used in the initial microarray experiment was small (n = 4) the between-animal variation did lead to genes being identified as differentially expressed erroneously. Some of the observed between-animal variation may be accounted for by sampling differences, as it is known that marked differences in gene expression can be detected between different areas of the liver (Irwin et al., 2005). Perhaps more significantly, the samples used for the Northern blot analysis were obtained from both male and female animals and also from animals of different ages (naïve resistant animals were 2 – 21 months of age, sporidesmin dosed resistant animals were 9 – 60 months of age, naïve susceptible animals were 2 – 36 months of age, and sporidesmin-dosed susceptible animals were 9 – 21 months of age, refer Table 2.1 and 2.2). The distribution of both males and females of different ages through these groups is likely to have had a significant impact on variation in gene expression. These animals are likely to have different circulating levels of sex hormones that are known to have a significant effect on hepatic gene expression in experimental models, for example rats (Stahlberg et al., 2005). But also the age related differences are likely to have a significant effect on gene expression, including genes of the immune system and hepatic drug detoxification system, as previously discussed (McClure et al., 2005; Mori et al., 2007). Further complicating this experiment, two of the resistant animals and one of the susceptible animals received higher doses of sporidesmin than the remaining sporidesmin exposed animals. The resistant animals were given 0.4 mg/kg (whereas the majority of animals were given 0.1 mg/kg) and the susceptible animal was given 0.15 mg/kg (whereas the majority of animals were given 0.1 mg/kg). In the case of the resistant animals this is a 4-fold higher exposure to sporidesmin, and this is likely to have a significant effect on hepatic gene expression.

Of the five genes that were determined to be differentially expressed two are of specific interest. The first of these is catalase, which was found to be expressed at higher levels in susceptible animals as compared with resistant animals. This was somewhat unexpected as the microarray experiment predicted the opposite result, that resistant animals expressed higher levels of the catalase mRNA. This difference
is readily explained; as examination of only the four animals used in the microarray experiment confirms that the naïve resistant animal (R10) does in fact express catalase higher than expected. However, the remaining resistant animals do not show this trend and overall, susceptible animals express catalase mRNA 1.5 fold higher than resistant animals.

The higher expression of catalase in the livers of susceptible animals goes against conventional thinking that higher levels of catalase would protect against sporidesmin toxicity by facilitating the detoxification of hydrogen peroxide produced by sporidesmin (Munday, 1982). Additionally, high hepatic catalase activity is known to protect against hepatotoxicity in other species (Yabe et al., 1999a; Yabe et al., 1999b; Ma et al., 2006).

Paradoxically, catalase activity has reported to be higher in the serum of resistant selection-line animals (Hohenboken et al., 2004) and while the increase is relatively modest, it is opposite to the observed hepatic mRNA expression obtained by Northern analysis. Catalase is highly expressed in erythrocytes, liver and kidney tissue and serum catalase activity is enhanced in haemolytic diseases, acute pancreatitis and liver disease (Goth et al., 1987; Goth, 1989). In humans serum catalase activity can also be influenced by a genetic variant in the catalase promoter which is correlated with higher serum catalase activity (Forsberg et al., 2001).

Despite the paradoxical differences in catalase gene expression and serum activity, the finding that catalase is differentially expressed in the livers of FE selection-line animals is extremely interesting. Markers surrounding the catalase locus have previously been shown to segregate with resistance or susceptibility (Phua et al., 1999). The segregation of markers surrounding this locus and the observed differential expression requires further investigation.

The second gene of interest is metallothionein II (MTII). The metallothionein family of proteins function as transporters and storage molecules for transition metals, in particular zinc and copper (Coyle et al., 2002). Both zinc and copper are thought to have a role in sporidesmin toxicity: copper promotes the generation of the superoxide
radical (refer to 1.3.4.1), and zinc is known to protect against sporidesmin toxicity (refer 1.5) by an unknown mechanism.

Northern hybridisation revealed that transcript B was expressed 1.4 fold higher in naïve resistant and sporidesmin-dosed susceptible animals as compared to sporidesmin-dosed resistant animals; and 1.6 fold higher in naïve susceptible animals (overall ANOVA \( P = 0.003 \)) (Fig. 5.6ii). Interpretation of the expression patterns reveals that sporidesmin-dosed resistant animals repress the expression of transcript B by 30% following exposure to sporidesmin, while susceptible animals maintain relatively high levels of metallothionein gene expression irrespective of exposure to sporidesmin.

Metallothionein is known to be transcriptionally repressed in two ways, the first involving alteration of the methylation state of the metallothionein promoter (Ghoshal et al., 2000; Majumder et al., 2006) and the second involving the transcription factor NFI which is a negative regulator of MTF1 (the major transcription factor responsible for metallothionein gene expression) (Majumder et al., 2001; Jacob et al., 2002). Comparative mapping for the metallothionein genes, the MTF1 and NFI genes was undertaken and the resulting map positions compared with pre-existing QTL data: no evidence for the involvement of any of these genes was indicated.

MTF1 activity is regulated by the levels of free zinc (Bittel et al., 1998) and, at least in vitro, sporidesmin can form a complex with zinc (Woodcock et al., 2001a; Woodcock et al., 2001b). Therefore it is possible that sporidesmin entering the cell may bind zinc either from the free zinc pool (Bozym et al., 2006), or by sequestering it from zinc binding proteins in the cell. This would reduce the levels of zinc in the cell and would repress the activity of MTF1. The repression of MTF1 would in turn reduce the transcription of MT2. The binding of sporidesmin by zinc would also reduce toxicity by preventing redox cycling and concomitant tissue damage (refer to 1.3.4). Alternatively if the decrease in metallothionein mRNA was mediated by another (as yet unknown) factor, this could result in an increase in the labile intracellular zinc pool which would then be able to interact with sporidesmin. It would be interesting to determine if the down-regulation of metallothionein in
resistant animals is a cause or consequence of the alteration of the intracellular free zinc pool. The differential expression of metallothionein detected using microarrays is the first experimental evidence to implicate this gene family as candidates for FE.

In conclusion, the microarray analysis presented was able to identify several candidate genes for resistance to facial eczema disease. These genes were identified despite the substantial limitations and problems discussed. Perhaps more significantly this study demonstrated the utility of microarray technology in FE research and by highlighting potential problems and pitfalls it will improve the design of future experiments in this area.
Chapter Six

An *in vitro* system to examine the mechanism of zinc protection against sporidesmin toxicity

6.1 INTRODUCTION

6.1.1 Epidithiodioxopiperazine (ETP) toxicity

ETP mycotoxins are a class of toxins that possess a diketopiperazine ring and an internal disulphide bridge (refer 1.3.1). ETP toxins, such as sporidesmin and gliotoxin are extremely cytotoxic both *in vivo* (Done *et al*., 1961; Towers and Stratton, 1978) and *in vitro* (Mortimer and Collins, 1968; Atherton *et al*., 1974; Munday, 1984a; Munday, 1984b; Jordan and Pedersen, 1986). Almost all of the work examining the toxic mechanisms of the ETP compounds has utilised either gliotoxin or sporidesmin and usually, in an *in vitro* system.

The toxicity of these ETP compounds is entirely mediated by the disulphide bridge (Trown and Bilello, 1972; Mullbacher *et al*., 1986), and there is evidence for two general mechanisms of toxicity (refer to 1.3.4). Briefly, the first mechanism involves redox cycling of the ETP compounds, which is facilitated by redox active metals such as copper. The redox cycling generates reactive oxygen species such as superoxide and hydrogen peroxide (Munday, 1984a; Munday, 1987), which can mediate widespread cellular damage including protein carbonylation, membrane peroxidation and DNA damage (reviewed in Marnett, 2002). The second mechanism of toxicity involves the formation of mixed disulphides between the ETP compound and cellular proteins. The formation of these ETP-protein complexes has been shown to occur for gliotoxin and alcohol dehydrogenase (Waring *et al*., 1995), creatine kinase (Hurne *et al*., 2000) and ras farnesyltransferase (Hara and Han,
1995). Recently, sporidesmin has been shown to form a mixed disulphide complex with glutaredoxin (Srinivasan et al., 2006).

6.1.2 Zinc, copper and facial eczema disease
Currently, the only successful control measure for FE is the prophylactic treatment of animals with high levels of zinc (up to 50 mg/kg) (Smith et al., 1977). However, the physiological and molecular mechanisms underlying this protection are unknown.

Zinc is the most abundant trace element in the cell with a physiological level of ~ 10 μM (Cousins, 1996). The intracellular function of zinc is varied and it has a key role in the catalytic mechanism, structure and regulation of many proteins (Cousins, 1996). In fact, zinc is a structural component or cofactor for over 300 metallo-enzymes (Kambe et al., 2004), including zinc finger transcription factors. Zinc can also modulate major cellular signalling pathways, including calcium signalling, cAMP metabolism, MAPK signalling and protein kinase C signalling (Beyersmann and Haase, 2001). More generally, zinc is required for cell proliferation, differentiation and apoptosis (Beyersmann and Haase, 2001; Truong-Tran et al., 2001): zinc also has a role in maintenance of the immune system (Ibs and Rink, 2003) and systemic responses to sepsis and inflammation include hepatic zinc uptake (Cousins et al., 2006). Recently, a zinc transporter (Zip14) was identified as being highly induced in the hepatocytes of mice during the acute phase response (Liuzzi et al., 2005).

6.1.2.1 Putative mechanisms for zinc protection against facial eczema disease
Zinc is generally regarded as a cytoprotectant and supplementation of cells with zinc is known to protect against several toxicities (Cousins, 1996), such as ethanol (Zhou et al., 2005), carbon tetrachloride (Itoh et al., 1997) and lipopolysaccharide (Shea-Budgell et al., 2006). However, the biochemical mechanism or mechanisms effecting this generalised cytoprotection is unclear.

As previously mentioned, zinc is known to modulate the function of the immune system in whole animals (Fraker et al., 2000). This modulation targets both the innate and acquired immune responses (Ibs and Rink, 2003). Zinc also suppresses the production of TNF-α in the liver, thereby acting as an anti-inflammatory agent.
Chapter Six: An in vitro system to examine zinc protection against sporidesmin toxicity

(Zhou et al., 2004). Correct activation of the immune response is important in the early stages of liver injury (Jaeschke and Hasegawa, 2006) and may facilitate toxin clearance, cell proliferation or regeneration.

Alternatively, zinc has the potential to protect animals against sporidesmin toxicity by modifying hepatic copper levels (Bremner et al., 1976). Copper is an essential trace element found in most tissues, but the highest concentration is in the liver (Gaetke and Chow, 2003). Copper has many functions, it is a structural component and cofactor for a variety of enzymes (Gaetke and Chow, 2003). However, copper is potentially deleterious due to its propensity to generate ROS. Copper also acts as a redox catalyst for a number of oxidases (Prožaska and Gybina, 2004). Sheep are very sensitive to copper toxicity (Winge and Mehra, 1990) and protection against this toxicity can be conferred by the administration of zinc (Bremner et al., 1976). It has been shown that zinc, iron, calcium and vitamin C all impair copper absorption (Jacob et al., 1987). However, the ability of zinc to modify hepatic copper stores in sheep has been disputed (Hatfield et al., 2001), but it is possible that the phenomenon is breed specific (van der Schee et al., 1983). The decrease in hepatic copper levels would confer protection against sporidesmin as copper is extremely redox active and it plays a key role in the generation of ROS by sporidesmin (refer to 1.3.4.1) (Munday, 1982). This hypothesis is supported by the fact that iron salts also protect against sporidesmin exposure (Munday and Manns, 1989). Intuitively, iron would be expected to exacerbate sporidesmin toxicity as it is also very redox active.

Although no protective effect was seen by zinc against sporidesmin in isolated rat hepatocytes (Cordiner and Jordan, 1983), it has been observed in several other in vitro systems, in particular in erythrocytes (Munday, 1984a), T-blasts and macrophages (Waring et al., 1990). As zinc is not thought to modulate intracellular copper concentrations in isolated cells in the same way it does in whole animals (Arredondo et al., 2006), in vitro protection by zinc is likely due to another mechanism.

Biochemically zinc is a strong Lewis acid, and is capable of binding to thiolate and amine electron donors. Sporidesmin is thought to form a thiolate anion intermediate (refer to Fig. 1.3) and it is possible that zinc can form a complex with sporidesmin
intracellularly. Indeed, a tetrahedral complex between sporidesmin and zinc has been detected \textit{in vitro} (Woodcock \textit{et al.}, 2001a; Woodcock \textit{et al.}, 2001b) (Fig. 1.5). A weak complex has also been detected \textit{in vivo} by treating macrophages with high concentrations of zinc. But the formation of this complex was not sufficient to prevent DNA damage (Waring \textit{et al.}, 1990), suggesting that the formation of a complex between sporidesmin and zinc does not completely abrogate its toxicity.

At the cellular level, there are several functions of zinc which could be relevant to the suppression of sporidesmin toxicity. Firstly, zinc is known to alter the activity of the P450 enzymes (Jeffery, 1983; Moreno \textit{et al.}, 1989; Tan \textit{et al.}, 1995), and there is evidence for the involvement of these enzymes in sporidesmin metabolism (Fairclough \textit{et al.}, 1978). Also zinc is a potent inhibitor of apoptosis, and sporidesmin is known to induce apoptosis in mouse macrophages and T-blasts (Waring \textit{et al.}, 1990). The inhibition of apoptosis by zinc may force cells exposed to sporidesmin into a pro-survival mode reducing the overall impact of the toxin (Rose \textit{et al.}, 2003).

However, one of the most tantalising prospects originates from the role that zinc plays in modulation of gene expression. One of the major functions of zinc, at a molecular level, is as a co-factor for zinc-containing DNA-binding proteins. Thereby zinc acts as a potent modulator of gene expression by activating transcription factors such as MTF1 (metal regulated transcription factor 1) (Rutherford and Bird, 2004). Several recent studies have analysed the effect of zinc on the global gene expression profile in various cell systems and found that zinc is capable of modulating the expression of hundreds of functionally diverse genes (tom Dieck \textit{et al.}, 2003). By modifying gene-expression zinc is integral to the cellular response to oxidative stress (Baldwin and Benz, 2002). Consequently zinc treatment could potentially induce the expression of genes that protect against sporidesmin toxicity. By inference these genes would be good candidate genes for conferring resistance to FE.

\subsection*{6.1.3 HepG2 cells as an in vitro model system}

As previously mentioned, \textit{in vitro} systems have been used extensively to examine sporidesmin toxicity; these systems include erythrocytes, mouse macrophages and mouse T-blasts. However, sporidesmin primarily affects the liver and bile ducts of
animals and it was therefore important to study sporidesmin toxicity in a biologically relevant context. For this reason HepG2 cells were chosen as an in vitro model system.

HepG2 cells are a human hepatocellular carcinoma cell line that retain many of the morphological features (Sormunen et al., 1993) and gene expression characteristics of hepatocytes (Knowles et al., 1980; Lu and Huang, 1994; Urani et al., 1998; Harris et al., 2004a). HepG2 cells are the most commonly used cell line for hepatotoxicity studies and have been used extensively to examine the toxicity of alcohol-induced hepatotoxicity (Neuman et al., 1993; Wu and Cederbaum, 1999; Zhuge and Cederbaum, 2006) and carbon tetrachloride induced hepatotoxicity (Holden et al., 2000; Beddowes et al., 2003). These cells have also been used to determine the effectiveness of protective compounds against both ethanol and carbon tetrachloride induced hepatotoxicity (Jimenez-Lopez and Cederbaum, 2004; Szuster-Ciesielska and Kandefer-Szerszen, 2005).

In addition, HepG2 cells have been used for gene expression studies. Specifically HepG2 cells have been used to investigate differences in gene expression in response to different toxins, including carbon tetrachloride (Holden et al., 2000), ethanol (Hong et al., 2003) and copper (Muller et al., 2007). The plethora of toxicology and gene expression data available for this cell line facilitates meaningful data comparisons.

6.1.3.1 Zinc and copper homeostasis in HepG2 cells
Given the critical role of zinc in many biochemical processes it is not surprising that its cellular levels are very tightly regulated. Intracellular zinc homeostasis is controlled by influx and efflux transporters, some of which mediate the sequestration of zinc into intracellular vesicles (Beyersmann and Haase, 2001). Zinc uptake from the extracellular fluid or from intracellular vesicles is mediated by the Zip family of proteins, of which there are 14 members in mammals. Conversely the extrusion of zinc from cells and sequestration in intracellular organelles is mediated by the ZnT family of transporters, of which there are 10 in mammals (reviewed in Cousins et al., 2006). Only ZnT1 has been localised to the plasma membrane and is thought to be the primary regulator of cellular zinc efflux (Palmiter and Findley, 1995). ZnT1 is
expressed in HepG2 cells and is responsive to exogenous zinc supplementation (Urani et al., 2003).

Inside the cell much of the cellular zinc is associated with proteins or complexed by anions. The major zinc binding proteins are metallothioneins which act as zinc exchangers (Maret et al., 1999). Levels of labile zinc are very low and are estimated to be 5 – 10 pM (Bozym et al., 2006) as compared with a total cellular concentration of approximately 100 μM (Vallee and Falchuk, 1993). It is known that exogenous zinc supplementation does increase intracellular zinc levels (Urani et al., 2001). However, it is not known if zinc supplementation is able to influence the labile zinc pool.

