Comparative gene expression studies of anthelmintic resistance in the parasitic nematode, *Teladorsagia circumcincta*

Alison Jane Dicker

BSc (Hons), MSc (Distinction)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy to the Faculty of Veterinary Medicine of the University of Glasgow

Research conducted at Moredun Research Institute, Edinburgh

July 2010

Abstract

Anthelmintic resistance in parasitic nematodes of small ruminants is widespread and, in some parts of the world, threatens the sustainability of sheep production. The mechanisms whereby parasitic nematodes become resistant to anthelmintics, particularly ivermectin, remain to be determined. The majority of studies to date have investigated target site mutations; relatively little attention has been paid to the role of gene expression changes. The present study focused on *Teladorsagia circumcincta*; the predominant parasitic gastrointestinal nematode species in the UK and the predominant resistant species. The role of changes in gene expression were investigated in an ivermectin-susceptible isolate (CVL) and a multidrug resistant isolate (MOTRI), utilising a range of molecular biological techniques.

In the first experiment, a panel of novel putative ivermectin resistance genes were identified from *T. circumcincta*, comprising 11 partial P-glycoprotein (Pgp) and 3 partial Cytochrome P450 (CYP) sequences. Both Pgps and CYPs have been implicated in the handling and metabolism of xenobiotics in other biological systems, but have not been investigated in *T. circumcincta* to date. Initial results, using semi-quantitative PCR identified changes in expression of this panel of genes between the CVL and MOTRI isolates.

Constitutive differences in expression of the Pgps and CYPs between CVL and MOTRI were determined using the ΔΔCt TaqMan® real-time PCR method. A statistically significant increase in expression was observed for *TeciPgp-9* NBD2 across all life-cycle stages but most notably in eggs (55-fold increase). A statistically significant reduction in expression of *TeciPgp-2* NBD2 was observed in all but the adult stages of MOTRI compared to CVL. Analysis of a 208 base pair sequence of *TeciPgp-9* NBD2 identified high levels of polymorphism, with at least four non-coding SNPs evident in the MOTRI isolate. These results merit further investigation.

Inducible changes in the expression of the Pgps and CYPs were investigated in MOTRI before and after ivermectin treatment, using real-time PCR. Statistically significant fold changes in expression in most of the genes occurred in at least one life-cycle stage. Inducible expression of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 was investigated further by comparing adult MOTRI parasites with those recovered three days after *in vivo* ivermectin exposure, and by exposing pools of MOTRI xL_3 to ivermectin in the larval migration inhibition test. The survivors of ivermectin exposure exhibited a statistically significant reduced 13.68-fold expression of *TeciPgp-2* NBD2 compared to MOTRI. Similarly, the MOTRI xL_3 able to migrate in the presence of ivermectin in the LMIT had a 1.88-fold reduction in *TeciPgp-2* NBD2 expression compared to MOTRI xL_3 unexposed to ivermectin. These results indicate that inducible changes in *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 expression can occur, but the experimental design is critical to being able to identify the changes.

In a more global approach, the transcriptomic response of MOTRI adults to *in vitro* ivermectin exposure was investigated using Roche 454 sequencing, generating 98,685 novel EST sequences, providing an important resource for a genome resource-poor organism. Objective bioinformatic analysis of the two datasets revealed statistically significant differences in the mean expression levels of the KEGG orthologous groups for 'translation', 'amino acid metabolism' 'carbohydrate metabolism' and 'xenobiotic degradation and metabolism'. On combining the two datasets, and through application of a novel statistical method, 16 clusters of ESTs were identified as containing statistically significant differences in the mean proportion of exposed reads compared to unexposed reads under the conservative model, whilst a further 355 clusters were found to have statistically significant differences under the liberal model.

One-way suppression subtractive hybridisation (SSH) was used to identify genes exhibiting increased expression in MOTRI adults compared to CVL adults. 28 contiguous sequences were identified from the SSH experiment; 6 contiguous sequences were selected for validation; 5 of these results were confirmed using semi-quantitative PCR. Each contig was BLAST searched against the Roche 454 dataset; contig SSH14 aligned most closely to one of the statistically significant clusters in the conservative model, SSHs 5, 6, 10 and 23

aligned most closely to statistically significant clusters in the liberal model. This suggests that changes in expression in these sequences occur both constitutively, between CVL and MOTRI isolates, and inducibly, following ivermectin exposure.

This work has shown that changes in gene expression, particularly the constitutively reduced expression in *TeciPgp-2* NBD2 and the constitutively increased expression in *TeciPgp-9* NBD2 (coupled with the presence of SNPs) could play a role in allowing multidrug resistant *T. circumcincta* to survive ivermectin exposure. Roche 454 sequencing and SSH approaches identified gene expression changes associated with *in vitro* ivermectin exposure and ivermectin resistance. These could form the basis of a novel panel of candidate resistance genes whose altered expression profiles may allow multidrug resistant *T. circumcincta* to survive ivermectin exposure by some, as yet identified, mechanism. Finally, we have also shown that a multidrug *T. circumcincta* isolate is affected by ivermectin exposure and that changes in gene expression could have a role to play in the ivermectin resistance phenotype in *T. circumcincta*. The genetic changes underpinning these changes in gene expression remain to be elucidated, and need to be investigated in other isolates. These changes could form the basis of an ivermectin resistance molecular marker, to monitor the spread of resistance, and to evaluate management practices aimed at delaying its spread.

Author's declaration

I declare that the work presented in this thesis is my own original work, except where otherwise stated, and has not been submitted for any other degree or professional qualification

Alison Jane Dicker

July 2010

Acknowledgements

I wish to thank the devolved bodies (Quality Meat Scotland, English Beef and Lamb Executive and Meat Promotion Wales) for funding my PhD studentship and Glasgow University and the Moredun Research Institute for enabling me to carry out my research. I am grateful to my supervisors, Dr Philip Skuce, Dr Alasdair Nisbet, Prof John Gilleard and Dr Collette Britton for the advice and support they have provided. I hope I wasn't that high-maintenance! Help and advice in bioinformatics and statistics was provided by Raja Yaga, Alex Lainson, Mintu Nath and Claus Mayer. The Parasitology department (both the upstairs and "big Parasitology" lab crowds –too numerous to name individually) have helped with various parts of this work and have also made me feel very welcome: I promise that this is the year when I don't have to leave early from the Christmas lunch. To all my friends in the PhD office and elsewhere in the institute, past and present; thanks for keeping me (slightly) sane. Thanks to Lynne Murrie, for letting me have walks and cuddles with Mij and Tess, for talking a lot of sense, and for being an awesome friend. I have made many friends white water kayaking, playing canoe polo and mountain biking in Scotland who have helped me prove that it is possible to work hard and play harder; here's to many more years of fun! Claire & Kieran, Lizzie & Steve, Madeleine & Al; one of these days I'll have a place where you can come and stay… Finally, but by no means least, I am indebted to my Mum & Dad, to my brothers and their wives; Simon & Anne, Tim & Julie, Nick & Sarah and to my nieces and nephew; Rachel, Laura & Ben, for all their love and support, and to Millie and Charlie for cuddles when in England.

"True understanding of how genes function requires knowledge of their expression patterns, their impact on all other genes and their effects on DNA structure and modifications"

A. Kahvejian, J. Quackenbush & J. Thompson (2008). What would you do if you could sequence everything? *Nature Biotechnology*, **26** (10) 1125-1133

Table of Contents

List of Tables

Chapter 2 Tables

Chapter 3 Tables

Chapter 4 Tables

Chapter 5 Tables

Chapter 6 Tables

List of Figures

Chapter 1 Figures

Chapter 2 Figures

Chapter 3 Figures

Chapter 4 Figures

Chapter 5 Figures

Chapter 6 Figures

Abbreviations

µg	micrograms
µL	microlitres
µm	micrometres
µM	micromolar
°C	degrees Celsius
AAD	amino-acetonitrile derivatives
ABC	ATP binding cassette
ABZ	albendazole
Ala	alanine
AMP	ampicillin
ASP	activation associated secreted proteins
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
bp	base pairs
BZ	benzimidazole
cDNA	complementary deoxyribonucleic acid
cDNA-AFLP	cDNA-amplified fragment length polymorphism
CET	controlled efficacy test
cm	centimetres
CO_2	carbon dioxide
CVL	*T. circumcincta* anthelmintic susceptible isolate
CYP	cytochrome P450
DDT	dichlorodiphenyltrichloroethane
dH_2O	distilled water
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide
DTT	dithiothreitol
ED_{50}	dose of drug which is effective against 50% of those exposed
EDTA	ethylenediaminetetraacetic acid
EHT	egg hatch test

E/S	excretory/secretory
EST	expressed sequence tag
FAMACHA	Faffa Malan chart
FBZ	fenbendazole
FDR	false discovery rate
FECRT	faecal egg count reduction test
FITC	fluorescein isothiocyanate
g	grams
GABA	γ-aminobutyric acid receptor
GI	gastrointestinal
GLM	generalised linear model
Gln	glutamine
Glu	glutamic acid
GluCl	glutamate-gated chloride channel
Gly	glycine
H_2O	water
HSP	heat shock protein
IgA	immunoglobulin A
IgE	immunoglobulin E
IL	interleukin
IPTG	isopropyl-1-thio-β-D-galactopyranoside
IVM	ivermectin
kb	kilo bases
kDa	kilo Dalton
kg	kilogram
L_1	first larval stage
L_2	second larval stage
L_3	third larval stage
L_4	fourth larval stage
L	litres
LB	Luria Bertani
LD_{50}	dose of drug at which 50% are killed
LDT	larval development test

LEV	levamisole
LFIT	larval feeding inhibition test
LMIT	larval migration inhibition test
M	molar
MALDT	micro agar larval development test
mb	mega bases
MDR	multidrug resistant
mg	milligram
$MgCl_2$	magnesium chloride
mL	millilitre
ML	macrocyclic lactone
mm	millimetre
mins	minutes
MOTRI	Moredun triple resistant *T. circumcincta* isolate
MOX	moxidectin
mRNA	messenger RNA
nAChR	nicotinic acetylcholine receptors
NBD	nucleotide binding domain
NJ	neighbour-joining
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PCR-RFLP	PCR- restriction fragment length polymorphism
PGE	parasitic gastroenteritis
Pgp	P-glycoprotein
pH	-log10 (hydrogen ion concentration)
Phe	phenylalanine
Post IVM MOTRI	MOTRI parasites which survived an *in vivo* therapeutic dose of IVM
PPR	peri-parturient rise
PYR	pyrantel
RACE PCR	rapid amplification of cDNA ends PCR
Rf	reading frame
RNA	ribonucleic acid
RNAi	ribonucleic acid interference

SCOPS	sustainable control of parasites in sheep
secs	seconds
SMART™	switching mechanism at 5’ end of RNA template
SNP	single nucleotide polymorphism
spp	species
SSH	suppression subtractive hybridisation
TAE	tris-acetic acid-EDTA buffer
TBZ	thiabendazole
Tm	melting temperature
TST	targeted selective treatment
Tyr	tyrosine
UK	United Kingdom
UPGMA	un-weighted pair group method with arithmetic mean
UV	ultra violet
V	volts
WAAVP	World Association for the Advancement of Veterinary Parasitology
x g	gravitational force
X-gal	5-bromo-4-chloro-3-inodyl-β-D-galactopyranoside
xL_3	exsheathed L_3

Chapter 1: General Introduction

1.1: Parasitic gastroenteritis (PGE)

Parasitic gastroenteritis (PGE) is caused by the presence of large numbers of gastrointestinal (GI) nematode parasites in the stomach and intestines of ruminants, leading to reduced liveweight gain, anaemia, and in some cases, death. Clinical signs of PGE include diarrhoea, lethargy and, in more severe cases, anaemia and submandibular oedema (Urquhart *et al.*, 1996; Sargison, Jackson, & Scott, 2002; Nieuwhof & Bishop, 2005; Taylor, Coop, & Wall, 2007). There is a range of costs associated with PGE in sheep, both for the farmers and the animals, such as loss of efficient production of meat, milk and wool and a reduction in the animals' welfare level (Boyne, Stott, & Gunn, 2006; Broughan & Wall, 2007). It is difficult to accurately quantify the economic costs of PGE as the effects of nematode infection can cause either clinical or sub-clinical disease, yet parasitism is seen as the single most important cause of production losses in small ruminants around the world (Molento, 2009). The cost of GI nematode infection in the UK sheep industry was estimated by Nieuwhof & Bishop (2005) to be £84 million per annum, with the range in estimates being £48M to £120M. This is calculated as the cost to the industry due to reduced lamb growth rates plus treatment and prevention measures. It has also been estimated that ovine nematodes inflict the greatest net cost to the Australian grazing industry (McLeod, 1995).

PGE in sheep is caused by individual or concurrent infections with the nematode parasites *Teladorsagia circumcincta*, *Haemonchus contortus*, *Trichostrongylus* spp, *Nematodirus* spp and *Cooperia* spp. In cool temperate regions, the principal cause of PGE is *T. circumcincta* whilst in tropical and subtropical regions, *H. contortus* is more important (Sargison, Jackson, & Scott, 2002; Sargison *et al.*, 2005; Stear *et al.*, 2006; Taylor, Coop, & Wall, 2007). PGE caused by *T. circumcincta* (teladorsagiosis) in sheep results in a hyperplastic gastritis which can be loosely described as two different types (analogous to type I and type II bovine ostertagiosis caused by *Ostertagia ostertagi*) (Urquhart *et al.*, 1996; Sutherland & Scott, 2010). Type I is found in lambs in their first grazing season usually between August and October, and is due to the ingestion and maturation of large numbers of infective larvae from the pasture. The rarer type II occurs during late winter and early spring and is caused by the delayed maturation of worms

which had been ingested the previous autumn and entered a period of arrested development in the host (Gibbs, 1986; Urquhart *et al.*, 1996). Alternatively, larvae that have overwintered on pasture may also cause type II teladorsagiosis (Sargison, Jackson, & Scott, 2002). Type I is a more acute disease whilst type II was generally seen as a sub-clinical disease leading to production losses (Abbott, Taylor, & Stubbings, 2004).

Surveillance reports have suggested that cases of PGE in sheep are increasing in Scotland, and are mostly due to *T. circumcincta* infection. It is thought that this is possibly due to a combination of climate change, an increase in anthelmintic resistance and extensive animal movement, especially post Foot and Mouth Disease (http://www.defra.gov.uk/vla/reports/rep_surv.htm). In South East Scotland there has been a 20% increase in annual rainfall and a 1ºC increase in average monthly temperatures between 1961 and 2004 which has led to an extended herbage growing i.e. grazing season of 4 weeks (Barnett *et al.*, 2006). This means that conditions for survival and development of the larval stages of nematodes such as *T. circumcincta* in the environment are becoming more favourable. Type II teladorsagiosis is now becoming more common in South East Scotland than in previous years; PGE outbreaks also appear to be becoming more severe (Sargison *et al.*, 2007a; Kenyon *et al.*, 2009a).

1.1.1: Immunity to GI nematodes

PGE is principally a disease of young animals. Regular exposure to moderate burdens of parasites allows the lamb to develop an immune response which prevents the parasite from establishing in the host in large numbers (Colditz *et al.*, 1996; Stear *et al.*, 2009). Repeated ingestion of larvae by lambs over a period of two to four months from the age of four to five months gradually allows a protective immunity to develop (Abbott, Taylor, & Stubbings, 2004). Immunity to *T. circumcincta* takes two grazing seasons to become fully established and adult sheep will generally only harbour a few adult worms (Urquhart *et al.*, 1996; Taylor, Coop, & Wall, 2007), suggesting the immunity which develops is not a completely sterile immunity (Abbott, Taylor, & Stubbings, 2004; Sutherland & Scott, 2010). The level of immunity appears to vary with age, the genotype

of the sheep and also the reproductive status of the ewe (McNeilly, Devaney, & Matthews, 2009).

GI nematode infection has been associated with an increase in IgE and peripheral and mucosal eosinophils (Miller, 1996; Huntley *et al.*, 1998a; Huntley *et al.*, 1998b; Sutherland & Scott, 2010). Infection with GI nematodes causes a specific humoral immune response leading to the production of the Th2 subset of CD4+ helper T cells characterised by the expression of the cytokines IL-4, -5, -10 and -13 (Miller, 1996; McNeilly, Devaney, & Matthews, 2009; Sutherland & Scott, 2010). This suggests that the worm burden is regulated, in part, in immune sheep by an immediate hypersensitivity reaction (Stear *et al.*, 1995; Urquhart *et al.*, 1996; McNeilly, Devaney, & Matthews, 2009). In abomasal lymph node cells from immune sheep there is a greater predominance of mRNA transcripts encoding the Th2 associated cytokines, IL-4, -5 and -13 compared to parasite naïve sheep (McNeilly, Devaney, & Matthews, 2009). A negative correlation between anti-larval excretory/ secretory specific mucus IgA levels and the number of GI nematodes which establish in the abomasum has been shown (Smith *et al.*, 2009). Increased mucosal IgA levels, in both natural and experimental infections, has been shown to negatively correlate with adult worm length, adult worm burden and faecal egg count (Stear *et al.*, 1995; Miller, 1996; Beraldi *et al.*, 2008; Stear *et al.*, 2009; McNeilly, Devaney, & Matthews, 2009). Overall, the immune response to GI nematode infection appears to affect the parasites in a number of ways: preventing the establishment of incoming infective third larval stages (L_3s); slowing or arresting larval development within the wall of the abomasum; affecting the ability for adult worms to grow and reproduce to their full potential and increasing the rate at which adult parasites are killed or expelled from the abomasum (Taylor, Coop, & Wall, 2007; McNeilly, Devaney, & Matthews, 2009; Sutherland & Scott, 2010).

1.2: *T. circumcincta*

T. circumcincta is a nematode parasite of small ruminants and is classified in the Order Strongylida, Superfamily Trichostrongyloidea and Family Trichostrongylidae alongside the other sheep nematode parasites; *H. contortus*, *Trichostrongylus* spp,

Nematodirus spp and *Cooperia* spp (Taylor, Coop, & Wall, 2007). *T. circumcincta* was previously classified as *Ostertagia circumcincta* but is also known by its common name, the brown stomach worm. *T. circumcincta* has a worldwide distribution (Taylor, Coop, & Wall, 2007) and in temperate regions is the predominant nematode species parasitizing sheep (Sargison, Jackson, & Scott, 2002; Bartley *et al.*, 2003; Sargison *et al.*, 2005; Broughan & Wall, 2007). *T. circumcincta* is a bursate nematode with the adults appearing as red-brown slender worms up to 10mm long in the mucosa on the surface of the abomasum (Urquhart *et al.*, 1996). *T. circumcincta* can be distinguished from other species by morphological features such as absence of proconus, bursal ray pattern and, in adult males, by the spicules at the posterior end of the worm (Denham, 1969; Lichtenfels & Hoberg, 1993; Taylor, Coop, & Wall, 2007).

1.2.1: *T. circumcincta* life-cycle

The direct life-cycle of *T. circumcincta*, as shown in Figure 1.1, is typical of trichostrongylid nematodes. Eggs are laid by female worms living on the surface of the abomasum in the definitive sheep host and are subsequently excreted onto the pasture in the faeces. Eggs develop in the faecal pat to the first larval stage (L_1) then moult to the second and third larval stages (L_2 and L_3) in approximately two weeks under optimal conditions. Ideal conditions for parasite development are 18-26°C and a relative humidity of approximately 60% (Gruner & Suryahadi, 1993; Urquhart *et al.*, 1996; O'Connor, Walkden-Brown, & Kahn, 2006). *T. circumcincta* is more tolerant of low temperatures than *H. contortus* and can complete development from the egg to L_3 stage between 1°C and 35°C. The increasing average monthly temperatures recorded between 1961 and 2004 means that there is now a longer period of the year where temperatures exceed 10°C, giving the parasites a longer period to develop through to the infective stage (Kenyon *et al.*, 2009a; van Dijk *et al.*, 2010). Once at the L_3 stage, the parasite is less susceptible to changes in climatic conditions due to the retained L_2 sheath and can over-winter on the pasture for up to two years and still be infective (Kerboeuf, 1985; O'Connor, Walkden-Brown, & Kahn, 2006; Stenhouse, 2007). The L_3 migrate from the faecal pat to the pasture where they are ingested by the host. Rainfall increases the infectivity of the pasture by allowing easier migration of L_3 from the faecal pat to the pasture and by reducing the chance of desiccation of the larval stages. However, immersion of the eggs in water is

detrimental to their viability (Gruner & Suryahadi, 1993; Abbott, Taylor, & Stubbings, 2004).

After ingestion, the L_3 exsheath in the rumen of the sheep and migrate to the abomasal glands where they can be found three to four days post infection (Michel, 1974). Once established, the parasite develops through the L_4 stage to the final adult stage. It is not possible to distinguish between male and female worms until the L_4 stage is reached (Denham, 1969). On emergence from the abomasal glands, the adult parasites become sexually mature on the mucosal surface. Emergence of parasites from the glands causes cytolysis and sloughing of the epithelium of the abomasum, which causes the normal secretory cells of the gastric glands to be replaced by hyperplastic cells. This results in the loss of the specialised secretory function and junctional integrity of the cells (Scott *et al.*, 1998), leading to leakage of proteins into the GI tract, resulting in hypoalbuminaemia and increased plasma pepsin levels (McKellar, 1993; Sutherland & Scott, 2010). The time taken from ingestion of L_3 to the parasites producing eggs, the pre-patent period, is around 18 days. The entire life-cycle of *T. circumcincta* can be completed in as little as three weeks (Urquhart *et al.*, 1996).

Once in the abomasal glands, *T. circumcincta* can undergo a process known as hypobiosis, or arrested larval development. This is defined as the temporary cessation of development, occurring in certain circumstances and times of the year until more suitable environmental conditions prevail. Hypobiosis often only affects a proportion of the worm population (Michel, 1974; Urquhart *et al.*, 1996). In *T. circumcincta*, a drop in temperature in the autumn coincides with an increase in the number of hypobiotic larvae in the abomasal gland of the sheep. The larvae develop to the early L_4 stage in the abomasal glands and then remain at that point for up to three months, or possibly longer (Michel, 1974; Urquhart *et al.*, 1996). The synchronised development of these hypobiotic larvae in the spring is a major cause of type II teladorsagiosis (Urquhart *et al.*, 1996). The numbers of larvae undergoing arrested/ inhibited development also seems to be correlated with the size of the local IgA immune response of the host and the size of the worm burden within the host (Stear *et al.*, 1995). Hypobiosis may also play a part in the peri-parturient rise (PPR) seen in ewes, when an increase in nematode eggs excreted in the faeces occurs

around lambing time. Three mechanisms appear to combine to cause the PPR; a temporary relaxation in the immunity to the parasites concurrent with increased prolactin levels in the ewe allows hypobiotic larvae to develop; a reduced loss of adult worms and an increased establishment of new worms; and an increased fecundity of the existing adult worms (Michel, 1974; Urquhart *et al.*, 1996; Taylor, Coop, & Wall, 2007). The eggs produced during the PPR are the source of virtually all the parasites acquired by the lambs; it is these parasites which cause type 1 PGE (Heath & Michel, 1969; Gibson, 1973). The PPR usually occurs from two weeks before lambing to six weeks post-lambing (Urquhart *et al.*, 1996) and ensures the pasture is heavily contaminated with eggs, and hence infective L_3s, at the point when the lambs are most susceptible to infection.

1.3: Control of parasitic gastroenteritis

Control of gastrointestinal infections is attempted through various means but the principal method of control of PGE in sheep flocks has been the regular strategic use of anthelmintics to prevent disease in lambs which have not yet developed natural immunity to the parasites (Abbott, Taylor, & Stubbings, 2004; Prichard *et al.*, 2007). The aim of anthelmintic treatment is to prevent clinical or sub-clinical disease, and hence some of the costs associated with PGE, whilst allowing the development of natural immunity (Urquhart *et al.*, 1996; Molento, 2009). An increase in anthelmintic resistance, together with the growing popularity of organically produced food and concerns about drug residues entering the food chain has generated interest in control strategies that do not solely rely on anthelmintics. However, none of these alternative control strategies discussed below offer an effective alternative to anthelmintics on their own (Wolstenholme *et al.*, 2004). Another method of controlling PGE, which has been the focus of much research, is the development of vaccines that target antigens excreted or secreted by the parasites, or molecules localised, or hidden, within the parasites. These alternative control strategies, including vaccination, will be discussed first followed by the anthelmintics.

1.3.1: Alternative control strategies

Other methods of control include breeding to increase host resilience to parasite infection. Resilience is the ability of the host to remain productive whilst infected with GI

nematodes. Certain breeds of sheep and individuals within a flock appear to be more able to sustain growth whilst infected with parasites. The majority of sheep in a flock will harbour few parasites whilst the minority will harbour the majority of the parasite population, this is known as over-dispersion (Stear *et al.*, 2006; Mitreva *et al.*, 2007; Stafford, Morgan, & Coles, 2009). Selective breeding for animals more resilient to parasite infection (i.e. those able to cope better with the worm burden) or animals relatively resistant to parasite infection (i.e. those that harbour fewer parasites than relatively susceptible animals) could be used to reduce the impact of GI nematodes (Stear *et al.*, 2000). However, this approach is slow and heritability values for traits associated with resilience are low and usually come at the expense of more desirable production traits such as growth, meat and wool quality (Bisset & Morris, 1996; Woolaston & Baker, 1996; Waller & Thamsborg, 2004; Stear, Doligalska, & Donskow-Schmelter, 2007).

Nutrition plays a role in the impact of GI nematodes on an animal's health status. Amino acids are required by the sheep for maintenance of condition, development of immunity, growth and reproduction. There is a loss of protein in the gut due to the damage to the epithelium caused by the parasites' emergence and a resulting reduction in the ability of the sheep to utilize metabolizable energy. This will, in turn, result in competition for resources between the different body systems (Sykes & Coop, 2001; Waller & Thamsborg, 2004). There is evidence that dietary supplementation with urea enhances resistance to *T. circumcincta* as measured by a lower establishment rate of worms in the abomasum (Stear *et al.*, 2000). Lambs will select for a high protein diet when infected with nematode parasites but those on a diet supplemented with urea have been shown to eat a higher proportion of the food offered (Stear *et al.*, 2000). Niezen *et al* (1998) and Heckendorn *et al* (2007) found that feeding plants containing condensed tannins resulted in reduced levels of GI nematode infection whilst also improving the performance of sheep. Other forages, not containing condensed tannins, also have an effect on performance levels under gastrointestinal nematode challenge. Lambs grazing chicory compared to a grass/ clover mix have been shown to have lower faecal egg counts and a faster growth rate (Waller & Thamsborg, 2004; Kidane *et al.*, 2010). There is increasing interest in the use of plants with natural anthelmintic properties, particularly for organic producers or in areas where there is not sufficient and reliable access to anthelmintics; studies are ongoing to identify

which plants and plant extracts may be of use. However trials to date on plants, such as *Albizia anthelmintica*, have not been that successful (Grade *et al.*, 2008).

Trials have been carried out on several nematophagous fungi but principally on *Duddingtonia flagrans*, which is fed, as a supplement to the host in the form of chlamydospores. The fungi are then excreted onto the pasture where they destroy the parasite larvae before they are able to migrate to the herbage, thus preventing infection of the host (Larsen, 1999). In principle this method has been shown to work but in field trials involving *D. flagrans* against naturally acquired infections in dairy ewes there was shown to be no significant parasitological benefit, although productivity was improved compared to control groups (Faessler, Torgerson, & Hertzberg, 2007).

Grazing management can be used to reduce the infectivity of the pastures and to slow the development of anthelmintic resistance. Suggested methods in the Sustainable Control of Parasites in Sheep (SCOPS) guidelines (Abbott, Taylor, & Stubbings, 2004), include grazing sheep and cattle together and grouping animals by age. Co-grazing sheep and goats is not recommended as the same parasites can infect both species. Other methods include reducing stocking densities, alternating pasture with crops to disrupt the transmission cycle and rotational grazing (Waller & Thamsborg, 2004; Stear, Doligalska, & Donskow-Schmelter, 2007).

1.3.2: Anti-parasite vaccines

There is a growing interest in the use of vaccines to control parasitic infections. Vaccines have some advantages over anthelmintics as they leave no chemical residues that would require withdrawal periods for food products and are also seen as more environmentally friendly as they are not toxic to dung and marine flora and fauna (Dalton & Mulcahy, 2001; Vercruysse *et al.*, 2004; Omura, 2008). The majority of commercially available anti-parasite vaccines are for parasitic protozoa and use live attenuated parasites (Dalton & Mulcahy, 2001; Smith & Zarlenga, 2006). There is only one commercially available nematode parasite vaccine, namely Bovilis® Huskvac (formerly known as

Dictol), for vaccinating calves against the lungworm *Dictyocaulus viviparus*. This uses irradiation-attenuated L_3s from donor calves and has been a financially viable vaccine, although there are some ethical concerns about the use of large numbers of donor animals (Vercruysse *et al.*, 2004). The majority of current nematode vaccine work concentrates on *H. contortus* (Sutherland & Scott, 2010).

Natural antigen candidates have been studied to determine whether the immunity that develops in sheep following GI nematode infection can be harnessed by exposing the sheep to antigens such as those excreted or secreted by the parasite (E/S products). The use of 15kDa and 24kDa proteins from E/S products from adult *H. contortus*, and thiol-binding fractions from *H. contortus* E/S were both found to give strong protective immune responses to challenge infections with *H. contortus* following vaccination (Smith & Zarlenga, 2006). However, large numbers of donor animals are required to obtain sufficient quantities of worms to provide E/S material, so attempts have been made to make recombinant versions (Smith & Zarlenga, 2006). The protective immune responses to native antigens have been difficult to replicate with recombinant antigens due to the complex nature of the molecules involved. Difficulties in replicating the correct conformation and post-translational modifications of the recombinant molecule in artificial expression systems like bacteria and yeast have proved to be too challenging (Smith & Zarlenga, 2006).

Current focus for vaccine discovery is on the use of "hidden antigens" that reside on the gut surface of the parasites and are not normally exposed to the host immune response (Knox & Smith, 2001). The hidden antigen approach is particularly effective against blood-feeding parasites such as the cattle tick, *Boophilus microplus* and in *H. contortus* (Willadsen *et al.*, 1995; Smith & Zarlenga, 2006). This approach aims to raise a circulating antibody response in the host which, once the antibodies are ingested by the parasite, interferes with the parasite's ability to feed normally, leading to starvation and expulsion. Molecules such as H11 from *H. contortus* have been identified as strongly immunogenic, giving greater than 90% reduction in worm burdens against challenge infection (Knox & Smith, 2001). Another candidate antigen from *H. contortus* is H-gal-GP, a complex of putative digestive enzymes extracted from a detergent-soluble, integral

membrane protein fraction. As H11 and H-gal-GP are hidden antigens, repeated doses of vaccine have to be given to maintain protection as the antigens are not immunologically recognised by the sheep following *H. contortus* infection (Smith, Van Wyk, & van Strijp, 2001). Recombinant versions of H11, expressed in bacterial and insect cell lines, and H-gal-GP expressed in bacterial, *Pichia* and insect cell lines, have, however, failed to elicit protection in animal trials (Smith & Zarlenga, 2006). Native homologues of H11 and H-gal-GP have had limited success in *T. circumcincta* vaccine trials to date as *T. circumcincta* is a mucosal browser so the parasites are not exposed to sufficiently high levels of the effector antibodies; *Tc*H11 and *Tc*H-gal-GP were shown to be protective against *H. contortus* proving their immunogenicity (Knox & Smith, 2001; Smith, Van Wyk, & van Strijp, 2001; Sutherland & Scott, 2010).

To compete against anthelmintics, a commercially viable vaccine would have to be developed with a broad spectrum of activity, giving protection against at least the three major parasites of sheep, namely *H. contortus*, *T. circumcincta* and *Trichostrongylus* spp (Dalton & Mulcahy, 2001). Modelling studies have suggested that a vaccine need only have 60% efficacy in 80% of the flock for substantial benefits in terms of reduced lamb mortality to be observed so a vaccine does not need to compete against anthelmintics in terms of efficacy (Barnes, Dobson, & Barger, 1995). Developing anti-parasite vaccines which are as efficacious as anthelmintics is proving to be problematic. This is due, in part, to the limited availability and complicated nature of the antigens required, difficulties in expressing these antigens in recombinant form, antigenic variation and diverse gene expression exhibited by the different life-cycle stages of GI nematodes (Vercruysse *et al.*, 2004). In the case of mucosal browsers such as *T. circumcincta*, it is considerably more difficult to deliver antibodies effectively as the parasites are not ingesting serum antibodies in sufficient quantities to have the desired effect (Vercruysse *et al.*, 2004; Sutherland & Scott, 2010). Also, delivering the antigens to surfaces such as the gut mucosa in sufficient quantities to elicit a mucosal immune response is proving to be difficult (Knox & Smith, 2001).