The intracellular levels of copper are also tightly regulated and are transported into the cell by the SLC31A1 protein. Very little of the transported copper is maintained as labile and most of it is bound to copper containing proteins (Prohaska and Gybina, 2004). In hepatocytes, copper is effluxed via exocytosis which is aided by ATP7B, and the transfer of copper into holoenzymes is facilitated by at least four copper chaperones, ATOX1, ATP7A/B, CCS and COX17 (Prohaska and Gybina, 2004). Copper treatment of HepG2 cells also results in increased intracellular accumulation (Urani et al., 2001)

6.1.4 Aims and Rationale

As previously stated, very little of the research on sporidesmin toxicity has been done in liver cells. The liver is the primary organ affected by sporidesmin toxicity, and it was important to establish a model system more closely approximating the target organ. Therefore the primary aim of this research was to evaluate and establish HepG2 cells as an in vitro model system for studying the effects of sporidesmin.

One of the unresolved questions in sporidesmin research concerns the mechanism of zinc protection against sporidesmin toxicity. Previous studies have shown that zinc can protect against sporidesmin exposure in both isolated erythrocytes and mouse thymocytes (Munday, 1984a; Waring et al., 1990). But the latter study used concentrations of zinc orders of magnitude higher than physiological levels (1000...
µM) and the cells were treated with zinc and sporidesmin simultaneously. Therefore the protection seen in that experiment may have been an artefact due to extracellular sequestration of a zinc-sporidesmin complex. Thus, the second aim of this research was to determine if the pre-treatment of HepG2 cells with zinc could protect against sporidesmin toxicity. The final aim of this work was to evaluate the role of gene transcription conferring this protection against sporidesmin toxicity. Conceptually, the genes that are modulated by zinc would be possible candidate genes for resistance to FE in sheep.
6.2 RESULTS

6.2.1 Sporidesmin cytotoxicity: a comparison of HepG2 cells with HeLa cells

HepG2 cells were selected as the model system for this study for the reasons outlined in 6.1.3. The primary objective of this chapter was to evaluate the response of this cell line to sporidesmin. To do this a dose response curve for sporidesmin was established using the methodology detailed in section 2.5. Briefly, HepG2 cells were plated at a density of \(2 \times 10^6\) cells per well (9.6 cm\(^2\)) in a six-well culture dish in 2 ml of \(\alpha\)-MEM, 10% foetal bovine serum, 50 U/ml penicillin G and 50 \(\mu\)g/ml streptomycin sulphate. The cells were cultured for 36 hours prior to being exposed to sporidesmin.

The HepG2 cells were exposed to a series of sporidesmin concentrations ranging from 2 to 10 \(\mu\)g/ml. In all cases a solvent-only control (ethanol) was included for comparison. Cells were incubated for 16 hours and the percentages of non-viable cells were determined using the LDH assay (refer to 2.6.4.1).

The data were plotted in SigmaPlot and a sigmoidal dose response curve was fitted \((r^2 = 0.99)\) (Fig. 6.1). The dose response curve covered a narrow range of concentrations \((0 - 10 \mu\)g/ml\) and the \(LC_{50}\) of sporidesmin was calculated to be 5.0

![Figure 6.1: Cytotoxic effect of sporidesmin on HepG2 cells. Cells were treated with sporidesmin at the indicated concentrations for 16 h. The percentages of non-viable cells were determined using the LDH assay. Each point represents the mean ± S.D. of three independent experiments each performed in triplicate. A sigmoidal curve (five parameter) was fitted using SigmaPlot \((r^2 = 0.99)\).](image)
μg/ml (10.6 μM).

The LC$_{50}$ value reported here (5.0 μg/ml) is orders of magnitude higher than previously reported for other cell lines (Done et al., 1961; Mortimer and Collins, 1968). In order to determine if the high LC$_{50}$ value obtained was a specific characteristic of the HepG2 cell line, a dose response curve was also established for HeLa cells. HeLa cells are a human cervical epithelial cell line and have been previously characterised for sporidesmin cytotoxicity with a LC$_{50}$ value of approximately 2.25 ng/ml (Done et al., 1961).

To compare the toxicity of sporidesmin in HeLa cells to that of HepG2 cells, the HeLa cells were exposed to a series of sporidesmin concentrations (0.001 - 10 μg/ml) and the percentages of non-viable cells were measured 16 hours after treatment using the LDH assay (Fig. 6.2). A much wider range of concentrations were tested in HeLa cells as the expected toxicity was in the ng/ml range (Done et al., 1961). Due to the broad range of concentrations used the graph in Fig. 6.2 is presented on a log scale.

![Graph](image.png)

**Figure 6.2:** Cytotoxic effect of sporidesmin in HeLa cells. Cells were treated with sporidesmin at the indicated concentrations for 16 h. The percentages of non-viable cells were determined using the LDH assay. Each point represents the mean ± S.D. of two independent experiments each performed in triplicate. A sigmoidal curve (three parameter) was fitted using SigmaPlot ($r^2 = 0.99$).
Overall, the dose response curve obtained for HeLa cells (Fig. 6.2) was very similar to that obtained for HepG2 cells (Fig. 6.1). Again a sigmoidal dose response curve was obtained over a limited range (1 - 10 µg/ml) (Fig. 6.2). From these data the LC50 can not be determined accurately but an estimate places the LC50 value around 5.0 µg/ml (10.6 µM). These values are approximately three orders of magnitude higher than those values previously published (Done et al., 1961; Mortimer and Collins, 1968). The differences between the LC50 values reported here and those previously published are likely due to an intrinsic difference in the experimental system, and are not a unique feature of the HepG2 cells.

6.2.2 HepG2 cells and sporidesmin toxicity

6.2.2.1 The effect of time on sporidesmin toxicity

The time course of sporidesmin toxicity was examined. Cells were treated with 3 - 5 µg/ml of sporidesmin and the percentages of non-viable cells measured every 4 hours, for a period of 28 hours (Fig. 6.3A).

Interestingly, the relationship between sporidesmin toxicity and time appeared to plateau after 20 hours (for 3 µg/ml and 4 µg/ml) and 24 hours (5 µg/ml). This plateau effect could be the result of sporidesmin toxicity nearing maximal levels and this is possibly the case for the highest exposure (5 µg/ml) where toxicity is approaching 80%. However, the plateau effect was also observed for the two lesser exposures (3 and 4 µg/ml) and is not therefore due to maximal toxicity, as the toxicity is only approaching 40% non-viable cells. It is also possible that the plateau effect could also be an artefact associated with the light lability of sporidesmin (Safe and Taylor, 1969).

At the previously determined LC50 exposure of 5 µg/ml a linear relationship of percent non-viable cells with time was observed up to 20 hours \( (r^2 = 0.97 \text{ for solvent control, } r^2 = 0.99 \text{ for sporidesmin treated}) \). The toxicity became significantly different from the control at 4 hours of treatment, and was significantly different from the control between 12 and 24 hours of treatment (Fig. 6.3B).
Chapter Six: An in vitro system to examine zinc protection against sporidesmin toxicity

A.

Figure 6.3: Time course of sporidesmin cytotoxicity in HepG2 cells. Cells were treated with 3, 4 and 5 µg/ml sporidesmin for a range of time points from 4 h to 28 h. Each point represents the mean ± S.D. of two independent experiments performed in triplicate. The percentages of non-viable cells were determined using the LDH assay. (A) Full time course data for 3, 4 and 5 µg/ml sporidesmin over the 28 hour time course. At time points following 20 hours the toxicity plateaus. (B) Time course data for the LC50 exposure of 5 µg/ml sporidesmin: a linear relationship for toxicity exists between 4 and 24 hours (r²=0.99). Asterisks indicate significant differences from solvent control based on a two sample t-test (*p<0.05, **p<0.01, ***p<0.001)
For all further experiments a time period of 16 hours was utilised for sporidesmin exposure. This time point was selected as the toxicity was significantly different from the control and fell in the linear range of the toxicity, thereby avoiding confounding problems associated with the plateau effect observed in Fig. 6.3A.

6.2.2.2 Morphological effects of sporidesmin exposure
In addition to quantitative measurements of cell death made using the LDH assay, there were many morphological differences in HepG2 cells that had been exposed to sporidesmin (Fig. 6.4Ai, ii). In particular, cells treated at high concentrations of sporidesmin (>5 μg/ml) rapidly lost their irregular morphology and became rounded in shape (Fig. 6.4A). During sporidesmin exposure some seemingly intact cells lost adherence to the cell culture plate. This phenomenon had been previously identified as one of the key cytopathological effects of sporidesmin in cultured rat liver cells (Jordan and Pedersen, 1986).

Using the LDH assay as a measure of cell death these non-adherent cells were classified as non-viable but as these cells appeared to be intact further work was undertaken to confirm the fate of these cells. Non-adherent cells were removed from the cell culture media and stained with acridine orange and ethidium bromide. Both ethidium bromide and acridine orange bind to the DNA of cells, but ethidium bromide is actively extruded from viable cells. Therefore non-viable cells appear orange and viable cells appear green using fluorescence microscopy. Using this assay cells that had lost their adherence were qualitatively determined to be mostly viable, at least initially. However, the cells did not reattach to the substrate (even in the presence of fresh media), and the viability of these non-adherent cells was lost in under 12 hours (data not shown). Therefore for the measurement and calculation of cell death using the LDH assay, the non-adherent cells were included in the culture medium sample and classified as non-viable. For this reason the toxicity of sporidesmin and other compounds is expressed as the % of non-viable cells (as opposed to % cell death).
A.

i. 

ii. 

B. 

Figure 6.4: Cytopathological effects of sporidesmin on HepG2 cells. (A) Light microscope images of untreated (i) and sporidesmin treated (5 μg/ml for 16 hours) (ii) HepG2 cells. (B) Demonstration of DNA fragmentation in HepG2 cells exposed to sporidesmin. Cells were treated with different concentrations of sporidesmin for 16 hours, and low molecular weight DNA was isolated and resolved on a 1.8% agarose gel. DNA laddering was used as a qualitative marker for the occurrence of apoptosis. The position of DNA size standards (base pairs, bp) are indicated.

Morphologically, cells exposed to sporidesmin become rounded and characteristic apoptotic morphology could be observed, in particular membrane blebbing and nuclear condensation (Fig 6.4Aii). Because LDH leakage is an indictor of necrosis and late apoptosis (O'Brien et al., 2000), the occurrence of apoptosis was determined semi-quantitativelly using DNA laddering (Fig.6.4B). Rather than isolating whole cell DNA, low molecular weight DNA was specifically isolated (refer 2.6.4.2). The DNA samples electrophoresed on the gel in Fig. 6.4 represent the entire sample of fragmented DNA present. This means that the amount of fragmented DNA could not be standardised between lanes. Therefore due to possible inconsistencies in sample
preparation or gel loading this experiment is not a strictly quantitative measure of apoptosis.

Treatment of HepG2 cells with sporidesmin resulted in a stereotypical DNA ladder with DNA fragments present in increasing increments of ~180bp. This laddering pattern indicated the presence of DNA fragments the size of single nucleosomes and oligonucleosomes (Duke et al., 1983) which are considered to be characteristic makers for apoptotic cells (Wyllie et al., 1980). The fragmentation pattern indicated that sporidesmin does in fact induce apoptosis in HepG2 cells (Fig. 6.4B) and apoptosis was apparent at the LC₅₀ sporidesmin concentration of 5 μg/ml. Despite the qualitative nature of this analysis, it appears that the maximal level of apoptosis occurred at the 10 μg/ml exposure to sporidesmin. At concentrations higher than 10 μg/ml a gradual decline in apoptosis was observed, with loss of the laddering pattern at the highest dose tested, 25 μg/ml. This suggests that at concentrations higher than 10 μg/ml (21.2 μM) there is a switch from apoptosis to necrosis.

6.2.3 HepG2 cells and metal toxicity (zinc and copper)

6.2.3.1 Zinc Sulphate
Zinc is known to protect against sporidesmin toxicity in vivo (Smith et al., 1977; Towers and Stratton, 1978; Munday et al., 1997; Munday et al., 2001). But there are conflicting reports for whether this protection is observed in vitro (Cordiner and Jordan, 1983; Munday, 1984b; Waring et al., 1990). While the overall aim of this study was to examine zinc protection against sporidesmin toxicity in HepG2 cells it was first necessary to characterise the toxicity of zinc to this cell line.

A dose response curve was established for zinc exposure in HepG2 cells. Briefly, the cells were treated at 20 hours post passage (cells were plated at a density of 1 × 10⁶ cells per ml) with 0 to 500 μM zinc sulphate for 16 hours. Following treatment the percentages of non-viable cells were measured using the LDH assay.
Figure 6.5: Cytotoxic effect of zinc sulphate on HepG2 cells. Cells were treated with various concentrations of zinc sulphate for 16 h. The percentages of non-viable cells were determined using the LDH assay. Each point represents the mean ± S.D. of two independent experiments each performed in triplicate. A sigmoidal curve (five parameter) was fitted using SigmaPlot ($r^2 = 0.99$).

The response of HepG2 cells to varying levels of zinc is shown in Fig. 6.5. A sigmoidal dose response curve was fitted to the data using SigmaPlot ($r^2 = 0.99$) and the LC$_{50}$ of zinc sulphate was calculated to be 320 μM (Fig. 6.5). The highest concentration of zinc sulphate that did not cause toxicity was determined to be 200 μM, using one way ANOVA with a Fisher’s LSD post-hoc test ($P > 0.05$). Therefore in testing the protective effect of zinc sulphate against sporidesmin toxicity, a range of zinc concentrations up to a maximum of 200 μM was used, so as to avoid confounding issues of intrinsic zinc toxicity.

6.2.3.2 Copper Sulphate
Small amounts of copper are thought to potentiate the toxicity of various compounds (reviewed in Gaetke and Chow, 2003) including sporidesmin (Munday, 1982). This potentiation is due, at least in part, to the ability of copper to promote redox reactions (refer to 6.4.4). Specifically, copper is thought to potentiate sporidesmin toxicity by enhancing the production of ROS via Haber Weiss chemistry (refer to 1.3.4.1). Prior to determining the effect of copper supplementation on sporidesmin toxicity, the toxicity of exogenous copper supplementation to HepG2 cells was examined. Cells were treated at 20 hours post passage (cells were plated at a density of $1 \times 10^6$ cells
Figure 6.6: Effect of copper sulphate on HepG2 cells. Cells were treated with various concentrations of copper sulphate for 16 h. The percentages of non-viable cells were determined using the LDH assay. Each point represents the mean ± S.D. of two independent experiments each performed in triplicate.

Surprisingly, no toxicity was associated with copper exposures up to 500 μM (Fig. 6.6). In fact the higher exposures tended to slightly protect the cells from normal cell turnover, in that the slope of the line of best fit is slightly negative (-0.0084); however, this was not a significant effect.

6.2.3.3 Synergistic toxicity of zinc sulphate and copper sulphate

Given that copper is predicted to exacerbate sporidesmin toxicity by promoting redox cycling and zinc is known to act as an antioxidant (Zago and Oteiza, 2001), it would be interesting to determine if copper supplementation of HepG2 cells could override the zinc protection against sporidesmin toxicity. If this were the case, this would raise the possibility that zinc supplementation was protecting cells via the antioxidant defence system. However, simultaneous treatment of cells with two compounds can be difficult to interpret as it is well known that the toxicity of two compounds in a mixture is often not simply the sum of the toxicity of the two individual compounds (Konemann and Pieters, 1996). Therefore it was imperative to examine the toxicity of mixtures of zinc and copper to HepG2 cells.
Cells were treated at 20 hours post passage (cells were plated at a density of $1 \times 10^6$ cells per ml) with the NOEL concentration of 200 μM zinc. The cells were simultaneously treated with various concentrations (0, 20, 50, 100, 200 and 500 μM) copper for 16 hours. The reciprocal experiment was also performed where the cells were treated with 200 μM copper and various concentrations (0, 20, 50, 100, 200 and 500 μM) zinc for 16 hours. Following treatment the percentages of non-viable cells measured using the LDH assay and the results are shown in Fig. 6.7.

When HepG2 cells are supplemented with zinc and copper simultaneously (Fig. 6.7) the toxicity of the compounds in a mixture is clearly altered from that of the individual compounds (Fig. 6.5, Fig. 6.6). When treated with zinc-alone, the LC$_{50}$ was calculated to be 320 μM (Fig. 6.5). Simultaneous supplementation of cells with 200 μM copper decreased the LC$_{50}$ of zinc approximately 4-fold to 82 μM. Conversely, when treated with copper-alone, the LC$_{50}$ was calculated to be >500 μM (Fig. 6.6) but this was reduced to 182 μM when cells were treated simultaneously with 200 μM zinc.

Figure 6.7: Cytotoxic effect of mixtures on zinc and copper on HepG2 cells. HepG2 cells were exposed to various concentrations of zinc sulphate (in the presence of 200 μM copper sulphate) or copper sulphate (in the presence of 200 μM zinc sulphate) for 16 hours. Non-viability of cells was assessed using the LDH assay. Each point represents the mean ± SD of three independent samples. A sigmoidal curve (four parameter) was fitted to the data using SigmaPlot ($r^2=0.99$).
The NOEL of zinc-alone was 200 µM (Fig. 6.5) but supplementation of cells with ≥ 100 µM copper reduced the NOEL to 50 µM (Fig. 6.7). Accordingly it can be argued that a reasonably small exogenous supplementation of HepG2 cells with copper is able to markedly increase the toxicity of zinc.

6.2.4 Zinc protection of HepG2 cells from sporidesmin toxicity

As previously outlined (refer to 6.1.2) sporidesmin toxicity can be abrogated *in vivo* by zinc pre-treatment (Smith *et al*., 1977; Towers and Smith, 1978; Munday *et al*., 1997; Munday *et al*., 2001). However, there is conflicting *in vitro* evidence for zinc protection against sporidesmin toxicity (Cordiner and Jordan, 1983; Waring *et al*., 1990). Significant protection has been observed for T-blasts and macrophages, but the zinc and sporidesmin was administered concurrently and the levels of zinc far exceeded physiological levels. As zinc can form a complex with sporidesmin it is possible that the protection seen in this experiment was due to the extracellular sequestration of sporidesmin by zinc. Therefore it was of interest to determine whether pre-treatment of HepG2 cells with zinc could protect them against sporidesmin toxicity.