1.3.3: Anthelmintics

Anthelmintics are the cornerstone of modern GI nematode control. Until recently, there have been three main chemical families of broad spectrum anthelmintics commercially available to treat GI nematode infections, namely; the benzimidazoles (BZs); levamisole and other imidazothiazoles (LEV); and the macrocyclic lactones (MLs). These are also known, respectively, as the "white drenches", "yellow drenches" and "clear drenches" (Abbott, Taylor, & Stubbings, 2004). A new class of drugs, the amino-acetonitrile derivatives (AADs) have recently been discovered (Kaminsky *et al.*, 2008). The AADs have been launched onto the small ruminant market in New Zealand and more recently in the UK. Each group of anthelmintics appears to have a different mode of action, as discussed below.

The BZs, such as albendazole (ABZ), thiabendazole (TBZ) and fenbendazole (FBZ), were introduced in the 1960s. BZs are effective against a broad range of parasites and also have wide safety margins, working at dosages of mg/kg bodyweight (McKellar & Jackson, 2004). Their mode of action appears to be mediated through binding to β-tubulin within the parasite, thus inhibiting the formation of microtubules that are central to the form and function of the parasite's cells. This prevents various essential cellular processes such as the transport of secretory granules and enzymes in the cell cytoplasm, resulting in cell lysis, with knock-on detrimental effects on motility and feeding (McKellar & Jackson, 2004; Mitreva *et al.*, 2007; von Samson-Himmelstjerna *et al.*, 2007).

A decade later, another broad spectrum anthelmintic, LEV, was released onto the market. Like the BZs, the dose is also in the mg/kg bodyweight range, however, LEV also acts on the host so care has to be taken with the dosage (McKellar & Jackson, 2004). LEV is a member of the imidazothiazole/ tetrahydropyrimidine class of anthelmintics, which also includes pyrantel (PYR) (McKellar & Jackson, 2004). LEV is a cholinergic agonist at the nicotinic neuromuscular junctions that works by first opening and then blocking the acetylcholine receptor-mediated cation channels. This causes a sustained neuromuscular depolarisation, leading to a rapid tonic paralysis of the parasite's somatic musculature, resulting in expulsion from the host (Dobson *et al.*, 1986; Martin *et al.*, 1998; Prichard,

2001; McKellar & Jackson, 2004; Rayes *et al.*, 2004; Kopp *et al.*, 2008; Prichard & Geary, 2008).

The ML, ivermectin (IVM) was isolated from *Streptomyces avermectinius* in 1974 and released onto the market in 1981, and is effective at doses of µg/kg bodyweight (McKellar & Jackson, 2004; Van Zeveren *et al.*, 2007a; Omura, 2008). IVM was the first commercially available endectocide, being effective against helminths, arachnids and insects (Geary, 2005; Omura, 2008). IVM was found to have high levels of potency, a broad spectrum of activity and good persistence, enabling novel application routes such as slow release devices (boluses), injectable formulations and topically applied (pour-on) IVM to be developed (Geary, 2005). When fully effective, IVM has an efficacy of almost 100%, as a result, it is now the largest selling anti-parasitic drug in livestock and has essentially revolutionised the animal health industry (McKellar & Jackson, 2004; Geary, 2005; Van Zeveren, 2009). IVM binds irreversibly to γ-aminobutyric acid- (GABA) and glutamate-gated chloride (GluCl) channels causing the hyperpolarisation and flaccid paralysis of pharyngeal and somatic muscle cells, leading to starvation and immobility of the worms (Martin *et al.*, 1998; Blackhall *et al.*, 1998a; Blackhall, Prichard, & Beech, 2003; Gilleard & Beech, 2007; Prichard & Roulet, 2007; Stenhouse, 2007; James & Davey, 2008). Another member of the ML family is moxidectin (MOX), which was also isolated from a species of *Streptomyces*. MOX appears to have a similar mode of action to IVM; the two ML subclasses, the avermectins (e.g. IVM) and milbemycins (e.g. MOX), have been shown to competitively displace each other from a GluCl channel in *C. elegans* (Ardelli *et al.*, 2009). MOX and IVM share the same central ML ring core, only differing at the side chains (Lespine *et al.*, 2007; Ardelli *et al.*, 2009). Compared to IVM, MOX is more lipophilic, allowing it to be absorbed faster and persist longer within the host (Lespine *et al.*, 2007; Lifschitz *et al.*, 2010).

The AADs are the first new anthelmintic class since the MLs and appear to have a novel mode of action, targeting nicotinic acetylcholine receptors that are unique to nematodes (Kaminsky *et al.*, 2008; Prichard & Geary, 2008). The AADs cause the hyper-contraction of nematode somatic muscle cells, leading to paralysis. They also cause the spasmodic contraction of the anterior portion of the pharynx. Both these effects ultimately

lead to the parasite's death (Kaminsky *et al.*, 2008). The AADs appear to have low toxicity against mammals whilst also being efficacious against both the larval and adult stages (Besier, 2009).

1.4: Anthelmintic resistance

Anthelmintic treatment applies a selection pressure on parasitic populations; an inevitable consequence of the regular prophylactic use of anthelmintics has been the parasites developing resistance. Anthelmintic resistance is defined as the ability of parasites to survive a dose of drug that would normally kill them; it is heritable and non-reversible (Barton, 1983; Jackson & Coop, 2000; Wolstenholme *et al.*, 2004). Phenotypic resistance can either be seen as a heritable decline in a drug's effectiveness or a reduction in the length of time that a drug is able to exert its effects on the parasites within the host (James, Hudson, & Davey, 2009). A parasite resistant to one anthelmintic in a drug class will usually be resistant to all anthelmintics within that class; this is known as side resistance (Kaplan *et al.*, 2007). The exception to this rule appears to be the ML, MOX, which appears to retain efficacy, for several treatment cycles, against parasites exhibiting IVM resistance, possibly due to longer persistency within the sheep (Sutherland *et al.*, 2002; Sutherland *et al.*, 2003; Lespine *et al.*, 2007; Wilson & Sargison, 2007; Lifschitz *et al.*, 2010). Drug resistance has developed rapidly to the three currently available classes of drugs, for example within three years of launch for IVM, and is a major threat to livestock production in many parts of the world (Van Wyk & Malan, 1988; Kaplan, 2004; Lespine *et al.*, 2008; Sutherland, Shaw, & Shaw, 2010). Anthelmintic resistance is a global phenomenon but is more prevalent in the Southern Hemisphere (Jackson & Coop, 2000; Bartley *et al.*, 2004; Kaplan, 2004). It has been estimated that carrying out a monthly preventative treatment regime, using a drug that the parasites are resistant to, results in the average price of the finished lamb being 14% lower compared to lambs which had been treated monthly with an effective drug (Sutherland, Shaw, & Shaw, 2010). In the absence of alternative control strategies, such as those discussed above, high levels of anthelmintic resistance may make it impossible to sustain economic sheep production on some farms, in some cases this has already occurred (Jackson & Coop, 2000; Sargison *et al.*, 2005).

In the UK, BZ resistant *T. circumcincta* were first reported in sheep in Cheshire in 1982 (Britt, 1982; Cawthorne & Whitehead, 1983), with the first report of LEV resistance in sheep in 1994. IVM resistance was only relatively recently identified in sheep in 2001 (Sargison, Scott, & Jackson, 2001), having previously been identified in goats in 1992 (Jackson, Jackson, & Coop, 1992). Resistance to IVM appears to be more common in small ruminants than in large ruminants, possibly due to different pharmacokinetics and the latter being able to mount a stronger immune response to the parasites and, therefore, requiring fewer anthelmintic treatments (Geary, 2005). The predominant nematode species exhibiting resistance in the UK appears to be *T. circumcincta* (Jackson & Coop, 2000; Bartley *et al.*, 2003). The true extent of anthelmintic resistance is not known (Coles, 2005), as few farmers routinely check the efficacy of the anthelmintics they use and resistance is often only diagnosed after the clinical failure of the drugs on the farm (Stubbings, 2003; Sargison *et al.*, 2007b). Large-scale anthelmintic surveys are labour intensive, expensive and logistically difficult to perform. Sixty-four percent of farms tested in a survey of Scottish sheep flocks had resistance to benzimidazoles (Bartley *et al.*, 2003), with the prevalence higher in lowland farms compared to upland or hill farms. This survey found no evidence for LEV and IVM resistance in Scottish sheep flocks at that time but a later survey, (Bartley *et al.*, 2006), found IVM resistant *T. circumcincta* in six out of seventeen flocks sampled. No routine surveillance for resistance is carried out, suggesting that anthelmintic resistance may be more widespread than stated in the literature (Sargison, Scott, & Jackson, 2001).

Multiple anthelmintic resistance is observed when a parasite population is resistant to more than one drug class. Triple resistance to the BZs, LEV and MLs was first recorded in the 1980s in *H. contortus* in South Africa (Van Wyk & Malan, 1988; Kaplan, 2004). Multiple anthelmintic resistance is more prevalent in the Southern Hemisphere (Bartley *et al.*, 2004). The first triple resistant parasites of ovine origin in the UK were reported in 2001 and triple resistance has now been reported in at least five UK sheep flocks (Sargison, Scott, & Jackson, 2001; Bartley *et al.*, 2004; Sargison *et al.*, 2007b). The inability to control *T. circumcincta* infection due to multiple anthelmintic resistance has been seen as a contributory factor in the abandonment of sheep farming on at least two UK farms (Sargison *et al.*, 2005; Blake & Coles, 2007).

The AADs were released onto the market in New Zealand in 2009 and in the UK in 2010, against a background of parasites already resistant to existing anthelmintics, as such, it will become increasingly important to monitor the efficacy of AAD anthelmintic treatment. Parasites that are already multi-drug resistant in the field, although not predisposed to AAD-resistance will inevitably develop resistance to the AADs; AAD-resistant *H. contortus* strains have been generated using *in vitro* and *in vivo* selection experiments whilst laboratory mutagenic experiments were used to generate AAD-resistant *C. elegans*. From these resources putative AAD resistance markers were identified (Kaminsky *et al.*, 2008). By identifying and monitoring the spread of AAD resistance using the putative AAD resistance markers when it inevitably occurs in the field, we will be able to better understand the mechanisms involved in AAD resistance and what causes its spread.

1.4.1: Factors affecting the development of anthelmintic resistance

The development of anthelmintic resistance has occurred relatively rapidly. It is facilitated by the large population size and inherent high genetic variability that is typical amongst nematode parasites, and compounded by the movement of their host species (Blouin *et al.*, 1995; Redman *et al.*, 2008). It is viewed as the inevitable consequence of routine anthelmintic use (Jackson & Coop, 2000; Abbott, Taylor, & Stubbings, 2004; James & Davey, 2008). Suppressive prophylactic parasite control strategies across the whole flock, utilising anthelmintics at regular intervals corresponding to the pre-patent period of the parasites, have been used to disrupt the parasite life-cycle and prevent clinical disease (Stubbings, 2003). The frequency of treatment determines how rapidly resistance will develop (Barton, 1983; Kwa, Veenstra, & Roos, 1994; Prichard, 2001). Suppressive prophylactic parasite control in young lambs has also been shown to prevent the lambs from developing a protective immune response to *H. contortus* meaning they are more at risk of developing haemonchosis once the parasites are no longer being controlled by the anthelmintics (Colditz *et al.*, 1996).

A study of the sequences of isotype 1 β-tubulin alleles of *H. contortus* and *T. circumcincta* from goat farms in France suggests that BZ resistance in small ruminants has

resulted from the selection of pre-existing resistance alleles in addition to spontaneous mutations resulting in novel resistance alleles (Silvestre & Humbert, 2002; Silvestre *et al.*, 2009). In an open sheep farm in the UK, analysis of the *T. circumcincta* BZ-resistant isolate present on the farm indicated that resistance to BZs has occurred through multiple mutations, spread between populations by gene flow as a result of the open nature of the UK sheep flock (Skuce *et al.*, 2010). Resistant parasites do not appear to revert to susceptibility once the selection pressure of anthelmintic treatment is removed, suggesting that there is no reduction in fitness in resistant parasites (Coles, 2005; James & Davey, 2008). No significant difference in egg production, development rate, establishment in the host or survival of adults or larvae was found when comparing BZ-resistant and -susceptible *T. circumcincta* (Elard, Sauve, & Humbert, 1998).

Some control methods advocated in the past have since been proven to be highly selective for anthelmintic resistance. One such recommendation was to dose sheep with anthelmintics and move them to clean pasture in an attempt to reduce the numbers of infective larvae available to re-infect the sheep (Boag & Thomas, 1973; Morley & Donald, 1980). This "dose and move" strategy has been shown to select more rapidly for resistance as all the parasites contributing to the next generation will be those which survived the anthelmintic treatment i.e. potentially resistant (Silvestre & Humbert, 2000; Stubbings, 2003; Waghorn *et al.*, 2009). This practice is now being actively discouraged. It is recommended that sheep are weighed prior to treatment to reduce the chance of under-dosing and that sheep are withheld from feed for 24 hours prior to treatment to slow the flow of digesta through the gut, which can improve drug efficacy. It is also important to ensure dosing guns are correctly calibrated and the anthelmintics are stored correctly (Barnes, Dobson, & Barger, 1995; Dobson, Le Jambre, & Gill, 1996; Silvestre, Cabaret, & Humbert, 2001; Abbott, Taylor, & Stubbings, 2004). The use of anthelmintics with persistent activity such as MOX can also hasten the development of anthelmintic resistance as the parasites are exposed to declining doses of the drug for a longer period, a process called tail selection (Dobson, Le Jambre, & Gill, 1996; Le Jambre *et al.*, 1999).

Another method to increase the bioavailability and, hence, the efficacy of the ML anthelmintics, but which has yet to get beyond research applications, is to co-administer an

inhibitor of P-glycoproteins (Pgps) alongside IVM. In the *in vitro* larval feeding inhibition test (LFIT), the use of Pgp interfering agents such as ketoconazole, verapamil and quercitin has been shown to drive BZ and IVM resistant *T. circumcincta* and *H. contortus* isolates towards a susceptible phenotype, as shown by a shift to the left on the dose response curve, indicating more larvae are affected by the drug at a lower concentration (Bartley *et al.*, 2009). *In vivo*, it has also been shown that ketoconazole increases the plasma concentration of IVM in sheep by reducing the Pgp mediated IVM efflux (Alvinerie *et al.*, 2008).

A certain proportion of the parasite population will not be exposed to the anthelmintic during treatment, these parasites are said to be *in refugia*, and include the environmental stages of the parasite (eggs to L_3), encysted larvae in the abomasal glands and parasites in sheep which are left untreated (Abbott, Taylor, & Stubbings, 2004). The principle of *refugia* is now seen as the most important factor in delaying the spread of anthelmintic resistance by providing a source of susceptible parasites to dilute out resistant individuals (Soulsby, 2007). Allowing part of the sheep flock to remain untreated, for example ewes which have built up natural immunity to parasites, or only treating those animals showing clinical symptoms of PGE will increase the numbers of parasites *in refugia* and slow the development of resistance (Abbott, Taylor, & Stubbings, 2004). The use of targeted selective treatment (TST) regimes is being investigated as a *refugia*-based management strategy aimed at maintaining drug efficacies whilst minimizing the loss of productivity caused by PGE. These strategies use indicators such as milk yield, weight gain, level of anaemia and faecal egg count to determine which animals in a flock require anthelmintic treatment. Liveweight gain in lambs has been shown to be a good indicator of individuals requiring treatment in areas where *T. circumcincta* predominates, and the FAMACHA (Faffa Malan Chart) system in tropical climates, using the level of anaemia, where *H. contortus* is predominant (Stafford, Morgan, & Coles, 2009; Kenyon *et al.*, 2009a). Preliminary results show that the TST approach maintains drug efficacy better than traditional suppressive treatment regimes whilst also having little effect on lamb liveweight gain and only requiring approximately half the amount of anthelmintic (Burgess, Bennett, & Kenyon, 2009; Greer *et al.*, 2009; Kenyon *et al.*, 2009b).

Anthelmintic resistance appears to be a genetic phenomenon; resistant worms give rise to resistant offspring. The means by which resistance is inherited will also affect how quickly resistance develops, for example, anthelmintics with multiple, genetically independent sites of action will cause the development of resistance in the parasites to occur more slowly compared to an anthelmintic with a single defined point of action (Martin *et al.*, 1998). If resistance is inherited as a dominant trait or is sex-linked, it will spread much more rapidly through a population than if it is a recessive trait. The genetics of anthelmintic resistance is complex, with little consensus in the literature; different species and sexes of parasites seem to exhibit different patterns of resistance. Furthermore, the different modes of action of the anthelmintic classes is likely to mean each class has a distinct mechanism of resistance. BZ resistance in *H. contortus* was initially described as incompletely dominant and subsequently as a multigenic, recessive autosomal trait whilst in *T. colubriformis* BZ resistance is an incompletely recessive trait (Dobson, Le Jambre, & Gill, 1996; Sangster, Redwin, & Bjorn, 1998; Le Jambre *et al.*, 1999; Sutherland & Scott, 2010). LEV resistance in *T. circumcincta* and in *T. colubriformis* females is a sex-linked recessive trait but in *H. contortus* LEV resistance is an autosomal recessive trait (Dobson, Le Jambre, & Gill, 1996; Sangster, Redwin, & Bjorn, 1998; Martin *et al.*, 1998; Wolstenholme *et al.*, 2004; Sutherland & Scott, 2010). LEV and TBZ resistance in an isolate of *H. contortus* was proposed to be controlled by more than one gene (Sangster, Redwin, & Bjorn, 1998). ML resistance in *H. contortus* and *T. circumcincta* appears to be controlled by dominant traits (Dobson, Le Jambre, & Gill, 1996; Le Jambre *et al.*, 1999; Le Jambre *et al.*, 2000; Sutherland *et al.*, 2002; Sutherland *et al.*, 2003; Wolstenholme *et al.*, 2004).

1.5: Detection of anthelmintic resistance

Currently, detection of anthelmintic resistance relies on *in vitro* and *in vivo* tests. However, these tests are limited in their use as not all drugs can be tested using each assay, they are time-consuming, expensive, labour intensive and in some cases require the use of test animals (Coles *et al.*, 1992). Some of the current diagnostic tests only work when at least 25% of the parasite population under test is phenotypically resistant (Martin, Anderson, & Jarrett, 1989; Coles *et al.*, 1992; McKellar & Jackson, 2004). The exceptions to this are DNA-based tests for detecting BZ resistance in which individual parasites and pooled samples can be assayed, giving much higher sensitivity. Details of how each

diagnostic test works are given below together with their limitations. There is a need for improved tests for detecting resistance to make the best use of existing anthelmintics, to investigate the genetic consequences of different management strategies and ultimately as surveillance tools for use in the field.

1.5.1: *In vivo* tests

The controlled efficacy test (CET) is seen as the gold standard test to calculate the true efficacy of anthelmintics (McKellar & Jackson, 2004; Coles *et al.*, 2006). Control animals are infected with a known number of L_3 and then dosed with anthelmintic at a range of concentrations. After a set time period, the animals are culled and worms recovered from the abomasum. Resistance is confirmed when the reduction in worm count is less than 90%, or more than 1000 worms survive treatment (Coles *et al.*, 1992; Taylor, Hunt, & Goodyear, 2002). An untreated group is also included as a control. Details of the standardised method for carrying out the CET are given in Wood *et al* (1995), having standardised methods allows drug efficacy trials from around the world to be easily compared. The use of test animals is expensive, time-consuming and labour intensive; there are also ethical concerns about the use of experimental animals.

The faecal egg count reduction test (FECRT) compares the faecal egg count of individual animals before and after anthelmintic treatment; it is relatively simple to perform and can be used to test all groups of anthelmintics but, like the CET, it is expensive and time-consuming (Taylor, Hunt, & Goodyear, 2002). A gap of 10 to 17 days, depending on the anthelmintic under test, between treatment and test days is required as a drop in egg production occurs directly after anthelmintic treatment. The inclusion of an untreated control group is recommended to identify any natural fluctuations in egg output during the test period. Resistance is confirmed when the reduction in faecal egg count post-treatment is less than 95% and when the lower 95% confidence interval of the reduction in faecal egg count is less than 90%. Resistance is suspected when only one of the two criteria is met (Coles *et al.*, 1992; Bartley *et al.*, 2006; Van Zeveren *et al.*, 2007a).

1.5.2: *In vitro* tests

The egg hatch test (EHT) only works with the BZs as LEV and the MLs are not ovicidal (Coles, 2005). Eggs collected from faeces are incubated in serial dilutions of TBZ and the percentage hatch rate observed, allowing the discriminating dose to be calculated from a dose response curve (Le Jambre, 1976; Taylor, Hunt, & Goodyear, 2002). The EHT performs best on species of nematodes with rapidly hatching eggs (Coles *et al.*, 1992). In mixed gastrointestinal nematode infections, species identification of the hatched larvae is possible. Recently, a standardised protocol has been determined to allow easy repeatable comparison between laboratories (von Samson-Himmelstjerna *et al.*, 2009b). An altered protocol for the egg hatch test is available for the detection of LEV resistance (Dobson *et al.*, 1986).

The micro-agar larval development test (MALDT) is used to identify resistance to BZs and LEV but cannot be used for the MLs as it is not reliable enough (Coles, 2005). The larval development test (LDT) can be used with any anthelmintic group and involves the development of L_1 into L_3 in the presence of the anthelmintic under test, with *Escherichia coli* as a food source (Taylor, Hunt, & Goodyear, 2002). A commercial larval development test called "Drenchrite" has been developed (Coles, 2005).

The LFIT is used to test for ML and LEV resistance. Eggs collected from faeces are allowed to hatch and the L_1 are cultured in serial dilutions of anthelmintic for 2 hours. The larvae feed on fluorescein isothiocyanate (FITC)-labelled *E. coli* and are incubated for a further 24 hours. The numbers of larvae which are fed or unfed at each anthelmintic concentration is determined by examining the larvae under a fluorescence microscope and this ratio compared to an untreated control group (Alvarez-Sanchez *et al.*, 2005). Another, related, assay for the detection of LEV and ML resistance is the larval migration inhibition test (LMIT) (Wagland *et al.*, 1992; Rabel, McGregor, & Douch, 1994). Like the LFIT, exsheathed L_3 are incubated in serial dilutions of anthelmintic and then placed, in solution, above a 25μm nylon mesh and incubated for a further two hours. Resistant worms, which have not been paralysed by the drug, will be able to actively migrate through the mesh. By counting the number of migrated and non-migrated L_3, the percentage migration can be

calculated and, hence, provide an idea of an isolate's ability to survive anthelmintic treatment.

In vitro assays are easier, faster and cheaper to perform than *in vivo* tests and do not require the use of animals, which also removes any inter-host variation (Dobson *et al.*, 1986; Dobson, Le Jambre, & Gill, 1996). The tests described above have different uses, primarily due to some of the tests being restricted to testing certain classes of drugs such as the MALDT and EHT not detecting ML resistance or the LFIT and LMIT not detecting BZ resistance. Like *in vivo* tests, *in vitro* tests only work when over 25% of the worm population under test is resistant (Coles *et al.*, 1992; McKellar & Jackson, 2004). One challenge to all the tests described is when mixed species infections are present; different species may have different sensitivities to a drug, making it difficult to interpret results.

1.5.3: Molecular based tests

As resistance has a genetic basis, either through qualitative changes (mutations) or quantitative changes (alterations in expression of genes), the identification of these changes could lead to the development of DNA based tests. Development of sensitive molecular-based diagnostic tests has enabled the detection of BZ resistance at lower levels compared to *in vitro* methods. The tests can allow the identification of individual parasites carrying resistance genes within a parasite population (Elard, Cabaret, & Humbert, 1999; von Samson-Himmelstjerna, 2006). The tests also allow species identification in mixed populations, and can be adapted to give the prevalence of resistance for each species (Silvestre & Humbert, 2000). Polymerase chain reaction (PCR)-based tests are only available for diagnosis of BZ resistance, where single nucleotide polymorphisms (SNPs), such as those discussed below, have been identified that are associated with resistance (Kwa, Veenstra, & Roos, 1994; Elard, Cabaret, & Humbert, 1999; Silvestre & Humbert, 2000). This PCR has been developed further to include PCR-RFLP, allowing the simultaneous species identification and determination of the resistance status of individual *T. circumcincta*, *H. contortus* and *T. colubriformis* L_3. This PCR has also been adapted for real-time PCR applications, which are more rapid than conventional PCR and allow the simultaneous detection of susceptible and resistant alleles, even in pooled samples (von

Samson-Himmelstjerna, 2006; Walsh *et al.*, 2007; von Samson-Himmelstjerna *et al.*, 2009a). When using molecular based tests, as for the *in vivo* and *in vitro* methods, the challenge is the correct identification of resistance in mixed parasite species infections. However, the reliance on only one mutation to diagnose resistance will not take into account other non-specific resistance mechanisms which may be enabling populations of parasites to exhibit phenotypic resistance (Coles, 2005). This could mean that a parasite not exhibiting a mutation associated with resistance in a molecular-based test may still have other gene expression level changes or mutations that enable it to be phenotypically resistant.

1.6: Anthelmintic resistance mechanisms

Resistance is thought to develop in a number of ways (Wolstenholme *et al.*, 2004). Firstly, a resistant nematode could have an altered drug target which results in the drug being unable to bind, or bind less effectively. Secondly, there could be a change in the metabolism of the drug causing the drug not to be metabolised into its active form, or to be removed from its target sites. Thirdly, a change in distribution of the drug in the parasite could prevent it reaching its target site. Fourthly, a change in the drugs target gene expression could overcome the drugs action. Alternatively, there could be non-specific resistance mechanisms (i.e. a mechanism not related to the precise mode of action of the drug) such as changes in the expression level of non-target proteins used by the parasite to handle drugs and toxins (Wolstenholme *et al.*, 2004). These mechanisms should be identifiable as quantitative and qualitative genetic changes.

Genetic studies have identified a point mutation, a substitution of tyrosine (Tyr) for phenylalanine (Phe) at codon 200 (Phe200Tyr or F200Y) in the isotype 1 β-tubulin gene, which is the major genetic determinant of BZ resistance in *T. circumcincta*, *Trichostrongylus colubriformis, H. contortus* and *Cooperia oncophora* (Kwa, Veenstra, & Roos, 1994; von Samson-Himmelstjerna *et al.*, 2007). A second much less common, Phe-Tyr polymorphism at codon 167 is also associated with BZ resistance, especially in the nematode parasites of equines (Silvestre & Cabaret, 2002; Wolstenholme *et al.*, 2004; Gilleard, 2006; Hodgkinson *et al.*, 2008). It appears that, in *H. contortus*, Tyr is required

at codon 200 for BZ resistance but in *T. circumcincta*, worms can survive BZ treatment when Phe/Phe homozygous at codon 200 and either Phe/Tyr heterozygous at codon 167 or Tyr/Tyr homozygous at codon 167 (Silvestre & Cabaret, 2002; von Samson-Himmelstjerna *et al.*, 2007). Some triple-resistant *T. circumcincta* with susceptible genotypes at codon 200, have been shown to survive BZ treatment *in vivo* and *in vitro*, suggesting that other resistance mechanisms may also be contributing to the expression of BZ resistance (Stenhouse, 2007). An alternative point mutation in *H. contortus* capable of conferring BZ resistance, was found at codon 198, a substitution of alanine (Ala) for glutamic acid (Glu) (Glu198Ala) (Ghisi, Kaminsky, & Maser, 2007; de Lourdes Mottier & Prichard, 2008). This mutation and Phe200Tyr appear to be mutually exclusive; the combination of the two may be lethal to the parasites. Alongside these mutations, it appears that binding of BZs to β-tubulin is reduced in resistant compared to susceptible isolates (von Samson-Himmelstjerna, 2006).

Resistance mechanisms to LEV and the other imidazothiazole/ tetrahydropyrimidines, including PYR, are less well researched (Kopp *et al.*, 2008). Three genes, *unc*-38, *unc*-29 and *lev*-1, identified in *C. elegans* appear to be involved in LEV resistance and were found to encode nicotinic acetylcholine receptors (nAChRs). Mutations in *unc*-29 or *unc*-38 caused a complete loss of sensitivity to LEV (Fleming *et al.*, 1997). A mutation from glutamic acid (Glu) to glycine (Gly) at amino acid 153 (Glu153Gly) in the *unc*-38 gene of *C. elegans* is required for LEV resistance (Rayes *et al.*, 2004; Martin & Robertson, 2007). LEV is a very weak agonist of mammalian muscle, resulting in low toxicity, however, the Glu153Gly mutation increases the efficacy of LEV as an agonist against mammalian muscle acetylcholine receptors (Rayes *et al.*, 2004). The mutation in the *unc*-63 gene, changing a glutamine (Gln) to glycine (Gly) at amino acid 57 (Gln57Gly) in *C. elegans*, appears to reduce the sensitivity of the receptors to PYR (Bartos, Rayes, & Bouzat, 2006). nAChRs are made up of five glycoprotein subunits arranged around a central ion channel; different subunits give the nAChR different pharmacological properties (Kopp *et al.*, 2008). In *C. elegans*, absence of one type of nAChR at the neuromuscular junction can cause LEV resistance whilst absence of the L subtype of nAChR from *Oesophagostomum dentatum* was associated with LEV resistance (Martin & Robertson, 2007). LEV resistant *H. contortus* and *T. colubriformis* membrane preparations were found to have decreased nAChR affinity for LEV (Sangster, Redwin, &

Bjorn, 1998). A gene fragment, HA17, was found to be differentially expressed between LEV-resistant and -susceptible parasites using cDNA-AFLP techniques and has been proposed as a potential marker for LEV resistance in *H. contortus* (Neveu *et al.*, 2007). Kopp *et al* (2008) showed in the dog hookworm, *Ancylostoma caninum*, that whilst there were no polymorphisms of significance in three putative PYR receptor subunits (AAR-29, AAR-38 and AAR-63), there was a significantly reduced level of transcription of these genes in a highly PYR resistant isolate compared to an isolate with low levels of PYR resistance. These subunits were orthologous to the *unc*-29, *unc*-38 and *unc*-63 genes in *C. elegans* (Kopp *et al.*, 2008).

Genes implicated in IVM resistance are the likely molecular targets of IVM and include Glutamate- and GABA- gated Cl channels (Njue *et al.*, 2004; Gilleard, 2006). Changes in allele frequencies of Glu- and GABA- Cl subunits in *H. contortus* have been observed but no single allele has been associated with resistance between different IVM resistant populations of *H. contortus* (Blackhall *et al.*, 1998a; Blackhall, Prichard, & Beech, 2003). Only one study has associated a change in an amino acid, from leucine to phenylalanine at codon 256, (Leu256Phe) in the GluClα3 subunit of *C. oncophora* to IVM resistance but this has not been found in *H. contortus*, *T. circumcincta*, *O. ostertagi* or other *C. oncophora* isolates (Njue *et al.*, 2004; Van Zeveren, 2009). In *C. elegans*, mutations in several GluCl subunit genes are required for ML resistance (McCavera, Walsh, & Wolstenholme, 2007). Concurrent mutations of three genes encoding GluCl α-type subunits, *avr*-14, *avr*-15 and *glc*-1 in *C. elegans* are required to confer high levels of IVM resistance; mutation of any of the two genes only confers low or no resistance (Dent *et al.*, 2000; Cook *et al.*, 2006). *Avr*-15 encodes a GluClα2 which is expressed in the pharyngeal muscle in *C. elegans*; *avr*-14 encodes a GluClα3 and is expressed in the extrapharyngeal nervous system: Significantly, one of the modes of action of IVM is starvation of nematodes caused by inhibition of the pharyngeal pump (Dent *et al.*, 2000; Cook *et al.*, 2006). Parasitic nematodes have a different set of GluCl subunit genes compared to *C. elegans* but do share some, such as *avr-14*, an orthologue of which, in *C. oncophora*, has a polymorphism making the subunit less sensitive to IVM (McCavera, Walsh, & Wolstenholme, 2007). The molecular basis of resistance to IVM in trichostrongyle parasites remains to be elucidated (Geary, 2005; Prichard *et al.*, 2007).