HepG2 cells were pre-treated with 0, 10, 50, 100 and 200 µM zinc sulphate for 16 hours, then washed with phosphate-buffered saline to minimise the extracellular zinc concentration. The cells were then treated with sporidesmin at 5 µg/ml for another 16 hours and the percentages of non-viable cells measured using the LDH assay.
Figure 6.8: Effect of zinc pre-treatment on sporidesmin toxicity in HepG2 cells. Following 16 h pre-treatment with various zinc concentrations, zinc-containing media were removed, the cells washed and fresh media containing sporidesmin were added, and the incubation continued for a further 16 h. The percentages of non-viable cells were determined using the LDH assay. Each bar represents the mean ± S.D. of three independent experiments each performed in triplicate. U: untreated (cells treated with neither zinc nor sporidesmin). Asterisks indicate statistically significant differences from the sporidesmin treated control based on one-way ANOVA with Fisher’s LSD post hoc test (***p < 0.001).

Pre-treatment of HepG2 cells with zinc did confer significant protection against sporidesmin toxicity (Fig. 6.8). This protection was concentration dependent with no observable protection at 10 μM. However, significant protection was seen at 50 μM zinc pre-treatment, where zinc conferred 49% protection compared to the control (sporidesmin only treatment). In this case the percentage of non-viable cells was reduced from 49.7% to 25.2%. The maximal protection was seen with 200 μM zinc which conferred 66% protection compared to the sporidesmin-only control and the percentage of non-viable cells was reduced from 49.7% to 16.9%. None of the concentrations tested could reduce the sporidesmin toxicity to the background level of 6.3% (untreated cells).

Exogenous zinc supplementation has been shown to increase intracellular zinc levels in a time dependent manner (Urani et al., 2001). To determine if there was any effect of the length of the zinc pre-treatment on the protection observed against sporidesmin toxicity, a 16-hour pre-treatment experiment was compared with a 2-hour pre-treatment experiment (Fig. 6.9).
It was found that there was no significant difference between the two zinc pre-treatment times for most of the zinc concentrations examined (Fig. 6.9); the one exception was a significant difference observed at the 200 μM zinc concentration, where the protection observed was slightly more effective after 2 hours of zinc pre-treatment as compared with 16 hours \((P < 0.05)\). This suggests that the protective effect induced by zinc occurred rapidly and possibly in response to small changes in intracellular zinc levels.

### 6.2.5 The role of transcription in zinc protection from sporidesmin toxicity

As discussed in the introduction to this chapter, one of the major biological functions of zinc is modulation of gene expression, both under normal physiological conditions and under conditions of oxidative stress (refer 6.1.2.1). Zinc is also known to modulate these changes rapidly and in response to relatively small increases in exogenous zinc supplementation (Cousins *et al.*, 2003). It was hypothesised that if zinc was protecting HepG2 cells by modulating gene expression, then the genes that were differentially expressed would be good functional candidate genes for resistance to FE.
Prior to attempting to identify differentially expressed genes it was important to determine if de novo gene transcription played a role in zinc protection against sporidesmin toxicity, and to determine the contribution of gene transcription to the overall protective mechanism. To do this the transcriptional inhibitor actinomycin D was used. HepG2 cells were treated with actinomycin D for 1 hour prior to and throughout the 2-hour zinc pre-treatment (refer to Fig. 2.1). Following the 2-hour zinc pre-treatment the cell culture media was removed, the cells were washed twice with sterile PBS, and fresh media containing sporidesmin (5 µg/ml) were added. The cells were then exposed to sporidesmin for a total of 16 hours and the percentages of non-viable cells measured using the LDH assay.

As shown in Table 6.1, actinomycin D treatment alone was significantly more toxic (27.4%) compared to the DMSO solvent control (10.6%, P < 0.05) (Table 6.1). In order to take into account this intrinsic toxicity of actinomycin D, the DMSO and actinomycin D control values were subtracted from the corresponding experimental values.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>% non-viable cells&lt;sup&gt;a&lt;/sup&gt;</th>
<th>% non-viable cells&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMSO solvent control</td>
<td>10.6 ± 2.7%</td>
<td>0.0</td>
</tr>
<tr>
<td>Zinc</td>
<td>10.4 ± 1.9%</td>
<td>-0.2 ± 1.1%</td>
</tr>
<tr>
<td>Sporidesmin</td>
<td>49.0 ± 3.3%</td>
<td>38.4 ± 5.7%</td>
</tr>
<tr>
<td>Sporidesmin + Zinc</td>
<td>18.5 ± 3.0%</td>
<td>7.9 ± 3.0%</td>
</tr>
<tr>
<td>Actinomycin D control</td>
<td>27.4 ± 4.7%</td>
<td>0.0</td>
</tr>
<tr>
<td>Actinomycin D + Zinc</td>
<td>25.9 ± 2.2%</td>
<td>-1.5 ± 3.3%</td>
</tr>
<tr>
<td>Actinomycin D + Sporidesmin</td>
<td>59.9 ± 4.0%</td>
<td>32.5 ± 4.6%</td>
</tr>
<tr>
<td>Actinomycin D + Zinc + Sporidesmin</td>
<td>36.7 ± 6.4%</td>
<td>9.3 ± 4.1%</td>
</tr>
</tbody>
</table>

<sup>a</sup> Mean ± SD of three independent experiments performed in triplicate.

<sup>b</sup> The relative percentages of non-viable cells was calculated by subtracting the percentage values of non-viable cells for the DMSO and actinomycin D controls. These data are presented in Fig. 6.10 for graphical comparison.
values to give the relative percent non-viable cells data (Table 6.1): these data are presented graphically in Fig. 6.10. These calculations enabled a direct comparison of zinc protection against sporidesmin in the absence and presence of actinomycin D.

In the absence of actinomycin D, zinc protected HepG2 cells from sporidesmin toxicity by 79% (the relative percentage of non-viable cells was reduced from 38.4% to 7.9%, Table 6.1). In the presence of actinomycin D the protection was reduced slightly to 71% (the relative percentage of non-viable cells was reduced from 32.5% to 9.3%, Table 6.1). However, the reduction from 79% to 71% was not statistically significant.

To assess the statistical significance between treatments in the presence or absence of actinomycin D, a two sample t-test statistic was used and the level of significance was calculated using the Bonferroni correction for multiple testing. There was a statistically significant difference ($P < 0.05$) between sporidesmin-treated versus sporidesmin plus zinc-treated samples (Fig. 6.10); this result is consistent with the previous findings (Figs. 6.8 and 6.9).

![Figure 6.10: Effect of actinomycin D on zinc protection against sporidesmin toxicity in HepG2 cells. The percent non-viable cells relative to the DMSO and actinomycin D controls (refer Table 6.1) are presented graphically here. Asterisk indicates statistically significant difference by two sample t-test ($*** P < 0.001$).](image)
More importantly there was no significant difference between the DMSO (without actinomycin D) and actinomycin D treated cells (Fig. 6.10). This result indicated that the actinomycin D treatment, and therefore the inhibition of RNA transcription, did not affect the mechanism of zinc protection against sporidesmin toxicity in HepG2 cells. Therefore zinc is protecting HepG2 cells from sporidesmin toxicity by a mechanism independent of *de novo* gene transcription.
6.3 DISCUSSION

The mycotoxin sporidesmin causes a pathological condition in ruminants called FE (reviewed in Morris et al., 2004). The liver and bile ducts are the organs primarily affected by sporidesmin toxicity and high levels of zinc supplementation are used as an effective prophylaxis against this disease (Smith et al., 1977; Towers and Smith, 1978). Zinc is a very abundant trace element and is functionally pleiotropic: it is involved in gene transcription and cell signalling as well as protein structure and function (refer 6.1.2). However, the mechanism by which zinc is able to protect animals against sporidesmin toxicity is unknown. Examination of such mechanisms in whole animals is problematic and in vitro systems have been used extensively in sporidesmin research. For this study HepG2 cells were chosen as a model system. This cell line is derived from a human hepatoma and has been used extensively in hepatotoxicity studies and HepG2 cells retain many of the morphological features (Sormunen et al., 1993) and gene expression characteristics of primary hepatocytes (Knowles et al., 1980; Lu and Huang, 1994; Urani et al., 1998; Harris et al., 2004a).

In this study HepG2 cells were established for their response to sporidesmin and then utilised to examine zinc protection against sporidesmin toxicity.

Treatment of HepG2 cells with increasing concentrations of sporidesmin (0 - 10 μg/ml) resulted in a classic sigmoidal dose response curve, but with a narrow linear range of toxicity (2 - 8 μg/ml) and an LC_{50} value of 5 μg/ml (Fig. 6.1). The narrow linear range of toxicity is consistent with published data using various other cell lines including HEp-2 (HeLa cell derivative), RKE (rabbit kidney epithelium) and VMK (Vernet monkey kidney cells) (Done et al., 1961; Mortimer and Collins, 1968). However the LC_{50} value for HepG2 cells reported here is three orders of magnitude higher than that published in the literature for sporidesmin toxicity in HeLa cells (Done et al., 1961). To elucidate this difference the toxicity of sporidesmin to HepG2 cells was compared to that of HeLa cells (Fig. 6.2). Treatment of HeLa cells with increasing concentrations of sporidesmin (0.0001 - 10 μg/ml) resulted in a similar dose response curve to that obtained for HepG2 cells.
In addition to cytotoxicity, low concentrations of sporidesmin (in the ng/ml range) have been reported to have a profound effect on cellular morphology in cell lines such as HEp-2, RKE and VMK (Mortimer and Collins, 1968), but these morphological changes were only observed in HepG2 and HeLa cells when exposed to sporidesmin in much higher concentrations (in the μg/ml range). The reason for the differences between cell types reported for sporidesmin sensitivity is not known, although the type of culture media used, varying serum concentrations (Mortimer and Collins, 1968), different preparations of sporidesmin and dissimilar methods of determining cytotoxicity could be contributing factors.

Following exposure to sporidesmin morphological analysis revealed that cells rapidly lost adherence to the cell culture plate, and displayed classic apoptotic morphology including membrane blebbing and nuclear condensation (Fig. 6.4A). The occurrence of apoptosis was confirmed using the DNA fragmentation assay (Fig. 6.4B). At the 5 μg/ml (10.6 μM) exposure, apoptosis could be detected but appeared to be maximal at 10 μg/ml. At concentrations higher than 10 μg/ml the levels of apoptosis appeared to decline steadily, indicating a switch from apoptotic cell death to necrotic cell death. This is characteristic of ETP compounds which are known to have a tendency to switch from apoptotic to necrotic modes of cell death at relatively high exposures (Beaver and Waring, 1994; Hurne et al., 2000). Although in HepG2 cells the switch was observed at concentrations higher than 21 μM, whereas previous studies have observed the switch occurring at concentrations above 10 μM (Waring et al., 1990; Beaver and Waring, 1994). It is known that the response of cells to apoptotic stimuli is cell-line specific, and it is likely that HepG2 cells respond differently to apoptotic stimuli those cell-lines previously tested.

The HepG2 cells were further characterised for their response to sporidesmin, in particular, the time dependence of sporidesmin toxicity in the cells (Fig. 6.3A). For all of the exposures tested there was an apparent plateau effect between 20 and 28 hours. There are several possible explanations for this phenomenon: for instance sporidesmin is light labile (Safe and Taylor, 1969) and although cells were treated in the dark, it is possible that prolonged exposure to the cell culture media, atmosphere or intermittent light resulted irreversible damage to the sporidesmin molecule,
negating its toxicity. Alternatively, the plateau effect may be biologically interesting, as it may indicate the induction of protective mechanisms in these cells, or active detoxification of the sporidesmin compound. Certainly, gene expression in HepG2 cells is known to change rapidly (in less than 6 hours) in response to cellular stress (Hockley et al., 2006), but the time period of induction is dependent on the source of the stress (Alia et al., 2005). It would therefore be interesting to determine if the plateau effect is stable at time points beyond 28 hours and if so, what genes or cellular mechanisms contribute to this effect. To avoid possible confounding factors associated with this plateau effect a sporidesmin-exposure time of 16 hours was utilised for all remaining experiments.

Once the toxicity of sporidesmin to HepG2 cells had been characterised the main focus of this research was to examine the effect of zinc and copper on this toxicity. As previously stated zinc is protective against sporidesmin toxicity, and copper is thought to exacerbate toxicity by promoting intracellular redox cycling. Prior to examining the effect of these compounds on sporidesmin toxicity, the toxicity of the individual compounds was examined. Treatment of HepG2 cells for 16 hours with concentrations of zinc ranging from 10 μM to 500 μM resulted in a classic sigmoidal dose response curve with an LC$_{50}$ of 320 μM and a no observed effect level of 200 μM (Fig. 6.5). However, treatment of HepG2 cells with concentrations of copper ranging from 10 μM to 500 μM did not result in any associated toxicity (Fig. 6.6). The lack of toxicity for copper observed in this study is markedly different to that reported for primary hepatocytes, where 50 μM copper supplementation caused severe cytotoxicity after a 2 hour exposure; the toxicity of copper in those cells was primarily due to the formation of reactive oxygen species (Pourahmad and O'Brien, 2000). The absence of copper toxicity is also different to previous reports for HepG2 cells; copper chloride supplementation (174 μM) of HepG2 cells has been shown to increase intracellular copper levels and decrease the viability of cells slightly (to 87% of control) (Urani et al., 2001). It is possible that the toxic effect observed is actually due to the use of copper chloride in these experiments, and it may be the chloride exerting the toxic effect rather than the copper per se. Therefore it is likely that copper is accumulating inside the HepG2 cells but for an undefined reason is not exerting any toxicity under the experimental conditions used.
Interestingly, concurrent treatment of HepG2 cells with copper and zinc (Fig. 6.8) resulted in a significant potentiation of toxicity between the two metals. There is precedent for the synergistic toxicity of copper and zinc in the literature. Specifically, treatment of HepG2 cells with 176 μM copper chloride and 174 μM zinc sulphate caused cell viability to be reduced dramatically to approximately 10% of that of the control (Urani et al., 2001). The reduction of viability observed in this study was slightly more modest, where 200 μM supplementation of HepG2 cells with zinc sulphate and copper sulphate resulted in a reduction in cell viability to approximately 20% of the control. A possible mechanism for this potentiation is provided by a previous study which found evidence that simultaneous treatment of HepG2 cells with copper and zinc for 4 hours increased the intracellular accumulation of zinc 9-fold (to 0.46 ppm/10^6 cells) while the copper concentration did not increase at all (Urani et al., 2001). It therefore appears that treatment of HepG2 cells with copper facilitates the uptake of zinc to toxic levels. This is partly challenged by another study which showed that intracellular zinc content actually decreased after a 48 hour-exposure with 100 μM copper (Tapia et al., 2003). For this study the primary purpose in using copper was to determine if copper treatment could abrogate any protection seen by zinc. However, on the basis of the confounding toxicity of mixtures of zinc and copper, copper was not investigated further.

After establishing the no observed effect level for zinc the ability of zinc pre-treatment to protect HepG2 cells against sporidesmin toxicity was examined. The zinc concentrations tested ranged from 10 μM, representing physiological plasma levels (Cousins, 1996), to 200 μM which was the maximum level at which zinc was not intrinsically toxic to HepG2 cells (Fig. 6.5). It was found that zinc pre-treatment is capable of providing significant protection at 50 μM, with maximal protection observed at 200 μM (Fig. 6.8). The results obtained are consistent with those seen by Waring et al. (1990) who found that zinc provided almost complete protection against sporidesmin toxicity in mouse macrophages using cell morphology and DNA laddering as markers for apoptosis. However, in the Waring et al. (1990) study, 1000 μM zinc treatment was used and the zinc treatment was concurrent with
sporidesmin treatment. Therefore, it was possible that the zinc could protect the cells by interacting directly with sporidesmin in the culture medium to form bivalent complexes, as such complex formation has been shown to occur in vitro (Woodcock et al., 2001a; Woodcock et al., 2001b). In this study, the HepG2 cells were pre-treated with zinc and the cells were washed prior to sporidesmin exposure to minimise such extracellular interactions.

Much of the data obtained were consistent with gene expression playing a role in zinc protection against sporidesmin toxicity. In particular, the protection by zinc only required a relatively small increase in exogenous zinc supplementation (~50 μM, Fig. 6.8) and the protection was effected rapidly (within 2 hours, Fig. 6.9) consistent with the effect of zinc supplementation (Cousins et al., 2003). Also there was no additional protection provided by a 16 hour zinc exposure versus a 2 hour exposure. Given that exogenous zinc supplementation in HepG2 cells results in a time-dependent increase in intracellular zinc (Urani et al., 2001), the results presented here suggest that not only is the protection effected rapidly, but that only a small increase in intracellular zinc is required to confer protection. The protective effect was also relatively long lasting, with protection being maintained for at least 16 hours after zinc withdrawal.

Therefore, the role of gene transcription in the protective effect of zinc sulphate on sporidesmin toxicity was examined. Conceptually, if gene transcription was altered in response to zinc pre-treatment, then these genes would be candidate genes for conferring genetic resistance to sporidesmin in sheep and could be identified by microarray analysis. To assess the role of transcription in zinc protection of HepG2 cells against sporidesmin toxicity, cells were treated with the transcriptional inhibitor, actinomycin D, for one hour prior to and throughout the zinc pre-treatment period. It is assumed that at the concentration used (10 μg/ml) that actinomycin D is inhibiting de novo gene transcription almost entirely. This assumption is justified as actinomycin D treatment was associated with significant cytotoxicity, and also previous studies using this concentration of actinomycin D in HepG2 cells have noted complete inhibition of de novo gene transcription (Zhou et al., 1996).
The results presented in Table 6.1 and Fig. 6.10 showed that actinomycin D did not alter the protective effect of zinc. This strongly indicates that zinc is protecting HepG2 cells from sporidesmin toxicity by a mechanism that is independent of de novo gene transcription.

Given that altered gene expression had been ruled out as an important mechanism for zinc protection, other mechanisms by which zinc could prevent or counteract the toxic effects of sporidesmin must be considered. It is not precisely known how sporidesmin and other ETP compounds mediate cytotoxicity (reviewed in Chai and Waring, 2000), although there is evidence that the oxidised disulphide ETP compounds can interact covalently with cellular macromolecules, such as proteins, and thus alter their functions (Rodriguez and Carrasco, 1992; Hara and Han, 1995; Waring et al., 1995; Pahl et al., 1996; Hume et al., 2000). Additionally, ETP compounds are also able to promote intracellular redox cycling and generate toxic reactive oxygen species (Munday, 1982; Munday, 1984a; Munday, 1987). Zinc has been shown to form a tetrahedral complex with sporidesmin (Woodcock et al., 2001a; Woodcock et al., 2001b). If this were to occur intracellularly, the formation of such a complex would abolish the reactivity of sporidesmin and limit its ability to interact with cellular macromolecules or to generate reactive oxygen species. The HepG2 system developed here provides a model system for analysing the role of reactive oxygen species formation in sporidesmin toxicity, and whether zinc is capable of abrogating the formation of these species.

The experiments described here using the HepG2 cells were not designed to determine the relative importance of apoptosis or necrosis in sporidesmin cytotoxicity, as the LDH assay is a measure of necrosis and late apoptosis (O'Brien et al., 2000). Analysis of the DNA fragmentation in HepG2 cells treated with sporidesmin revealed that apoptosis was occurring, similar to that observed for mouse macrophages (Waring et al., 1990) and LLC-PK1 cells (Zhou et al., 2000). Zinc is a potent inhibitor of apoptosis, capable of inhibiting multiple steps in the pro-apoptotic pathway including caspase activation (Stennicke and Salvesen, 1997; Segal and Beem, 2001; Wei et al., 2004), Bax translocation and cytochrome c release (Ganju and Eastman, 2003; Wei et al., 2004). Apoptosis initiated by Fas ligand mediated signalling is also suppressed by zinc (Lambert et al., 2003). Therefore zinc
Chapter Six: An in vitro system to examine zinc protection against sporidesmin toxicity

Protection of HepG2 cells from sporidesmin toxicity may be mediated through the inhibition of apoptosis.

Further, zinc may also protect against sporidesmin toxicity by modulating the activities of cytochrome P450 enzymes. Acute zinc administration has been shown to markedly decrease cytochrome P450 activity, both in rats (Tan et al., 1995) and in isolated microsomes (Jeffery, 1983; Moreno et al., 1989). Sporidesmin appears to be a substrate for the P450 enzymes (refer to 1.3.3 and 1.7.2.1) and the product of these reactions are hydroxylated derivatives of sporidesmin (Fairclough et al., 1978). Products of such phase I reactions tend to be more toxic than the parent compound due to an increase in solubility of the metabolite. Hydroxylation of sporidesmin would increase the polarity and solubility and perhaps toxicity. Therefore inhibition of the hydroxylation may abrogate any potential toxicity associated with the hydroxylated derivative.

This work has established the use of HepG2 cells as a model system for sporidesmin toxicity studies. Additionally, it was the first study to demonstrate that zinc pretreatment protected cells from sporidesmin toxicity. Using this model, it was shown that zinc protects cells from sporidesmin in a manner independent of de novo gene transcription, highlighting the importance of other mechanisms, such as inhibition of apoptosis, which zinc may utilise to protect against sporidesmin. The HepG2 cell-line could be used to examine these alternative mechanisms of protection.
Chapter Seven
General Discussion and Future Work.

7.1 Project Overview
FE is a common hepatogenous mycotoxicosis of ruminants. It is caused by the sporidesmin toxin and costs the New Zealand economy around $100 million annually (Anonymous, 1990). Exposure to sporidesmin causes severe liver disease in susceptible animals, resulting in secondary photosensitisation. There is a significant genetic component to resistance to FE and realistically, the best long-term management strategy for FE is using marker-assisted selection to aid selective breeding of resistant animals (Morris et al., 2004).

Marker-assisted selection requires a genetic marker or series of markers linked to the disease trait of interest. To be useful, these markers should explain the majority of the variation observed within the trait of interest. The discovery of such markers is problematic with diseases such as FE, as resistance is known to be multigenic, with an unknown number of genes contributing to resistance. In such complex disease traits the effect provided by each gene may be small.

Because of the complexity of discovering genes associated with a polygenic trait, using multiple approaches to gene discovery can be expected to yield the most information. Whole genome linkage studies are the gold-standard approach to identifying areas of the genome that are associated with the trait of interest. Such studies have been undertaken for FE resistance with moderate success (S. H. Phua, pers. comm.). However, whole genome linkage studies have a limited power to detect small or recessive effects. Aside from whole genome linkage studies, the candidate gene approach is the most common approach to gene discovery. Using this approach genes or classes of genes that are perceived to be functionally relevant to the trait of interest are selected for investigation. Previously the candidate-gene
approach has been used to evaluate several genes within the context of FE, these include antioxidant genes (Phua et al., 1998) and an ABC-transporter protein (Longley, 1998).

The overriding aim of this PhD project was to identify genes associated with resistance to FE, in order to advance the goal of using marker-assisted selection as a long-term solution for FE. To do this several methods were undertaken. Firstly, phylogenetic comparisons of a yeast ABC transporter known to transport sporidesmin (PDR5), combined with data on function, expression and comparative mapping was used to identify a novel candidate gene for resistance to FE (ABCG2). Second, gene-expression profiling of liver tissue from resistant and susceptible animals was undertaken in order to identify candidate genes and pathways involved in conferring resistance. This was the first time that such analysis had been undertaken for FE. Finally, an in vitro system for examining sporidesmin toxicity was characterised. This in vitro system was used to determine if zinc could protect against sporidesmin, and if so whether zinc was able to confer this protection by modulating gene transcription. The rationale was that if zinc was conferring protection by modulating gene transcription, the genes targeted by zinc would be candidate genes for conferring resistance to FE. The findings from each of these approaches are discussed in more detail below.

7.2 Quantification of ABCB1 Gene Expression
ABCB1 (P-glycoprotein, MDR1) had been previously identified as a candidate gene for FE (Longley, 1998). ABCB1 is an ABC-transporter that is able to confer a multiple drug resistance phenotype in humans. A previous study showed that microsatellite markers surrounding the ovine ABCB1 locus did not segregate with resistance to FE, but did show that following exposure to sporidesmin the levels of ABCB1 transcript were up-regulated approximately 4-fold in the livers of resistant, but not susceptible animals (Longley, 1998). The implication from these data is that a cis- or trans- acting factor may be influencing the transcription of ABCB1 and in-turn be a good candidate gene for FE.

Prior to investigating the genetic mechanism of this induction, an attempt was made to reproduce the original observation concerning differences in ABCB1 mRNA
expression. Using qRT-PCR no induction of ABCB1 was observed, contradicting the original findings (Longley, 1998). However, there were several problems with the qRT-PCR protocol, not the least of which was that, by necessity (in order to ensure specificity), the qRT-PCR assay was designed towards the 5’ region of the gene. This introduces potential bias due to the inefficiency of the reverse transcriptase reaction. The use of GAPDH as an endogenous control was also not optimal as it not only necessitated DNAse treatment of the RNA but also was present at much higher levels than ABCB1, therefore small variations in the expression of GAPDH could mask any true differences in ABCB1 expression. This work highlights the care that needs to be taken when designing probes and primers for quantification of mRNA especially for homologous gene families, and the necessity of selecting an appropriate endogenous-control gene.

Nevertheless, examination of the original competitive RT-PCR primers showed that they were likely to be able to amplify a second gene product, ABCB4. The presence of a second amplicon was confirmed by restriction digest of the competitive RT-PCR product, but repeated efforts to clone this second product failed. Melt curve analysis of the competitive RT-PCR product, generated in the presence of SYBR green, did not show a difference in melting temperature for the two products, suggesting that the two products are very similar in sequence and hence melting temperature. This was consistent with the second product being ABCB4. Further work is needed to confirm the identity of this second product as it is possible that another member of the ABCB family, or another gene altogether, is actually the gene responsible for the differential expression observed by Longley (1998).

A final attempt was made to confirm the differential expression previously observed by Longley (1998). To do this, the competitive RT-PCR product was used as a probe for Northern analysis. Due to the similarity of the ABCB1 and ABCB4 sequences, it was expected that the competitive RT-PCR probe would be able to bind and hybridise to at least these two mRNA species. Using this analysis significant differences in expression were observed, specifically differences were detected between both naïve susceptible and sporidesmin exposed resistant animals as compare with naïve resistant animals. However, these were restricted to a minor hybridisation band which, based on its size, is likely to be a half transporter and is
therefore neither ABCB1 nor ABCB4. Additionally, the differences in expression observed were more modest than that previously reported. Here, the differences in gene expression were around 1.2 fold, versus the 4-fold difference previously reported (Longley, 1998).

Subsequent to this work the bovine genome sequence has been released. The availability of ruminant sequence could facilitate the design of specific qRT-PCR assays to the 3' UTR of both ABCB1 and ABCB4. Together with a more suitable endogenous control that was expressed at relatively low levels in liver tissue and had sequence amenable to qRT-PCR design, such as porphobilinogen deaminase, this would definitively determine whether ABCB1 or ABCB4 are induced 4-fold in resistant animals following exposure to sporidesmin, as reported in the previous study (Longley, 1998).

Due to the methodological difficulties outlined it is impossible to rule out the involvement of ABCB1 in resistance to FE, although based on the genetic data available (Longley, 1998) any effect of ABCB1 is likely to be minor.

7.3 Identification and evaluation of a mammalian ortholog of the yeast PDR5 protein.

Sporidesmin has previously been shown to be a substrate for the yeast ABC-transporter, PDR5 (Bissinger and Kuchler, 1994). The release of several key genome sequences has shown that the ABC transporter family is more diverse than expected and has 48 members in humans (Dean et al., 2001). The aim of this section of work was to use these fully sequenced genomes to identify the likely mammalian ortholog of the yeast PDR5 protein.

To do this the ABC domain sequences were systematically extracted from the two fully sequenced genomes of interest, S. cerevisiae and H. sapiens. Three additional and evolutionary intermediate genomes were also examined; these were C. elegans, D. melanogaster and M. musculus. This analysis extracted a total of 348 domain sequences which were then compared using two different computational methods, neighbour-joining and maximum parsimony. Using both of these methods the yeast PDR5 sequence was found to be most closely related to the mammalian ABCG sub-
family of proteins. This was somewhat surprising as PDR5 is a full-transporter, consisting of twelve transmembrane domains and two ABC domains, while the ABCG sub-family consists entirely of half transporters which consist of six transmembrane domains and a single ABC domain. The homology seen between these sequences is suggestive of a common evolutionary origin or, alternatively, convergent evolution.

The mammalian ABCG sub-family has six members one of which, ABCG3, appears to be rodent specific. Analysis of synteny surrounding this locus in rodents indicates that ABCG3 has arisen as a result of chromosomal breakage and repair during genome evolution in the rodent lineage. The remaining five ABCG transporters are all involved in cholesterol homeostasis and lipid metabolism, with one notable exception, ABCG2. ABCG2 localises to the basolateral membrane of the hepatocyte and is involved in xenobiotic transport. Two of the five ABCG transporters, ABCG5 and ABCG8, map to a region of the sheep genome which shows a reasonably strong association with FE, thus making them excellent positional candidates. However, following this work, polymorphisms in these two genes were examined for segregation with resistance to FE. No evidence for an association between ABCG8 was found and ABCG5 is currently still being investigated (S. H. Phua, pers. comm.). This study focussed on ABCG2 as it is a xenobiotic transporter and is therefore a good functional candidate for FE. Additionally, mice deficient in ABCG2 display a diet-dependent phototoxicity similar to FE afflicted animals. Finally, ABCG2 is also a positional candidate for resistance to FE as comparative mapping localised this gene to an area of the sheep genome that had been weakly associated with FE resistance identified from a whole genome linkage study (S. H. Phua, pers. comm.).

In general, the use of phylogenetic comparisons to infer function across species, particularly in species as diverse as yeast and mammals, is problematic and can only be confidently inferred from analysis of orthologous pairs. In this case the functional inference made from the phylogenetic data was strengthened using functional, expression and mutation data from various species. These data were taken into consideration and a testable hypothesis was formed, which is that ABCG2 is a candidate gene for FE.
The first step in evaluating ABCG2 as a candidate gene for resistance to FE was to obtain the coding-domains sequence. Only two synonymous SNPs were detected in four animals sequenced (two resistant and two susceptible selection-line animals), confirming no differences in the primary structure of the ABCG2 protein between resistant and susceptible animals. A recent study has shown that synonymous mutations in ABCB1 can affect the activity of the active transporter (Kimchi-Sarfaty et al., 2007). However, as there are no differences in the allele frequencies of either of these polymorphisms between the resistant and susceptible selection lines, these polymorphisms are highly unlikely to contribute to resistance.

These two SNP polymorphisms together with an additional SNP in intron 4 of the gene were genotyped across a panel of resistant, susceptible and control selection-line animals. Interestingly, a significant difference in allele frequencies for the intronic SNP was observed. This finding of an association between alleles of an ABCG2 SNP with FE validates the use of phylogenetic and functional data to identify candidate genes. However, the significance level of the association reported here is fairly weak, especially compared to the published association of the catalase locus with FE resistance (Phua et al., 1999).

The observed association could be explained if the actual gene of interest was in linkage disequilibrium with ABCG2. Therefore, it would be of interest to determine if other markers in this region (either SNP or microsatellite markers) could strengthen or abolish the association of this region of the genome with resistance to FE. It would also be essential to examine the segregation of these marker alleles in flocks of animals that have been independently selected for resistance to FE. This would substantiate their involvement in conferring resistant to FE and rule out the possibility of association due to a strong founder effect in this genetically isolated population.

The association of the ABCG2 SNP with FE and the absence of any non-synonymous mutations in the coding region warranted further experimental studies into ABCG2. In particular, the expression levels of ABCG2 were examined, partly because differences in the expression of ABCG2 are known to be associated with drug resistance (reviewed in Sarkadi et al., 2006), and also because variation in the
promoter region of ABCG2 was not specifically examined by either sequencing or genotyping. Therefore, Northern hybridisation was used to determine if there were any differences in the expression of ABCG2 in the livers of resistant and susceptible animals, both pre- and post-exposure to sporidesmin.

Northern hybridisation revealed no transcriptional difference in livers of resistant and susceptible animals, and that ABCG2 was induced in all animals in response to sporidesmin exposure, with the major transcript being induced 1.6 fold in animals exposed to sporidesmin. This is the first report of modulation of ABCG2 expression in ruminants by a xenobiotic, although the mechanism of this induction remains undefined. It is possible that ABCG2 is induced in response to cellular damage caused by sporidesmin. This could possibly be mediated through HIF-1 activation (Krishnamurthy et al., 2004) in response to a cytotoxic hypoxic state generated by sporidesmin. Some xenobiotics, in particular microsomal enzyme inducers, have been shown to enhance the expression of various ABC transporters via activation of nuclear receptors (Klaassen and Slitt, 2005). Although this is not the case for ABCG2 (Han and Sugiyama, 2006), elucidation of the molecular pathways of induction may yield insight into the toxic mechanism and cellular response to sporidesmin.

Interestingly, the knowledge that ABCG2 is induced in both resistant and susceptible sheep by sporidesmin indicates that sporidesmin activates a common pathway in these sheep that leads to the induction of ABCG2. This induction may be functionally interesting as ABCG2 is known to transport phylloerythrin (Robey et al., 2006), which is a photo-reactive breakdown product of chlorophyll and is thought to be responsible for the photosensitisation observed in FE. It is therefore possible that the increased expression of ABCG2 may enhance the excretion of phylloerythrin and hence prevent or delay the onset of photosensitivity.

Besides ABCG2, it is also possible that another gene in linkage disequilibrium with ABCG2 may be the causative gene and therefore the gene of interest. Alternatively, the expression or function of ABCG2 in tissues other than the liver may be more relevant, as ABCG2 is known to modulate both hepato-biliary excretion and intestinal absorption of substrates (reviewed in Staud and Pavek, 2005). Therefore, it
may be of interest to study ABCG2 expression in intestinal tissue from FE resistant and susceptible sheep.

The work presented here represents the first genetic association between an ABC transporter and resistance to FE. It is likely that other ABC transporters will show similar associations, and as a super-family, they remain excellent putative candidate genes. Of particular interest are ABCC1 and ABCC2, both of these proteins are well characterised xenobiotic transporters in humans (Hoffmann and Kroemer, 2004) and are excellent functional candidate genes.

In conclusion, although ABCG2 is not a major contributor to FE resistance and is therefore not useful for marker-assisted selection, the finding of genetic association with FE validates the approach taken to identify ABCG2 as a candidate gene. Therefore when identifying candidate genes it is wise to incorporate information from as many sources as possible, including comparative genomics, phylogenetics, tissue expression patterns, and mutation data.