Even though AADs are a recent discovery and are yet to reach a wider market, research has already been carried out to determine if and how resistance to the AADs will develop. AAD resistant populations of *C. elegans* have been selectively bred by exposing them to the drug; this has allowed the AAD target gene to be identified as a nAChR, called *acr*-23 (Prichard & Geary, 2008; Kaminsky *et al.*, 2008; Zarlenga & Gasbarre, 2009). This nAChR is unique to nematodes and is different from those associated with LEV resistance (Prichard & Geary, 2008). Three AAD-resistant *H. contortus* mutant lines have been generated which all appear to have lost at least part of the *H. contortus des*-2 homologue, which is part of the nAChR group. The two fully AAD-resistant *H. contortus* mutants had also lost part of the *H. contortus* homologue of the *acr*-23 gene (Kaminsky *et al.*, 2008).

1.7: Candidate ivermectin resistance genes

Most of the anthelmintic resistance mechanisms identified to date have used a candidate gene approach to identify qualitative changes conferring resistance. The limited success of this approach is partly because classical genetic linkage studies to identify 'resistance genes' are difficult, especially where the mode of action of the anthelmintics is still not clear. This makes it difficult to propose candidate resistance genes for further study. Alternative mechanisms of resistance such as changes in gene expression level remain largely unexplored, particularly in *T. circumcincta* and will be the focus of this thesis. Differences in expression levels of genes may be constitutive, where a gene is always expressed differentially between anthelmintic susceptible and resistant strains of parasites at rest, or inducible, where a change in gene expression is observed between parasites that have been treated with or exposed to anthelmintic and those that have not. Changes in gene expression pattern may be caused by either up or down -regulation of the gene or increased or decreased gene copy number. Details of two gene families, the cytochrome P450s and P-glycoproteins, which may be linked to the IVM-resistance phenotype and appear to play a pivotal role in the handling and metabolism of drugs, are given below.

1.7.1: P-glycoproteins

Pgps are large ~170kDa transmembrane proteins which are part of the ATP binding cassette (ABC) superfamily of genes. These are found throughout the phyla *Archaea*, *Eubacteria* and *Eukarya* (Valverde *et al.*, 1992; Kerboeuf, Guegnard, & Le Vern, 2003; Jones & George, 2005; Riou *et al.*, 2005). Pgps are involved in a range of cellular processes, including the transport of endogenous and exogenous hydrophobic molecules (Valverde *et al.*, 1992; Sangster *et al.*, 1999a; Kerboeuf, Guegnard, & Le Vern, 2003). Pgps comprise two symmetrical halves, but they can also exist as half-transporter molecules (Sheps *et al.*, 2004; Ardelli, Guerriero, & Prichard, 2006a). Each half contains a transmembrane domain, comprised of six transmembrane α-helices, and a cytoplasmic nucleotide binding domain (NBD) at the carboxy end of each half molecule which contains an ATP binding site (Sangster, 1994). Within the cytoplasm of cells, the two halves of the molecule are linked by a hydrophilic region of approximately 60 amino acids; together, the two transmembrane domains form a membrane channel that acts as a substrate binding site (Jones & George, 2005). The two NBDs are structurally non-identical or functionally interchangeable but are highly conserved and show an amino acid homology of approximately 70% (Sangster, 1994; Higgins *et al.*, 1997; Sangster *et al.*, 1999a). Within the NBD there is the so-called "Walker motif", a characteristic amino acid sequence, LSGGQ, which is the signature sequence of the NBD in ABC transporters (Higgins *et al.*, 1997; Jones & George, 2005). A diagram of a typical Pgp is shown in Figure 1.2.

Compared to the amount of information about Pgps in humans and *E. coli*, the amount of data on parasites is much smaller and appears to concentrate on human parasites such as *Plasmodium falciparum,* trypanosomes and *Onchocerca volvulus.* Nematode Pgp studies have concentrated on *C. elegans, O. volvulus* and *H. contortus* (Sangster, 1994; Jones & George, 2005). Nematodes appear to possess more Pgps than mammals, possibly due to parasites living in a more hostile environment (Lespine *et al.*, 2008). For example the eggs and larval stages are exposed to a range of toxins and conditions whilst in the faeces or on the pasture, such as xenobiotics from fungi, and the adults are exposed to an acidic environment and digestive enzymes in the abomasum of the host (Sutherland & Scott, 2010). Pgps in nematode parasites appear to be expressed in a range of tissues, including the pharynx and intestinal tract in *H. contortus* and *C. elegans*, sheaths of L_3, eggshells and cuticles of *H. contortus* and microfilariae and adults of *O. volvulus* (Lincke

et al., 1993; Broeks *et al.*, 1995; Kwa *et al.*, 1998; Huang & Prichard, 1999; Smith & Prichard, 2002; Kerboeuf, Guegnard, & Le Vern, 2003; Kerboeuf *et al.*, 2003; Riou *et al.*, 2005; Prichard & Roulet, 2007; Riou *et al.*, 2007; Lespine *et al.*, 2008). Pgps are thought to play a pivotal role in handling xenobiotics as the MLs, such as IVM, appear to be absorbed into, distributed around and eliminated from parasites under the control of multidrug resistant transporters such as Pgps (Lespine *et al.*, 2008).

Over-expression of Pgps is linked to drug resistance in cancerous tumours, malaria parasites and the human immunodeficiency virus (Beugnet, Gauthey, & Kerboeuf, 1997; Zhang *et al.*, 1998; Loo & Clarke, 1999; Jones & George, 2005). In human tumours, such as sarcomas and neuroblastomas, Pgp expression correlates with a multidrug resistant (MDR) phenotype in the cancer cells, leading to a lower response to chemotherapy and reduced patient survival. Interestingly, IVM, which is a known substrate of Pgp, can be used as a MDR reversing agent (Pouliot *et al.*, 1997; Loo & Clarke, 1999). An increase in Pgp transcription has also been observed in *P. falciparum* and is associated with chloroquine resistance (Ekong *et al.*, 1993), whilst the use of verapamil, a calcium channel blocker that binds to Pgp, reverses chloroquine resistant *P. falciparum* (Martin, Oduola, & Milhous, 1987).

As stated previously, Pgps appear to be abundant in nematodes. Various PCR and probe-based approaches have identified at least 2 Pgp genes from *O. volvulus*, at least 14 Pgp genes in *C. elegans* and at least 12 Pgp genes in *H. contortus* (Kwa *et al.*, 1998; Huang & Prichard, 1999; Le Jambre, Lenane, & Wardrop, 1999; Sangster *et al.*, 1999a; Kerboeuf *et al.*, 2003; Sheps *et al.*, 2004; Blackhall, Prichard, & Beech, 2008). To date, no Pgps have been identified in *T. circumcincta*; this is despite ongoing EST and genome sequencing projects. An increase in Pgp mRNA level has been observed in IVM-selected *H. contortus* (Xu *et al.*, 1998) and changes in Pgp allele frequencies and/ or reductions in genetic diversity have been observed between IVM- or MOX-selected isolates of *O. volvulus* and *H. contortus* (Blackhall *et al.*, 1998b; Ardelli, Guerriero, & Prichard, 2005; Ardelli, Guerriero, & Prichard, 2006a; Ardelli, Guerriero, & Prichard, 2006b). These observations would suggest that the anthelmintics, particularly the MLs, are placing the Pgps under selection pressure. Expression of OvPgp-1 in *O. volvulus* negatively correlates

with stage-specific sensitivity to IVM as microfilariae express less Pgp and are more sensitive to IVM compared to adults (Huang & Prichard, 1999). Pgps do appear to be implicated in ML resistance, for example, co-administration of verapamil with MOX or IVM significantly increases the efficacy of the anthelmintic against a MOX-selected strain of *H. contortus* (Xu *et al.*, 1998). Resistance to IVM in *C. elegans* is associated with increased expression of Pgps and multidrug resistance proteins, which are members of the ABC superfamily, like Pgps (James & Davey, 2008). No SNPs in Pgps associated with ML resistance have been identified thus far, possibly due to the size and number of the molecules. Van Zeveren (2009) showed that there was a 3.4 fold constitutive increase in expression of a Pgp, designated PGP2a, in adult *O. ostertagi* from a laboratory selected isolate that was demonstrably resistant to IVM.

Treatment of *H. contortus* laboratory- and field- isolated populations with BZs also appears to exert a selection pressure on Pgp alleles as none of the populations of worms were in Hardy-Weinberg equilibrium at the Pgp locus (Blackhall, Prichard, & Beech, 2008). There may be an element of "cross-talk" between the two resistance mechanisms; Pgps may offer an alternative, or additional, mechanism of resistance to BZs in addition to the aforementioned SNPs in β-tubulin (Blackhall, Prichard, & Beech, 2008). The use of Pgp inhibitors, such as *Lens culinaris* lectin, causes an enhanced susceptibility of *H. contortus* eggs to TBZ (Kerboeuf, Guegnard, & Le Vern, 2002) and the use of verapamil increases the toxicity of TBZ and ABZ to *H. contortus* eggs (Beugnet, Gauthey, & Kerboeuf, 1997). Inhibitors of Pgps cause a reversal towards a susceptible phenotype of resistant strains of *T. circumcincta* and *H. contortus* when exposed to IVM in the LFIT. These inhibitors, many of which are already used therapeutically, although not for parasitic nematode control, included pluronic P85, verapamil (used to treat hypertension and cardiac arrhythmia), quercitin, ketoconazole (a broad spectrum anti-fungal agent) and valspodar (Virkel *et al.*, 2009; Bartley *et al.*, 2009).

1.7.2: Cytochrome P450s

Cytochrome P450s (CYPs) are a large superfamily of haemproteins which are found throughout the Animal Kingdom, suggesting an ancestral gene existed before the

divergence of prokaryotes and eukaryotes (Nelson *et al.*, 1996; Berge, Feyereisen, & Amichot, 1998; Mansuy, 1998; Omura, 1999). CYPs are the largest gene superfamily identified to date and can be classified into over 70 individual families of genes, each identified by a number, which have at least 40% amino acid identity. These can be further classified into subfamilies and, finally, genes and alleles (Nelson *et al.*, 1996; Gong *et al.*, 2005). Subfamilies are designated by a letter and share at least 55% amino acid identity, whilst the numeral at the end of the CYP name denotes the gene (Berge, Feyereisen, & Amichot, 1998). The divergence in protein sequence of alleles is less than or equal to 3% (Nelson *et al.*, 1996). Families of CYPs can be loosely grouped into clans which probably represent genes that diverged from a single common ancestor (Nelson, 1998; Nelson, 1999).

CYPs are single domain proteins which protrude into the cytoplasm and are anchored to membranes by an N terminal signal sequence linked to the catalytic domain by a highly conserved proline rich region, as shown in Figure 1.3 (Chen & Kemper, 1996; Gotoh, 1998; Graham & Peterson, 1999). In eukaryotes, CYPs are bound to the membrane of the endoplasmic reticulum or mitochondria whilst, in most bacteria, CYPs are water soluble (Chen & Kemper, 1996; Omura, 1999). Studies have shown that the structure of CYPs generally consists of four β-sheets and thirteen α-helices (Graham & Peterson, 1999). Only three amino acid positions are conserved across the whole CYP superfamily, two form the amino acid motif E-x-x-R in one of the α-helices which make up the structural core of the protein whilst the third conserved amino acid is a cysteine that is the fifth ligand for the haem ion (Graham & Peterson, 1999). CYPs catalyse the oxidative and reductive metabolism of a wide range of, predominantly hydrophobic, exogenous and endogenous molecules, such as sterols, fatty acids, prostaglandins, drugs and environmental toxins (Guengerich, 1991; Nelson *et al.*, 1996; Amichot *et al.*, 1998; Barrett, 1998; Mansuy, 1998; Graham & Peterson, 1999; Menzel, Bogaert, & Achazi, 2001). One protein is able to catalyse different reactions depending on the substrate presented. Also, there is redundancy in the CYP pathways as more than one CYP can metabolize the same substrate (Guengerich, 1991; Menzel *et al.*, 2005).

In *C. elegans*, approximately 80 CYPs have been identified as a result of the *C. elegans* genome sequencing project, and are classified into clans 2, 3 and 4 containing 7, 3 and 3 gene families, respectively (Barrett, 1998; Gotoh, 1998; Menzel, Bogaert, & Achazi, 2001). Within clan 3, the CYP13 family appears to be specific to *C. elegans* (Gotoh, 1998; Nelson, 1998; Nelson, 1999). It is possible that *C. elegans* has such a large number of CYPs to enable the nematode to cope with environmental toxins. Few studies to date have identified CYPs from parasitic nematodes. Adult helminth enzyme extracts from a range of species, including *H. contortus*, *Heligmosomoides polygyrus* and *Fasciola hepatica*, were unable to oxidise known CYP substrates and attempts using spectrophotometry or CYP-inducing agents failed to identify CYPs in adult nematodes (Barrett, 1998). The methods used at the time may have lacked the sensitivity to detect CYPs which are only present in very low abundance or in the free-living larval stages (Barrett, 1998). However, CYPs have since been found in all life-cycle stages of *C. elegans*, including adults (Menzel, Bogaert, & Achazi, 2001).

There is some evidence that *H. contortus* larval stages may possess CYPs, through the identification of microsomal CYP activity against the substrates aldrin and 7-ethoxycoumarin that could be inhibited by the CYP inhibitor, piperonyl butoxide, and induced by phenobarbitol (Kotze, 1997; Berge, Feyereisen, & Amichot, 1998). There is also evidence that the efficacy of BZs can be improved by the use of the CYP inhibitors piperonyl butoxide, ketoconazole and metyrapone and the administration of BZs has been shown to increase the catalytic activity of CYP1A in rat hepatocytes (Baliharova *et al.*, 2003; McKellar & Jackson, 2004; Virkel *et al.*, 2009). It has also been suggested that CYP activity in *H. contortus* may be restricted to the free-living stages where oxygen is readily available (Kotze, 1997). The CYP activity in adult *H. contortus* is 10,000-fold lower compared to rat liver microsomes, so the assays used may not have been sensitive enough to detect this (Kotze, 1997). Approximately 60 CYP tags have been identified from the *H. contortus* genome sequencing project database but it is anticipated this will reduce in number due to the presence of numerous polymorphisms or as tags become assembled (R. Laing, Pers. Comm.).

It appears that CYPs may also play a role in drug resistance in a range of species other than nematode parasites. Insect CYPs are involved in the metabolism of the majority of insecticides, either activating the molecule or as a detoxification mechanism (Berge, Feyereisen, & Amichot, 1998). In insects, over-expression or over-transcription of CYPs appear to enable *Drosophila* spp, *Anopheles minimus*, *Anopheles gambiae* and *Culex pipiens pallens* to exhibit insecticide resistance (Amichot *et al.*, 1998; Daborn *et al.*, 2002; Le Goff *et al.*, 2003; Nikou, Ranson, & Hemingway, 2003; Gong *et al.*, 2005; Rodpradit *et al.*, 2005; Vontas *et al.*, 2005). For example, *Drosophila* spp over-expressing Cyp6g1 are resistant to various insecticides such as organophosphates, DDT and growth regulators (Ffrench-Constant, Daborn, & Le Goff, 2004). Chloroquine resistance in the malaria parasites, *P. berghei* and *P. falciparum*, has also been shown to correlate with an increase in CYP activity (Ndifor, Ward, & Howells, 1990; Barrett, 1998).

1.8: Summary

T. circumcincta is the most important nematode parasitizing sheep in the UK and contributes significantly to the economic and welfare costs associated with PGE. In the UK, it is also the dominant resistant species yet the full extent of anthelmintic resistance is not known due to insensitive diagnostic tests and the absence of active surveillance (Coles, 2005). Sheep do eventually build up immunity to *T. circumcincta* but lambs are very vulnerable to infection and the ability of lambs to reach slaughter weight in an economically viable length of time relies heavily on the use of anthelmintics, without which the sheep industry would not be able to survive in its present form (Sargison *et al.*, 2005; Blake & Coles, 2007). As such, it is important that the mechanisms of anthelmintic resistance, particularly to the MLs are investigated, molecular markers associated with resistance are found and improved techniques for the detection of resistance are developed (Prichard & Roulet, 2007). This would allow enhanced anthelmintic usage management strategies to be developed with the aim of prolonging the effective life of the available classes.

Candidate gene studies looking at specific genes which are assumed to be the target site for anthelmintic drugs have only managed to identify SNPs associated with resistance

to the BZs (Gilleard, 2006; von Samson-Himmelstjerna *et al.*, 2007). As discussed in this literature review, other mechanisms of resistance such as changes in gene expression in non-target genes like the CYPs and Pgps remain to be investigated, particularly in *T. circumcincta*. Alongside these classical candidate gene studies focussing on gene expression changes as opposed to mutations, the potential of alternative genetic techniques such as next generation sequencing and suppression subtractive hybridisation, which do not rely on initial assumptions about a gene's association with a drug's mode of action, need to be explored. Furthermore, a greater understanding of the drug handling mechanisms of MDR *T. circumcincta* will potentially enable the development of molecular markers for resistance which will enable a more effective control strategy for *T. circumcincta* to be developed whilst reducing the rate at which anthelmintic resistance develops.

1.9: Aims and objectives

The overall aim of this research was to gain an insight into how MDR *T. circumcincta* processes anthelmintics, in particular IVM, and using molecular biology techniques, to identify the role non-specific drug handling mechanisms, such as CYPs and Pgps, play in the expression of resistance with a view to developing sensitive genetic markers for resistance detection. Comparisons were made between a *T. circumcincta* isolate resistant to BZs, LEV and IVM (MOTRI) and the non-related CVL isolate which was isolated before exposure to LEV and IVM and is phenotypically BZ susceptible as measured by the EHT (D. Bartley, Pers. Comm.). Alongside this other comparisons were made using *in vitro* and *in vivo* IVM exposed isolates; further details of the isolates used in the different experiments are given in Chapter 2.

- The first objective was to identify a panel of novel Pgp and CYP homologues from *T. circumcincta* for a more detailed genetic analysis. This was performed on cDNA and EST datasets using bioinformatics approaches, and PCR using degenerate primers based on available sequences from related nematode species. At the time of writing, no complete genome sequence and only a limited EST dataset was available for *T. circumcincta*. If *T. circumcincta* follows the same pattern as *C.*

elegans and *H. contortus*, then the numbers of Pgps and CYPs is likely to be large. It was not anticipated that all Pgps and CYPs will be identified by this approach.

- The second objective was to develop semi-quantitative or real-time quantitative PCR assays based on the selected panel of Pgp and CYP genes to investigate comparative gene expression in three different *T. circumcincta* isolates, namely susceptible (CVL), triple-resistant (MOTRI) and triple-resistant survivors of *in vivo* IVM treatment (Post IVM MOTRI) and over five life-cycle stages. The first part of this objective was to identify constitutive changes in the expression of the CYPs and Pgps between IVM- susceptible and resistant *T. circumcincta* isolates. The second part of this objective was to identify inducible changes in the expression of CYPS and Pgps by stressing different isolates of *T. circumcincta* with IVM to identify genes which display altered expression patterns in response to anthelmintic exposure. Two experiments were carried out, to investigate this: (i) exposing the parasite isolates to IVM *in vitro*, utilising a modified LMIA method, and (ii) comparing the Post IVM MOTRI isolate with the MOTRI isolate *in vivo*.

- The third objective used a more global, non-candidate gene approach, to investigate gene expression changes in the MOTRI isolate in response to *in vitro* IVM exposure by comparing an unexposed pool of adult parasites to an IVM exposed pool. Using the next generation Roche 454 sequencing approach, a large amount of novel sequence data for *T. circumcincta* was generated. Bioinformatic analysis of this sequence data allowed the identification of clusters of sequences which exhibited altered expression profiles in response to IVM exposure.

- The final approach was to use a non-hypothesis driven one-way suppression subtractive hybridisation method, followed by sequencing of the cloned subtracted products to investigate constitutive gene expression differences between the CVL and MOTRI adult parasites at rest. This was to identify which genes in the MOTRI adults had increased expression compared to the CVL adults, and could potentially

be used as markers for anthelmintic resistance, or gene expression changes allowing the parasite to exhibit resistance.

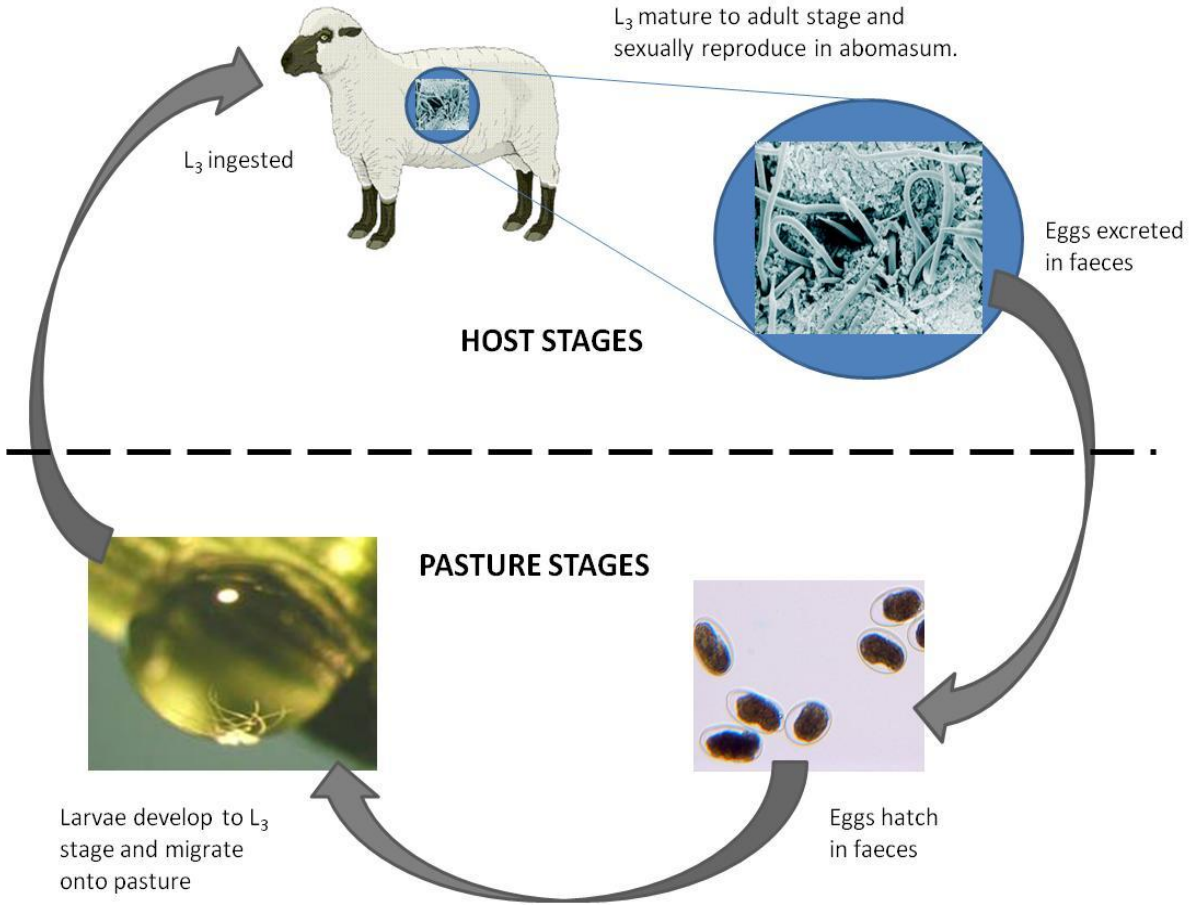


Figure 1.1: Life-cycle of *T. circumcincta*.

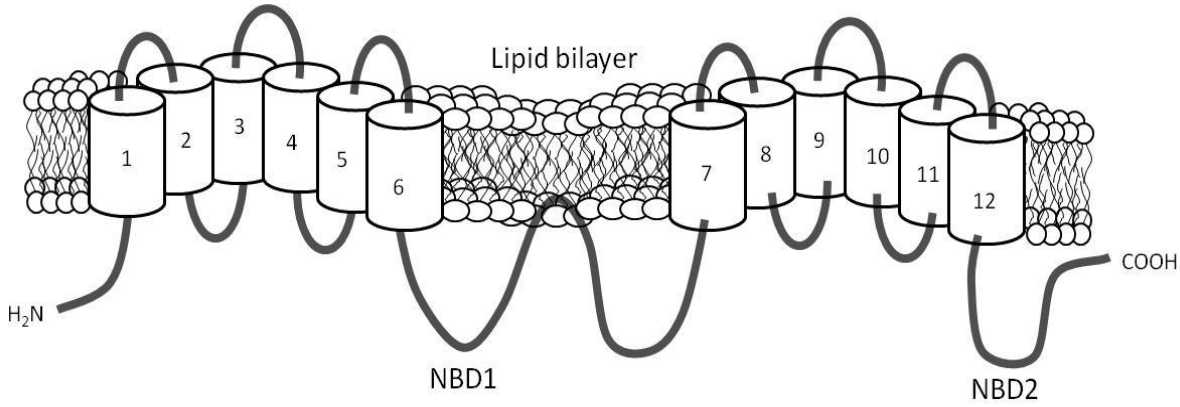


Figure 1.2: Structure of a typical P-glycoprotein. The P-glycoprotein has been expanded to aid visualisation as it normally forms a cylindrical channel within cell membranes. The transmembrane α helices are numbered 1 -12, the two nucleotide binding domains labelled NBD 1 and 2 and the hydrophobic linker region is shown at the centre of the diagram, protruding into the lipid bilayer.

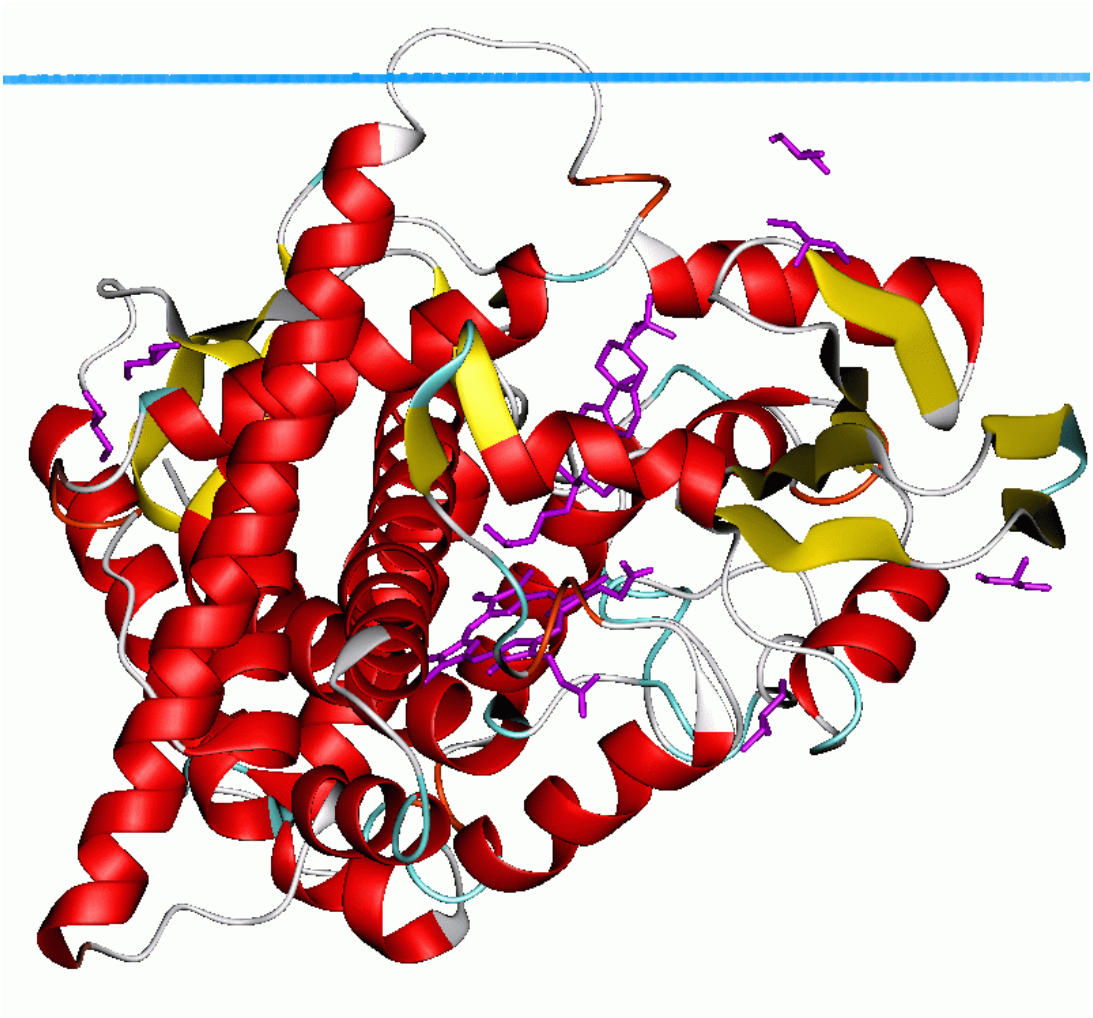


Figure 1.3: Structure of a Cytochrome P450. Human CYP46a1 showing it binding to a membrane (blue line). Alpha helices are shown as red coils and beta sheets as yellow ribbons. From: http://opm.phar.umich.edu/ protein.php?pdbid=2q9f

Chapter 2: General Parasitological and Molecular Biology Materials and Methods

This chapter describes general parasitological and molecular biology techniques used in the completion of more than one experiment. Specific materials and methods for real-time PCR, Roche 454 sequencing and Suppression Subtractive Hybridisation can be found in the relevant chapters.

2.1: General parasitological techniques

Sheep were euthanized according to Home Office guidelines by individuals holding a Personal Licence permitting them to do so. All animal experimental procedures listed below had been previously reviewed and approved by the Moredun Internal Ethics Committee prior to commencement.

2.1.1: *T. circumcincta* isolates

The cDNA initially used for identification of genes was L_4 and xL_3 SMART cDNA, and single-stranded adult cDNA, kindly provided by Dr A. J. Nisbet and from the known IVM susceptible *T. circumcincta* MTci2 (CVL) isolate. The CVL isolate was isolated prior to the use of LEV and IVM and, in an EHT for BZ resistance status, was shown to have an estimated ED_{50} of 0.09µg/mL where the cut-off for resistance is 0.1µg/mL, indicating that it is also BZ-susceptible (D. Bartley, Pers. Comm.). For the synthesis of SMART cDNA, *T. circumcincta* L_3 parasites of the same IVM susceptible CVL strain were used. For use in the real-time PCR experiments, three isolates were used; the CVL isolate, a triple IVM, BZ and LEV resistant isolate (MOTRI or MTci5) and finally a drug-exposed isolate. Controlled efficacy and faecal egg count reduction tests revealed that the efficacies of FBZ, LEV and IVM against the MOTRI isolate were 59%, 88% and 60%, respectively (Bartley *et al.*, 2004). To obtain the drug-exposed isolate, the MOTRI isolate was exposed to a full therapeutic dose of IVM (Oramec®; Merial) at a dose of 0.25mg/kg body weight two days after infection. This isolate was subsequently named Post IVM MOTRI. Eggs, L_1, xL_3, L_4 and adults were obtained from each of the three isolates. MOTRI adult parasites were used in the *in vitro* drug exposure experiment prior

to next generation sequencing. For the suppression subtractive hybridisation work, MOTRI and CVL adults which had previously been collected and stored in liquid nitrogen were used. *In vivo* IVM survivors were obtained by infecting adult sheep with 30,000 MOTRI xL_3 and then giving the sheep a full therapeutic dose of IVM (Oramec®; Merial) at a dose of 0.25mg/kg body weight 28 days post infection. Three days after IVM treatment the sheep were euthanized and adult parasites collected using the method described below.

2.1.2: Recovery of eggs from faeces

Sheep were orally infected with 10,000 to 15,000 *T. circumcincta* L_3. Faeces were collected by attaching a collection bag via a harness to the rump of male sheep. Eggs were extracted from the faeces by emulsifying the faeces in 5mL tap water per gram of faeces. The resulting suspension was washed through a series of sieves with mesh sizes of 1mm, 212µm, 68µm and 38µm, respectively. The retentate, containing the parasite eggs, was collected and washed from the 38µm sieve, transferred to centrifuge tubes and centrifuged at 200 x g for 2 mins. The supernatant was removed and the pellet re-suspended in saturated sodium chloride solution before being centrifuged for a further 2 mins at 200 x g. Using forceps, the meniscus of the tubes was clamped off and poured into a beaker then transferred to the 38µm sieve and washed with tap water. The retentate was collected using tap water and centrifuged for 2 mins at 200 x g with the supernatant being removed. Tap water was used to return the volume in the centrifuge tube to 10mL and an egg count performed. The eggs were centrifuged, the supernatant removed and the eggs re-suspended in 1X Phosphate Buffered Saline (PBS) before being stored in liquid nitrogen.

2.1.3: Collection of L_1

Eggs extracted from faeces as described above were used to obtain L_1 parasites by culturing the eggs according to the following procedure. Eggs were incubated for 18-24 hours at 25°C and then transferred into the top of a mini Baermann apparatus with 20-25µm mesh and cultured for a further 2 hours at 25°C. Hatched L_1 were collected from the well of a 6 well plate following their migration through the Baermann apparatus during the 2 hour incubation period. L_1 were stored in 1X PBS in liquid nitrogen.

2.1.4: Recovery of L_3 from faeces

L_3 were obtained by coproculture. Faeces were collected as described above for recovery of eggs. Macerated faecal material from *T. circumcincta* infected sheep was mixed with vermiculite at a ratio of 10g faeces to 4g vermiculite, placed in a sample container and covered with a plastic bag containing holes to allow gas exchange. The faeces were incubated at 20-27°C for 10 days, the plastic was removed and the vermiculite faeces mix flooded with tepid water and incubated at room temperature for 4 hours. The supernatant was then removed and L_3 extracted using a Baermann apparatus. The L_3 were exsheathed by addition of a 1% sodium hypochlorite solution (Milton sterilising fluid) and observed under a microscope until 90% exsheathment had occurred. Sodium hypochlorite was removed from the solution by repeatedly re-suspending the larvae in water, centrifuging at 200 x g for 2 mins and removing the supernatant. This procedure was repeated at least three times and the parasites were re-suspended in 1X PBS, then stored in liquid nitrogen.

2.1.5: Recovery of L_4 from abomasum

Sheep orally dosed with 150,000 L_3 were used to obtain L_4 from the abomasum. Seven days after infection, the sheep was euthanized and the abomasum removed, cut along the greater curvature, and the contents washed out by rinsing three times with warm saline. The abomasum was then pinned out onto a polystyrene board and placed in a large funnel containing warm saline so that the internal abomasal surface hung down into the saline. The abomasum was left at 37°C for several hours allowing the L_4 to migrate to the bottom of the funnel where they were collected, re-suspended in 1X PBS and transferred to liquid nitrogen for long term storage.

2.1.6: Recovery of adult parasites from abomasum

Adult *T. circumcincta* were collected from the abomasum of sheep 21 days post infection with 30,000 to 50,000 L_3. The abomasum was removed from an infected sheep and opened up along the greater curvature of the abomasum, collecting the abomasal contents. The surface of the abomasum was washed with as little warm saline as possible,

with the abomasal washings in saline being collected with the abomasal contents. 600mL of the abomasal contents and washings was mixed with 600mL of 1.8% agarose diluted in saline, and allowed to set over a wire mesh at the bottom of a bucket. Once set, the bucket with the agarose-abomasal mix was placed in a large funnel containing warm saline and incubated at 37°C for several hours to allow the adults to migrate out of the agarose-abomasal mix into the saline where they were collected and transferred to liquid nitrogen for long term storage 1X PBS.

2.2: *In vitro* bioassays

2.2.1: Larval migration inhibition test (LMIT): Standard technique

The LMIT works by incubating L_3 in serial dilutions of IVM and determining the percentage of L_3 which are capable of migrating through a mesh submerged in a solution containing IVM, as shown in Figure 2.1, part A. Isolates of parasites most resistant to the paralysing effects of IVM will be able to migrate at a greater percentage in a higher dose of drug compared to those less resistant isolates (Wagland *et al.*, 1992; Rabel, McGregor, & Douch, 1994).

Approximately 2400 L_3 parasites, obtained via coproculture as described previously, were allowed to migrate through a Baermann chamber for 1 hour at 26°C using a double thickness piece of 25µm mesh, attached to a section of a syringe, placed into a well of a 6-well plate containing 10mL tap water. After the 1 hour incubation was completed, the tap water containing the migrated L_3 was transferred to a centrifuge tube and centrifuged for 2 mins at 200 x g. The volume in the tube was reduced, the L_3 were re-suspended and a second count made to determine the number of L_3. The volume in the centrifuge tube was adjusted to give between 10 and 15 L_3 per 10µL. A serial dilution of IVM in Dimethyl Sulphoxide (DMSO) was prepared in darkened Eppendorf tubes from a stock solution (3000µg/mL) to give four working concentrations of 300µg/mL, 60µg/mL, 20µg/mL and 5µg/mL. In clean darkened Eppendorf tubes, 10µL of each drug concentration or neat DMSO (control) was added to 190µL of the L_3 suspension to give final concentrations of IVM of 150µg/mL, 15µg/mL, 3µg/mL, 1µg/mL, 0.25µg/mL and 0µg/mL. The L_3 were incubated for 2 hours at 26°C in the above IVM concentrations.

A 24-well plate was set up to contain the migration chambers made from small plastic tubes covered at one end with a 20μm mesh. The layout of the plate is shown in Figure 2.1, part B. 210μL of the stock or working concentrations of IVM was mixed with 3801μL tap water in a darkened Bijou to provide 1910μL of solution to fill each of the duplicate wells for each drug concentration. This was calculated to give the same final concentrations of IVM as previously described. 90μL of the drug exposed L_3 were added down the side of the migration chamber into the well above the mesh and the plate incubated for a further 2 hours in the dark at 26°C. Migrated L_3 were washed off the bottom of the migration chamber back into the well, the migration chamber was tipped upside down into the corresponding empty well below it on the plate and non-migrating L_3 washed back through the mesh. The L_3 were stained with helminthological iodine and counted under an inverted microscope at x100 magnification. Percentage migration for each well was determined and an average for each drug concentration calculated.

2.2.2: Larval migration inhibition test (LMIT): Modified technique

The LMIT technique described above was modified to enable the protocol to be scaled up to give sufficient quantities of L_3 for use in molecular experiments. Following several experiments using the standard LMIT method, a dose of 2.5μg/mL IVM was chosen as the previous results suggested this would give a migration percentage of approximately 10%. However, using the protocol described below this resulted in an average migration of 0.41%, so subsequently a dose of 1μg/mL IVM was used, which had been shown to give an average migration of 24% in the standard migration technique.

The first Baermannisation step (1 hour at 26°C using a 25μm mesh) was not performed, instead the L_3 were exsheathed by addition of a 1% sodium hypochlorite solution (Milton sterilising fluid) into a flask containing the L_3 in 70mL water. A subsample of the L_3 were observed under a microscope, and when 90% had exsheathed, the L_3 were washed as described in Section 2.1.4, counted and re-suspended in 240mL warm tap water.

A serial dilution of IVM in DMSO was prepared in darkened Eppendorf tubes from a stock solution (3000µg/mL) to give a final working concentration of 50µg/mL. In clean darkened Falcon tubes, 2.5mL of IVM or neat DMSO (control) was added to 47.5mL of the L_3 suspension to give final concentrations of IVM of 1µg/mL, and 0µg/mL. The L_3 were incubated in the dark for 2 hours at 26°C. Following incubation, the L_3 were centrifuged for 5 mins at 1500 x g with the brake off to allow the L_3 to pellet to the bottom of the Falcon tubes without vortexing back into suspension. The volume in each Falcon tube was reduced to 10mL and the L_3 were inverted to re-suspend them.

Five migration chambers were set up using clean glass jam jars of identical size, capable of holding 320mL. The Baermann chamber was made using two sample pots with their bottoms cut out, placed one inside the other, holding a piece of 20µm mesh in place between them. These were then placed in the jam jars. Four of the migration chambers were to provide replicates of the drug concentration whilst the fifth was for the no drug control. 65.6mL of the working concentration of IVM was mixed with 1205.4mL warm tap water in a darkened 2L Duran bottle to provide the solution to fill the jam jars. Each IVM jam jar was filled with 310mL of the drug whilst the control jam jar was set up by mixing 294mL of tap water and 16mL DMSO. 10mL of the drug exposed L_3 were added down the side of the sample pot above the mesh and the jam jars incubated for 2 hours at 26°C in the dark.

Migrated L_3 were washed off the bottom of the migration chamber back into the jam jar, the migration chamber was tipped upside down onto a Petri dish and non-migrating L_3 washed back through the mesh after removing the mesh from between the two sample pots to ensure no non-migrating L_3 were caught in the mesh folds. The contents of the jam jars and Petri dishes were sequentially spun down in the centrifuge (2 mins at 200 x g) until each of the migrating and non-migrating samples were re-suspended in a total volume of 10mL. Five 10µL subsamples from the migrating L_3 samples were taken and the number of L_3 present determined under a microscope. From this the percentage migration of each jam jar was calculated. Each sample was reduced in volume and stored in 1X PBS in liquid nitrogen.

2.3: General molecular biological techniques

2.3.1: Extraction of RNA

The Trizol® method was used to extract total RNA from parasite material. Parasites removed from storage in liquid nitrogen were placed in a pre-chilled pestle and mortar (-80°C) and ground to a fine powder using liquid nitrogen to ensure parasites remained frozen. The liquid nitrogen was allowed to evaporate and 8mL Trizol® reagent (Invitrogen) added. Once solidified, the Trizol-parasite mix was ground up until liquid then transferred into 1mL aliquots in Eppendorf tubes and allowed to incubate at 15-30°C for 5 mins. 200µL chloroform was added to each Eppendorf tube and the Eppendorf tubes shaken vigorously and incubated at 15-30°C for 2 mins, then centrifuged at 12000 x g for 15 mins at 2-8°C. The RNA was precipitated from the Trizol mix by transferring the colourless upper aqueous phase from the Eppendorf tube into a clean Eppendorf tube, adding 500µL isopropanol, mixing and incubating for 10 mins at 15-30°C. A second centrifugation at 12000 x g for 10 mins at 2-8°C was performed and the supernatant removed without dislodging the RNA pellet. The RNA was washed by adding 1mL 75% ethanol, vortexing and centrifuging at 7500 x g for 5 mins at 2-8°C. The ethanol was removed and the pellet allowed to air dry to remove all traces of the ethanol. The RNA was re-suspended in up to 75µL nuclease free water and stored at -80°C. Confirmation of RNA quality was obtained by running 5µL on a 1% agarose gel to identify the 18S band and by determining the RNA concentration using the NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific Inc). RNA extracted later in the project had 1µL RNAse OUT (Invitrogen) added prior to storage at -80°C to prevent possible degradation of RNA.

2.3.2: Synthesis of cDNA

To enable the identification of candidate resistance genes, single-stranded cDNA was synthesised from RNA using the Invitrogen Superscript reverse transcriptase kit with oligo dT primers, according to the manufacturer's instructions. Briefly 2µL oligo dT primer was mixed with 5µg total RNA and the reaction made up to 12µL with PCR-grade water, then incubated at 70°C for 10 mins and chilled on ice. 4µL 5x first strand buffer, 2µL 0·1M DTT and 1µL 10mM dNTP mix was added before the reaction mix was

incubated at 42°C for 2 mins. 1µL Superscript II enzyme was added and the reaction mix incubated at 25°C for 10 mins then at 42°C for 50 mins. The enzyme was heat inactivated at 70°C for 10 mins and the resulting cDNA stored at -20°C.

The Switching Mechanism At 5' End of RNA Template (SMART™) Rapid Amplification of cDNA ends (RACE) cDNA amplification kit (Clontech) and the Advantage® 2 PCR Enzyme System (Clontech) were used to manufacture RACE-ready cDNA. Two separate reactions were performed to make 5' and 3' RACE-ready cDNA. To manufacture 5' RACE-ready cDNA 1µL of RNA was mixed with 1µL 5'-CDS primer and 1µL SMART II A oligo and the final volume made up to 5µL with sterile water. 3' RACE-ready cDNA was prepared by adding 1µL of RNA to 1µL of 3'-CDS primer A and the final volume made up to 5µL with sterile water. Both reactions were incubated at 70°C for 2 mins and then cooled on ice for 2 mins. To both the 5' and 3' reaction mixes, 2µL of 5X First-Strand buffer, 1µL 20mM DTT, 1µL 10mM dNTP Mix and 1µL MMLV Reverse Transcriptase was added and the reaction mix incubated at 42°C for 90 mins. The starting concentration of the total RNA was greater than 200ng/µL so 100µL of Tricine-EDTA Buffer was added to each reaction and the tubes incubated at 72°C for 7 mins. Aliquots of the RACE-ready cDNA were made and stored at 20°C.

cDNA for use in real-time PCR was manufactured using the Invitrogen Superscript reverse transcriptase kit as described above with the following modifications: The concentration of RNA added in step 1 was 1.5µg total RNA and 2µL random primers were used instead of oligo dT primers in the first strand synthesis reaction. Triplicate cDNA synthesis reactions were set up for each isolate and life-cycle stage with the three reactions pooled and mixed prior to the cDNA concentration being determined using the NanoDrop® ND-1000 spectrophotometer. The cDNA was diluted to 50ng/µL and aliquots made prior to storage at -20°C.

2.3.3: Standard PCR

Four different types of primers were used: housekeeping gene primers (Table 2.1), degenerate primers to identify Pgps and CYPs (Table 2.2), specific primers to amplify Pgp and CYP sequence (Table 2.3) and standard primers for amplification of sequence in cloning vectors (Table 2.4). Primers were designed using the services on MWG Biotech and Primer3 (www.eurofinsdna.com and http://frodo.wi.mit.edu/, respectively) to calculate Tm. Primers were ordered from MWG Biotech and diluted to a stock concentration of 100μM. The working concentration of all primers for PCR was 10μM.

Standard PCR used the Platinum® *Taq* DNA polymerase kit (Invitrogen) and dNTPs (Roche). Standard PCR reactions were carried out by combining 12·75μL dH_2O with 2·5μL 10X PCR buffer, 0·75μL 50mM $MgCl_2$, 1·0μL 10mM dNTP, 0·5μL Platinum® *Taq* DNA Polymerase and 2·5μL each of the forward and reverse primers and template to give a total reaction volume of 25μL. All PCR reactions were run on an Applied Biosystems 2720 thermal cycler as follows: 5 mins at 94°C for a denaturing step followed by 40 cycles of 94°C for 30 secs, X°C for 30 secs and 72°C for 30 secs, (where X is the Tm of the primers minus 5°C), followed by 10 mins at 72°C and then a hold at 4°C. The PCR results were visualised by running the whole PCR reaction mix on a 1% agarose gel as described below.

2.3.4: Rapid amplification of cDNA ends (RACE) PCR

RACE PCR primers (Table 2.5) were designed, as described above, to extend the 5' gene sequence of specific Pgps. RACE PCR reactions were set up by combining 34·5μL PCR grade water with 5μL 10X Advantage 2 PCR buffer, 1μL 10mM dNTP mix, 1μL 50X Advantage 2 Polymerase Mix, 2·5μL 5' RACE ready cDNA, 5μL 10X Universal Primer Mix and 1μL of the specific primer to give a total reaction volume of 50μL. A touchdown PCR protocol was used as follows: 5 cycles of 94°C for 30 secs and 72°C for 3 mins followed by 5 cycles of 94°C for 30 secs, 70°C for 30 secs and 72°C for 3 mins followed by 25 cycles of 94°C for 30 secs, 68°C for 30 secs and 72°C for 3 mins followed by a hold at 72°C for 7 mins and then a hold at 4°C. The results of the RACE PCR were visualised as for the standard PCR, as described below. PCR product sizes on agarose gels were

estimated by comparing the position of the band on the gel to the sizes of bands of the DNA Molecular Weight Marker X (Roche) which has a size range of 0.07 to 12.2 kbp.

2.3.5: Semi-quantitative PCR

Semi-quantitative PCR was performed by setting up a PCR using the gene-specific primers designed for CYPs 1, 2 and 3 and PGPs 2, 3, 5, 6, 7 and 9 (Table 2.3), the actin and β-tubulin control primers (Table 2.1) and cDNA generated using random primers from eggs, L_1, xL_3, L_4 and adults of the CVL and MOTRI *T. circumcincta* isolates as follows: 36µL of nuclease-free water was combined with 5µL 10X PCR buffer, 1.5µL 50mM $MgCl_2$, 1·0µL dNTP, 0·5µL Platinum® *Taq* DNA Polymerase, 1µL each of the forward and reverse primers, and 4µL template cDNA diluted to 250ng/µL to give a total reaction volume of 50µL. The PCR reaction conditions were 5 mins at 94°C followed by 30 cycles of 94°C for 30 secs, 55°C for 30 secs and 72°C for 30 secs, followed by 10 mins at 72°C and then held at 4°C. After 15, 20, 25 cycles and upon completion of the PCR, 5µL of the reaction mix was removed from each tube and stored at 4°C prior to all the samples being visualised on an agarose gel as described below.

2.3.6: Visualisation of PCR products:

To visualise PCR products between 5µL and the full volume of the PCR reaction was loaded onto a 1% agarose gel (containing TAE and Gel Red as described below) which had been submerged in an electrophoresis tank filled with 1 x TAE buffer. Electrophoresis was carried out on the 1% agarose gels at 70-110V for between 30 to 60 mins to allow separation of any PCR products according to molecular weight. The gel was examined under UV light and the size of the bands present determined against DNA Molecular Weight Marker X (Roche) which has a size range of 0.07 to 12.2 kb or against the TrackIt™ 1Kb Plus DNA Ladder (Invitrogen) which has a size range of 0.1 to 12 kb. Reagents used were prepared as follows:

Tris-acetic acid-ethylenediaminetetraacetic acid (TAE) buffer

A 50X stock solution of TAE was prepared by dissolving 242g Tris base in 500mL distilled water (dH_2O), to which 57.1mL glacial acetic acid and 100mL 0·5M ethylenediaminetetraacetic acid (EDTA) (pH8) was added and the volume topped up to 1L with dH_2O. For use as an agarose gel tank buffer the 50X stock solution was diluted to 1X in dH_2O.

Agarose gel

1% agarose gel was prepared by dissolving 4g agarose and 8mL 50X TAE in 392mL of dH_2O using a microwave. 20µL 10000X Gel Red (Biotium Inc) was added to enable visualisation of DNA under UV light.

2.3.7: Cloning of PCR products

Bands visualised under UV on 1% agarose gels were excised using a clean scalpel and purified following the protocol for the QIAquick Gel Extraction Kit (Qiagen). Briefly, the gel slice was dissolved in Buffer QG at a ratio of 300µL per 100mg of gel in an Eppendorf tube placed in a water bath at 50°C. The DNA in Buffer QG was bound to the spin column by centrifugation at 17900 x g for 1 minute. The DNA on the column was washed by addition of 500µL of buffer QG, centrifuged for 1 minute, addition of 750µL of Buffer PE and then centrifuged for 1 minute. Waste was removed from the collection tube between each step and a final centrifugation step of 1 minute carried out to remove all residues of Buffer PE. Purified DNA was eluted by adding 30µL dH_2O to the spin column, which was allowed to stand for 1 minute before a final centrifugation for 1 minute. The eluted DNA was stored at -20°C.

Purified PCR product was ligated into pGEM®-T vector on ice as follows: 5µL 2X ligation buffer was mixed with 1µL pGEM®-T vector, 1µL T4 DNA ligase and 3µL of purified PCR product. The ligation mixture was incubated at 4°C overnight. On day two, JM109 competent cells (Stratagene) were transformed with the pGEM®-T vector

containing the purified PCR product as follows: On ice, 3µL of the ligation mix was added to 27µL of JM109 competent cells and incubated on ice for 30 mins. The cells were heat-shocked at 42°C in a water bath for 1 minute and returned to ice for 2 mins. 200µL of SOC medium was added to each Eppendorf tube and the cells placed in a shaking incubator at 37°C for two hours. After two hours, 150µL of the cell suspension was spread onto LB AMP XI plates (manufactured as described below) using an ethanol-sterilised glass spreader and the plates placed in a 37°C incubator for overnight culture. White colonies, containing the pGEM®-T vector with purified PCR product insert, were selected from the plates using sterile pipette tips and placed in glass universals, containing 10mL LB broth and 20µL 25mg/mL AMP, and cultured overnight at 37°C in the shaking incubator.

Overnight cultures were pelleted by pouring off the supernatant following centrifugation of the universals at 3000 x g for 15 mins. The plasmids were extracted from the cells following the protocol for the Wizard® Plus SV Miniprep DNA purification kit (Promega). Briefly, the cells were re-suspended in 250µL Cell Resuspension Solution and transferred to an Eppendorf tube. The cells were lysed by addition of 250µL Cell Lysis solution, mixing, addition of 10µL Alkaline Protease Solution, mixing and incubating at room temperature for 5 mins. The reaction was neutralized by addition of 350µL of Neutralization Solution and then centrifuged at 14000 x g for 10 mins at room temperature to pellet the cell debris. The supernatant was carefully removed from the Eppendorf tube and placed into the spin column. Binding of the plasmid to the column was achieved by centrifugation of the spin column at 14000 x g for 1 minute, with the waste from the column being removed from the collection tube after each step. Two washes of the column were carried out by adding 750µL of Column Wash Solution, centrifuging for 1 minute, adding 250µL Column Wash Solution and centrifuging for 1 minute, with the waste being removed from the collection tube after each step. A final centrifugation step for 2 mins at 14000 x g removed all traces of Column Wash Solution. The purified plasmid was eluted from the column by adding 50µL dH_2O and centrifuging at 14000 x g for 1 minute. The extracted plasmid DNA concentration was determined using the NanoDrop® ND-1000 spectrophotometer and the purified plasmid stored at -20°C. Reagents used for cloning were prepared or obtained as shown below:

Cells

JM109 competent cells (Stratagene) and the pGEM®-T vector system (Promega) were used to clone PCR products.

Luria Bertani (LB) broth

LB broth was prepared by dissolving in 1L dH_2O, 10g Bacto®-tryptone, 5g Bacto®-yeast extract and 5g sodium chloride. Sterilisation was achieved through autoclaving at 121°C for 15 mins.

Luria Bertani (LB) agar

LB agar was prepared from LB broth by dissolving 15g Bacto®-agar in 1L LB broth and autoclaving.

SOC medium

SOC medium (pH7) was prepared as follows. 1g Bacto®-tryptone, 0·25g Bacto®-yeast extract, 0·5mL 1M sodium chloride and 0·125mL 1M potassium chloride was dissolved in 48·5mL dH_2O, autoclaved and allowed to cool. 0·5mL 2M filter-sterilised Mg^{2+} and 0·5mL 2M filter-sterilised glucose was added and the final volume increased to 50mL with dH_2O. The pH was adjusted to pH7 and the SOC medium divided into 5mL aliquots and stored at -20°C.

Ampicillin solution

A 25mg/mL stock solution of ampicillin (AMP) was prepared by dissolving a 25mg ampicillin tablet (Stratagene) in 1mL dH_2O and was stored at -20°C.

Isopropyl β-D-1-thiogalactopyranoside (IPTG)

1M IPTG was made by adding 238mg IPTG to 1mL dH_2O, vortexing and storing at -20°C.

5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal)

50mg/mL X-gal in dimethylformamide was obtained from Promega.

LB agar plates containing ampicillin, X-gal and IPTG

LB AMP XI plates were prepared by melting LB agar and cooling to 50°C in a water bath. 1mL of ampicillin stock solution was added per 500mL LB agar, mixed and then poured into 9cm diameter Petri dishes and left to set. Plates were stored upside down at 4°C until required, for a maximum of 2 weeks. Immediately prior to use, 40μL of 1M IPTG and 40μL of X-gal was spread on the surface of the LB agar using an ethanol-sterilised glass spreader and allowed to absorb into the LB agar at 37°C.

2.3.8: Kits for purification of PCR and cloning products

Purification of PCR products from agarose gels was performed using the QIAquick® Gel Extraction Kit (Qiagen). Purification of the remaining PCR products, which had not been run on agarose gels, was performed using the QIAquick® PCR purification kit (Qiagen). Extraction of plasmids from JM109 cell cultures was performed using the Wizard® Plus SV Miniprep DNA purification system (Promega).

2.3.9: Sequencing

Purified plasmids were submitted for sequencing to either MWG or GATC according to the instructions specified on the companies' websites (www.eurofinsdna.com and www.gatc-biotech.com, respectively). The primers chosen were generally either M13

For, M13 Rev, M13 Uni, SP6 or T7, as shown in Table 2.4, which amplify the pGEM®-T vector across the insertion site. The majority of sequencing used the M13 Uni, SP6 and T7 primers. Occasionally, purified PCR products were sequenced directly without first being cloned into the pGEM®-T vector. In these cases, the primer chosen for the sequencing reaction was the gene specific forward primer for that PCR product.

2.3.10: Analysis of sequencing data

Sequence data received in ABI format following cloning and sequencing was analysed using SeqMan (DNASTAR Lasergene version 8) to enable the accuracy of base calls to be checked, to align sequences into contigs and to remove plasmid sequence. Consensus sequence was used as the query in BLAST X search engines on EMBL-EBI (www.ebi.ac.uk/blast2/parasites.html) and NCBI (www.ncbi.nlm.nih.gov/blast/Blast.cgi) (Altschul *et al.*, 1990). Novel sequences generated using the degenerate primer, cloning and sequencing approach with homology to Pgps or CYPs from these databases were numbered sequentially. Translated protein sequences and open reading frames were generated using The Sequence Manipulation Suite (http://www.ualberta.ca/~stothard/javascript/index.html). Clustal W analysis, to generate alignment files for both nucleotide and protein sequence, was carried out using EMBL-EBI (http://www.ebi.ac.uk/Tools/clustalw2/).

Electronic searches to identify potential Pgp and CYP ESTs in *T. circumcincta* were carried out on the following webpages: Nembase2; (http://xyala.cap.ed.ac.uk/nematodeESTs/search.php) (Parkinson *et al.*, 2004a), NCBI; (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore&itool=toolbar) (Wheeler *et al.*, 2008) and EMBL-EBI (http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-page+top). Keywords such as P-glycoprotein, Pgp, Cytochrome P450, CYP450 and *Teladorsagia circumcincta* were used as search terms.

Individually, the novel Pgp nucleotide sequences, generated using the degenerate primers as described above, were aligned against the 15 full-length *C. elegans* Pgp

nucleotide sequences using the very accurate alignment option in CLC DNA Workbench version 5.6.1 (www.clcbio.com). The names and GenBank accession numbers of the protein sequences of the 15 *C. elegans* Pgps are shown in Table 2.6 below and were obtained from Sheps *et al* (, 2004). The link to the protein sequence was used to trace back to the nucleotide sequence. The alignments were visually inspected and trimmed to only display the part of the alignment containing the *T. circumcincta* Pgp sequence and then used to generate phylogenetic trees to enable the closest matching *C elegans* Pgp sequence to be identified for each individual *T. circumcincta* Pgp. This was to enable the correct naming of the Pgps as suggested in the guidelines on the *Caenorhabditis* genetics centre website (http://www.cbs.umn.edu/CGC/nomenclature/index.html). Phylogenetic trees were generated using the neighbour-joining (NJ) and Un-weighted Pair Group Method with Arithmetic mean (UPGMA) methods using a bootstrap value of 1000.

2.3.11: Identification of single nucleotide polymorphisms (SNPs)

Using cDNA generated from pools of adult *T. circumcincta*, SNPs were identified by using gene-specific primers (Table 2.3) to selectively amplify cDNA for PGPs 3, 5, 6, 7 and 9, using the SuperScript™ III One-Step RT-PCR System with Platinum® *Taq* DNA Polymerase. The actin primers (Table 2.1) were used as positive controls. The template was CVL and MOTRI adult RNA extracted as described in Section 2.3.1, and the reaction was set up as follows: 25µL of the 2X Reaction Mix was combined with 1µL of the Reverse Transcriptase/ Platinum® *Taq* Mix, 1µL each of the forward and reverse primers and 21µL dH2O. 1µL RNA was added to give a total reaction volume of 50µL and the PCR reactions run on an Applied Biosystems 2720 thermal cycler programmed as follows: 50ºC for 30 mins, 94ºC for 2 mins followed by 40 cycles of 94ºC for 15 secs, 55ºC for 30 secs and 72ºC for 1 minute, followed by 72ºC for 10 mins and a hold at 4ºC. Following the identification of bands on an agarose gel loaded with 10µL PCR product (Section 2.3.6 above), the remaining 40µL PCR product was cleaned up using the QIAquick® PCR purification kit (Qiagen), eluting into 30µL water. The RNA concentration was determined using the NanoDrop® ND-1000 spectrophotometer and the samples submitted for sequencing as described above (Section 2.3.9).

Analysis of the sequence data was carried out by first confirming the sequence was for the correct Pgp by aligning it against the Pgp consensus sequences in Lasergene Seqman. Subsequently, the pairs of sequences for each Pgp in raw .abi file format (CVL and MOTRI) were entered into the online SeqDoC programme (http://research.imb.uq.edu.au/seqdoc/) to identify the presence of SNPs by comparing the sequence chromatograms. Where SNPs were observed, the individual nucleotide sequences from both CVL and MOTRI were translated into protein sequence using the Sequence Manipulation Suite (http://www.bioinformatics.org/sms2/). The nucleotide and translated protein sequences were used to determine whether any of the SNPs represented coding changes by carrying out Clustal W analysis using EMBL-EBI (http://www.ebi.ac.uk/Tools/clustalw2/).

Table 2.1: Primers used to amplify *T. circumcincta* housekeeping genes.

Primer name	**Primer direction**	**Sequence**	**Melting temp (°C)**	**Target**
Tci actin F	Sense	GTT GCT GCT CTT GTG GTT GA	57.3	*T. circumcincta* actin
Tci actin R	Antisense	GGA GAG CAC AGC CTG GAT AG	61.4	*T. circumcincta* actin
Tci 60S F	Sense	AAA CAA TGG GTC GCC GGA	56.0	*T. circumcincta* 60S
Tci 60S R	Antisense	GCA GCA ACA ACG GTG CT	55.2	*T. circumcincta* 60S
Tci_BTubF	Sense	TTCCATTCCCTCGTCTTCAC	60.0	*T.circumcincta* β-tubulin
Tci_BTubR	Antisense	AGCCATTTTCAATCCACGAG	60.0	*T.circumcincta* β-tubulin

Table 2.2: Degenerate primers used to identify Pgps and CYPs from *T. circumcincta* cDNA. * means the Tm was calculated approximately using W instead of I.

Primer name	Primer direction	Sequence	Melting temp (°C)	Target	Source
1S	Sense	AGY GGI TGY GGN AAR AGY AC	58.3*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
2S	Sense	AGY GGI TGY GGN AAR TCN AC	58.3*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
3S	Sense	TCI GGI TGY GGN AAR AGY AC	57.3*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
4S	Sense	TCI GGI TGY GGN AAR TCN AC	57.3*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
1A	Antisense	AGI GCI GAN GTN GCY TCR TC	59.4*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
2A	Antisense	AGI GCR CTN GTN GCY TCR TC	60.4*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
3A	Antisense	AAI GCI GAN GTN GCY TCR TC	57.3*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
4A	Antisense	AAI GCR CTN GTN GCY TCR TC	58.3*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
5A	Antisense	ATI GCD ATI CGY TGY TTY TGN CC	59.5*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
6A	Antisense	ATI GCD ATY CTY TGY TTY TGN CC	58.6*	*H. contortus* Pgp	(Sangster *et al.*, 1999b)
13A-1	Sense	CAA GAM GAA GTR GAY ARA GAA TG	57.1	*C. elegans* CYP13A	(Menzel, Bogaert, & Achazi, 2001)
13A-2	Antisense	CCC ATC TYT CIG GYY TRA AC	57.3		
31A-1	Sense	TCA AGC GGA ACT GGA TGA AG	57.3	*C. elegans* CYP31A	(Menzel, Bogaert, & Achazi, 2001)
31A-2	Antisense	GAG CCA TGA TGA CCT TCT CT	57.3		
33C-1	Sense	ATG CAK GAA AAI ATY CTD ATD GA	52.9	*C. elegans* CYP33C	(Menzel, Bogaert, & Achazi, 2001)
33C-2	Antisense	TAM GGW AGA TYA TTY TTA TCA	49.1		
35A-1	Sense	ATA TGG GRG TTG GDA ARG AT	53.9	*C. elegans* CYP35A	(Menzel, Bogaert, & Achazi, 2001)
35A-2	Antisense	GCA TGR CGT TGA ATY TCT CC	57.3		
DegenTcP450For	Sense	TTG AYY TWT GGA TWG CWG GTC A	56.5	*T. circumcincta* CYP EST	
CypGrp2 Deg For	Sense	TGG NTN GCT GGT MWD GAR AC	58.0	*T. circumcincta* CYP clan2	
CypGrp2 Deg Rev	Antisense	TCC TKW CGA GCN ANC CAH ARA TC	61.2		

Primer name	Primer direction	Sequence	Melting temp (^{o}C)	Target	Source
CypGrp3 Deg For	Sense	GGT GTR GGW CCW MGR CAD TGT	61.4	*T. circumcincta* CYP clan3	
CypGrp3 Deg Rev	Antisense	ACA CAH TGY CKW GGW CCY ACA	59.5		
CypGrp4 Deg For	Sense	TGG ADA VRT CGV CGA AAR ATG	57.2	*T. circumcincta* CYP clan4	
CypGrp4 Deg Rev	Antisense	TGG DGT NAD CAT YTT TCG BCG A	59.0		

Table 2.3: Gene specific *T. circumcincta* primers.