7.4 cDNA microarrays as a tool for identifying candidate genes for resistance to FE

The use of microarrays to investigate complex disease traits in humans is well established (Shai, 2006), but the application to such traits in livestock is relatively recent. Microarray data can be used in combination with QTL studies to identify novel candidate genes or pathways involved in the disease process. This approach has been used successfully to identify a genetic variant of the MHC class II gene, DQA1, associated with parasite resistance in sheep (Keane et al., 2007).

In this study, a relatively small bovine cDNA microarray was used to detect differences in hepatic gene expression between resistant and susceptible animals. This was the first time that gene expression profiling had been used to examine differences between sheep resistant and susceptible to FE. This work was completed very early in the development of the AgResearch microarray programme and as such the array used was based on a bovine EST library, and contained only 1550 probe sequences.
Chapter Seven: General Discussion and Future Work

Using this technology a number of genes were identified as being differentially expressed between resistant and susceptible animals, both in naïve and sporidesmin-exposed states. However, interpretation of these data was complicated by several factors. Most importantly, the weight that can be placed on these data is limited as only four individual tissue samples were available for microarray analysis. Subsequently more samples became available and were included in the Northern hybridisations that were used to confirm some of the microarray observations. Analysis of the Northern data indicated that, for the majority of genes examined, variation between animals exceeded the variation associated with resistance or susceptibility to FE. Therefore many of the genes identified as differentially expressed by microarray analysis were not confirmed by Northern hybridisation. This variation is at least partly due to differences in the ages and genders between treatment groups, both factors are known to have an affect on drug metabolism and gene expression in the liver (McClure et al., 2005; Stahlberg et al., 2005; Mori et al., 2007). In addition, quantification of the Northern blots was problematic due to issues associated with sample degradation and perhaps more importantly the limited linear range of the film. However, the limited linear range of the film is likely to underestimate true differences in gene expression, so it is likely that the trends (rather than the strict numerical relationships) observed in this study are robust. Further to this, there were significant issues with the quality of the array with an estimated error of 30% in the EST sequences or positions. Therefore, the inferences made from this data as a whole, including the gene ontology and biomarker analyses, are likely to be spurious.

Nevertheless, six of the ten genes selected for confirmation by Northern analysis were shown to be differentially expressed, but not necessarily in the way predicted by the microarray results. One of these genes was catalase, which is involved in the detoxification of hydrogen peroxide to water. The microarray analyses predicted that catalase would be expressed 1.2-fold higher in naïve resistant animals. However, Northern analysis revealed that susceptible animals express catalase approximately 1.5 fold higher than resistant animals, irrespective of sporidesmin exposure. This differential expression is interesting as catalase has previously been shown to be involved with FE (Phua et al., 1999). This raises the possibility that the genetic association previously observed may functionally result in differential hepatic
expression of catalase. It would therefore be interesting to determine the levels of catalase mRNA expression across the genotype groups, to determine if there is a correlation between gene expression and the genotypes associated with catalase. To do this, the expression of catalase would have to be quantified in a much larger group of animals, as several of the marker alleles are quite rare. If such a correlation did exist this would provide evidence for a functional association of catalase with FE. Such a finding would validate the use of microarray expression profiling to identify candidate genes for resistance to FE.

The most marked difference in expression observed in the microarray data was for metallothionein and this gene also showed the most marked differential expression of the ten genes examined by Northern analysis. Metallothionein is the key protein involved in intracellular zinc homeostasis and as zinc is known to have a role in protecting against sporidesmin toxicity in sheep, metallothionein appears to be a relevant functional candidate gene. Based on the Northern analysis, metallothionein mRNA expression is repressed by ~30% in sporidesmin-dosed resistant animals. This repression could be either a cause or a consequence of an alteration in cellular zinc metabolism. Sporidesmin is known to form a tetrahedral complex with zinc (Woodcock et al., 2001a; Woodcock et al., 2001b) (Fig. 1.5) and the formation of such a complex would not only decrease the reactivity of sporidesmin but would also deplete the pool of intracellular zinc. The depletion of this pool could facilitate the down-regulation of metallothionein gene transcription as the transcriptional activator MTF1 requires zinc for activity. Alternatively, metallothionein may be down-regulated by another, perhaps genetic, mechanism. If the repression of metallothionein mRNA resulted in a decrease of metallothionein protein, it is likely that the intracellular levels of labile zinc would increase. Theoretically, this labile zinc could then bind to sporidesmin directly or influence changes in gene expression that could protect against sporidesmin toxicity. Before further investigation it would be imperative to determine if a ~30% reduction in metallothionein mRNA results in a concomitant decrease in the levels of metallothionein protein. If a decrease in the levels of metallothionein protein were observed, it would be possible to test the two hypotheses outlined above using the in vitro model system characterised in chapter six of this thesis.
The first possibility, as outlined above, is that sporidesmin may be acting directly to modulate metallothionein gene expression via depletion of the intracellular zinc pool and subsequent inhibition of the transcription factor MTF1. This could be examined using zinc fluorophores to measure flux in the intracellular zinc pool, but the expected differences are quite small and perhaps beyond the limits of detection using standard fluorophores. An alternative would be to measure MTF1 activity using a reporter gene expression fused to a MTF1 binding site. Using this system the activation of this promoter in the presence and absence of sporidesmin could be determined. A decrease in reporter gene expression, while not conclusive, would provide evidence for this hypothesis. The other alternative outlined above is that the decrease in metallothionein is actually modulated by an intrinsic, possibly genetic, factor. Using an in vitro system it would be possible to modulate the levels of metallothionein mRNA and protein by using RNAi to repress expression of metallothionein. It would then be possible to determine if the RNAi mediated repression of metallothionein is able to affect the sensitivity of these cells to sporidesmin. Additionally, it would also be possible to determine if, repression of the metallothionein protein is able to alter the labile zinc pool. If an effect on either sporidesmin sensitivity or zinc metabolism is observed this would provide functional evidence for metallothionein as a candidate gene for FE.

The use of microarray analysis in identifying candidate genes yields the most information when incorporating genetic data such as that generated by QTL studies or alternatively utilising gene ontology or transcriptional motif analysis to identify biochemical pathways of interest. The latter two approaches are currently difficult in ruminants as the information is usually extrapolated from human or rodent data and as such requires functional conservation between these diverse species for accuracy. Therefore, the use of microarrays in non-model systems also requires a systematic way to evaluate the functionality of these differentially expressed genes. In FE it may be of interest to determine the effects of such differential expression on sporidesmin sensitivity in an in vitro model system such as the HepG2 model developed in this thesis.

In conclusion, while microarrays clearly have some utility in candidate gene identification, it is critical to include samples from enough age and gender matched
animals to minimise false positives associated with inter-animal variation. Additionally, it is important to have a good quality array, preferably one that incorporates a number of ESTs or oligonucleotides corresponding to the same gene, to act as an intra-experiment control, lending more confidence and robustness to the results.

7.5 HepG2 cells as an *in vitro* system for studying sporidesmin toxicity

Much of what is known about the toxicity of sporidesmin and other ETP compounds is derived from the use of *in vitro* models. However, very little of this work has been carried out using hepatic cells despite the knowledge that the liver is the primary organ affected by sporidesmin toxicity. In this study, the human hepatoma cell line HepG2 was characterised as an *in vitro* model system for studying sporidesmin toxicity. HepG2 cells are an established cell-line that has been used routinely for toxicity studies and are known to retain some of the characteristics of primary hepatocytes.

The characterisation of such an *in vitro* model also provides a unique opportunity to facilitate gene discovery. For example, little is known about how the cell responds to sporidesmin, knowledge about this response could identify genes or pathways that are of interest in the detoxification, metabolism or recovery of cells after exposure to sporidesmin. In this system sporidesmin was shown to be toxic to HepG2 cells with an LC$_{50}$ of 5 μg/ml, and apoptosis was readily detectable at this concentration using DNA laddering.

One of the outstanding questions in FE research concerns the prophylactic administration of zinc, and the protection that this provides against sporidesmin exposure. This treatment is relatively successful, but the molecular mechanism underlying this protection is unknown. Zinc is a very abundant trace element and has roles in many cellular and physiological processes (Cousins, 1996). Perhaps its most integral role in cellular biochemistry is via the modulation of gene transcription, where zinc plays a critical role in zinc finger transcription factors. Theoretically, if zinc is acting to protect against sporidesmin by the modulation of gene transcription then the genes or pathways being altered by zinc may provide new candidate genes or pathways of interest.
Chapter Seven: General Discussion and Future Work

This study was the first to show that pre-treatment of HepG2 cells with zinc resulted in dose-dependent protection against sporidesmin toxicity. This protection was observed at zinc concentrations higher than physiological levels, but appeared to be effected rapidly and was not disrupted by the co-administration of the transcriptional inhibitor, actinomycin D. This data excludes *de novo* gene transcription as the major mechanism of zinc protection against sporidesmin toxicity.

The exclusion of gene transcription as a mechanism of zinc protection against sporidesmin draws attention to several other possible mechanisms. For instance, zinc is known to be a potent inhibitor of apoptosis (Stennicke and Salvesen, 1997; Ganju and Eastman, 2003; Lambert *et al.*, 2003). Hypothetically, the inhibition of apoptosis *in vivo* may force cells down a pro-survival pathway, and this may reduce the damage inflicted by sporidesmin by reducing the level of immune cell infiltration and inflammation associated with cell death. Using the *in vitro* model developed, it would be possible to test this hypothesis by determining if inhibitors of apoptosis, such as Boc-D-FMK, can protect HepG2 cells from sporidesmin toxicity. However, to fully characterise this response it would be necessary to quantify the levels of apoptosis relative to the levels of necrosis at the sporidesmin concentrations tested. In this study, LDH was used to measure cell death, but LDH is a marker for late apoptosis and necrosis and cannot differentiate between the two modes of cell death. However, using the qualitative method of DNA laddering, apoptosis was readily detectable at the LC_{50} exposure of 5 μg/ml. A more quantitative approach, such as fluorescence activated cell sorting (FACS) with 6-carboxyfluorescein diacetate and annexin V stained cells, would yield more accurate quantification.

Another possible mechanism for protection by zinc against sporidesmin toxicity involves the cytochrome P450 enzymes. There is evidence that these enzymes are involved in modifying sporidesmin, specifically to hydroxy and dihydroxy sporidesmin (Fairclough *et al.*, 1978). Additionally, zinc exposure is known to decrease the activity of several of the key P450 isoforms (Tan *et al.*, 1995). It is hypothesised that the modification of sporidesmin to either the mono- or dihydroxy forms of sporidesmin would increase the polarity of the molecule and thus increase the solubility. As the disulphide bridge is likely to remain unmodified, it is expected...
that the hydroxylated sporidesmin would be more toxic than the parent compound. Using the *in vitro* system characterised here it would be possible to determine if the hydroxylated sporidesmin derivatives are in fact more toxic to the cells than the parent compound, and secondly to determine the rate of conversion of sporidesmin to hydroxy-sporidesmin in the presence of various concentrations of zinc. It would also be interesting to use inhibitors and inducers of various P450 enzymes to determine if the toxicity of sporidesmin was enhanced or inhibited, thus implicating particular P450 isoenzymes.

The HepG2 model system would also facilitate screening of possible therapeutic compounds. In the case of FE a novel therapeutic is unlikely to be more effective or cost-effective than the current zinc treatment, but understanding how these compounds protect the cells against sporidesmin may highlight possible molecular mechanisms of genetic resistance.

Finally, although discovering genes associated with a complex disease trait like FE is difficult, it is even harder to demonstrate the role of the genetic variation in the disease pathology and thus proving it is the causative mutation and not in tight linkage with the real polymorphism. In sheep this unequivocal relationship has really only been shown for single gene traits such as Inverdale (Galloway *et al.*, 2002) and Booroola (Wilson *et al.*, 2001). Both of these genes result in an alteration in BMP signalling, and defects in the same pathway cause the fertility phenotype seen in both of these unrelated families of sheep. But in the absence of such strong genetic information, such as with a rare or complex disease, the alternative is to use model systems. Model systems can include knock-in or knock-out mouse models, or *in vitro* systems such as cell culture. In this study HepG2 cells were shown to have utility as such an *in vitro* model. The expression of different genes can be easily modulated using RNAi or over-expression and the sensitivity of these cells to sporidesmin can be easily and reproducibly measured, as shown in this work. In addition, as HepG2 cells are a human cell line the microarray resources are more advanced than those currently available for livestock. Using microarray technology with these cells would facilitate understanding of sporidesmin toxicity and also of protection against this toxicity.
7.6 Final remarks

Ultimately the aim of the FE research program is to develop and use marker-assisted selection to breed sheep that are resistant to FE. Some of the work presented in this thesis has directly contributed to this goal, in particular the identification of SNPs within the ABCG2 locus that segregate with susceptibility to FE. Although further work is needed to confirm this association, it provides the first genetic association between an ABC-transporter and FE.

Additionally, the work included in this thesis has assisted the broader FE research program by introducing alternative approaches to gene discovery, such as gene-expression profiling. Several new avenues of research have also been identified, in particular, the development and characterisation of an in vitro model will not only facilitate gene discovery in FE, but also allow the function of gene variants to be tested.

FE has been the subject of active research for several decades. It is an important disease, not only as a model of fungal toxicosis, but also because of the economic and animal health ramifications of the disease. It is likely that FE is going to become a more significant problem over time as both temperatures and rainfall are predicted to increase significantly over the next 50 years (Ministry for the Environment, 2006). This will generate environmental conditions that favour the growth of *P. chartarum*, which will result in an increase in the incidence of FE. Therefore, the research into the genetic basis of FE resistance in livestock is of paramount importance, and it is hoped that the work presented in this thesis will help advance this goal.
References


References


References


References


References


Eisenstein, E., Gilliland, G. L., Herzberg, O., Moult, J., Orban, J., Poljak, R. J., Banerjei, L., Richardson,


References


Harris, M. A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S.,


References

31, 1296-1304.


References


References


References


References


References


286
References


References


Smith, B. L., Embling, P. P. and Gravett, I. M. (1987). ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. Molecular Cell 10, 139-149.


References


induction of metallothioneins and Hsp70 and presence of the membrane transporter ZnT-1 in HepG2 cells exposed to copper and zinc. *Toxicology In Vitro* 17, 553-559.


References


294
References


Appendix A
Oligonucleotide Primer Sequences Used in this Study.

Table A1: Oligonucleotide primers common to several chapters

<table>
<thead>
<tr>
<th>Fragment Name</th>
<th>Primer Name</th>
<th>Primer Sequence (5' – 3')</th>
<th>Tm (°C)</th>
<th>Fragment Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7</td>
<td>TAA TAC GAC TCA CTA TAG</td>
<td>54</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>SP6</td>
<td>ATT TAG GTG ACA CTA TAG AAT</td>
<td>52</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

Table A2: Oligonucleotide primers used in the quantification of ABCB1 gene expression (Chapter Three)

<table>
<thead>
<tr>
<th>Fragment Name</th>
<th>Primer Name</th>
<th>Primer Sequence (5’ – 3’)</th>
<th>Tm (°C)</th>
<th>Fragment Size (bp)</th>
</tr>
</thead>
</table>
| TaqMan qRT-PCR primer and probe sequences
| ABCB1_qRT-PCR | LR* F       | AGA AAG AAA AGA GAC CAA CTG TCA GTA CT | 55.2 | 89 |
|               | LR_R        | CAA AGT CCC CAG CAC CAT ACA | 54.8 | |
|               | LR_Probe    | 6FAM- AAC CTA TCG AGC CAA TTT GAA TAG CGA AAC ATT- TAMRA | 65.7 | |
| GAPDH_qRT-PCR | GAPDH1_F    | GGG CCA TCC ACA GTC TTC TG | 54.2 | 128 |
|               | GAPDH1_R    | AAG ATT GTC AGC AAT GCC TCC T | 54.5 | |
|               | GAPDH_Probe | TET- CAC TGT CCA CGC CAT CAC TGC CA -TAMRA | 64.7 | |
Appendix A: Oligonucleotide Primer Sequences Used in this Study

Sequencing Primers / Probes for Northern blotting

<table>
<thead>
<tr>
<th>Competitive RT-PCR</th>
<th>Con7</th>
<th>TGA GCG ATC ACA ATG CAG</th>
<th>60.6</th>
<th>452</th>
</tr>
</thead>
<tbody>
<tr>
<td>Con8</td>
<td>CTC CTG GAG CGG TTC TAC</td>
<td>61.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CGA CC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* LR refers to the leader region (the sequence preceding the first transmembrane region in the ABCB1 sequence)
† Primer sequences obtained from Longley (1998)

Table A3: Oligonucleotide primers used in the sequencing and analysis of ABCG2 (Chapter Four)

<table>
<thead>
<tr>
<th>Fragment Name</th>
<th>Primer Name</th>
<th>Primer Sequence (5' – 3')</th>
<th>Tm</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 1 – 6</td>
<td>E1.1-F</td>
<td>TGCTCATAAAATGTCTTTCCAATTAGTTACGA</td>
<td>58</td>
<td>717</td>
</tr>
<tr>
<td></td>
<td>E6.1-R</td>
<td>AGGCTGTAGAATGGAAGAAGATGTTGT</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Exon 5 - 9</td>
<td>E5.1-F</td>
<td>CAGAGTATTAGCTGTTAAATGTCAGAT</td>
<td>60</td>
<td>640</td>
</tr>
<tr>
<td></td>
<td>E9.1-R</td>
<td>TGAGCTGATGACAGAAGGAGGTGG</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Exon 9 - 13</td>
<td>E9.1-F</td>
<td>TATGTCACCTCCCTTTCTCAAGGAACAC</td>
<td>59</td>
<td>701</td>
</tr>
<tr>
<td></td>
<td>E13.1-R</td>
<td>TGAAGCCATGACAAACCAAGGCAC</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Exon 12 - 16</td>
<td>E12.1-F</td>
<td>GGTGGCTTATTCACGTATGTCAGATGG</td>
<td>60</td>
<td>356</td>
</tr>
<tr>
<td></td>
<td>E16.1-R</td>
<td>CCAAGGCTACGTGATTCTTCCACAG</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

5' and 3' RACE

5' RACE

| GI            | CCGACAAATGTAGAAGGACAGTCTTCAC | 60  |      |
| GO            | GATTTCCTCCACTGTGGGTC | 59  |      |
| adapter*      | GCUGAUGGCCAUGAAUGCACUGCG | 80  |      |
| Outer*        | GCTGATGCGGTGAAATGAAACTG | 60  |      |
| Inner*        | CGCGGATCCGAAACTGCGTTTCTGCTTTGATG | 80  |      |

3' RACE

| GI            | CTGGGAGAAGATCAACGATTTGGG | 58  |      |
| GO            | TTCTGACAGGACTCAAGTAACA | 60  |      |
| Adapter*      | GCAGCGACAGAATTACGACTACAACAGATTAGGT*2VN | 80  |      |
| Outer*        | GCGACGACAGAATTTAAACGACT | 52  |      |
| Inner*        | CGCGGATCCGAAACTGCGTTTCTGCTTTGATG | 80  |      |
### Appendix A: Oligonucleotide Primer Sequences Used in this Study

#### Probe for Northern hybridisation

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Tm</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAR_F</td>
<td>AGCTAAGGAGACCGAAGAGCC</td>
<td>55</td>
<td>767</td>
</tr>
<tr>
<td>OAR_R</td>
<td>TACTGAGCCACGACAGCCAA</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>GAPDH_F†</td>
<td>TGAAGGTCCGTGTGAACCGAATTGGC</td>
<td>67</td>
<td>987</td>
</tr>
<tr>
<td>GAPDH_R†</td>
<td>CATGTAGGCCATGAGGCCACCAC</td>
<td>61</td>
<td></td>
</tr>
</tbody>
</table>

#### Intron Amplification

<table>
<thead>
<tr>
<th>Intron</th>
<th>Primer</th>
<th>Sequence</th>
<th>Tm</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>I4.1-F</td>
<td>TATCTGGAGATGTTTTGATCAATGGAGCAC</td>
<td>62</td>
<td>916</td>
<td></td>
</tr>
<tr>
<td>I4.1-R</td>
<td>CATAGTTGTTGGAAGCCGAGGG</td>
<td>60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Mass extend assay for intron 4 position 671†

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Tm</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>ss65824074 A (For)</td>
<td>[ACGTTGGATG]TTATCCACAAAAACAGATCC</td>
<td>69</td>
<td>120</td>
</tr>
<tr>
<td>ss65824074 (AgR ref AS-000343) B (Rev)</td>
<td>[ACGTTGGATG]GCTACAGTGCCATAGAATGCAA</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>ss65824074 E (Probe)</td>
<td>GTGTTGGTGCTCAGTCATG</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

#### Mass extend assay for SNP c555†

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Tm</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>ss65824075 A (For)</td>
<td>[ACGTTGGATG]GTTCGAACTCAGTTATCCGTG</td>
<td>68</td>
<td>118</td>
</tr>
<tr>
<td>ss65824075 (AgR ref AS-000346) B (Rev)</td>
<td>[ACGTTGGATG]GCTACAGTCCATAGAATGCAA</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>ss65824075 E (Probe)</td>
<td>GTGTTGGTGCTCAGTCATG</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

#### Mass extend assay for SNP c973†

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Tm</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>ss65824134 A (For)</td>
<td>[ACGTTGGATG]GAAGCTAAGGAGACCGAAGAG</td>
<td>67</td>
<td>112</td>
</tr>
<tr>
<td>ss65824134 (AgR ref AS-000347) B (Rev)</td>
<td>[ACGTTGGATG]TCCTGGAAGAGGAGGTGG</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>ss65824134 E (Probe)</td>
<td>GACCGAAGAGCCCTCCAAAAAA</td>
<td>55</td>
<td></td>
</tr>
</tbody>
</table>

† Primers supplied with Ambion FirstChoice™ RLM-RACE kit.

† Primer sequences obtained from Charlier et al. (2001)

† Sequence within square brackets indicates the 10mer tag placed on the 5'-end of the Mass Extend primers.
Appendix A: Oligonucleotide Primer Sequences Used in this Study
Appendix B
Microarray Data

In accordance with MIAME guidelines (Brazma et al., 2001) a full description of the microarray slide is provided. This file (Array description file_MIAME.xls) contains EST clone information, position on the array, NCBI accession numbers and best human RefSeq hit.

In addition all original GenePix data and image files are provided as are the processed and normalised data files.

1. The user reads and agrees to the conditions of use on the publisher’s license agreement, if any, before using the software and content of this disc. Look for a license agreement in printed material accompanying the disc or in preliminary material on the disc itself.

2. The user should assume that, if permission in the license agreement allows copying of software and content to the user’s computer, it is temporary and the user should remove software and content from their computers when returning the disc to the Library, unless license conditions allow open distribution of copies of software and content.

3. The University of Otago Library can take no responsibility for any damage to software or hardware or data incurred through the use of this item. We recommend that you consult the system requirements of the disc before use.

University of Otago Library
Appendix C

Publications arising from this work


Cloning, mapping and association studies of the ovine ABCG2 gene with facial eczema disease in sheep

E. J. Duncan*, K. G. Dodds†, H. M. Henry*, M. P. Thompson‡ and S. H. Phua*
*AgResearch Molecular Biology Unit, Department of Biochemistry, University of Otago, PO Box 56, Dunedin 9054, New Zealand.
†AgResearch Invermay, Private Bag 50034, Mosgiel 9053, New Zealand. ‡Department of Biochemistry, University of Otago, PO Box 56, Dunedin 9054, New Zealand

Summary
Facial eczema (FE) is a hepato-mycotoxicosis in sheep caused by the fungal toxin sporidesmin. Resistance to FE is a multigenic trait. To identify QTL associated with this trait, a scan of ovine chromosomes was performed. In addition, ABCG2 was investigated as a possible positional candidate gene because of its sequence homology to the yeast PDR5 protein and its functional role as a xenobiotic transporter. The sequence of ovine ABCG2 cDNA was obtained from liver mRNA by RT-PCR and 5' and 3' RACE. The predicted protein sequence shares >80% identity with other mammalian ABCG2 proteins. SNPs were identified within exon 6, exon 9 and intron 4. The intron 4 SNP was used to map ABCG2 to ovine chromosome 6 (OAR6), about 2 cM distal to microsatellite marker OarAE101. Interestingly, this chromosomal region contains weak evidence for a FE QTL detected in a previous genome-scan experiment. To further investigate the association of ABCG2 with FE, allele frequencies for the three SNPs plus three neighbouring microsatellite markers were tested for differences in sheep selected for and against FE. Significant differences were detected in the allele frequencies of the intronic SNP marker among the resistant, susceptible and control lines. No difference in the levels of ABCG2 expression between the resistant and susceptible animals was detected by Northern hybridisation of liver RNA samples. However, significantly higher expression was observed in sporidesmin-dosed sheep compared with naïve animals. Our inference is that the ABCG2 gene may play a minor role in FE sensitivity in sheep, at least within these selection lines.

Keywords ABC transporter, liver, mycotoxin, photosensitivity, quantitative trait locus.

Introduction
Facial eczema (FE) disease is a hepato-mycotoxicosis of ruminants caused by sporidesmin, which is a toxin produced by the saprophytic fungus Pithomyces chartarum. This toxin causes severe necrotising inflammation of both the liver and bile ducts in FE-susceptible animals. Secondary photosensitisation occurs due to the inability of the affected animal to excrete phylloerythrin, a photo-reactive breakdown product of chlorophyll. As a consequence, phylloerythrin accumulates in the blood leading to localised cellular necrosis when activated by UV-light at exposed skin surfaces (reviewed by Mortimer & Ronaldson 1983).

There is a significant genetic component in resistance to FE, with an estimated heritability of 0.42 (Campbell et al. 1981). Several groups of genes have been examined for their involvement in conferring resistance to sporidesmin (Phua et al. 1999; Hohenboken et al. 2004). One gene family of particular interest is the ABC (ATP-binding cassette) transporter superfamily, which has 48 members in humans and are divided into eight sub-families (A–H; reviewed by Dean et al. 2001). This super-family was first highlighted as a potential group of candidate genes when it was found that levels of PDR5 (pleiotropic drug resistance protein 5, STS1) expression modulated the sporidesmin sensitivity of Saccharomyces cerevisiae (Bissinger & Kuchler 1994). At that time the closest known mammalian ortholog to the yeast PDR5 protein was ABCB1 (P-glycoprotein, MDR1), as they shared significant overlap in their substrate-specificity profile (Balzi et al. 1994; Bissinger &
Association of ovine ABCG2 gene with facial eczema disease

Kuchler 1994). Subsequent studies of ovine ABCB1 ruled out its possible involvement in FE (Longley et al. 1999; Morris et al. 2004).

Identification of a closer mammalian ortholog of the yeast PDR5 protein is now possible because of the availability of complete sequences of several genomes (Ponting 2001). Our phylogenetic analyses identified the mammalian ABCG sub-family as the most likely homologs of the yeast PDR5 protein, which agrees with published data (Sheps et al. 1999). Subsequently more appropriate genome-wide significance thresholds were set by permutation testing (Churchill & Doerge 1994) using 1000 replicates. These latter values were used in Fig. 1.

Romney facial eczema selection lines

The selection lines were as described previously (Phua et al. 1999; Morris et al. 2004). Briefly, genetic selection for FE resistance and susceptibility in Romney sheep started in 1975, with the control line introduced in 1982 (Morris et al. 1989). The selection response was assessed from

Materials and methods

Detection of quantitative trait loci in outcross families

Four outcross families were generated as described in Phua et al. (1999). Briefly, four F1 Romney rams were produced from reciprocal crosses of FE resistant and susceptible selection-line animals and mated with 130–170 Romney ewes to generate four outcross pedigrees, with 124–167 progeny per family. The severity of liver damage in sporidesmin-dosed progeny was assessed by measuring serum levels of the liver-specific enzymes gamma-glutamyl transferase (GGT) and glutamate dehydrogenase (GDR) 3 weeks post-dosing. Additionally, there was a cohort of very tolerant animals that did not respond to the initial sporidesmin dose (0.13 mg/kg live-weight); these animals were given a second dose of sporidesmin (0.17 mg/kg live-weight) and their GGT levels were measured 3 weeks after dosing. Data from the first initial dosing were designated GGT and GDR, and combined GGT data from the two dosings were designated GGT21.

The QTL analysis using the outcross families was conducted via primary and secondary screens. Ten evenly spaced microsatellite markers on OAR6 that were heterozygous in at least one of the four sires were used. In the primary screen, inheritance of the sire alleles was determined for 22 of the most resistant and 22 of the most susceptible progeny (based on GGT data). If a suggestive or significant QTL was detected, then all progeny (n = 124–167/sire) of the heterozygous sires were genotyped and analysed.

The GGT and GDH data together with the genotype information were analysed using the Haley-Knott method (Knott et al. 1996). In our original analyses, genome-wide suggestive and significant linkage probabilities were determined to be $P = 0.0016$ and $P = 0.000054$ respectively (Lander & Kruglyak 1995). Subsequently more appropriate genome-wide significance thresholds were set by permutation testing (Churchill & Doerge 1994) using 1000 replicates. These latter values were used in Fig. 1.

Figure 1 A region detected on OAR6, associated with facial eczema (FE) in outcross pedigrees. Selective genotyping was used in the primary genome screen (a). In the secondary screen (b), all progeny of informative sires were genotyped and analysed. The vertical axis plots the F-value for the allele-within-sire term (1 numerator d.f. for each of the four families). The three horizontal lines indicate the genome-wide suggestive levels as determined by permutation analysis for the three FE traits analysed (log$_2$GGT, log$_2$GGT and log$_2$GGT). The positions of microsatellite markers used in the genome scan are indicated by the arrows below each graph: A = OarCP125, B = BM9058, C = MCM53, D = BM1329, E = OarAE101, F = BM143, G = BM4621, H = BM4371, I = OarMP8 and J = BL1038.

© 2007 AgResearch, Journal compilation © 2007 International Society for Animal Genetics, Animal Genetics, 38, 126–131
changes in the FE breeding value for log eGGT. When ani­
mals were sampled to test for FE differences (birth years 1991–1995), the resistant and susceptible lines differed by a factor of ×3.7 (1991) and ×6.9 (1995) (Morris et al. 2004).

Isolation of sheep liver samples

Liver samples as a source for total RNA were obtained from 19 resistant- and 18 susceptible-line animals born in years 1995–1997 and 2000–2002. Naïve animals were undosed. Dosed animals were animals orally challenged with sporidesmin at 0.1 or 0.4 mg/kg live-weight. Sporidesmin-dosed animals used in the Northern hybridisation experiment all received 0.1 mg/kg, with the exception of R8 and R9 (Fig. 2) which received 0.4 mg/kg. Animals were killed 24 h after dosing for collection of livers: liver samples were snap-frozen in liquid nitrogen and stored at −80 °C.

Amplification and sequencing of ovine ABCG2 cDNA

Primers were designed using vector NTI suite 7.1 (InforMax Inc.) based on the human ABCG2 gene (NM_004827). Primer sequences, melting temperatures and the sizes of the amplified fragments are shown in Table S1. Sequencing was done on cDNA obtained after PCR amplification, and 5’ and 3’ RACE.

Single nucleotide polymorphism (SNP) discovery and genotyping

Human and mouse genome sequences for ABCG2 were found using the UCSC genome browser (http://genome.ucsc.edu/), and megaBLAST (Zhang et al. 2000) was used to extract equivalent bovine genomic sequences. The genome sequences were then aligned against the ovine cDNA sequence using the snoopy algorithm (Wheelan et al. 2001) set to adjust for divergent sequences. Primers were then designed to flank intron 4 (about 1 kb) of the ABCG2 gene (Table S1). Approximately 100 ng of genomic DNA from

![Image](https://via.placeholder.com/150)

**Figure 2** Northern hybridisation showing expression of the ABCG2 transcripts in livers of resistant and susceptible selection-line animals, both naïve and sporidesmin-dosed. The gender of individual animals is indicated below the lanes (M = male, F = female).

the four AgResearch International Mapping Flock (IMF) sires (Crawford et al. 1995) was used in a PCR reaction, and SNPs were detected by sequencing the PCR products.

The intronic SNP marker (ss65824074:G>A) was geno­typed in both the full IMF population (n = 125) and FE selection-line animals (n = 176), including 66 resistant, 66 susceptible and 44 control animals. Two other SNPs (ss65824075:G>A and ss65824134:C>T), discovered in one of the selection-line animals during cDNA sequencing, and three microsatellite markers (JL36, CSAP14E, OarAE101), which were within 2.4 cM of the ABCG2 locus, were genotyped in the selection-line animals. All three SNPs were genotyped using primer extension with analyte detection in a MALDI-TOF mass spectrometer (Sequenom Inc.) (Tang et al. 1999) (Table S1).

Allele frequencies of markers in FE selection lines

Allele frequencies for each marker were compared between the resistant, susceptible and control selection lines using Chi-squared statistics. The significance level for these statistics was found using the peddrift method (Dodds & McElwan 1997), which accounts for genetic drift, founder and sampling effects. The method finds the distribution of the Chi-squared statistics under the null hypothesis of no selection on the marker by simulating genotypes through the actual pedigree data.

Northern hybridisation

Northern blots containing 20 μg of total liver RNA per animal were hybridised with (α-32P)dCTP labelled probes. The ABCG2 probe consisted of part of the linker region plus the first five transmembrane spanning domains.

The autoradiographs were scanned using an ImageScanners™ II and LabScan™ (Amersham Biosciences), and quantified using ImageQuant TL™ (Amersham Biosciences). To correct for variation in loading or transfer, the total absorbance of each band was determined and normalised to the total absorbance units for a GAPDH control (Charlier et al. 2001) in the same sample. Absorbance values were log-transformed to account for the increased variance scaled to the mean and then analysed by least squares (ANOVA) methods. The initial model included the main effects of sporidesmin exposure and genetic lines along with their interaction. The interaction was dropped from subsequent models as it was found to be non-significant.

Results and discussion

Facial eczema quantitative trait loci experiment

As part of a whole genome-scan QTL experiment, 10 microsatellite markers evenly spaced on OAR6 were geno­typed in four outcross families in the primary screen. The
analysis revealed weak evidence for a QTL between markers MCM53 and BM4621, with its peak around OarAEIO1. Initial analysis using the method of Lander & Kruglyak (1995) showed significant linkage in this region for log_gGT21 and a suggestive linkage for log_gGTT. However, subsequent analyses using permutation thresholds (Churchill & Doerge 1994) gave a F-ratio for log_gGTT21 near the suggestive level (Fig. 1a). In addition, the 95% confidence interval spanned all 10 markers tested. When the four families were considered individually, one family (R × S Sire 3) showed suggestive linkage for log_gDH, log_gGT and log_gGTT21 in this region (data not shown).