Primer name	Primer direction	Sequence	Melting temp (°C)	Target	Amplicon length (bp)	Specificity
Pgp1Tc Rev	Antisense	GCT CAA TTT CCT CCT GAG AAC	57.9	*T. circumcincta* PGP1		
Pgp1Tc IntFor	Sense	GAA GGT GGC ATG TAC CGG	61.4	*T. circumcincta* PGP1		
Pgp1Tc IntRev	Antisense	GAG GTT CCT GTG CTG CAG CAC	63.7	*T. circumcincta* PGP1		
P1TcF	Sense	GTG CTG CAG CAC AGG AAC	58.2	*T. circumcincta* PGP1	676	Failed to amplify
P1TcR	Antisense	GCG GCG TGC TCA ATT TCC	58.2			
Pgp2Tc Int For	Sense	CGA AAA CAT CAG AAA CAT GA	51.2	*T. circumcincta* PGP2		
Pgp2Tc Int Rev	Antisense	AAT GTT CTC AAG GAT CGT ACA	54	*T. circumcincta* PGP2		
P2TcF	Sense	GAA AGA GCC TCA TCG GGA G	58.8	*T. circumcincta* PGP2	649	PGP2 specific
P2TcR	Antisense	GGA AGG TTT AGG GTC GTC C	58.8			
P3TcF	Sense	CCA ATG CAT CCC TGG CTC	58.2	*T. circumcincta* PGP3	618	PGP3 specific
P3TcR	Antisense	CCT CGT CTG TGG CTG AC	57.6			
Pgp4Tc Int For	Sense	ACG GTG TGG AAA TCG ACA AG	57.3	*T. circumcincta* PGP4		
P4TcF	Sense	GGA AGA CGG CAA ATC GAC	56	*T. circumcincta* PGP4	662	Not specific
Pgp4Tc Int Rev	Antisense	CCG TGA TTT CAG CAT CTG TG	57.3			
P5TcF	Sense	CAA GGT TTG AGG CAG CGT TG	59.4	*T. circumcincta* PGP5	688	PGP5 specific
P5TcR	Antisense	GCT TCC TGG ACC TCT TCC	58.2			
P6TcF	Sense	GCY TTG GAC GGT TCT GTG G	59.9	*T. circumcincta* PGP6	259	PGP6 specific
P6TcR	Antisense	GCG CTC CCT TTT CTC CTG	58.2			
P7TcF	Sense	GTG CGT ACG CCG ACA AG	57.6	*T. circumcincta* PGP7	699	Not specific, redesigned
P7TcR	Antisense	CGT TTG ACT GCT CTG CAG C	58.8			

Primer name	Primer direction	Sequence	Melting temp (^{o}C)	Target	Amplicon length (bp)	Specificity
P7TcF2	Sense	CGT ACG CCG ACA AGA TTT ACG	59.8	*T. circumcincta* PGP7	663	PGP7 specific
P7TcR2	Antisense	CCT CGG TGA TTT TAT CGG GG	59.4			
P8TcF	Sense	CTG TGG CTG GTG CTG TG	57.6	*T. circumcincta* PGP8	315	Not specific, redesigned
P8TcR	Antisense	CTT TCG GGT TAC GGA TCA C	56.7			
P8TcF2	Sense	GTT CGA TGA AGT GGA TGC TC	57.3	*T. circumcincta* PGP8	246	Not specific
P8TcR2	Antisense	CCA CCG GAC AGC ATG CTA	58.2			
P9TcF	Sense	CTT GTT GGA CTA CTT CTT CGC	57.9	*T. circumcincta* PGP9	352	PGP9 specific
P9TcR	Antisense	GCA GAA TTC GCG GAT TTC TC	57.3			
P10TcF	Sense	CGC CAG TTT ATT GAT GGG C	56.7	*T. circumcincta* PGP10	331	Not specific, redesigned
P10TcR	Antisense	CTC GAC CAG CAC GCG C	59.4			
P10TcF2	Sense	TAT TGA TGG GCT ATT ACC CGC	57.9	*T. circumcincta* PGP10	283	Not specific
P10TcR2	Antisense	CCA ACT GAG AGA TTA TTC CCC	57.9			
P11TcF	Sense	CGT TGT GTA AAC CTG CTG G	56.7	*T. circumcincta* PGP11	347	Not specific, redesigned
P11TcR	Antisense	CAC AGC AAT ACT TTG GGA TTG C	58.4			
P11TcF2	Sense	CAG GAA CTG ACC ACG CTG	58.2	*T. circumcincta* PGP11	213	Not specific
P11TcR2	Antisense	GTA GCC ATC ATG CTT ATC GC	57.3			
Cyp1For	Sense	CAG CAC GGG TAG AGG AG	57.6	*T. circumcincta* CYP1	366	CYP1 specific
Cyp1Rev	Antisense	GTT CTG CYT GAG CAA GTG	54.8			
Cyp2For	Sense	CAA AAA TCC TGA GGT CAT GC	55.2	*T. circumcincta* CYP2	323	CYP2 specific
Cyp2Rev	Antisense	GCT GTT CCA GTG TCT TAC C	56.7			
Cyp3For	Sense	CGT ATG GTT GTT CTA TGC G	54.5	*T. circumcincta* CYP3	467	CYP3 specific
Cyp3Rev	Antisense	CCT GTA ATG GTC TTC GTA G	54.5			

Table 2.4: Standard primers used to amplify sequence from cloned cDNA.

Primer name	**Primer direction**	**Sequence**	**Melting temp (^{o}C)**	**Target**
T7	Sense	TAA TAC GAC TCA CTA TAG GG	53.2	pGEM®-T vector
SP6	Antisense	CAT TTA GGT GAC ACT ATA G	50.2	pGEM®-T vector
M13 For	Sense	AGC GGA TAA CAA TTT CAC ACA GGA	47.0	pGEM®-T vector
M13 Rev	Antisense	CGC CAG GGT TTT CCC AGT CAC GAC	48.0	pGEM®-T vector
M13 Uni (-21)		TGT AAA ACG ACG GCC AGT	53.7	pGEM®-T vector
TrplX for	Sense	TCG GGA AGC GCG CCA TT	57.6	Phagemid vector
TrplX rev	Antisense	TGC GGC CGC ATG CAT AAG	58.2	Phagemid vector

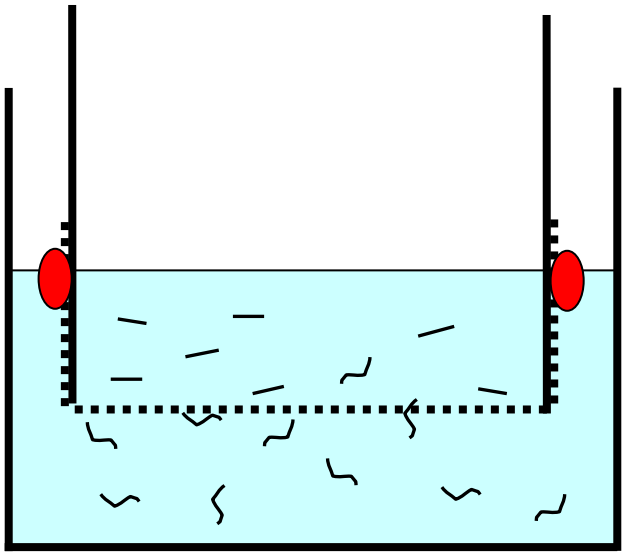
Table 2.5: RACE PCR primers used to extend 5’ sequence of Pgps.

Primer name	**Primer direction**	**Sequence**	**Melting temp (^{o}C)**	**Target**
Pgp1Tc 5'race	Antisense	GTG CGG CGT GCT CAA TTT CCT CCT GAG AAC	70.9	*T. circumcincta* PGP1
Pgp2Tc 5'race	Antisense	CTC GTA GGA AGG TTT AGG GTC GTC CAG TCC	70.9	*T. circumcincta* PGP2
Pgp3Tc 5'race	Antisense	CTC GTC TGT GGC TGA CGG GCG ACC AAA GC	73.7	*T. circumcincta* PGP3
Pgp4Tc 5'race	Antisense	CGG AGT GCC GCC GTG ATT TCA GCA TCT G	71.0	*T. circumcincta* PGP4
Pgp5Tc 5'race	Antisense	GCA ATA CGA CAA GCT TCC TGG ACC TCT TCC	69.5	*T. circumcincta* PGP5
Pgp6Tc 5'race	Antisense	CTA ACG GAW CCT GGW GGR AGR CCG TAG TGG	72.2	*T. circumcincta* PGP6
Pgp7Tc 5'race	Antisense	CGA GAG CGT TTG ACT GCT CTG CAG CCT C	71.0	*T. circumcincta* PGP7
Pgp8Tc 5'race	Antisense	GGG CGT TGG CCA GTC TAG CGG CAT CCT CG	75.2	*T. circumcincta* PGP8
Pgp9Tc 5'race	Antisense	GCC GCG AAA TGT CCA CCC GTC CCA AGC GG	75.2	*T. circumcincta* PGP9
Pgp10Tc 5'race	Antisense	CGC GTT CTT CGG AGA TAT CCC GCC CCA GCG	75.0	*T. circumcincta* PGP10
Pgp11Tc 5'race	Antisense	CCA GCT CCA GCG GCA GAG CCA CGT TGC C	75.4	*T. circumcincta* PGP11

Table 2.6: *C elegans* Pgp genes and accession numbers used for phylogentic analysis. All were obtained from Sheps *et al.*, 2004

Gene name	**GenBank Accession number**
Pgp-1	CAB01232
Pgp-2	AAB52482
Pgp-3	CAA91467
Pgp-4	CAA91463
Pgp-5	CAA94202
Pgp-6	CAA94220
Pgp-7	CAA94219
Pgp-8	CAA94203
Pgp-9	CAB03973
Pgp-10	AAC48149
Pgp-11	CAA88940
Pgp-12	CAA91799
Pgp-13	CAA91800
Pgp-14	CAA91801
Pgp-15	CAA91802

A



B

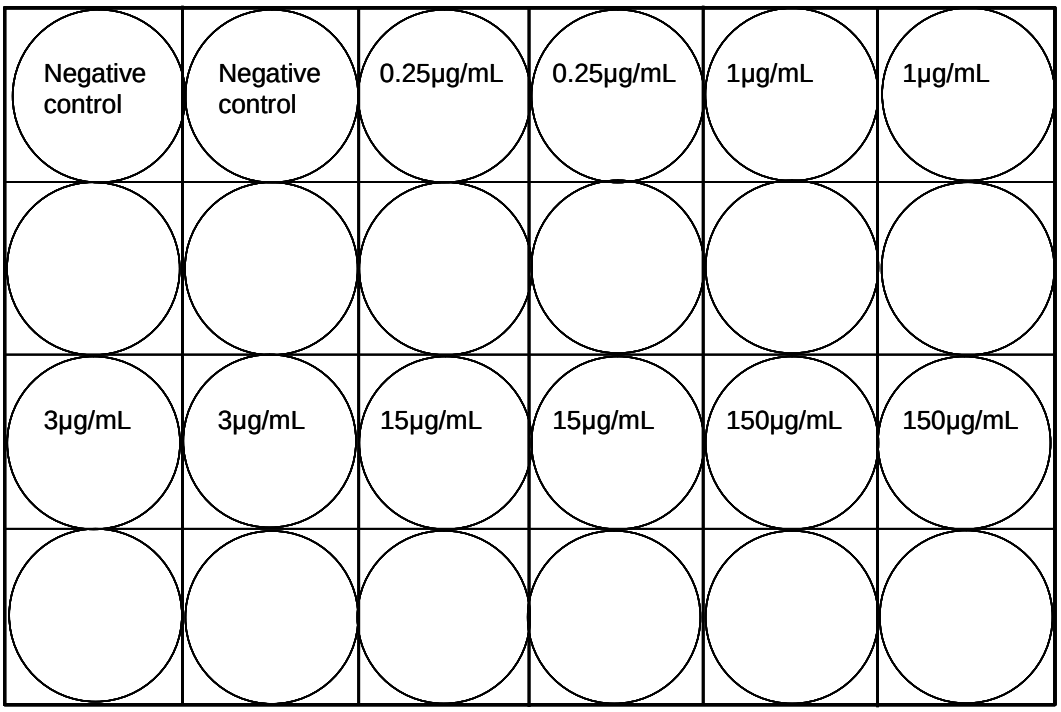


Figure 2.1: Larval migration inhibition test **A**: Diagram showing a side view of the migration chamber used in the LMIT. The dotted line represents the mesh which L_3 have to actively migrate through whilst exposed to a solution of IVM. Larvae affected by IVM are represented by a straight line and larvae resistant to IVM by a curved line. **B**: Layout of the plate in the standard LMIT. Non migrating L_3 are washed into the empty wells below each drug migration well prior to counting.

Chapter 3: Identification of Candidate Ivermectin Resistance Genes

3.1: Introduction

Macrocyclic lactone anthelmintics, and IVM in particular, have been used to treat sheep infected with parasitic nematodes like *T. circumcincta* for over twenty-five years, yet the mechanisms of resistance to IVM have yet to be determined (Geary, 2005; Omura, 2008). The majority of research on drug resistance in parasitic nematodes has been focused on identifying target site mutations associated with resistance, for example, the Phe200Tyr SNP which confers resistance to BZs (Wolstenholme *et al.*, 2004; von Samson-Himmelstjerna *et al.*, 2007). IVM binds to GABA and GluCl channels and causes the hyperpolarisation and flaccid paralysis of the pharyngeal and somatic muscle cells (Blackhall *et al.*, 1998a; Blackhall, Prichard, & Beech, 2003), leading to starvation and immobility of the worms. Changes in allele frequencies of these genes in *H. contortus* have been associated with IVM resistance, but no single allele has been shown to confer resistance in more than one isolate or population (Blackhall *et al.*, 1998a; Blackhall, Prichard, & Beech, 2003). In *C. elegans*, three concurrent mutations in GluCl subunits (namely, *avr*-14, *avr*-15 and *glc*-1) are required for high levels of IVM resistance. The orthologue of *avr*-14 in *C. oncophora* (the GluClα3 subunit) also carries a functional SNP (Leu256Phe), which makes the channel less sensitive to IVM (Dent *et al.*, 2000; Njue *et al.*, 2004; McCavera, Walsh, & Wolstenholme, 2007).

As has been shown in Chapter 1, candidate IVM resistance genes such as Pgps and CYPs have been studied in a range of organisms, although, as yet, there is very little information on these genes in *T. circumcincta*. They could offer alternative, non-target gene mechanisms for parasites to survive drug exposure, either through increased drug efflux or metabolism (Wolstenholme *et al.*, 2004). Pgps are transmembrane proteins involved in the active transport of endogenous and exogenous hydrophobic molecules, whilst CYPs catalyse the metabolism of a wide range of molecules such as drugs, environmental toxins and sterols (Mansuy, 1998; Sangster *et al.*, 1999a). As such, although not definitive targets of IVM, changes in the expression levels of these genes (and their gene products) might enable parasites to survive IVM exposure. Pgps have been identified in the model nematode *C. elegans* and the parasitic nematode *H. contortus*, along with some parasitic nematodes of humans such as *O. volvulus*, yet none have been

identified in *T. circumcincta* to date (Huang & Prichard, 1999; Jones & George, 2005). However a number of experiments have suggested they are implicated in IVM resistance. For example, the use of Pgp inhibitors has been shown to cause a shift towards an IVM-susceptible phenotype in an IVM-resistant *T. circumcincta* isolate using the LFIT (Bartley *et al.*, 2009). Also, changes in the expression levels or allelic frequencies of Pgps have been described in IVM-selected *H. contortus*, *O. volvulus* and *C. elegans* (Xu *et al.*, 1998; Blackhall *et al.*, 1998b; James & Davey, 2008), indicating that changes in the expression levels of non-target genes could play a role in expression of an IVM-resistance phenotype.

CYPs have been identified in all life-cycle stages of *C. elegans*, with approximately 80 identified through the *C. elegans* genome project (Barrett, 1998; Menzel, Bogaert, & Achazi, 2001). Identification of CYPs in parasitic nematodes has not been so fruitful, in Barrett (1998), the author went as far as to claim there was no experimental evidence for CYPs in adult helminths, based on mono-oxygenase enzyme activity, but that they may be present in low abundance in larval stages. However, the *H. contortus* genome sequencing project has identified 60 CYP ESTs to date, although these could reduce in number as the sequence is further assembled (R. Laing, Pers. Comm.). No CYPs have been identified in *T. circumcincta* in the literature to date. In the LFIT experiment described above, the *T. circumcincta* IVM-resistant isolate also showed a shift towards susceptibility when CYP inhibitors were used, suggesting the presence of CYPs in *T. circumcincta* (Bartley *et al.*, 2009).

In this Chapter, the candidate gene approach was used to investigate the role of differential gene expression of Pgps and CYPs in the IVM-resistant phenotype in *T. circumcincta*. Firstly, a panel of novel Pgp and CYP homologues from *T. circumcincta* were identified prior to a more detailed genetic analysis. This was performed using degenerate PCR and sequencing interfaced with bioinformatics approaches applied to the available EST datasets, as described in Chapter 2. At the time this work was carried out, no complete genome sequence was available for *T. circumcincta* and there was only a limited EST dataset available, thus degenerate primers designed to amplify Pgps from *H. contortus* were employed from (Sangster *et al.*, 1999a) and degenerate primers to amplify CYPs were designed utilising other nematode CYP sequences. If *T. circumcincta* follows

the same pattern as *C. elegans* and *H. contortus*, then the estimated numbers of Pgps and CYPs is likely to be large. It was not anticipated that all Pgps and CYPs would be identified by this approach.

3.2: P-glycoproteins

3.2.1: Identification of novel P-glycoprotein sequences

Identification of Pgps was performed using the degenerate *H. contortus* PCR primers which target the NBD of the Pgp molecule (as shown in Table 2.2, sequences obtained from Sangster *et al* (1999a)). These were employed in a standard PCR reaction followed by cloning and sequencing, as described in Sections 2.3.3 and 2.3.6 to 2.3.10. The template for these reactions was L_4 and xL_3 SMART cDNA, and single-stranded adult cDNA (all kindly provided by Dr A. J. Nisbet) and from the known IVM susceptible *T. circumcincta* CVL isolate. Eleven novel sequences, ranging in size from 395bp to 407bp, with BLAST homology to Pgps from *H. contortus*, *O. volvulus*, *Cyathostomum* spp and *Caenorhabditis briggsae* were identified. None of the BLAST hits aligned to *T. circumcincta* sequence suggesting that all 11 Pgp sequences are novel sequences, which are not yet in the public domain, even in EST databases. Sequences were numbered sequentially as they were identified (i.e. PGP1 to PGP11). The nucleotide sequence was translated into the corresponding protein sequence and identification of the correct reading frame made. As the primers used targeted the NBD of Pgps, the presence of the characteristic Walker motif (LSGGQ) of ABC transporters in the protein sequence was used to confirm the BLAST results (Higgins *et al.*, 1997; Jones & George, 2005). Additional cloning and sequencing of PCR products was carried out to derive consensus sequences, to improve the accuracy of the sequencing results and to resolve any multiple base calls.

3.2.2: Elongation of P-glycoprotein sequence using RACE PCR

None of the sequences amplified using the degenerate primers encoded a full-length coding sequence for any given gene; potentially two sequences could be the two separate NBDs of the same gene. RACE-ready cDNA was synthesised in order to extend the

sequences in the 5' and 3' directions, to provide longer sequence reads and to determine if any of the NBDs were parts of the same gene. The nucleotide sequences identified above were aligned in Clustal W (http://www.ebi.ac.uk/Tools/clustalw/index.html) and used to design gene-specific RACE primers as shown in Table 2.5. All the RACE primers were designed for 5' RACE; NBDs are situated more closely to the 3' end so 5' RACE would be more likely to generate significantly longer products compared to 3' RACE. Use of the RACE primers followed by cloning and sequencing enabled the sequences for PGPs 1, 2, 3, 4, 5 and 7 to be extended whilst also improving the sequence accuracy. The RACE primers for PGPs 6, 8, 9, 10 and 11 were found to be non Pgp-specific (i.e. none of the selected clones aligned with Pgp sequence) or to produce no PCR product. The RACE product for PGP1 was cloned into the pGEM®-T vector and, when sequenced from both ends using the M13 uni and SP6 primers, did not produce a contiguous sequence so primers Pgp1Tc IntFor and Pgp1Tc IntRev (Table 2.3) were designed and used in a sequencing reaction to enable a complete sequence of the PGP1 RACE product to be obtained. The lengths of the Pgps and the numbers of sequences making up each Pgp consensus sequence are shown in Table 3.1.

Individually, each Pgp protein sequence was aligned using Clustal W against *H. contortus* PgpA (Accession number AF003908) and against *O. volvulus* PGP1 (AF083642) to discover whether the region amplified aligned more closely to NBD1 or NBD2. All, apart from PGP1, aligned more closely to either NBD1 or NBD2 whilst PGP1 was found to contain both NBDs as shown in Table 3.1. As eight of the consensus sequences appear to align to NBD1 (PGPs 1, 3, 4, 5, 7, 8, 9 & 11), this suggests that there could be at least eight Pgps in *T. circumcincta*, and probably more.

On translating the nucleotide sequence for PGP7, it was discovered that a frame shift occurs approximately between amino acids 272 and 277, equating to between nucleotides 814 and 829. Analysis of the nucleotide sequence was unable to identify if a sequencing error had caused a deletion or insertion of a nucleotide in the sequence at this point. As a result, no full length protein sequence in one continuous reading frame could be made; thus the Clustal W protein alignment (Appendix 2) contains two entries for PGP7, one in each of the reading frames. The nucleotide and protein alignments of the 11

T. circumcincta Pgps are shown in the Appendix. Even though there are some sequences for each of the NBDs and one sequence that spans both NBDs, as described above, when aligned together, all 11 Pgp sequences align to NBD2.

3.2.3: Design and validation of P-glycoprotein specific primers

In order to determine the gene expression profiles of the 11 Pgps in the different isolates and life-cycle stages of *T. circumcincta*, pairs of gene-specific primers were designed for each Pgp using the Pgp nucleotide alignment (Appendix 1). The primers were designed to give as large an amplicon as possible to increase the coverage of any sequence data generated whilst also providing sufficient sequence data to confirm the identity of any Pgp sequences generated. The gene-specific primers and size of amplicons are shown in Table 2.3.

The primers were firstly validated against a panel of known Pgp-containing purified plasmids to determine whether the primers were specific for the Pgp they were designed against. The gels showing the outcome of this validation are shown in Figure 3.1. The sizes of each of the bands visualised on the gel was estimated by comparing to the Mx marker lane and were found to be of the anticipated size. The primers designed against PGP 1, 2, 3, 4, 5, 6, 7 and 9 exclusively amplified their corresponding Pgp plasmids, whilst the PGP8 primers also weakly amplified PGPs 6 and 9, and the PGP 10 and 11 primers were completely non-specific, amplifying all the PGP-containing plasmids. Attempts were made to optimise the respective reactions by altering the reaction conditions for the PGP 8, 10 and 11 primers; however, this failed to improve their specificity. These gene-specific primers were redesigned as shown in Table 2.3 but on repeated validation against the Pgp-containing plasmids, the same pattern of non-specificity was observed (data not shown), so these Pgps were excluded from further analysis.

The next stage of the validation of the Pgp-specific primers was to determine whether the primers would selectively amplify their particular Pgp from *T. circumcincta* cDNA. PCRs on IVM-susceptible *T. circumcincta* exsheathed L_3 and L_4 SMART cDNA

and single-stranded adult cDNA were set up using the primer pairs for PGPs 1, 2, 3, 4, 5, 6, 7, 9 (Table 2.3) and *T. circumcincta* actin (Table 2.1) as a positive control. The results of this validation are shown in Figure 3.2, part A. Only one weak band was identified from the PGP1 primers against L_3 and no bands were amplified from PGP4 or PGP9. All the main bands identified on the gel were of the anticipated size with the exception of the lower band from PGP5 versus adult cDNA and the lowest arrowed band from PGP7 versus L_4 cDNA. A second PCR was set up for PGP4 and PGP9 using IVM susceptible adult *T. circumcincta* 5' RACE-ready cDNA as template. The results of this validation are shown in Figure 3.2, part B. Both the PGP4 and PGP9 primers resulted in a single strong band on the agarose gel of the expected size.

The bright bands and the three weak bands (arrowed) were excised from the gels in Figure 3.2, parts A and B with a clean scalpel, cleaned up using the QIAquick Gel Extraction Kit (Qiagen), and cloned into pGEM-T vector in JM109 cells and sequenced, as described previously, to determine if the bands were the correct Pgp. The results of this cloning experiment are shown in Table 3.2. The specificity of the primers was determined after sequence analysis by firstly excluding any clones which did not contain an insert. Next, the remaining sequences were trimmed in SeqMan (DNASTAR Lasergene, version 8) to remove any vector sequence and the resulting sequence aligned against the consensus sequences for the 11 Pgps. Any sequences failing to align with the Pgp sequences were BLAST searched to give a putative identity. A primer pair was determined to be non-specific if a clone which should have aligned to that particular Pgp was found to not have that identity, and the primer sequences were found at the start and end of the sequence. After several attempts at validation, the PGP1-specific primers were discarded as PCR products were either not amplified or failed to clone successfully. The original primers for PGP7 (P7Tc F&R) were not sufficiently specific and so these were redesigned, validated against the plasmids and against the life-cycle stage cDNA from the CVL isolate as described above. Although one of the subsequent clones from the xL_3 PCR product did not align to PGP7 (Table 3.2), these primers were determined to be specific as neither the P7TcF2 or P7TcR2 primer sequence was present at the ends of the cloned sequence. The identity of this clone was a ribosomal protein of only 241bp in length compared to the PCR product size of ~650bp and, along with the absence of the primer sequences, this suggested

that the insert was a contaminant of the cloning process. The primer pairs for PGPs 2, 3, 5, 6 and 9 were also shown to be gene-specific as shown in Table 3.2.

3.2.4: Semi-quantitative PCR using P-glycoprotein specific primers

Using the gene-specific primers for PGPs 2, 3, 5, 6, 7 and 9 (Table 2.3) and for both the actin and β tubulin control genes (Table 2.1), a semi-quantitative PCR, using cDNA from eggs, L_1, xL_3, L_4 and adults from the CVL and MOTRI isolates, was set up with aliquots of the PCR product collected after 15, 20, 25 and 30 cycles. The PCR products were run on a 1% agarose gel and visualised under UV as shown in Figure 3.3. The gel shows that the expression of the control genes (actin and β tubulin) is approximately equal between the CVL and MOTRI isolates, however, there is some variation in the cycle number where the bands became visible between the life-cycle stages. No bands were present in either isolate, across all life-cycle stages for PGP2. The expression of PGPs 3 and 7 showed no discernable difference between the two isolates although there was some variation in band intensity between the life-cycle stages. The band intensity of PGP5 for the adult stage was also approximately equal between CVL and MOTRI but in the egg, L_1 and xL_3 stages the CVL isolate showed a greater intensity compared to the MOTRI isolate at 30 cycles. In the L_4 stage for PGP5, the MOTRI band at 30 cycles was more intense than the CVL band. Across all life-cycle stages for PGP6, the intensity of the band from MOTRI cDNA was greater than the band from CVL cDNA. No bands were present using the PGP9 primers in the egg, xL_3 and adult life-cycle stage whilst in the L_1 and L_4 stages the bands from MOTRI cDNA were more intense compared to the bands from CVL cDNA.

3.2.5: Correct nomenclature of *T. circumcincta* P-glycoproteins

Phylogenetic analysis of the contiguous *T. circumcincta* Pgp sequences against *C. elegans* Pgp homologues, as described in Section 2.3.10, was carried out to determine the correct nomenclature for the 11 Pgp genes. The results of this analysis are summarised in Table 3.3. PGP1 aligned most closely to *CePgps- 3* and *-4* with bootstrapping values of 1000 for both trees (Figure 3.4). However, because it aligned equally well to both *C. elegans* sequences this sequence will still be referred to as PGP1. PGP2 aligned most

closely to NBD2 in *CePgp-2* with bootstrapping values of 1000 for both trees (Figure 3.5) and so from now on will be referred to as *TeciPgp-2* NBD2. PGP4 aligned most closely to the NBD1 in *CePgp-9* (NJ bootstrap value 887, UPGMA bootstrap value 1000; Figure 3.6) and from now on will be referred to as *TeciPgp-9* NBD1. Like *TeciPgp-2* NBD2, PGP5 aligned most closely with *CePgp-2* with bootstrapping values of 1000 for both trees (Figure 3.7) however, PGP5 aligned against the NBD1. As a result, PGP5 will now be known as *TeciPgp-2* NBD1. Like *TeciPgp-9* NBD1, PGP6 aligned most closely with *CePgp-9* (NJ bootstrap value 850, UPGMA bootstrap value 983; Figure 3.8) however, PGP6 aligned against the NBD2. As a result, PGP6 will now be known as *TeciPgp-9* NBD2. Like PGP1, PGP9 aligned closest to two *C. elegans genes*; *CePgps -1* and *-9* (NJ bootstrap value 529, UPGMA bootstrap value 390) and so this sequence will still be called PGP9 as it could not be determined which it aligned more closely to. The alignments generated for PGPs 3, 7, 8, 10 and 11 were quite poor, resulting in the *T. circumcincta* sequences being broken up into lots of small sections spread across both NBD areas of the *C. elegans* Pgp nucleotides sequences. As a result, the trees generated did not indicate which *C. elegans* Pgp was the closest match so these genes will still be referred to by their sequence names. Interestingly, the major sections of the PGP8 sequence aligned in the NBD2 region as opposed to the NBD1 region of *H. contortus* PgpA (Accession number AF003908) and *O. volvulus* PGP1 (AF083642) it had aligned to previously, as described in Section 3.2.2.

3.2.6: Identification of single nucleotide polymorphisms

Using the gene-specific primers as described in Table 2.3 and the one-step PCR protocol described in Section 2.3.11, gene-specific PCR products were obtained from adult CVL and MOTRI *T. circumcincta* pooled RNA for PGP3, *TeciPgp-2* NBD1, *TeciPgp-9* NBD2, PGP7 and PGP9. The resultant PCR products were purified using the QIAquick® PCR purification kit (Qiagen) and submitted for direct sequencing as described in Section 2.3.9. Aligning the sequences using the SeqDoC programme, as described in Section 2.3.11, showed that for PGP3, *TeciPgp-2* NBD1, PGP7 and PGP9 the sequences were highly conserved with no significant SNPs over the length of the sequence available for analysis. However, *TeciPgp-9* NBD2 proved to be highly polymorphic, with several SNPs visible between the CVL and MOTRI sequences. Figure 3.9 shows a representative section of PGP3 and a section of *TeciPgp-9* NBD2 where at least four major SNPs are

present. Figure 3.10 shows the nucleotide and protein Clustal W alignments for *TeciPgp-9* NBD2 comparing the CVL and MOTRI sequences. This shows that, although there are SNPs in this gene, they are silent or non-coding and do not result in any amino acid changes, although this does not preclude the presence of coding SNPs elsewhere in the sequence. The alignments in Figure 3.10 only show the section of the *TeciPgp-9* NBD2 which was amplified using the gene-specific primers. As such, the four SNPs are (CVL to MOTRI) as follows; A to G at nucleotide 138, C to T at nucleotide 177, A to T at nucleotide 201 and A to C at nucleotide 216. Other smaller peaks also indicate polymorphic sites in the *TeciPgp-9* NBD2 gene.