In the secondary genome screen where all progeny of heterozygous sires were analysed, similar QTL results were obtained with the combined families’ data (Fig. 1b), except that the 95% confidence interval for log_gDH was reduced to the 20–170 cM interval on OAR6. When the four families were analysed individually, the R × S Sire 3 family gave a significant QTL (genome-wise P < 0.05) in this same region (between MCM53 and BM4621) for log_gDH, with an estimated allele substitution effect of 0.84 phenotypic SD and a 95% confidence interval spanning the 44.7–86.7-cM region (data not shown).

Ovine ABCG2 cDNA sequence

Using primers designed to the human ABCG2 sequence (Table S1), the full coding region of the ovine ABCG2 gene was obtained. The ovine sequence showed conserved motif structure, domain organisation and sequence similarity to the human ABCG2 gene.

Several SNPs have been identified in the human ABCG2 gene, and three of these SNPs are known to affect the function of the ABCG2 protein. The Q141K mutation affects the transport efficiency of the protein (Matsurai et al. 2004), whereas the S441N mutation is known to alter the localisation of the mature protein from the cell membrane to an intracellular location. However, at the time of our study, the only known mutation that altered ABCG2 protein function was R482X, which was discovered in drug-selected cell lines and altered substrate specificity (Honjo et al. 2001). Subsequent studies failed to detect the R482X mutation in human populations (Honjo et al. 2002).

The entire coding region of the ovine ABCG2 gene from two resistant and two susceptible selection-line sheep was sequenced and two synonymous SNPs were detected. The first SNP was a G>A transition at position 554 of the cDNA sequence (exon 6, ss65824075:G>A), and the second SNP was a C>T transition at position 973 (exon 9, ss65824134:C>T). Additionally, the 1059–1800 bp coding region encompassing amino acids 353–600 (exons 9–13) was sequenced from eight resistant- and eight susceptible-line animals, and no polymorphisms were detected. These data suggest that there are no differences in the primary structure of the ABCG2 gene with facial eczema disease.

Association of ovine ABCG2 gene with facial eczema disease

Linkage mapping of the ABCG2 locus

To find ABCG2 SNPs for linkage mapping in the AgResearch IMF animals, intron 4 was sequenced for four IMF sires. Intron 4, which is 813 bp in length (DQ886529), has the classical GT donor and AG acceptor splice boundaries. Six SNPs and a single insertion were discovered within this intron. Three of the four IMF sires were heterozygous for a SNP at position 671 (ss65824074:G>A). The ss65824074:G>A mutation was linked to markers on OAR6 (Fig. S1) with a maximum two-point LOD score of 16.1 with LSCV43 on the framework map of Maddox et al. (2001) and a recombinant fraction of 0.01. Multipoint mapping placed ABCG2 2 cM distal to OarAEIO1, consistent with that predicted from comparative mapping between human, cattle and sheep. The ABCG2 locus falls within the peak of a FE QTL on OAR6 (Fig. 1).

Association studies of markers in facial eczema selection lines

Association studies of ABCG2 and nearby microsatellite markers with FE were carried out in FE selection lines. Only the ABCG2 SNP ss65824074:G>A showed a significant difference in allele frequencies between the selection lines (P = 0.044): the susceptible line (allele A frequency = 0.17) differed from the resistant (A = 0.52) and control (A = 0.58) lines. In addition, haplotypes composed of the three ABCG2 SNPs showed significant association within FE selection lines (P = 0.03). Therefore in addition to ABCG2 being a positional and functional candidate gene for the OAR5 FE QTL region, this association suggests its possible involvement in FE disease.

Northern analysis of the ABCG2 gene

Northern hybridisation was carried out to determine if there was any differential expression of ABCG2 in liver between resistant and susceptible sheep under naïve and sporidesmin-dosed conditions. Three RNA bands were detected (Fig. 2); band B (2.1 kb) was the most predominant transcript with a size consistent with the full-length cDNA sequence (2191 bp). The Northern autoradiograph was scanned and quantified, and the expression of the 2.1 kb ABCG2 transcript was calculated relative to the GAPDH control. Statistical analysis showed no differential expression of the ABCG2 gene between resistant and susceptible sheep, in both naïve and sporidesmin-dosed animals. No difference in expression was detected between male and female sheep despite a number of reports concerning gender differences in ABCG2 expression (Krishnamurthy & Schuetz 2005).
Interestingly, there was a 1.56-fold increase of ABCG2 expression following exposure of animals to sporidesmin when compared with naïve animals (P = 0.017). As expression of ABCG2 in stem cells is known to be modulated by the transcription factor HIF-1 (hypoxygen-inducible factor 1) (Krishnamurthy et al. 2004) and HIF-1 is also expressed in liver, it is possible that the observed increase in ABCG2 transcription may be via HIF-1 activation in response to a cytotoxic hypoxygen state induced by sporidesmin. Moreover, sporidesmin-mediated induction of ABCG2 may enable it to act as a phase III defence mechanism by accelerating removal of the xenobiotic via enhanced excretion into the bile (reviewed by Krishnamurthy & Schuetz 2005).

Although only liver expression was studied, ABCG2 is known to be expressed in other tissues and its expression or function in those tissues may be of more relevance to FE. For instance, in mice the ABCG2 protein modulates both hepatic biliary excretion and intestinal absorption of substrates (reviewed by Staud & Pavek 2005). Therefore, it would be of interest to study ABCG2 in the intestines of FE sheep.

Acknowledgements

The authors thank K. Paterson for designing the Mass Extend assays, A. Beattie and T. Van Stijn for CRIS-MAP analysis, B. Auvray for haplotype estimation and Dr R. McLaren and Dr O. Kene for providing the GAPDH clone. E. J. Duncan was supported by a MeatNZ grant to S. H. Phua.

References


Supplementary Material
The following supplementary material is available for this article online from http://www.blackwell-synergy.com/doi/full/10.1111/j.1365-2052.2007.01557.x
Figure S1 Localisation of the ABCG2 locus on the sex-averaged map of ovine chromosome 6 (OAR6).
Table S1 PCR primer sequences for ovine ABCG2.

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors.
Zinc protection of HepG2 cells from sporidesmin toxicity does not require de novo gene transcription

Elizabeth J. Duncan\textsuperscript{a}, Mary P. Thompson\textsuperscript{b}, Sin H. Phua\textsuperscript{a, *}

\textsuperscript{a} AgResearch Molecular Biology Unit, Department of Biochemistry, University of Otago, P.O. Box 56, Dunedin, New Zealand
\textsuperscript{b} Department of Biochemistry, University of Otago, Dunedin, New Zealand

Received 28 February 2005; received in revised form 19 May 2005; accepted 19 May 2005
Available online 7 July 2005

Abstract

Sporidesmin is an epidithiodioxopiperazine mycotoxin secreted by the saprophytic fungus \textit{Pithomyces chartarum}. Ingestion of sporidesmin by ruminants grazing on the saprophyte infested pasture causes severe liver and bile duct damage leading to secondary photosensitisation. Zinc supplementation is used as an effective prophylaxis against sporidesmin toxicity in ruminants, however, the mechanism by which zinc protects is unknown. This study used the human hepatoma cell line, HepG2, as a model to examine the mechanism of zinc protection against sporidesmin toxicity. Treatment of cells with various concentrations of sporidesmin (0–10 \(\mu\)g/ml) resulted in a sigmoidal dose response curve with an LC\textsubscript{50} of 5 \(\mu\)g/ml. Cells were protected from sporidesmin toxicity by pre-treatment for 2 h or 16 h with zinc sulphate in a concentration dependent manner, with significant protection at 50 \(\mu\)M zinc and maximal protection at 200 \(\mu\)M zinc. To determine whether zinc protection required de novo gene transcription, cells were treated with the transcriptional inhibitor actinomycin D for one hour prior to and throughout the zinc pre-treatment. The presence of actinomycin D did not significantly reduce the zinc protection against sporidesmin cytotoxicity (80% protection without actinomycin D versus 71% protection with actinomycin D). Therefore, de novo gene transcription does not play a major role in the mechanism of zinc protection against sporidesmin toxicity in HepG2 cells.

© 2005 Elsevier Ireland Ltd. All rights reserved.

Keywords: Sporidesmin; Epidithiodioxopiperazine; HepG2; Zinc; Hepatotoxin; Actinomycin D

1. Introduction

Epidithiodioxopiperazine (ETP) mycotoxins such as sporidesmin (Scheme 1) and gliotoxin are extremely cytotoxic both in vitro (Mortimer and Collins, 1968; Mortimer et al., 1974; Cordiner and Jordan, 1983; Monday, 1984a,b; Jordan and Pedersen, 1986) and in vivo (Cunningham et al., 1942; Done et al., 1961; Towers and Stratton, 1978). Sporidesmin, which is produced by the saprophytic fungus \textit{Pithomyces chartarum}, causes a pathological condition in ruminants called facial eczema disease (Mortimer and Taylor, 1952; reviewed in Morris et al., 2004). Gliotoxin is...
involved in the aetiology of invasive *Aspergillus* infections in humans (Bodey et al., 1992).

Facial eczema is a significant problem in New Zealand as 86% of its *P. chartarum* strains are toxigenic (Collin et al., 1998). Ruminants ingest sporidesmin by grazing on the fungal infected pasture. Sporidesmin is excreted primarily in the bile and causes severe necrotising inflammation of both the liver and bile ducts in sensitive animals (Atherton et al., 1974; Mortimer and Ronaldson, 1983). Secondary photosensitisation occurs due to the inability of the animal to excrete phylloerythrin, a photoreactive breakdown product of chlorophyll. As a consequence phylloerythrin accumulates in the blood leading to localised cellular necrosis when activated by UV light at exposed skin surfaces (Mortimer and Ronaldson, 1983).

Dosing of ruminants with high levels of zinc (up to 50 mg/kg) is an effective prophylactic treatment for facial eczema (Smith et al., 1977; Towers and Smith, 1978; Munday et al., 1997, 2001). However, the mechanism by which zinc protects animals against sporidesmin toxicity is unknown. Zinc is the most abundant trace element in the cell and has numerous biological roles (Cousins, 1996). It is involved as a key structural component or as a cofactor for over 300 metallo-enzymes (Kambe et al., 2004). It is also required for zinc-containing DNA-binding proteins, thereby acting as a potent modulator of gene expression by activating transcription factors such as MTF1 (Cousins, 1996). Several recent studies have analysed the effect of zinc on the global gene expression profile in various cell systems and found that zinc is capable of modulating the expression of hundreds of functionally diverse genes (Blanchard and Cousins, 1996; Blanchard et al., 2001; Moore et al., 2001; tom Dieck et al., 2003; Cousins et al., 2003; Kindermann et al., 2004).

In the present study we characterised the cytotoxicity of sporidesmin on HepG2 cells. We also examined the role of zinc in protecting HepG2 cells against sporidesmin toxicity, and evaluated a possible mechanism effecting that protection.

2. Materials and methods

2.1. Materials

The human hepatoma cell line, HepG2, was originally sourced from ATCC. Alpha-MEM, actinomycin D and foetal bovine serum (New Zealand origin) were obtained from Invitrogen – Gibco (Grand Island, NY, USA). Sodium pyruvate and zinc sulphate were obtained from Sigma Chemical Co. (St. Louis, MO, USA). NADH was purchased from Applichem (Darmstadt, Germany). Purified sporidesmin was a gift from the AgResearch Toxinology Group (Ruakura, NZ).

2.2. Cell culture

HepG2 cells (passage 21) were plated at a density of $2 \times 10^6$ cells per well (9.6 cm$^2$) in a six-well culture dish in 2 ml of alpha-MEM, 10% foetal bovine serum, 50 U/ml penicillin G and 50 μg/ml streptomycin sulphate. Cells were cultured for at least 20 h (37°C, 5% CO$_2$/95% air) prior to being subjected to experimental treatments as detailed below (refer to Scheme 2A and B).

Unless indicated media were not changed prior to experimental treatment, and all media used were alpha-MEM containing 10% foetal bovine serum, 50 U/ml penicillin G and 50 μg/ml streptomycin sulphate. All experiments included appropriate solvent controls. Experiments with zinc sulphate included water as the solvent control. In experiments with sporidesmin the solvent control was ethanol. Experiments with actinomycin D included DMSO as the solvent control. In all cases the volume of the individual solvent added did not exceed 0.5% of the total medium volume.

To assess the cytotoxicity of sporidesmin, cells were cultured for 36 h post-passage before treatment with various concentrations of sporidesmin (0–10 μg/ml). Cells were harvested 16 h later (total of 52 h post-
passage) and the percentages of non-viable cells were determined using the lactate dehydrogenase (LDH) assay. A time course study was also performed where cells were treated with 5 μg/ml of sporidesmin at 36 h post-passage, and cells were harvested every 4 h for 24 h (from 40 to 60 h post-passage) for determination of percent non-viable cells.

To assess zinc cytotoxicity, the HepG2 cells were treated with different zinc sulphate concentrations (0–500 μM) at 20 h post-passage. After the 16 h incubation with zinc, the percentages of non-viable cells were measured.

In experiments on zinc protection against sporidesmin toxicity, cells were incubated at either 20 h or 34 h post-passage with different concentrations of zinc sulphate (0–200 μM). The incubation periods were for either 16 h or 2 h, respectively. At 36 h post-passage, the zinc-containing media were removed and the cells were washed with phosphate-buffered saline (PBS) before being treated with sporidesmin-containing media for a further 16 h. At 52 h post-passage, the percentages of non-viable cells were determined using the LDH assay (refer to Scheme 2A and B).

The effect of transcriptional inhibitor, actinomycin D, on zinc protection was determined by treating the cells at 33 h post-passage with 10 μg/ml of actinomycin D (Scheme 2B), then at 34 h post-passage with 200 μM zinc sulphate for a further 2 h (total of 36 h post-passage). The media containing zinc sulphate and actinomycin D were then removed and replaced with fresh medium containing sporidesmin (5 μg/ml) for a further 16 h of incubation. The percentages of non-viable cells were measured at 52 h post-passage.

2.3. Lactate dehydrogenase assay

The lactate dehydrogenase (LDH) assay was used to measure the percentage of non-viable cells as a marker for cytotoxicity. This assay measures the proportion of the cytosolic enzyme LDH released into the culture media by dead and dying cells as compared to the total LDH in the sample well. In practice, the culture medium was removed and its LDH level measured. The adherent cells that remained in the sample well were washed briefly with PBS before lysing with 0.1% Triton X-100 in PBS; this lysate was then analysed for LDH activity. Routinely, 200 μl of assay buffer (0.2 mM NADH, 1.6 mM sodium pyruvate, 200 mM NaCl, 80 mM Tris–HCl pH 7.5) was aliquoted into each well of a 96-well plate, and 10 μl of either culture medium or cell lysate was then added. Data were collected using an ELx808tm ultramicroplate reader (Bio-Tek Instruments, Winooski, VT, USA). The absorbance at 340 nm was measured over 3 min, and the initial rate of the oxidation of NADH to NAD⁺ was calculated as a measure of LDH activity. The percentage of non-viable cells was then calculated as follows: % non-viable cells in the well = percentage of LDH activity in culture medium/total LDH in the well (i.e. LDH activity in both culture medium and cell lysate).

2.4. Statistical analysis

Statistical significance of differences among treatments was assessed with MINITAB® (release 14, Minitab Inc.) using either a two sample t-test with the
Bonferroni correction for multiple testing, or one way ANOVA with Fisher's LSD post hoc test, as indicated in the text. Graphs were plotted in SigmaPlot® (version 7.0, SPSS Inc.), and curves were fitted using either sigmoidal or linear regression as stated in the text.

3. Results

3.1. The effect of sporidesmin on the viability of HepG2 cells

The toxicity of various concentrations (0–10 µg/ml) of sporidesmin on HepG2 cells was measured. The percentages of non-viable cells were measured 16 h after sporidesmin treatment using the LDH assay. A sigmoidal dose response curve was obtained over a narrow range (0–10 µg/ml) (Fig. 1A) and the LC50 of sporidesmin was calculated to be 5.0 µg/ml (10.6 µM).

In order to examine the time course of sporidesmin toxicity, cells were treated with 5 µg/ml of sporidesmin and the percentages of non-viable cells determined every 4 h for 24 h (Fig. 1B). A linear relationship of percent non-viable cells with time was observed ($r^2 = 0.97$ for solvent control, $r^2 = 0.99$ for sporidesmin treated), with the toxicity becoming significant as compared to the control after 12 h of treatment. At time points beyond 24 h the graph plateaued (data not shown).

![Graph](https://example.com/graph1.png)

Fig. 1. (A) Cytotoxic effect of sporidesmin on HepG2 cells. Cells were treated with sporidesmin at the indicated concentrations for 16 h. The percentages of non-viable cells were determined using the LDH assay. Each point represents the mean ± S.D. of three independent experiments performed in triplicate. A sigmoidal curve (five parameter) was fitted using SigmaPlot ($r^2 = 0.99$). (B) Time course of sporidesmin cytotoxicity in HepG2 cells. Cells were treated with 5 µg/ml sporidesmin for a range of time points from 4 h to 24 h. Each point represents the mean ± S.D. of two independent experiments each performed in triplicate. The percentages of non-viable cells were determined using the LDH assay. Asterisks indicate statistically significant differences from solvent control based on a two sample t-test (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

3.2. The effect of zinc sulphate on the viability of HepG2 cells

Cells were treated with 0–500 µM zinc sulphate for 16 h and the percentages of non-viable cells measured using the LDH assay. A sigmoidal dose response curve was fitted using SigmaPlot ($r^2 = 0.99$) and the LC50 of zinc sulphate was calculated to be 320 µM (Fig. 2). The highest concentration of zinc sulphate that did not cause toxicity ("no observed effect level" or NOEL) was determined to be 200 µM, using one-way ANOVA.