3.3: Cytochrome P450s

3.3.1: Identification of novel Cytochrome P450 sequences

The first attempt at amplifying CYPs from *T. circumcincta* utilised the degenerate CYP primers from Menzel *et al*, (2001) as shown in Table 2.2. These primers failed to amplify any CYPs from *T. circumcincta* cDNA, possible because the primers were not sufficiently degenerate, so EST CYP sequences from *Ancylostoma ceylanicum, A. caninum* and *O. ostertagi* were aligned and used to design a new degenerate primer, DegenTcP450For. The use of this primer (Table 2.2) in conjunction with the TrplX Rev vector primer (Table 2.4) followed by cloning into the pGEM®-T vector and sequencing of the resultant product, identified a CYP from *T. circumcincta* L_4 SMART cDNA. Whilst designing the DegenTcP450For primer, a second sequence from an xL3 enriched SSH EST dataset generated at Moredun Research Institute (Nisbet *et al.*, 2008) was found to have BLAST similarities to CYPs from other nematodes. Finally, a bioinformatics search identified a third EST (accession number CB037767) from the Nembase cluster TDC01265 from a mixed sex *T. circumcincta* adult cDNA library. A summary of the CYPs found in *T. circumcincta* is shown in Table 3.4, again sequences were numbered sequentially as they were identified (CYP1 to CYP3).

Three further pairs of degenerate CYP primers were designed using a panel of CYP EST sequences from a range of nematodes species, such as *A. caninum, C. elegans, O. ostertagi* and *T. circumcincta*, with each pair of primers being specific for CYP clans 2, 3

and 4 (Table 2.2), respectively, in an attempt to increase the chance of finding further *T. circumcincta* CYP sequences. Again, these primers failed to amplify further CYPs from *T. circumcincta* cDNA.

Alignments of the nucleotide sequences and translated protein sequences for the three *T. circumcincta* CYPs were performed using Clustal W and are shown in Appendices 3 and 4. The nucleotide sequences for the three *T. circumcincta* CYPs were aligned against a panel of CYP sequences from *H. contortus, C. elegans, A. caninum* and *O. ostertagi* using CLC DNA Workbench version 5.6.1 (www.clcbio.com). The alignment was used to generate phylogentic trees as shown in Figure 3.11. Part **A** shows a tree generated using the UPGMA distance matrix with a bootstrap value of 1000 whilst part **B** shows a tree generated using the NJ distance matrix with a bootstrap value of 1000.

3.3.2: Design and validation of Cytochrome P450 specific primers

Specific primers were designed, using the nucleotide alignment in Appendix 3, for each of the three CYPs to enable their amplification from *T. circumcincta* cDNA. The primer sequences, melting temperature and expected size of PCR product are shown in Table 2.3. These three primer pairs were used in PCRs on IVM susceptible *T. circumcincta* exsheathed L_3 and L_4 SMART cDNA and single-stranded adult cDNA alongside the *T. circumcincta* actin primers as a positive control. Figure 3.12 shows the PCR products run on a 1% agarose gel, showing the presence of single bands of the expected sizes. The bands were excised from the gel with a clean scalpel, the gel product cleaned up and sent for sequencing. The results of the sequencing are shown in Table 3.5 below. These results show that the primers designed successfully amplified the CYPs from cDNA and are specific, only amplifying the CYP they were designed against.

3.3.3: Semi-quantitative PCR using CYP specific primers

As described previously for the Pgp work, semi-quantitative PCR was set up to investigate the relative expression of CYPs 1, 2 and 3 in five life-cycle stages of the CVL

and MOTRI isolates of *T. circumcincta* compared to actin and β-tubulin controls. Samples of the PCR reaction mix were taken after 15, 20, 25 and 30 cycles. The agarose gels, visualised under UV, are shown in Figure 3.13. As discovered previously, the expression levels of actin and β tubulin were approximately equal between the two isolates. The expression of all three CYP genes was low with the majority of the life-cycle gene combinations only giving identifiable bands after 30 cycles. For CYP1, only three very weak bands were apparent in the 30 cycle samples, these indicated that expression of this gene in the xL_3 stage was greater in the CVL isolate compared to the MOTRI isolate and the expression in the L_4 stage was approximately equal. No bands were visible in the egg and L_1 stages for CYP2 and only very weak bands of equal intensity were visible in the L_4 and adult stage. Bright bands were visible in the xL_3 stage but, like those in the L_4 and adult stage, there was no difference between the CVL and MOTRI samples. No bands were present for CYP3 in the egg, L_1 or adult life-cycle stages. Weak bands were present in the xL_3 and L_4 stages for CYP3; in both cases the MOTRI band was more intense than the CVL band.

3.4: Discussion

Candidate IVM resistance genes such as the Pgps and CYPs have not been investigated previously in *T. circumcincta* despite extensive investigation in other nematodes such as *C. elegans*, *H. contortus* and *O. volvulus* (Nelson, 1998; Huang & Prichard, 1999; Sangster *et al.*, 1999a; Sheps *et al.*, 2004; Ardelli, Guerriero, & Prichard, 2005). Both have generalised roles within cells; Pgps are involved in the transport of endogenous and exogenous hydrophobic molecules whilst CYPs catalyse the metabolism of hydrophobic molecules, such as drugs and environmental toxins (Barrett, 1998; Mansuy, 1998; Sangster *et al.*, 1999a; Kerboeuf, Guegnard, & Le Vern, 2003).

Pgps have been shown to be involved in drug resistance in cancerous tumours, malaria parasites and the human immunodeficiency virus (Beugnet, Gauthey, & Kerboeuf, 1997; Zhang *et al.*, 1998; Loo & Clarke, 1999; Jones & George, 2005), where an increase in expression of Pgps allows a drug-resistant phenotype to develop. It is also known that the anthelmintic, IVM, is a substrate of Pgps (Pouliot *et al.*, 1997). Changes in mRNA

levels, allele frequencies and expression of Pgps in *H. contortus, O. ostertagi* and *O. volvulus* have been observed following IVM selection, or in isolates resistant to IVM (Xu *et al.*, 1998; Blackhall *et al.*, 1998b; Ardelli, Guerriero, & Prichard, 2005; Ardelli, Guerriero, & Prichard, 2006a; Ardelli, Guerriero, & Prichard, 2006b; Van Zeveren, 2009). Inhibitors of Pgps have been shown to cause a reversal towards a susceptible phenotype in resistant strains of *T. circumcincta* and *H. contortus* when exposed to IVM, as measured by the LFIT (Bartley *et al.*, 2009).

In this Chapter, the use of degenerate PCR primers, previously utilised to identify Pgps in *H. contortus* (Sangster *et al.*, 1999a) against cDNA generated from IVM-susceptible *T. circumcincta* has allowed the identification of eleven novel partial Pgp sequences. These partial gene sequences are the first time Pgps have been identified in *T. circumcincta*. Out of the eleven genes identified, seven aligned to the NBD1 from *H. contortus* PgpA and *O. volvulus* PGP1, three aligned against NBD2 and one sequence (PGP1) contained both NBDs, as shown in Table 3.1. This originally suggested that the potential number of Pgps in *T. circumcincta* was at least eight. By aligning each individual *T. circumcincta* against the fifteen *C. elegans* Pgp sequences from Sheps *et al* (2004), it was determined that the two genes that had originally been called PGPs 2 and 5 were both the closest homologues of *CePgp-2*. As a result, these were re-named *TeciPgp-2* NBD2 and *TeciPgp-2* NBD1, respectively. *CePgp-2* has been shown to be homologous to PgpA in *H. contortus*, as such, *TeciPgp-2* NBD2 and *TeciPgp-2* NBD1 can be assumed to be homologous to this gene. The two partial genes, PGPs 4 and 6 were shown to align most closely to *CePgp-9* and were subsequently renamed *TeciPgp-9* NBD1 and *TeciPgp-9* NBD2, respectively. If these four partial gene fragments prove to be each half of the respective genes, this would indicate *T. circumcincta* has at least nine Pgps. In *H. contortus* at least twelve Pgp genes have been identified, whilst at least two have been identified in *O. volvulus,* and at least fifteen in *C. elegans*; the identification of eight or nine in *T. circumcincta* seems to fit this pattern (Kwa *et al.*, 1998; Huang & Prichard, 1999; Le Jambre, Lenane, & Wardrop, 1999; Sangster *et al.*, 1999a; Kerboeuf *et al.*, 2003; Blackhall, Prichard, & Beech, 2008). Unequivocally defining the final number of Pgps requires the completion of the genome sequence of *T. circumcincta*.

Although not a main component of this thesis, the identification of SNPs in *TeciPgp-9* NBD2 using SeqDoc analysis is interesting. *TeciPgp*-9 NBD2 appears to be considerably more polymorphic than the other Pgps identified. Four main SNPs were identified in the 208bp of sequence analysed, although the sequence chromatogram for *TeciPgp-9* NBD2 does show other minor polymorphisms (Figure 3.9, part **B**). The four main SNPs are A138G, C177T, A201T and A216C; the numbering of these SNPs is based on their position in the gene sequence currently available; when the full length sequence is obtained these numbers will change. All of these SNPs are silent; as shown in Figure 3.10, part **B**, the protein sequence remains unchanged between CVL and MOTRI. However, as previously stated, this does not preclude the presence of coding SNPs elsewhere in this gene. Other polymorphisms in *TeciPgp-9* NBD2 could also be SNPs but neither CVL nor MOTRI are clonal populations; the presence of individual worms exhibiting a greater degree of resistance in the CVL isolate, or the presence of individual worms exhibiting a lesser degree of resistance in the MOTRI isolate, could result in some of the nucleotide positions being identified as polymorphic within an individual isolate when they are, in fact, a SNP between resistant and susceptible worms. Identifying the presence of a highly polymorphic region in *TeciPgp-9* NBD2 (the left hand portion of Figure 3.9, part B) could also explain why the RACE primer for *TeciPgp-9* NBD2 failed to work; the primer, Pgp6Tc 5'race, aligns to nucleotides 180 to 208 which includes the silent SNP A201T alongside other polymorphisms. The presence of SNPs, whether silent or coding, could be a useful marker to distinguish between susceptible and resistant isolates. They may only prove to be isolate-specific so it would be interesting to determine whether these silent SNPs are also present in other isolates of *T. circumcincta*. Interestingly Kimchi-Sarfaty *et al* (2007) have also found silent, or synonymous, SNPs in a haplotype of the *Multidrug Resistance* 1 (*MDR*1) gene expressed in a range of cell lines; the *MDR*1 gene is associated with multidrug resistance in cancer cells. The Pgp encoded by *MDR*1 exhibits altered drug and inhibitor interactions, yet the mRNA and protein levels remain unchanged. The authors hypothesise that the silent SNPs result in the use of rare codons during translation that could lead to altered protein translation kinetics which, in turn, affects co-translational folding and the structure of substrate and inhibitor interaction sites (Kimchi-Sarfaty *et al.*, 2007; Komar, 2007).

As a first step to determining whether there are changes in the expression of the Pgps, as measured by their mRNA levels, which correlate to the resistance status of the *T. circumcincta* isolate, a semi-quantitative PCR approach was adopted (Figure 3.3). This experiment showed that the expression of the six Pgp genes, to which there were validated gene-specific primers available, was not uniform across all life-cycle stages. Pgp expression was evident in all life-cycle stages, which correlates with results from other nematodes, as reported in the literature. For example, the expression of Pgps in *H. contortus* has been shown in eggs, L_1, L_3 and adults using gene probes and *in situ* hybridisation (Kwa *et al.*, 1998; Smith & Prichard, 2002; Riou *et al.*, 2005). *TeciPgp-2* NBD2 (PGP2) was not detected in the semi-quantitative PCR experiment (Figure 3.3), however, this may be related to the low expression level of Pgps in general and this one in particular. In the original PCR experiment to identify Pgps, expression of these genes was often only evident after 40 PCR cycles. There were also discernable differences in the expression levels between the MOTRI and CVL isolates. There was no uniform pattern of expression of the Pgps; some appeared to be expressed equally between the CVL and MOTRI isolates (PGPs 3 and 7) whilst *TeciPgp-9* NBD2 (PGP6) exhibited increased expression in MOTRI compared to CVL. Caution needs to be applied when analysing these results, however, as inaccuracies in the method such as pipetting error or amount of starting cDNA added to each well prior to PCR cycling could have caused some of the slight differences in band intensity observed. In the case of *TeciPgp-2* NBD2 and other genes, where no expression was observed at certain life-cycle stages, it may be necessary to repeat the experiment with more PCR cycles before it is assumed that these genes are not expressed in these life-cycle stages. Alternatively, the use of methods such as real-time PCR could allow the quantification of these genes more accurately.

Like the Pgps, the CYPs have also been implicated in drug resistance. In mosquito species such as *An. minimus*, *An. gambiae* and *C. p. pallens*, over-expression or over-transcription of CYPS enables the insects to exhibit resistance to insecticides (Nikou, Ranson, & Hemingway, 2003; Rodpradit *et al.*, 2005; Gong *et al.*, 2005). In malaria parasites such as *P. berghei* and *P. falciparum*, an increase in CYP activity correlates with chloroquine resistance (Ndifor, Ward, & Howells, 1990). Over 80 CYPs have been identified in *C. elegans* in all life-cycle stages but their identification in parasitic nematodes has not been so straightforward, until the increase in genome sequence

availability in recent years (Menzel, Bogaert, & Achazi, 2001). Barrett (1998) claimed that there was no evidence of the presence of CYPs in adult parasitic nematodes, based on biochemical measurements of mono-oxygenase enzyme activity and that they would only be present at low abundance in larval stages, if present at all. In a LFIT experiment, Bartley *et al* (2009) were able to show that CYP inhibitors caused a shift towards a susceptible phenotype in *T. circumcincta*, and so CYPs may be implicated in the mechanism of IVM resistance in this species. The efficacy of BZs can also be improved by the use of the CYP inhibitors, piperonyl butoxide and metyrapone (McKellar & Jackson, 2004), possibly hinting at some generic (i.e. not linked to mutations in β-tubulin) resistance mechanisms.

Using a combination of degenerate PCR, bioinformatic searches of NEMBASE, and an EST dataset at Moredun, three distinct CYPs were identified, one each from L_3, L_4 and adult *T. circumcincta* (Table 3.4). None of the sequences obtained were full-length coding sequences. The presence of a CYP in a mixed sex adult cDNA library contradicts the statement from Barrett (1998) that adult parasitic nematodes do not possess CYPs. Unlike some of the Pgps, it has not been possible to identify which *C. elegans* CYP gene is the nearest homologue. The CYP gene superfamily is very large and assigning the correct nomenclature to a newly identified CYP is carried out through the P450 Nomenclature Committee; as such the sequences identified here remain named as CYP 1, 2 and 3 (Nelson *et al.*, 1996). Identifying the three partial CYPs from *T. circumcincta* was problematic and is possibly due to the low levels at which they are expressed. Kotze (1997) measured the CYP activity towards aldrin and 7-ethoxycoumarin in *H. contortus* microsomes and found it to be 10,000-fold lower compared to rat liver microsomes. Gene-specific primers were designed for each of the three CYPs identified and validated by sequencing the PCR products obtained from amplification of L_3, L_4 and adult cDNA (Figure 3.12). This proved that the primers designed solely amplified their specific CYP target.

As with the Pgps, the gene-specific primers were used to determine the expression profile of the three CYPs across the life-cycle stages (eggs, L_1, xL_3, L_4 and adults) of the CVL and MOTRI isolates, compared to the actin and β-tubulin control genes in a semi-quantitative PCR. This experiment showed that the level of CYP expression was low;

most of the bands obtained were very weak, compared to the bands for actin and β-tubulin, even after 30 cycles. In fact, none of the three genes were amplified from egg cDNA; however, it is not possible to conclude at this stage that they are not expressed in eggs. The indication that CYP1 was expressed more in CVL xL_3 than in MOTRI xL_3 and CYP3 was expressed more in MOTRI L_4 compared to CVL L_4 (Figure 3.13) suggests that the CYP genes are worthy of further investigation in the context of IVM resistance in *T. circumcincta*. Using real-time PCR would allow a more accurate measure of whether these genes do exhibit an altered gene expression profile between CVL and MOTRI; semi-quantitative PCR only allows a subjective measure of expression differences.

The identification of 11 novel partial Pgp gene fragments and 3 partial CYP gene fragments from *T. circumcincta* gives a good panel of candidate resistance genes for use to investigate whether altered expression patterns of these generic drug handling molecules have a role to play in the expression of the IVM-resistance phenotype. Semi-quantitative PCR has revealed demonstrable differences in mRNA expression levels for both the CYPs and Pgps between a multi-drug resistant and susceptible *T. circumcincta* isolates; more accurate quantification of these genes using real-time PCR is required. The differences observed in the expression of *TeciPgp-9* NBD2, indicating a constitutive increase in expression across all life-cycle stages in the MOTRI isolate compared to CVL coupled with the observation of SNPs in the same gene, make this the priority for further investigations.

Table 3.1: Length and number of sequences aligning to each of the 11 novel Pgp sequences identified from *T. circumcincta* cDNA.

Pgp name	**Size of nucleotide sequence (bp)**	**Number of sequences in alignment**	**Predicted region of Pgp**
PGP1	2322	13	NBD 1 & NBD 2
PGP2	1041	14	NBD 2
PGP3	945	17	NBD 1
PGP4	1501	18	NBD 1
PGP5	935	5	NBD 1
PGP6	405	4	NBD 2
PGP7	1892	5	NBD 1
PGP8	398	2	NBD 1
PGP9	398	2	NBD 1
PGP10	395	3	NBD 2
PGP11	395	2	NBD 1

Table 3.2: Results of cloning of Pgp gene-specific primers to determine their specificity. The sequence of the primers is shown in Table 2.3. * Remaining clone was a contaminant as no P7TcF2 and P7TcR2 primers were present in the sequence and sequence was not correct length compared to original band size on agarose gel.

Band name	**Template**	**Band size (approx bp)**	**Primers**	**Clone results (# correct ID)**	**Specificity**
P1L3	CVL xL_3 SMART cDNA	650	P1TcF & P1TcR	0/0	Cloning failed
P1	CVL adult 5' RACE	650	P1TcF & P1TcR	0/0	Cloning failed
P2L3	CVL xL_3 SMART cDNA	600	P2TcF & P2TcR	2/2	PGP2 specific
P2L4	CVL L_4 SMART cDNA	600	P2TcF & P2TcR	3/3	
P2Ad	CVL single-stranded adult cDNA	600	P2TcF & P2TcR	3/3	
P3L3	CVL xL_3 SMART cDNA	600	P3TcF & P3TcR	3/3	PGP3 specific
P3L4	CVL L_4 SMART cDNA	600	P3TcF & P3TcR	3/3	
P3Ad	CVL single-stranded adult cDNA	600	P3TcF & P3TcR	3/3	
P4	CVL adult 5' RACE	600	P4TcF & P4TcR	6/7	Not specific
P5L4	CVL L_4 SMART cDNA	600	P5TcF & P5TcR	3/3	PGP5 specific if band of correct size
P5Ad 1	CVL single-stranded adult cDNA	600	P5TcF & P5TcR	3/3	
P5Ad 2	CVL single-stranded adult cDNA	500	P5TcF & P5TcR	0/3	
P6L3	CVL xL_3 SMART cDNA	250	P6TcF & P6TcR	3/3	PGP6 specific
P6L4	CVL L_4 SMART cDNA	250	P6TcF & P6TcR	3/3	
P7L3	CVL xL_3 SMART cDNA	650	P7TcF & P7TcR	3/3	Not specific
P7L4 1	CVL L_4 SMART cDNA	700	P7TcF & P7TcR	1/1	
P7L4 2	CVL L_4 SMART cDNA	650	P7TcF & P7TcR	3/3	
P7L4 3	CVL L_4 SMART cDNA	250	P7TcF & P7TcR	1/3	
P7Ad	CVL single-stranded adult cDNA	650	P7TcF & P7TcR	2/3	

Band name	**Template**	**Band size (approx bp)**	**Primers**	**Clone results (# correct ID)**	**Specificity**
P7L3	CVL xL_3 SMART cDNA	650	P7TcF2 & P7TcR2	2/3*	PGP7 specific
P7L4	CVL L_4 SMART cDNA	650	P7TcF2 & P7TcR2	1/1	
P7Ad	CVL single-stranded adult cDNA	650	P7TcF2 & P7TcR2	3/3	
P9	CVL adult 5' RACE	350	P9TcF & P9TcR	9/9	PGP9 specific

Table 3.3: Results of the phylogenetic analysis of the 11 *T. circumcincta* Pgp genes to determine their correct gene identification. * denotes sequence which were reduced in size slightly when the alignment was trimmed due to a small proportion of the *T. circumcincta* sequence aligning a very large distance away from the remaining part of the sequence.

Gene name	**Sequence size**	**NBD**	**UPGMA result**	**UPGMA bootstrap value**	**Neighbour-joining result**	**Neighbour- joining bootstrap value**	**Notes**	**Accession number**
PGP1	2273*	1&2	*CePgp-3*or *4*	1000	*CePgp-3* or *4*	1000	Keeps temporary name	FR691845
PGP2	936*	2	*CePgp-2*	1000	*CePgp-2*	1000	Gene now called *TeciPgp-2* NBD2	FR691846
PGP3	945	1	Not clear	-	*CePgp-10*	1000	Alignment poor quality, keeps temporary name	FR691847
PGP4	1501	1	*CePgp-9*	1000	*CePgp-9*	887	Gene now called *TeciPgp-9* NBD1	FR691848
PGP5	935	1	*CePgp-2*	1000	*CePgp-2*	1000	Gene now called *TeciPgp-2* NBD1	FR691849
PGP6	405	2	*CePgp-9*	983	*CePgp-9*	850	Gene now called *TeciPgp-9* NBD2	FR691850
PGP7	1892	1	Not clear	-	*CePgp-10*	646	Alignment poor quality, keeps temporary name	FR691851
PGP8	398	1	Not clear	334	*CePgp-10*	166	New alignment shows it to be NBD2	FR691852
PGP9	398	1	*CePgp-9*or *1*	390	*CePgp-9* or *1*	529	Keeps temporary name	FR691853
PGP10	395	2	Not clear	-	*CePgp-10*	-	Alignment poor quality, keeps temporary name	FR691854
PGP11	395	1	Not clear	-	*CePgp-10*	-	Alignment poor quality, keeps temporary name	FR691855

Table 3.4: A summary of *T. circumcincta* CYPs found using PCR and bioinformatics approaches.

CYP Name	**Identification method**	**Life cycle stage**	**Size of nucleotide sequence (bp)**
CYP 1	Degenerate PCR	L_4	724
CYP 2	EST dataset	L_3	647
CYP 3	Bioinformatics	Mixed sex adult	681

Table 3.5: Results of cloning of CYP PCR products to determine the specificity of the specific *T. circumcincta* CYP primers.

Primers	**Template cDNA**	**No. clones submitted**	**Expected size of product**	**Actual size of product**	**Notes**	**Primer specificity**
CYP 1	L_3	6	365	366	100-95% sequence identity to CYP1	Yes
CYP 1	L_4	6	365	366	100-95% sequence identity to CYP1	
CYP 2	L_3	1	323	323	97% identity to CYP2	Yes
CYP 3	L_3	1	467	468	95% identity to CYP3	Yes
CYP 3	L_4	2	467	468	95% identity to CYP3	

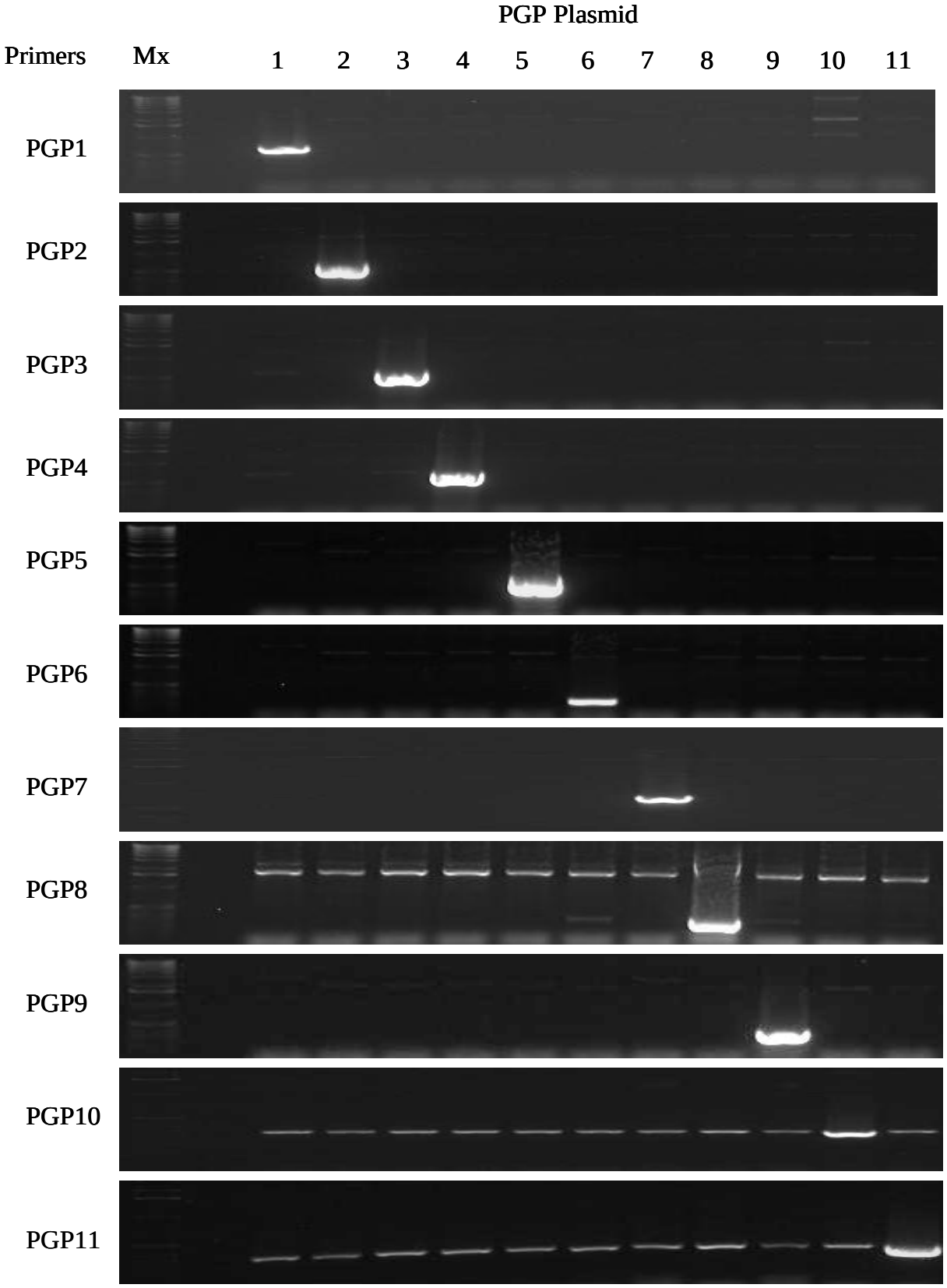
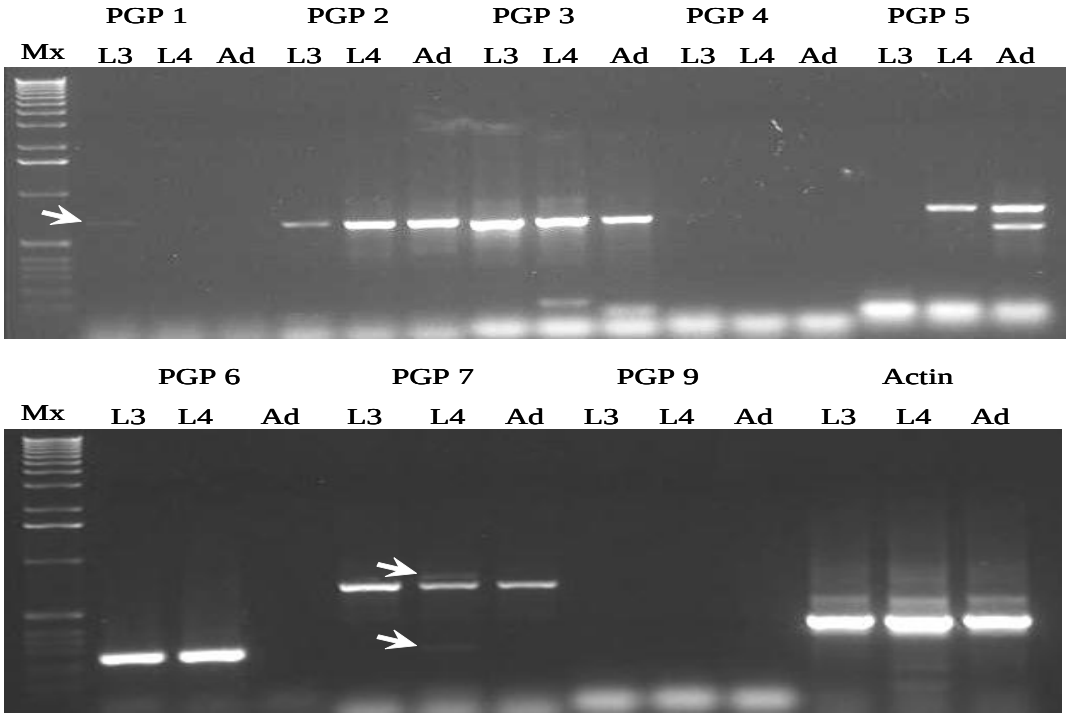
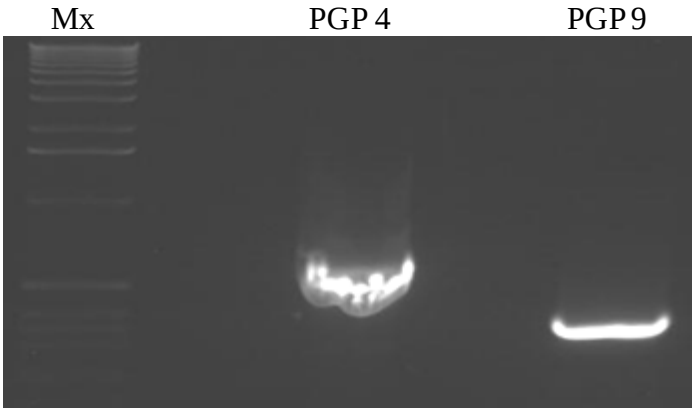


Figure 3.1: 1% agarose gel electrophoresis of PCR products to determine whether the *T. circumcincta* Pgp-specific primers only amplified their target Pgps using plasmids containing the respective partial Pgp sequences as template DNA. The primers designed against PGPs 1, 2, 3, 4, 5, 6, 7 and 9 were specific for their target whilst PGP 8, 10 and 11 primers were amplifying non target Pgps. Each of the bands on the agarose gel was the expected size of the PCR product as stated in Table 2.3. Positive and negative controls are not shown.



A



B

Figure 3.2: 1% agarose gel electrophoresis of PCR products to determine whether the Pgp specific primers only amplified their target Pgps. **A**: IVM susceptible *T. circumcincta* exsheathed L_3 and L_4 SMART cDNA and single-stranded adult cDNA used as template against primers for PGPs 1, 2, 3, 4, 5, 6, 7 and 9. Weak bands have been arrowed to aid identification. **B**: IVM susceptible adult 5' RACE-ready cDNA used as template against primers for PGPs 4 and 9.

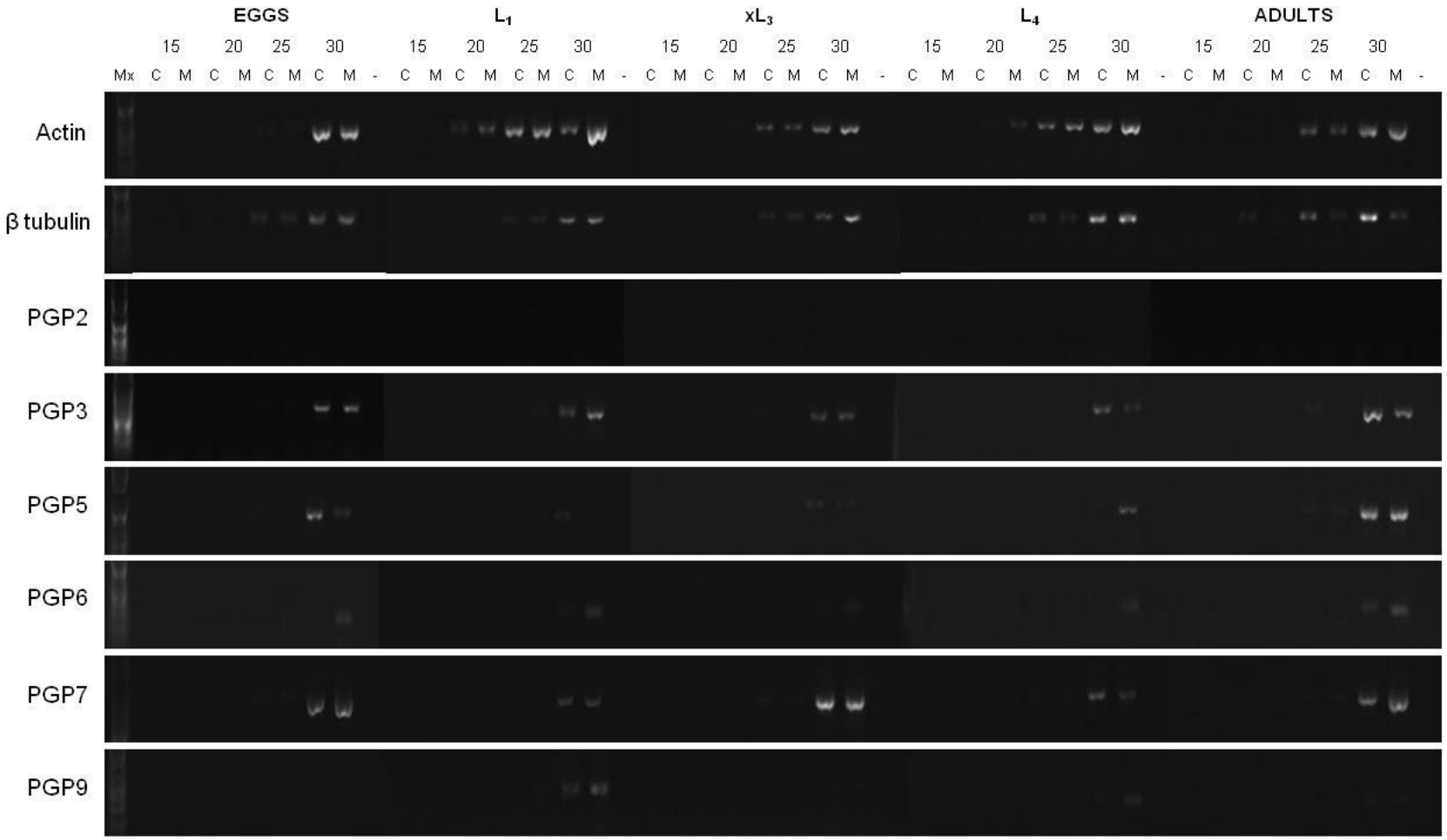
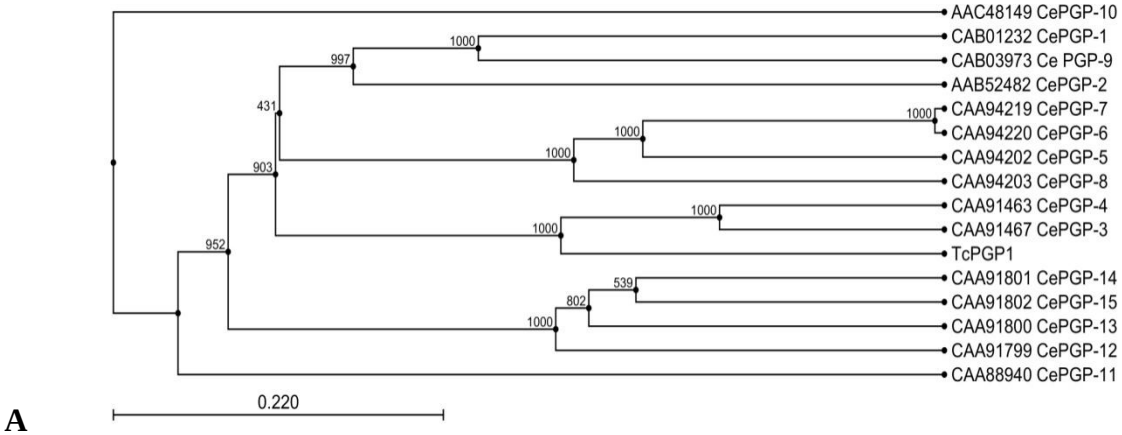
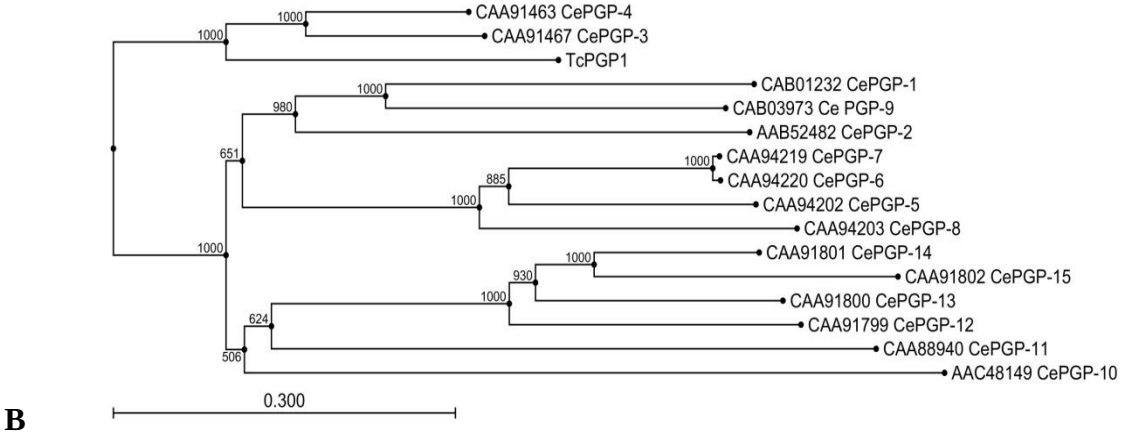


Figure 3.3: 1% agarose gel of the semi-quantitative PCR products using the Pgp gene specific primers. Each set of primers was used to amplify cDNA from eggs, L_1, xL_3, L_4 and adult *T. circumcincta* of either the CVL (C) or MOTRI (M) isolate. 5µL of PCR product was removed after 15, 20, 25 and 30 cycles. A negative control (-) was included for each lifecycle stage – primer combination and the gel marker lane (Mx) containing TrackIt™ 1Kb Plus DNA Ladder (Invitrogen) is on the left.

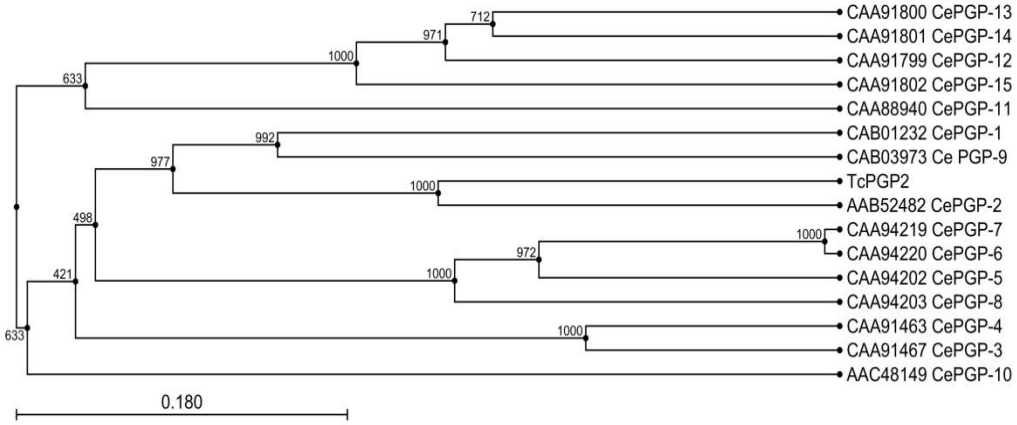


A

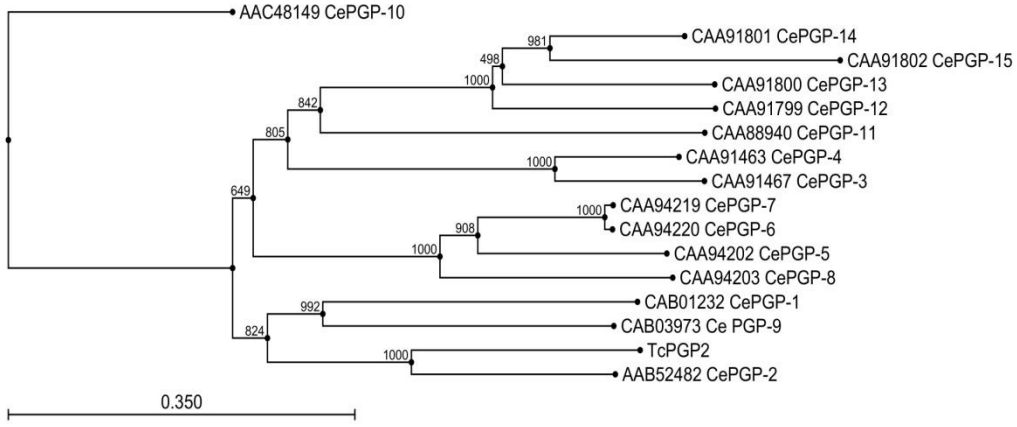


B

Figure 3.4: Phylogenetic trees generated using the CLC DNA Workbench package showing the relationship of *T. circumcincta* PGP1 to 15 *C. elegans* Pgp genes. **A**: Tree generated using the UPGMA distance matrix with a bootstrap value of 1000. **B**: Tree generated using the Neighbour-joining distance matrix with a bootstrap value of 1000. Both trees indicate *T. circumcincta* PGP1 is closest to *CePgp-3* and *CePgp-4*.

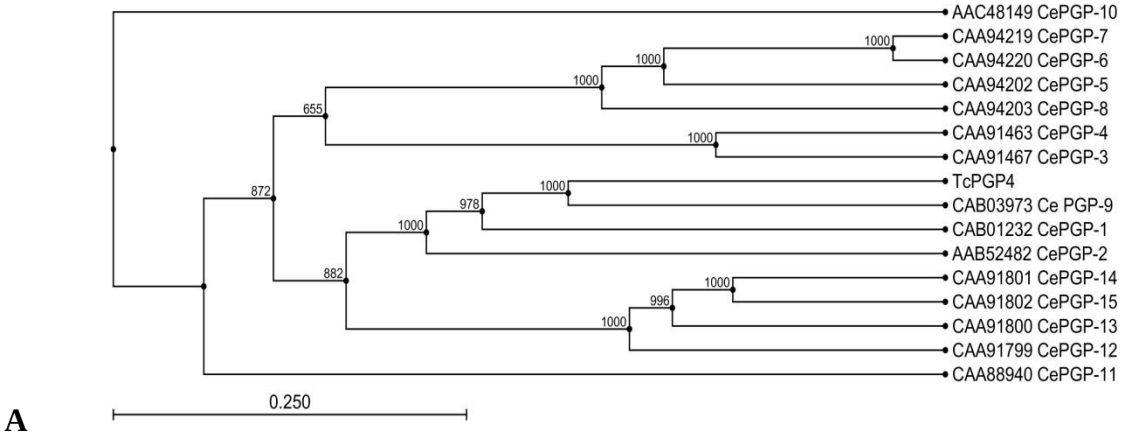


A

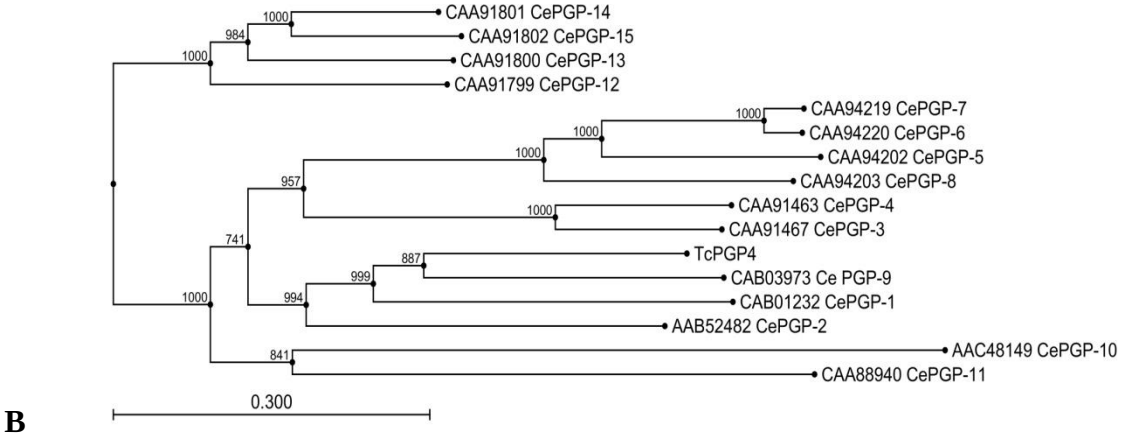


B

Figure 3.5: Phylogenetic trees generated using the CLC DNA Workbench package showing the relationship of *T. circumcincta* PGP2 to 15 *C. elegans* Pgp genes. **A**: Tree generated using the UPGMA distance matrix with a bootstrap value of 1000. **B**: Tree generated using the Neighbour-joining distance matrix with a bootstrap value of 1000. Both trees indicate *T. circumcincta* PGP2 is closest to *CePgp*-2. PGP2 will now be known as *TeciPgp-2* NBD2.

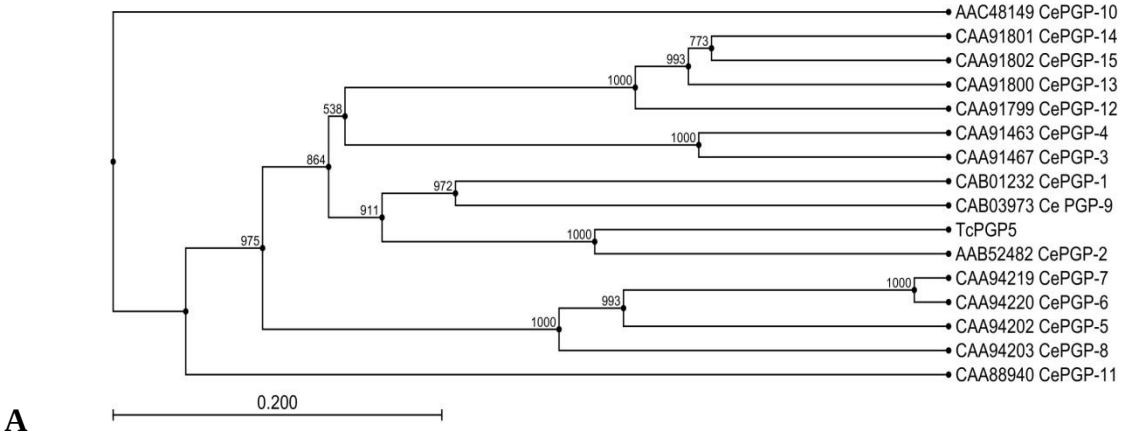


A

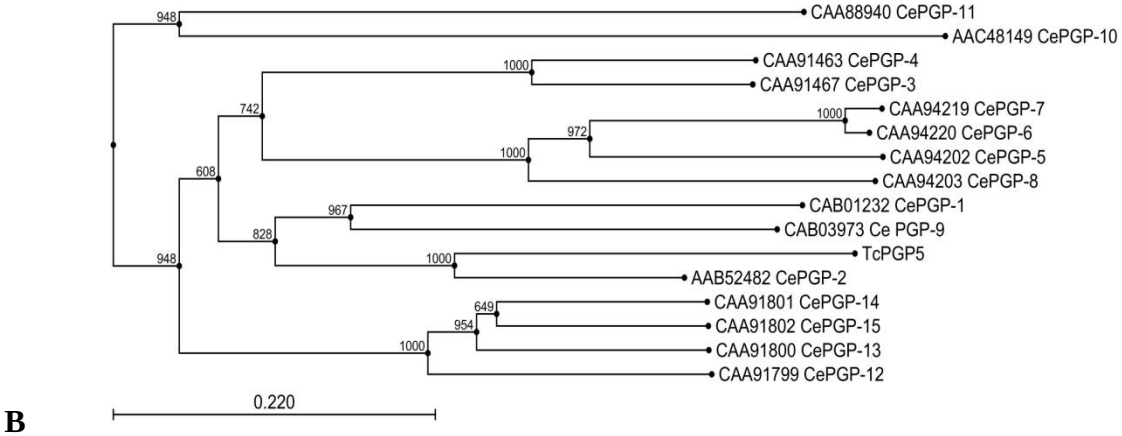


B

Figure 3.6: Phylogenetic trees generated using the CLC DNA Workbench package showing the relationship of *T. circumcincta* PGP4 to 15 *C. elegans* Pgp genes. **A**: Tree generated using the UPGMA distance matrix with a bootstrap value of 1000. **B**: Tree generated using the Neighbour-joining distance matrix with a bootstrap value of 1000. PGP4 will now be known as *TeciPgp-9* NBD1 as it aligns most closely in both trees to *CePgp-9*.

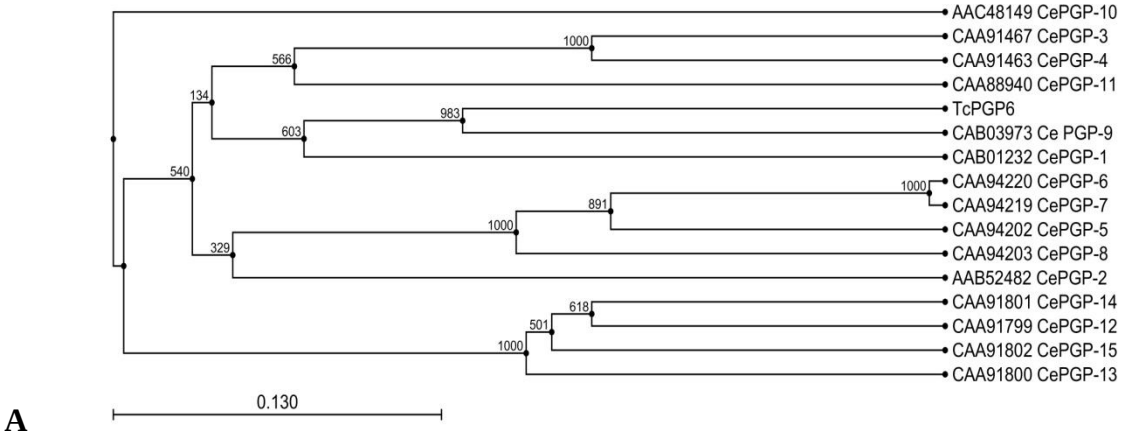


A

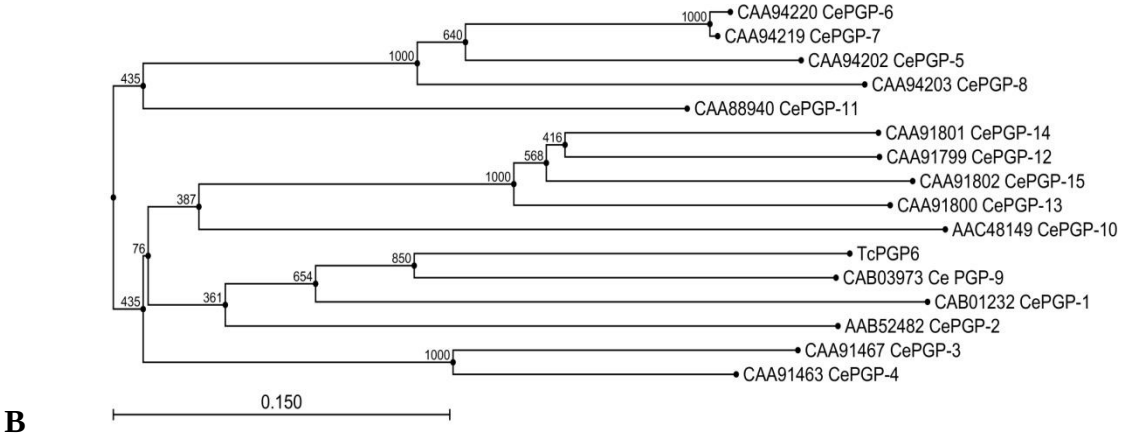


B

Figure 3.7: Phylogenetic trees generated using the CLC DNA Workbench package showing the relationship of *T. circumcincta* PGP5 to 15 *C. elegans* Pgp genes. **A**: Tree generated using the UPGMA distance matrix with a bootstrap value of 1000. **B**: Tree generated using the Neighbour-joining distance matrix with a bootstrap value of 1000. Due to this analysis PGP5 will now be known as *TeciPgp-2* NBD1 as it aligns most closely to *CePgp-2*.

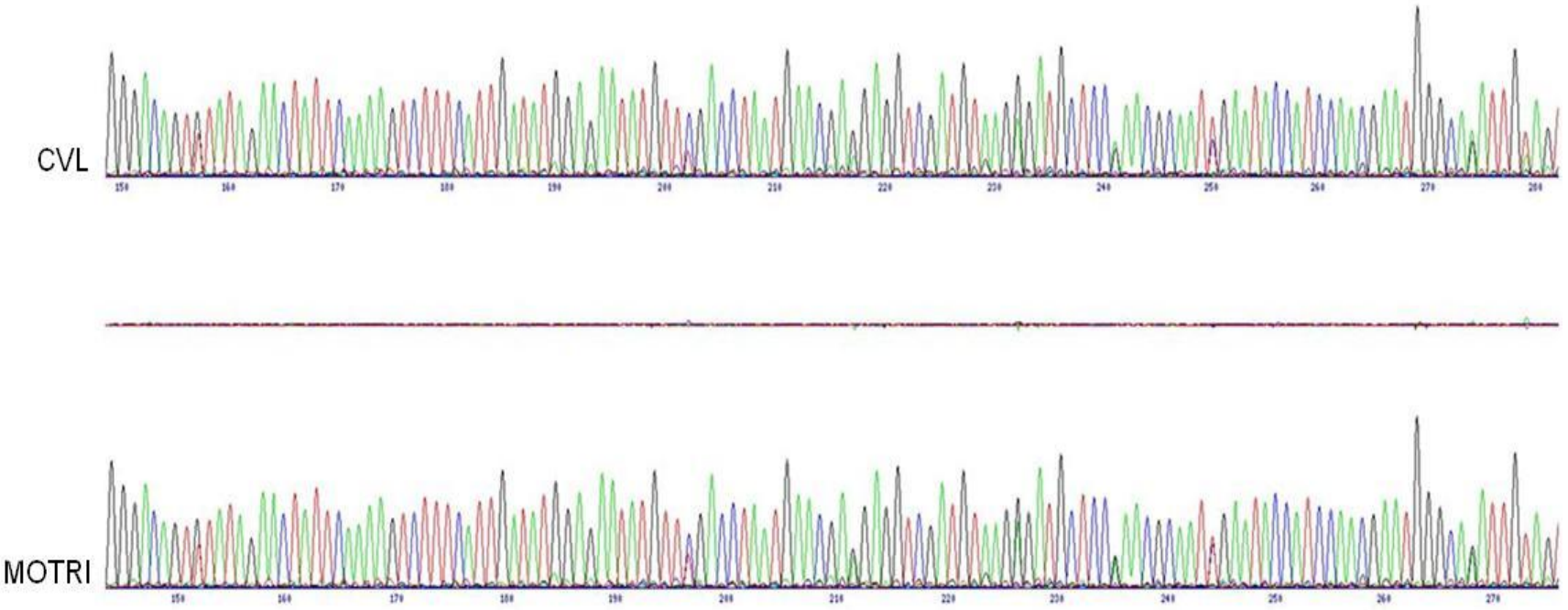


A

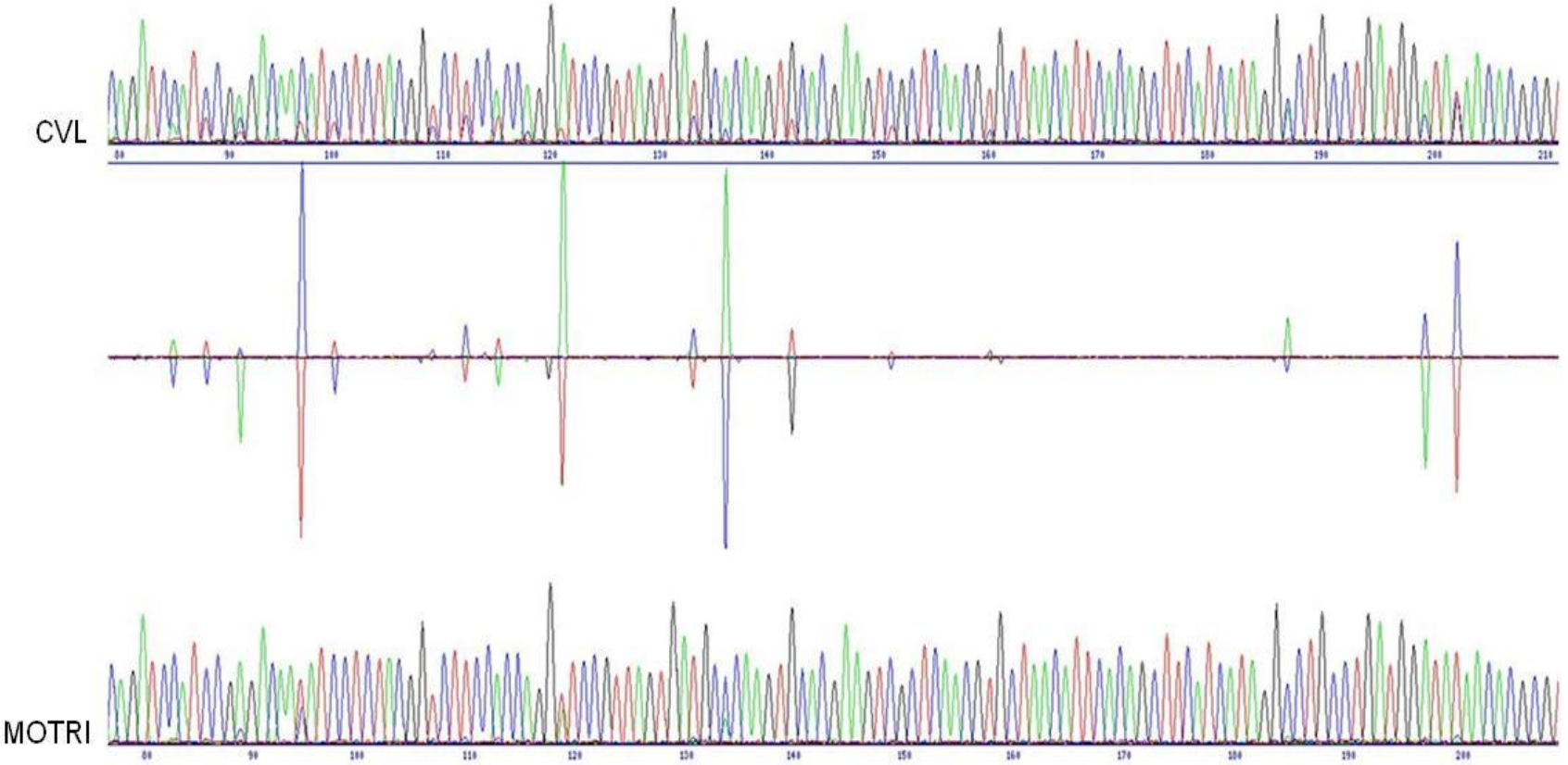


B

Figure 3.8: Phylogenetic trees generated using the CLC DNA Workbench package showing the relationship of *T. circumcincta* PGP6 to 15 *C. elegans* Pgp genes. **A**: Tree generated using the UPGMA distance matrix with a bootstrap value of 1000. **B**: Tree generated using the Neighbour-joining distance matrix with a bootstrap value of 1000. PGP6 will now be known as *TeciPgp-9* NBD2 as it aligns most closely to *CePgp-9*.



A



B

Figure 3.9: SeqDoC alignments comparing sequence generated from CVL and MOTRI adult *T. circumcincta* for PGP3 and *TeciPgp-9* NBD2. **A**: Partial SeqDoc alignment for PGP3. **B**: Partial SeqDoc alignment for *TeciPgp-9* NBD2. In each alignment the top chromatogram was the CVL sequence and the lower chromatogram the MOTRI sequence. The middle trace indicates where a SNP has occurred between the two sequences (difference profile). The regions shown in this figure are not the same parts of the Pgp molecule.

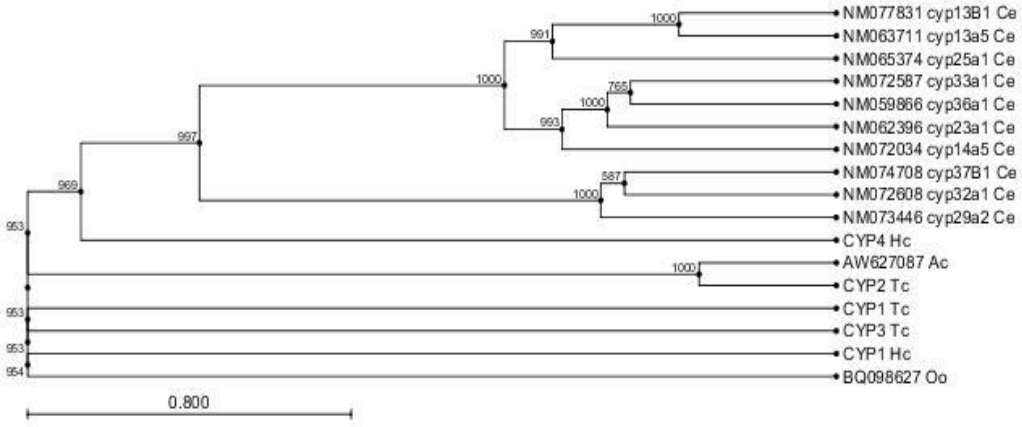
A

```
  CVL TeciPgp-9 NBD2  CCTACTCATC TACGTGCCCA TATAGCTTTG GTATCACAAG AGCCGATTCT 50
MOTRI TeciPgp-9 NBD2  CCTACTCATC TACGTGCCCA TATAGCTTTG GTATCGCAAG AGCCGATTCT 50

  CVL TeciPgp-9 NBD2  TTTCGACAGA TCCATCCGAG ACAACATCCT CTACGGTCTT CCACCAGGAT 100
MOTRI TeciPgp-9 NBD2  TTTCGACAGA TCCATCCGAG ACAATATCCT CTACGGTCTT CCACCAGGTT 100

  CVL TeciPgp-9 NBD2  CCGTTAGTGA TGCACAAGTG CACGAAGTCG CTCAACGTGC TAACATTCAC 150
MOTRI TeciPgp-9 NBD2  CCGTTAGTGA TGCCCAAGTG CACGAAGTCG CTCAACGTGC TAACATTCAC 150

  CVL TeciPgp-9 NBD2  AGCTTCATCA TAGGCCTGCC TGATGGATAT AACACGCGTG CAGGAGAAAA 200
MOTRI TeciPgp-9 NBD2  AGCTTCATCA TAGGCCTGCC TGATGGATAT AACACGCGTG CAGGAGAAAA 200

  CVL TeciPgp-9 NBD2  GGGAGCGC 208
MOTRI TeciPgp-9 NBD2  GGGAGCGC 208
```

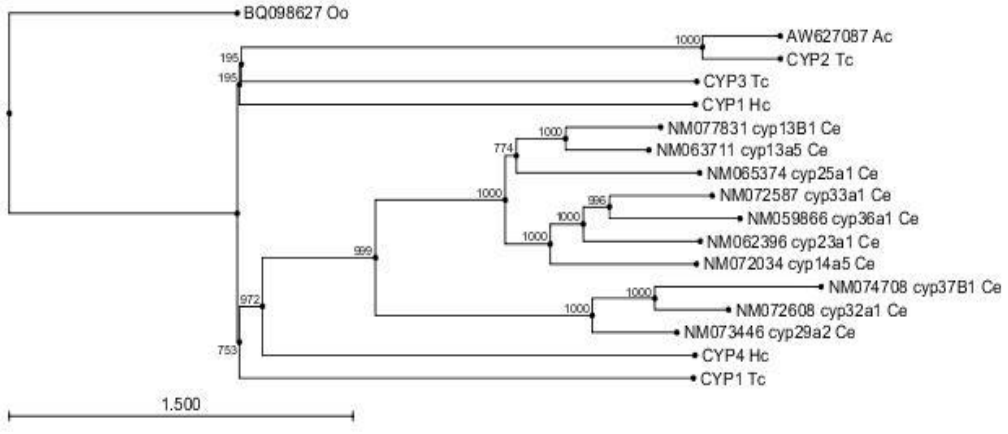
B

```
  CVL TeciPgp-9 NBD2  PTHLRAHIAL VSQEPILFDR SIRDNILYGL PPGSVSDAQV HEVAQRANIH
MOTRI TeciPgp-9 NBD2  PTHLRAHIAL VSQEPILFDR SIRDNILYGL PPGSVSDAQV HEVAQRANIH

  CVL TeciPgp-9 NBD2  SFIIGLPDGY NTRAGEKGA
MOTRI TeciPgp-9 NBD2  SFIIGLPDGY NTRAGEKGA
```

Figure 3.10: Clustal W alignments of the nucleotide and protein sequences of *TeciPgp-9* NBD2 comparing the sequences derived from adult *T. circumcincta* from the MOTRI and CVL isolates. **A**: Nucleotide alignment with the SNPs highlighted in red text. **B**: Protein alignment of translated coding sequence showing that the respective SNPs identified are synonymous.