![Graph](https://example.com/graph2.png)

Fig. 2. Cytotoxic effect of zinc sulphate on HepG2 cells. Cells were treated with various concentrations of zinc sulphate for 16 h. The percentages of non-viable cells were determined using the LDH assay. Each point represents the mean ± S.D. of two independent experiments each performed in triplicate. A sigmoidal curve (five parameter) was fitted using SigmaPlot ($r^2 = 0.99$).
with a Fisher’s LSD post hoc test ($p > 0.05$). Therefore, in testing the protective effect of zinc sulphate against sporidesmin toxicity, a range of zinc concentrations up to a maximum of 200 µM was used.

3.3. Protective effect of zinc sulphate on HepG2 cells against sporidesmin cytotoxicity

Cells were pre-treated with 0, 10, 50, 100 and 200 µM zinc sulphate for 16 h, then washed with phosphate-buffered saline to minimise the extracellular zinc concentration. The cells were then treated with sporidesmin at 5 µg/ml for another 16 h (Scheme 2A). As shown in Fig. 3, the cells were protected from sporidesmin toxicity by zinc in a concentration dependent manner. None of the concentrations tested could reduce the sporidesmin toxicity to the background level of 6.3% (untreated cells). Significant protection was seen at 50 µM zinc as assessed by one-way ANOVA: zinc conferred 49% protection compared to the control (sporidesmin only treatment) where the percentage of non-viable cells was reduced from 49.7% to 25.2%. The maximal protection was seen with 200 µM zinc which conferred 66% protection compared to the control where the percentage of non-viable cells was reduced from 49.7% to 16.9%.

Fig. 3. Effect of zinc sulphate pre-treatment on sporidesmin toxicity in HepG2 cells. Following 16 h pre-treatment with various zinc concentrations, zinc-containing media were removed, the cells washed and fresh media containing sporidesmin were added, and the incubation continued for a further 16 h. The percentages of non-viable cells were determined using the LDH assay. Each bar represents the mean ± S.D. of three independent experiments each performed in triplicate. U: untreated (cells treated with neither zinc nor sporidesmin). Asterisks indicate statistically significant differences based on one-way ANOVA with Fisher’s LSD post hoc test (**p < 0.001).

To determine if there was any effect of zinc pre-treatment time on the protection observed against sporidesmin toxicity, a 16 h pre-treatment experiment was compared with a 2 h pre-treatment experiment (Fig. 4). Within each zinc pre-treatment time, there were significant differences between zinc-treated samples compared to the water control in terms of protection against sporidesmin toxicity; these results are consistent with the data presented in Fig. 3. However, it was found that there was no significant difference between the two zinc pre-treatment times for most of the zinc concentrations examined (Fig. 4); the one exception was a significant difference observed at the 200 µM zinc concentration, where the protection observed was slightly more effective after 2 h of zinc pre-treatment as compared with 16 h ($p < 0.05$ using two sample t-test).

3.4. Effect of actinomycin D on zinc protection of HepG2 cells against sporidesmin toxicity

To determine whether gene transcription played a role in the mechanism of protection by zinc against sporidesmin toxicity, cells were treated with actinomycin D for 1 h prior to and throughout the 2 h zinc pre-treatment. As summarised in Table 1, actinomycin D treatment alone was significantly more toxic (27.4%) compared to the DMSO solvent control (10.6%), as
Table 1

Effect of actinomycin D on zinc protection of HepG2 cells against sporidesmin toxicity

<table>
<thead>
<tr>
<th>Treatment</th>
<th>% Non-viable cellsa</th>
<th>Relative % non-viable cellsb</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMSO solvent control</td>
<td>10.6 ± 2.7</td>
<td>0.0</td>
</tr>
<tr>
<td>DMSO + zinc</td>
<td>10.4 ± 1.9</td>
<td>-0.2 ± 1.1</td>
</tr>
<tr>
<td>DMSO + sporidesmin</td>
<td>49.0 ± 3.3</td>
<td>38.4 ± 5.7</td>
</tr>
<tr>
<td>DMSO + zinc + sporidesmin</td>
<td>18.5 ± 3.0</td>
<td>7.9 ± 3.0</td>
</tr>
<tr>
<td>Actinomycin D control</td>
<td>27.4 ± 4.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Actinomycin D + zinc</td>
<td>25.9 ± 3.2</td>
<td>-1.5 ± 3.3</td>
</tr>
<tr>
<td>Actinomycin D + sporidesmin</td>
<td>59.9 ± 4.0</td>
<td>32.5 ± 4.6</td>
</tr>
<tr>
<td>Actinomycin D + zinc + sporidesmin</td>
<td>36.7 ± 6.4</td>
<td>9.3 ± 4.1</td>
</tr>
</tbody>
</table>

a Mean ± S.D. of three independent experiments each performed in triplicate.

b The relative percentages of non-viable cells were calculated by subtracting the percentage values of non-viable cells for the DMSO and actinomycin D controls. This data is presented in Fig. 5 for graphical comparison.

To assess the statistical significance between treatments in the presence or absence of actinomycin D, a two sample t-test statistic was used and the level of significance was calculated using the Bonferroni correction for multiple testing. There was a statistically significant difference (p < 0.05) between sporidesmin-treated versus sporidesmin plus zinc-treated samples (Fig. 5); this result is consistent with the previous findings (Figs. 3 and 4). However, there was no significant difference between the DMSO (without actinomycin D) and actinomycin D treated experiments (Fig. 5). This result indicates that the actinomycin D treatment, and therefore the inhibition of zinc protection of HepG2 cells against sporidesmin toxicity.

4. Discussion

The hepato-mycotoxin sporidesmin causes a pathological condition in ruminants called facial eczema disease (reviewed in Morris et al., 2004), and high levels of zinc supplementation are used as an effective prophylaxis against this disease (Smith et al., 1977; Towers and Smith, 1978). However, the mechanism by which zinc protects animals against sporidesmin toxicity is unknown. HepG2 cells are a hepatocellular carcinoma cell line that retains many of the morphological features (Sarmunen et al., 1993) and gene expression characteristics of hepatocytes (Knowles et al., 1980; Lu and Huang, 1994; Urani et al., 1998; Harris et al., 2004), thus making them useful for hepatotoxicity studies. This study used HepG2 cells as an in vitro model to...
characterise sporidesmin cytotoxicity in order to examine the mechanism of protection by zinc.

Treatment of HepG2 cells with increasing concentrations of sporidesmin (0–10 µg/ml) resulted in a classic sigmoidal dose response curve, with a narrow linear range of toxicity (2–8 µg/ml) and an LC50 value of 5 µg/ml (Fig. 1A). The narrow linear range of toxicity is consistent with published data using various other cell lines including HEp-2 (HeLa cell derivative), RKE (rabbit kidney epithelium) and VMK (Vernet monkey kidney cells) (Done et al., 1961; Mortimer and Collins, 1968). However, the LC50 value for HepG2 cells reported here is three orders of magnitude higher than that published in the literature for sporidesmin toxicity in HeLa cells (Done et al., 1961). Similar low concentrations of sporidesmin (ng/ml) have been reported to have a profound effect on cellular morphology in other cell lines such as HEp-2, RKE and VMK (Mortimer and Collins, 1968). The reason for the differences reported for sporidesmin sensitivity is not known, although the type of culture media used, varying serum concentrations (Mortimer and Collins, 1968), different preparations of sporidesmin and dissimilar methods of determining cytotoxicity could be contributing factors.

This is the first in vitro study to show that zinc pre-treatment protects cells from sporidesmin toxicity. The zinc concentrations tested ranged from 10 µM, representing physiological plasma levels (Cousins, 1996), to 200 µM which is the maximum level at which zinc was not intrinsically toxic to HepG2 cells (Fig. 2). It was found that zinc pre-treatment is capable of providing significant protection at 50 µM, with maximal protection observed at 200 µM (Fig. 3). The results obtained are consistent with those seen by Waring et al. (1990) who found that zinc provided almost complete protection against sporidesmin toxicity in mouse macrophages using cell morphology and DNA laddering as markers for apoptosis. However, in the Waring et al. (1990) study, 1000 µM zinc treatment was used and the zinc treatment was concurrent with sporidesmin treatment. Therefore, it was possible that the zinc could protect the cells by interacting directly with sporidesmin in the culture medium to form bivalent complexes, as such complex formation has been shown to occur in vitro (Woodcock et al., 2001a,b). In this study, the HepG2 cells were pre-treated with zinc and the cells were washed prior to treatment with sporidesmin to minimise such extracellular interactions.

The observation that there was equivalent protection provided by zinc pre-treatment for either 2 h or 16 h suggests that intracellular zinc levels may be saturated after a 2 h incubation. However, studies on the uptake of zinc into HepG2 cells by Urani et al. (2001) would suggest that uptake of zinc can continue and is not saturated at 2 h. This suggests that only a relatively small intracellular concentration of zinc is required to confer protection against sporidesmin, and that the protective mechanism is effected rapidly.

As gene expression in response to zinc treatment or zinc depletion is known to be altered rapidly, at least in human THP-1 mononuclear cells (Cousins et al., 2003), the role of gene transcription in the protective effect of zinc sulphate on sporidesmin toxicity was examined. HepG2 cells were treated with the transcriptional inhibitor, actinomycin D, for one hour prior to and throughout the zinc pre-treatment period. The result presented in Table 1 and Fig. 5 indicates that actinomycin D did not alter the protective effect of zinc. This suggests that zinc is protecting HepG2 cells from sporidesmin toxicity by a mechanism that is independent of de novo gene transcription. This is an important finding as zinc is known to be a potent modulator of gene expression (Cousins, 1996).

Since altered gene expression has been ruled out as an important mechanism for zinc protection, this leaves open other possibilities by which zinc could prevent or counteract the toxic effects of sporidesmin. It is not yet clear how sporidesmin and other ETP compounds mediate cytotoxicity (reviewed in Chai and Waring, 2000), although there is evidence that the oxidised disulphide ETP compounds (Scheme 1) can interact covalently with cellular macromolecules, such as proteins, and alter their functions (Rodriguez and Carrasco, 1992; Hara and Han, 1995; Waring et al., 1995; Pahl et al., 1996; Hurne et al., 2000). Additionally, ETP compounds are also able to promote intracellular redox cycling and generate toxic reactive oxygen species (Munday, 1982, 1987, 1984a). Zinc has been shown to form a tetrahedral complex with sporidesmin in vitro (Woodcock et al., 2001a,b); if this was to occur intracellularly, the formation of such a complex would abolish the reactivity of sporidesmin and limit its ability to interact with cellular macromolecules or to generate reactive oxygen species.
The experiments described here using the HepG2 cells were not designed to determine the relative importance of apoptosis or necrosis in sporidesmin cytotoxicity, as the LDH assay is a measure of necrosis and late apoptosis (O’Brien et al., 2000). However, it is known that both gliotoxin and sporidesmin induce apoptosis in mouse macrophages by an unknown mechanism (Waring et al., 1990) and in LLC-PK1 cells in a caspase dependent manner (Zhou et al., 2000). Zinc is a potent inhibitor of apoptosis, capable of inhibiting multiple steps in the pro-apoptotic pathway including caspase activation (Stenmicke and Salvesen, 1997; Segal and Beem, 2001; Wei et al., 2004). Bax translocation and cytochrome c release (Ganjoo and Eastman, 2003; Wei et al., 2004). Apoptosis initiated by the ligand mediated signalling is also suppressed by zinc (Lambert et al., 2003). Therefore, zinc protection of HepG2 cells from sporidesmin toxicity may be mediated through the inhibition of apoptosis.

Acknowledgements

E.J. Duncan’s PhD fellowship was supported by a MeatNZ grant to S.H. Phua. The authors thank the AgResearch Toxinology Group (Ruakura Research Centre, Hamilton, New Zealand) for supplying the purified sporidesmin, the Cancer Genetics Laboratory (University of Otago, Dunedin, New Zealand) for providing the HepG2 cells, and Ken Dodds (AgResearch, Invermay Agricultural Centre, Dunedin, New Zealand) for statistical advice.

References


INTRODUCTION
Facial eczema (FE) is a liver disease caused by the mycotoxin sporidesmin. The disease affects sheep, cattle and deer, and costs the New Zealand livestock industries NZ$63-126M annually. Investigations into the genetic bases of FE resistance in sheep have included a genome QTL scan (Phua et al., 1998a) and candidate genes approach (Phua et al., 1998b). This report describes an alternative, but complementary, method using microarray technology to examine gene expression profiles in livers of FE resistant- and susceptible-line sheep in response to sporidesmin challenge.

MATERIALS AND METHODS
Microarray and probes. A joint bovine EST project by Genesis company and AgResearch generated about 5500 liver cDNA clones, of which 1500 were selected, PCR amplified and arrayed onto polylysine-coated slides. One susceptible- and two resistant-line sheep were challenged with sporidesmin (0.15 and 0.4 mg/kg liveweight, respectively) and their livers collected 24 hours post dosing. In probe preparation, total RNA was extracted from a liver using TriZol reagent (Life Technologies), and further purified using RNAeasy columns (Qiagen). Fluorescent labelling (with either Cy3- or Cy5-dCTP) of the RNA sample was performed as in Hedge et al. (2000).

Slide hybridisation. Probes from a resistant and a susceptible animals were pooled and hybridised to the arrayed slides. Following hybridisation and washing, the slides were scanned with a ScanArray 5000 (Packard) scanner and analysed with QuantArray software (Packard).

Microarray analysis. The method of Callow et al. (2000) was used for statistical analysis. Data was log2 transformed and normalised using the housekeeping gene, β-2 microglobulin. Each gene was then considered a univariate testing problem and student t-statistics were calculated; multiple testing was corrected for using adjusted p-values.

Cluster analysis. Expression data was clustered using two iterations of the J-Express clustering algorithm (http://www.ii.uib.no/~bjarted/jexpress/). The first used principal components analysis (PCA) to determine the major axes of variation within the data. The ESTs within each axis were then subjected to hierarchical clustering. Genes in each cluster were examined for trends in gene function and regulation.
Quantitative RT-PCR. The semi-quantitative method of Hoen et al. (2000) was used. It involved PCR amplification of the cDNA of interest multiplexed with 18S rRNA with $^{32}$P end-labelled primers. The template cDNAs were generated from reverse transcription of total RNA with random hexamer primers. The linear phase of PCR amplification for both 18S rRNA and fibrinogen-γ were first determined, and different samples were then tested at a cycle which fell within the determined linear amplification range.

RESULTS AND DISCUSSION

Microarray result. Four replicate experiments, against the same susceptible sheep reference, were performed: three were with one resistant sheep and the fourth was with a different resistant animal. A scanned slide with pseudo-coloured images is shown in Fig. 1A. Fig. 1B is an enlarged section showing differential expression of some genes: the red spots indicate over-expression in resistant animal, the green spots were over-expressed genes in susceptible animal, and the yellow dots indicate equal expression in both animal types.

The distribution plot of log$_2$ expression ratios against ordered EST ID is shown in Fig. 2: this gives an overall view of trends in the expression profiles. There are a number of genes in the resistant animals which have more than 1.5-fold higher expression than the susceptible animal.

Figure 1. A: A microarray image generated by QuantArray. B: An image enlargement showing differentially expressed genes

Figure 2. Distribution plot of global expression patterns across all four microarray experiments. Normalisation is to β-2 microglobulin

The microarray data indicated differential expression of 133 genes between resistant and susceptible sheep livers. Twenty-four of these were differentially expressed at the 99% confidence interval, thirty-one at the 97.5% and seventy-eight at the 95% confidence level. Of the 133 differentially expressed genes, 24% were of unknown function: that is the ESTs have no hit to either known protein or nucleotide sequences. This highlights an advantage of using microarray technology as it allows rapid screening of genes with unknown function for involvement in the disease trait.
**Cluster analysis.** Cluster analysis was used in an attempt to place the observed differential expression patterns into a biological context; the underlying assumption of this analysis was that genes with similar expression levels share commonality in regulation. The J-Express clustering algorithms using principal components analysis indicated four major axes of variation in the microarray data. The four axes were further analysed using the hierarchical clustering algorithm, which gave ten sub-clusters of genes (Fig. 3). Examination of the clustered data revealed weak trends as shown in Fig. 3. The general theme across all clusters indicates the differential expression of several immune mediated and cell cycle regulated genes. This is the first time that these genes have been implicated in sheep resistance to FE disease.

![Cluster analysis diagram](attachment:cluster_analysis.png)

**Figure 3. Summary of clustering result.** The four main clusters were defined by PCA analysis, the smaller clusters by hierarchical clustering.

Several functionally important classes of genes can be identified. One of these classes was of antioxidant enzymes, including catalase. Interestingly, the catalase gene has been previously found to be associated with FE resistance in sheep (Phua et al., 1999).

**Quantitative RT-PCR.** Confirmation of the microarray results is essential prior to candidate gene selection. Fig. 4 shows the multiplexed PCR results used to determine the linear amplification phases of fibrinogen-γ and 18S rRNA genes, and the densitometry graph is plotted in Fig. 5. The exponential phases of the two genes overlap at cycle 12; thus 12 cycles of PCR was used for the measurement of fibrinogen-γ in all further liver samples. Fig. 6 shows that fibrinogen-γ is significantly over expressed in FE resistant animals irrespective of exposure to sporidesmin.

![PCR results](attachment:pcr_results.png)

**Figure 4. Autoradiograph of multiplexed PCR products**

Session 13. Disease resistance
CONCLUSION
Using microarray technology, at least twenty-four genes were detected to be differentially expressed between livers of FE resistant and susceptible sheep in response to sporidesmin challenge. Fibrinogen-γ gene was independently validated, using a semi-quantitative RT-PCR technique, to be over-expressed in both dosed and undosed FE-resistant sheep.

REFERENCES