A



B

Figure 3.11: Phylogenetic trees showing the relationship of *T. circumcincta* CYPs to CYP sequences from *H. contortus* (Hc), *C. elegans* (Ce), *A. caninum* (Ac) and *O. ostertagi* (Oo). The *T. circumcincta* (Tc) sequences are named in the order they were discovered. The *C. elegans* CYPs are given their official family name whilst *A. caninum* and *O. ostertagi* sequences are named according to their accession number. **A**: Tree generated using the UPGMA distance matrix with a bootstrap value of 1000. **B**: Tree generated using the CLC DNA Workbench package using the Neighbour-joining distance matrix with a bootstrap value of 1000.

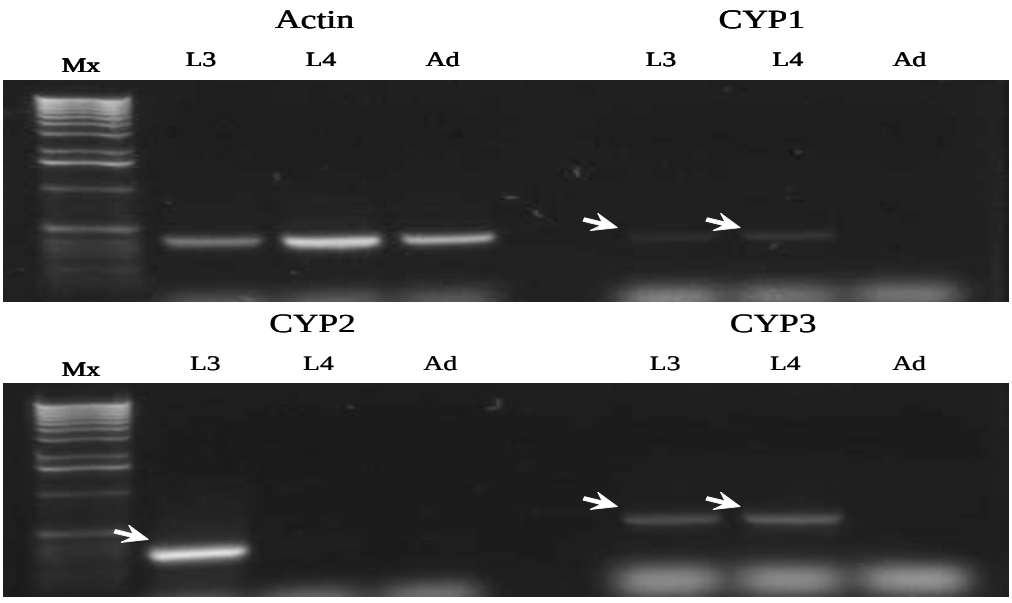


Figure 3.12: 1% agarose gel showing the specific bands for the three *T. circumcincta* CYP specific primers. The white arrows denote bands which were analysed further. A very weak band could also be observed from CYP2 versus L_4

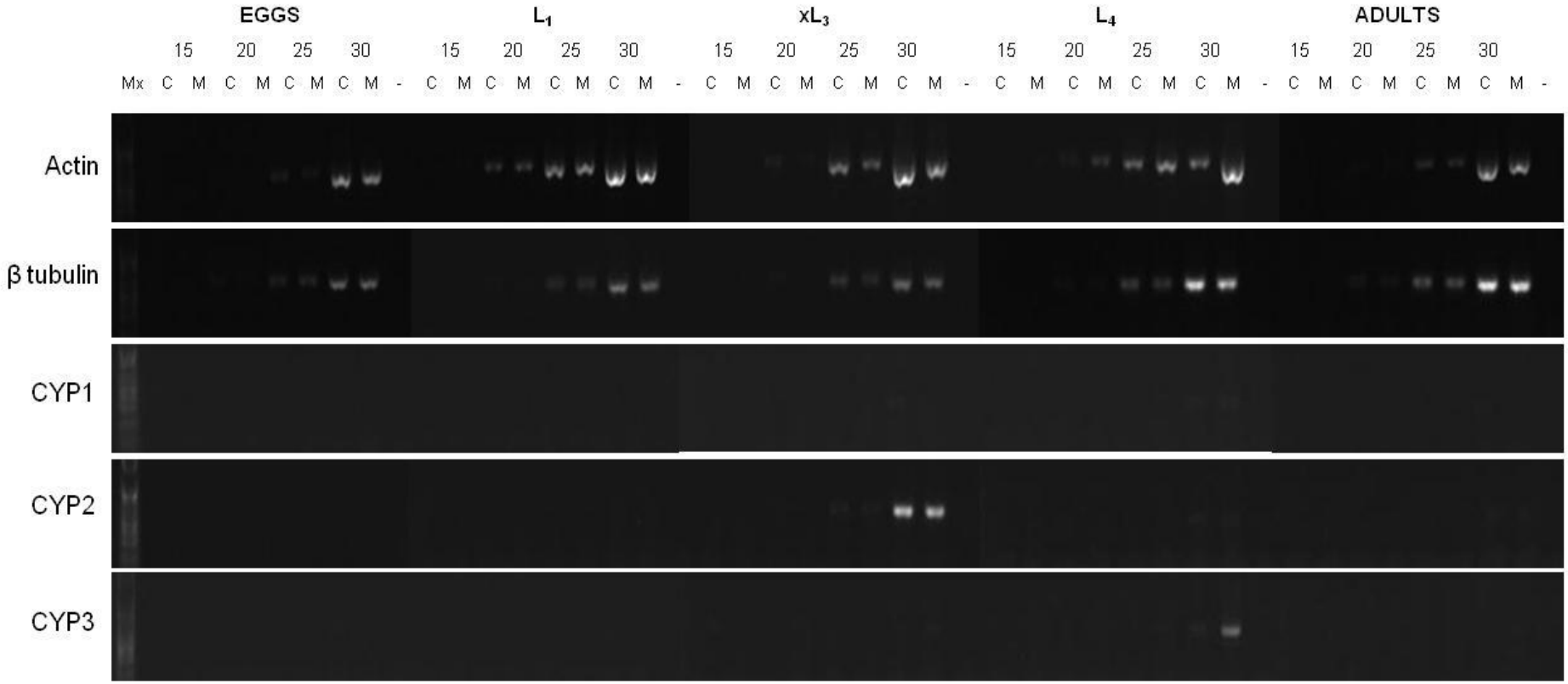


Figure 3.13: 1% agarose gel of the semi-quantitative PCR products using the CYP gene specific primers. Each set of primers was used to amplify cDNA from eggs, L_1, xL_3, L_4 and adult *T. circumcincta* of either the CVL (C) or MOTRI (M) isolate. 5µL of PCR product was removed after 15, 20, 25 and 30 cycles. A negative control (-) was included for each lifecycle stage – primer combination and the gel marker lane (Mx) containing TrackIt™ 1Kb Plus DNA Ladder (Invitrogen) is on the left.

Chapter 4: Quantitative TaqMan® Real-Time PCR to Measure the Expression of Candidate Ivermectin Resistance Genes in *Teladorsagia circumcincta* Isolates

4.1: Introduction

The role of candidate IVM resistance genes, such as the Pgps and CYPs, in the expression of resistance to the anthelmintics, has not been investigated to date in *T. circumcincta*. This is despite these genes being implicated in drug resistance in other systems, for example, over-expression of Pgps linked to drug resistance in cancerous tumours and over-expression of CYPs associated with drug-resistance in mosquitoes (Loo & Clarke, 1999; Nikou, Ranson, & Hemingway, 2003). Alongside this, work carried out in related parasitic nematodes such as *H. contortus* and *O. volvulus* has also suggested that differences in expression of these genes could play a role in the IVM-resistant phenotype, as described below.

The likely molecular targets of IVM appear to be the GABA and GluCl channels; concurrent mutations in three GluCl α-type subunits in *C. elegans* confer high levels of IVM resistance (Dent *et al.*, 2000; Njue *et al.*, 2004; Gilleard, 2006). However, more generic drug handling mechanisms, such as the Pgps also seem to play a role in the expression of an IVM resistance phenotype as shown by various studies utilising Pgp inhibitors. The co-administration of verapamil (a Pgp inhibitor) with MOX or IVM has been shown to significantly increase anthelmintic efficacy against MOX-selected *H. contortus* and the use of Pgp inhibitors in the LFIT caused reversion towards susceptibility in IVM-resistant *H. contortus* and *T. circumcincta* (Xu *et al.*, 1998; Bartley *et al.*, 2009). IVM or MOX selection in *H. contortus* and *O. volvulus* appear to place the Pgps under selection pressure as indicated by changes in allele frequencies and/ or loss of genetic diversity (Blackhall *et al.*, 1998b; Ardelli, Guerriero, & Prichard, 2005; Ardelli, Guerriero, & Prichard, 2006b). An increase in PgpA mRNA level has also been observed in IVM-selected *H. contortus* (Xu *et al.*, 1998).

CYPs are a large family of membrane-anchored proteins involved in the metabolism of a range of predominantly hydrophobic, exogenous and endogenous compounds, including drugs (Guengerich, 1991; Mansuy, 1998). CYPs appear to play a role in insecticide resistance in mosquitoes and *Drosophila* spp (Daborn *et al.*, 2002; Nikou, Ranson, & Hemingway, 2003; Le Goff *et al.*, 2003), whilst chloroquine resistance in malaria parasites also appears to be associated with increased CYP activity (Ndifor, Ward, & Howells, 1990; Barrett, 1998). Their role in drug resistance, particularly to IVM, which is hydrophobic, has not been extensively investigated in parasitic nematodes. This is possibly due to the enzymatic assays previously used not being sufficiently sensitive to detect CYP activity (Kotze, 1997; Barrett, 1998). The use of the CYP inhibitors, piperonyl butoxide and metyrapone has also been shown to increase the efficacy of BZ anthelmintics (McKellar & Jackson, 2004), providing further evidence of their role in generic xenobiotic metabolism.

If Pgps and CYPs are involved in generic drug handling mechanisms, for example, by preventing IVM from reaching its target sites in the pharyngeal and somatic muscle cells or by metabolising IVM into an inactive form, then changes in the expression profile of these genes would be expected when comparing drug–resistant and –susceptible isolates (Gilleard & Beech, 2007; Prichard & Roulet, 2007; Blackhall, Prichard, & Beech, 2008). Increased or reduced expression levels, reflected in mRNA levels, could be identified depending on the specific role of that particular candidate resistance gene in the parasite. A panel of candidate resistance genes, comprising 11 partial Pgp and 3 partial CYP gene sequences was identified from *T. circumcincta*, as described in Chapter 3. Initial semi-quantitative PCR results, also described in Chapter 3, indicated that differences in expression of these genes were evident when comparing well-characterised *T. circumcincta* isolates with known IVM resistance status.

The expression of the Pgp and CYP genes in the CVL, MOTRI and Post IVM MOTRI *T. circumcincta* isolates was investigated further, using the relative quantification (or $\Delta\Delta$Ct) real-time PCR method, as described below. In this method, the Ct for the test gene is normalised to the actin Ct values for the same isolate and life-cycle stage on the same reaction plate. Next, the expression of the gene in MOTRI is expressed relative to

the expression in CVL and the expression in Post IVM MOTRI expressed relative to the expression in MOTRI. The value obtained gives a relative fold change in expression. The first comparison (CVL vs. MOTRI) investigated constitutive expression differences i.e. those found between the IVM–susceptible and –resistant isolates "at rest" whilst the second comparison (MOTRI vs. Post IVM MOTRI) investigated inducible expression differences i.e. those found between the pre- and post-IVM exposure isolates.

4.2: Materials and methods

4.2.1: Design of primers and probes

TaqMan® primers and TaqMan® Minor Groove Binder (MGB) fluorogenic probes were designed using the Primer Express® software version 3.0 (Applied Biosystems). The consensus sequence for each gene, generated using standard and RACE PCR as described previously, was entered into Primer Express, and 10 combinations of forward and reverse primers and probes designed by the software. Default parameters were chosen to give short amplicons of less than 150bp, primers with a Tm of 58- 60°C, the difference in Tm between the primers to be less than 2°C and the primers to have a high G-C content. Default parameters were also chosen for the probes, briefly, the probe Tm was designed to be 8-10°C higher than the primers, to have no G on the 5' end, and to have a maximum of 30 bases between the primer and the probe. The primer and probe sets chosen for each gene are shown in Table 4.1.

The specificity of the primers and probes was validated by BLASTN searching the sequences against the EBI database (http://www.ebi.ac.uk/Tools/blast2/parasites.html) to determine if they would align to any other *T. circumcincta* sequence. The primer and probe sequences were also searched against the Lasergene EditSeq (DNASTAR Lasergene version 8) files generated for each of the Pgp and CYP specific gene sequences from the Seqman sequence alignments to confirm they were specific for the gene they were designed to target.

Primers and probes were ordered from Applied Biosystems (http://www3.appliedbiosystems.com). The primers were supplied lyophilised and were re-suspended in HPLC water (Sigma) to give a stock concentration of 100µM. The working dilution was 10µM and the primers were stored at -20°C in individual reaction aliquots. FAM labelled MGB probes were supplied in solution at 100µM, again a working dilution of 10µM was made, aliquots made and stored at -20°C.

Real-time PCR reagents

TaqMan® Gene Expression Master Mix, MicroAmp™ Optical 96-well plate and MicroAmp™ Optical adhesive film were obtained from Applied Biosystems.

4.2.2 Real-time PCR protocol

Real-time PCR reactions were set up by combining 10µL TaqMan® Gene Expression Master Mix, 1µL each of the forward and reverse primers, 0.25µL of the MGB probe and 5.75µL of water to give a total master mix volume of 18µL per reaction. Due to the light sensitivity of MGB probes, these were allowed to thaw in the dark and only briefly exposed to light whilst making up the master mix. The master mix was kept in the dark whilst being transferred to the PCR workstation area where it was dispensed into the 96-well PCR plate in 18µL aliquots. To each well of the plate, 2µL template, either cloned purified PCR product or *T. circumcincta* cDNA (diluted previously to 50ng/µL) was added. Duplicate wells were set up for each template used. Following the addition of the template, the plate was sealed with an adhesive film, ensuring that the film was pressed down firmly to prevent contamination and evaporation of the PCR reactions.

All real-time PCR runs were carried out on an ABI 7000 sequence detection system (Applied Biosystems). The reaction conditions were as follows: 95°C for 12 mins followed by 40 cycles of 95°C for 15 secs and 60°C for 60 secs. Collection of data, by detection of the FAM dye attached to the probe, occurred during the final phase of each cycle; this enabled a crossing threshold (Ct) value to be determined for each well. This

value shows at how many PCR cycles the linear phase of the PCR reaction crossed a pre-determined threshold. Triplicate real-time PCR runs were carried out for each gene under test. The baseline and threshold for each of the three runs for each gene were set identically and the Ct values recalculated to enable direct comparisons between the run data to be made, as described below.

4.2.3: Validation of efficiency of real-time primers and probes

To determine the efficiency of the primers and probes, individual gene fragments were cloned into pGEM®-T vectors, amplified and purified to provide sufficient plasmid template for the real-time PCR reactions. The gene sequences used to determine the efficiency of the primer and probe sets are shown in Table 4.2. Using the concentration of the plasmid and the plasmid size, the copy number for each gene specific plasmid was determined. A 10-fold dilution series ($10^{\wedge 8}$ to $10^{\wedge 3}$ copies) for each plasmid was set up to enable the efficiency of the primers and probes to be calculated.

Initial PCR reactions using the plasmid standards as template were carried out as described previously. From these reactions, a standard curve was generated for each primer and probe set. From the standard curve, a correlation coefficient (R^2) and value representing the slope of the standard curve were generated. An R^2 value of >0.99 was determined as showing the individual data points on the standard curve fitted the line sufficiently. From the slope value, the efficiency of the reaction could be calculated and expressed as a percentage as follows:

$$\text{Efficiency} = (10^{\wedge(-1/\text{slope})} - 1) \times 100$$

Ideally, the efficiency of the reaction should be 100%, this is achieved if the value of the slope is -3.33. If the efficiency was found to be too low or too high then that specific gene was removed from further analysis.

4.2.4: Analysis of real-time data

Relative quantification of the gene expression level was carried out using the comparative Ct (ΔΔCt) method. The raw Ct values for the Pgp or CYP (target) gene and the actin (endogenous control) gene for the three parasite isolates (CVL, MOTRI, Post IVM MOTRI) and five life-cycles stages (eggs, L_1, xL_3, L_4 and adults) were exported from the ABI 7000 SDS program into Excel (Microsoft Windows XP). The relative fold change in expression of a particular gene, as reflected in cDNA level, between two different isolates (either CVL vs. MOTRI or MOTRI vs. Post IVM MOTRI) was calculated for each life-cycle stage by following the three steps below:

Firstly, for each isolate and life-cycle stage combination, the target gene was normalised to the endogenous control:

$$\Delta Ct = Ct_{\text{target gene}} - Ct_{\text{endogenous control}}$$

Next, the sample was normalised to the calibrator sample: In CVL vs. MOTRI comparison, the calibrator was CVL and the sample MOTRI. In the Post IVM MOTRI vs. MOTRI comparison, the calibrator was MOTRI and the sample Post IVM MOTRI.

$$\Delta\Delta Ct = \Delta Ct_{\text{sample}} - \Delta Ct_{\text{calibrator}}$$

Finally, the relative fold change in expression between the two isolates, normalised to the endogenous control, was calculated using the formula:

$$\text{Fold change} = 2^{-\Delta\Delta Ct}$$

4.2.5: Statistical analysis of the real-time results

Triplicate real-time runs were carried out for each gene. To analyse the multiple real-time runs, the data were converted to a log scale to control for occasions when one replicate could be up-regulated and the next down-regulated. If the fold changes were averaged, a false reading would result. For example: Replicate 1 = 5-fold up-regulation, replicate 2 = 5-fold down-regulation (≈0.2). Average fold change = (5 + 0.2) / 2 = 2.6 average fold up-regulation, which is incorrect. Taking the log of the two fold changes shown above gives the following equation (C.-D. Mayer, Pers. Comm.), which calculates the correct average:

$$\text{Log (5)} + \text{log (1/5)} = \text{log (5)} - \text{log (5)} = 0 \text{ fold change in expression.}$$

The ΔCt values were log transformed as follows, where A is the sample and B the calibrator:

$$\text{Expression A / Expression B} = 2\text{^ (-(}\Delta A - \Delta B)$$

Applying $\log_2$ to each side gives:

$$\text{Log}_2 \text{ (Expression A)} - \text{Log}_2 \text{ (Expression B)} = (-\Delta A) - (-\Delta B)$$

This means to log transform the ΔCt values, a positive ΔCt value becomes negative and vice versa. Hence, the log transformed values for $2^{-\Delta\Delta Ct}$ were calculated as follows: Firstly, the average of the triplicate results was obtained by calculating the average log transformed ΔCt value for both the sample and calibrator. The average log ratio was calculated by:

Average log ratio = Average log expressionA – Average log expressionB

The average fold change was calculated as:

$2^{\wedge \text{Average log ratio}}$

A two tailed T-Test was performed in Excel (Microsoft Windows XP) on the triplicate log transformed ΔCt values to generate a P-value.

4.3: Validation of real-time primers and probes

Each primer and probe set (Table 4.1) was validated against a duplicate 10-fold serial dilution, calculated to give 10^8 to 10^3 gene copies, of pGEM®-T plasmids containing the target gene. The standard curve, generated using the threshold crossing values (Ct) for each plasmid copy number, was used to calculate the efficiency of the reaction using the efficiency calculation, as described earlier. A representative plot of raw data obtained for PGP7 from the ABI 7000 sequence detection system is shown in Figure 4.1 alongside the standard curve generated from the Ct values for PGP7. Each gene's primer and probe sets (PGP1, *TeciPgp-2* NBD2, PGP3, *TeciPgp-9* NBD1, *TeciPgp-2* NBD1, *TeciPgp-9* NBD2, PGP7, PGP8, PGP9, PGP10, PGP11, CYP1, CYP2 and CYP3) were quantified on a reaction plate alongside the actin containing plasmids. This allowed the relative efficiency of the amplification of the genes to be compared to actin both by visual inspection of the standard curves and also by comparing the efficiency values. The efficiency values for the test genes are shown in Table 4.3; actin has more than one set of values as it was validated several times against the other genes. Efficient primer and probes sets have an efficiency value close to 100%, indicating that the amount of product doubles with each real-time cycle (Bustin *et al.*, 2009). The *TeciPgp-9* NBD1 primer and probes were validated twice as the efficiency value for the first run was so low, compared

to the other genes. An example of the difference in standard curves for *TeciPgp-9* NBD1 compared to actin is shown in Figure 4.2. Due to the relatively low efficiency, the expression of this transcript in *T. circumcincta* was not investigated further.

Each experimental plate quantified the expression of one of the test genes and actin for each isolate and life-cycle stage combination in duplicate alongside a 10-fold serial dilution of the actin plasmids. The serial dilution was made up fresh from a frozen aliquot of the actin plasmid for each experimental run. The efficiency of the actin standards was calculated for each run to monitor if there were any significant changes over time. Alongside this, the duplicate actin Ct values for each isolate and life-cycle stage were recorded and a graph (Figure 4.3) produced showing the average actin Ct values and the standard deviation. To be able to trust the fold changes obtained using the relative quantification method, the Ct values for the control gene between the sample and calibrator need to be comparable. From Figure 4.3, it is possible to see that comparisons between life-cycle stages were not feasible due to the variation in actin Ct values. Also, the comparison between MOTRI and Post IVM MOTRI in the L_4 stage cannot be made due to the large difference in actin Ct values between the two samples.

4.4: Relative quantification of expression of P-glycoproteins in *T. circumcincta*

4.4.1: Constitutive Pgp expression: CVL vs. MOTRI

The expression of Pgps in CVL and MOTRI was compared to investigate the constitutive expression of these genes in *T. circumcincta*. Using CVL as the calibrator, the expression of each gene in MOTRI was calculated as a fold change in expression. The expression of each gene in the CVL isolate was, therefore, always 1, whilst a value greater than 1 for MOTRI indicated that the gene had increased expression in MOTRI and a value between 0 and 1 for MOTRI indicated that the gene had reduced expression in MOTRI. The closer the relative fold change in expression is to 1, the more equal the level of expression between the MOTRI and CVL isolate. Statistical analysis of triplicate runs for each gene generated an average fold change in expression together with a P-value; these results are shown in Table 4.4.

These results show that there was a statistically significant constitutive increase in expression of *TeciPgp-9* NBD2 across all life-cycle stages, the most dramatic being a 55.27-fold increase ($P<0.01$), a 5.06-fold increase ($P<0.05$), a 17.49-fold increase ($P<0.01$), a 14.04-fold increase ($P<0.05$) and a 6.75-fold increase ($P<0.01$) of *TeciPgp-9* NBD2 in MOTRI compared to CVL eggs, L_1, xL_3, L_4 and adults, respectively. The average fold change in expression of *TeciPgp-9* NBD2 in MOTRI relative to CVL is shown in Figure 4.4; the lines for each bar of the graph show the range in relative expression of *TeciPgp-9* NBD2 in MOTRI from the triplicate real-time runs.

There was a statistically significant constitutive reduction in expression of *TeciPgp-2* NBD2 of 6.02-fold ($P<0.05$), 8.50-fold ($P<0.01$), 7.61-fold ($P<0.05$) and 7.12-fold ($P<0.01$) in MOTRI eggs, L_1, xL_3 and L_4 compared to CVL, respectively, whilst the reduced expression of *TeciPgp-2* NBD2 in adults was not statistically significant. However, the expression of *TeciPgp-2* NBD1 did not follow the same expression pattern as *TeciPgp-2* NBD2 as shown in Figure 4.5, parts A and B. The relative fold change in expression of *TeciPgp-2* NBD1 was a 1.04-fold reduction, a 1.37-fold reduction, 1.07-fold increase, 1.37-fold increase and 1.78-fold increase in MOTRI eggs, L_1, xL_3, L_4 and adults compared to CVL, respectively, however none of these changes in expression were statistically significant, although the change in expression in the adult stage was approaching significance at the 5% level.

Alongside the changes in gene expression across all life-cycle stages observed for *TeciPgp-9* NBD2 and *TeciPgp-2* NBD2, there were also changes in the constitutive expression of other genes in individual life-cycle stages, as shown in Table 4.4. For example, the expression of PGP1 in MOTRI eggs was 3.9-fold greater than in CVL eggs ($P<0.05$) whilst in MOTRI L_4, the expression of PGP1 was 1.46-fold lower than in CVL L_4. In PGP3, only one life-cycle stage had a statistically significant difference ($P<0.05$) in expression between MOTRI and CVL; MOTRI eggs were shown to have a 2-fold increase in expression compared to CVL eggs. The expression of PGP9 in MOTRI L_1 was 1.36-fold higher than in CVL L_1 ($P<0.05$) whilst in MOTRI L_4 the expression was 2-fold higher than CVL L_4 ($P<0.01$). The expression of PGPs 10 and 11 in MOTRI varied between life-cycle stages. In the xL_3 stage, there was a 1.88-fold increase in PGP10 expression in

MOTRI compared to CVL ($P<0.05$) whilst in the L_4 stage there was a 3.99-fold reduction in PGP10 expression in MOTRI compared to CVL ($P<0.01$). In the L_4 stage there was a 2.68-fold reduction in PGP11 expression in MOTRI compared to CVL ($P<0.05$), whilst in the adult stage, there was a 1.6-fold increase in PGP11 expression in MOTRI compared to CVL ($P<0.05$).

4.4.2: Confirmation of primer specificity for *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2

As the most significant results for the constitutive gene expression differences, showing statistically significant expression differences across all the life-cycle stages, the identity of the real-time PCR products for *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 needed to be confirmed. Two real-time PCR products from each of the *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 real-time runs were selected at random, cloned into pGEM®-T vector and transformed into JM109 cells as described in Section 2.3.7. Two white colonies were selected for each PCR product and, following overnight culture in LB broth, purified plasmids were submitted for sequencing. The sequences obtained were trimmed to remove vector sequence and aligned against the consensus sequences for the eleven partial *T. circumcincta* Pgps identified in Chapter 3, using SeqMan (DNASTAR Lasergene version 8), as shown in Table 4.5. Of the four sequences from *TeciPgp-2* NBD2 real-time PCR, three aligned exclusively to the *TeciPgp-2* NBD2 consensus sequence whilst the fourth was found not to contain an insert. Of the three sequences from *TeciPgp-9* NBD2, all three aligned exclusively to *TeciPgp-9* NBD2. This confirmed the specificity of the *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 primer and probe sets.

4.4.3: Biological replicates of constitutive *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 expression

Biological replicates, comprising different pools of *T. circumcincta* MOTRI and CVL eggs, xL_3 and adults to those used in the original experiments were obtained, processed as described previously, and used to quantify the expression of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2, to confirm the original results. The actin Ct values for these biological replicates were compared to the original actin Ct values as is shown in Figure 4.6. This comparison indicated that the actin Ct values for the biological replicates of the

CVL and MOTRI eggs were not comparable to each other and also were different to the original average actin Ct values. As such, the constitutive gene expression of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 in eggs in the biological replicates was excluded from further analysis. The actin Ct values for the xL3 and adult stages between MOTRI and CVL were sufficiently close to each other to allow the relative fold changes in expression of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 in the biological replicates to be calculated, as is shown in Table 4.6. The expression of *TeciPgp-2* NBD2 in MOTRI xL_3 was 14.71-fold lower than in CVL xL_3 ($P<0.05$), compared to the 7.61-fold decrease found in the original triplicate results. The expression of *TeciPgp-2* NBD2 in MOTRI adults was 13.28-fold lower than in CVL adults ($P<0.01$), compared to the original 2.72-fold decrease in the original results. There was a 14.93-fold increased expression of *TeciPgp-9* NBD2 in MOTRI xL_3 compared to MOTRI ($P<0.01$) which was similar to the 17.49-fold increased expression in this gene found in the original results. The result for *TeciPgp-9* NBD2 for adult constitutive expression in the biological replicate samples was not statistically significant.

4.4.4: *In vivo* inducible Pgp expression: MOTRI vs. Post IVM MOTRI

Inducible expression of Pgps was investigated by carrying out the relative quantification of these genes in the *T. circumcincta* Post IVM MOTRI isolate compared to the MOTRI isolate. This was to see whether adult parasites which survived *in vivo* exposure to IVM, as described in Chapter 2, expressed the Pgp genes at a different level compared to unselected parasites of the same isolate. In the same way that the CVL isolate was used as a calibrator in the constitutive gene expression comparison above, the MOTRI isolate was used as the calibrator in this comparison, such that the expression of the Pgps in the Post IVM MOTRI samples was described as a fold change relative to the MOTRI sample. The results of this comparison are shown in Table 4.7. Due to the difference in actin Ct values between MOTRI and Post IVM MOTRI L_4 (Figure 4.3), this comparison was excluded from any further analysis.

Unlike the constitutive gene expression comparison, there were no changes in expression of any one gene across all life-cycle stages. All the statistically significant changes in the larval stages reflected increases in expression, whilst the statistically

significant expression changes observed in the adult stage all showed reduced expression in the Post IVM MOTRI isolate compared to the MOTRI isolate. In the L_1 stage, there was a 3.42-fold increase in *TeciPgp-2* NBD2 expression ($P<0.05$), a 1.58-fold increase in PGP3 expression ($P<0.05$) and a 3.16-fold increase in *TeciPgp-2* NBD1 expression in Post IVM MOTRI compared to MOTRI ($P<0.01$). In the xL_3 stage, there was a 1.67-fold increase in PGP1 expression ($P<0.01$), a 1.54-fold increase in *TeciPgp-9* NBD2 expression ($P<0.01$) and a 3.84-fold increase in PGP8 expression in Post IVM MOTRI compared to MOTRI ($P<0.01$). In the adult stage, there was a 1.78-fold reduction in PGP7 expression ($P<0.05$), a 2.12-fold reduction in PGP8 expression ($P<0.05$) and a 2.27-fold reduction in PGP11 expression in Post IVM MOTRI compared to MOTRI ($P<0.05$).

4.4.5: *In vivo* inducible *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 expression: IVM survivors

Surviving adult parasites of the MOTRI isolate were collected from the abomasum of donor sheep three days after a full therapeutic dose of IVM, as described in Sections 2.1.1 and 2.1.6. This was to investigate whether the expression of Pgps was transiently altered during the period of IVM exposure *in vivo*. The average actin Ct values for the IVM survivors, original CVL and MOTRI adults and the biological replicates CVL and MOTRI adults are shown in Figure 4.7. The biological replicate MOTRI adults were chosen as the calibrator sample due to the similarity of their actin Ct values, as shown in Figure 4.7, and because they were of the same isolate to the IVM survivors. The relative fold change in expression of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 in the IVM survivors was calculated compared to the MOTRI adult biological replicate sample, as shown in Figure 4.8. In the adult IVM survivors there was a 13.68-fold reduction in expression of *TeciPgp-2* NBD2 relative to the expression in the biological replicates MOTRI adults ($P<0.05$) whilst the 1.82-fold increase in expression of *TeciPgp-9* NBD2 in the IVM survivors compared to the biological replicates MOTRI adults was not statistically significant.

4.4.6: *In vitro* inducible *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 expression: Larval migration inhibition test

The *in vitro* inducible expression of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 was investigated in MOTRI L_3 by utilising the LMIT. Using the standard LMIT method, as described in Section 2.2.1, a concentration of 2.5µg/mL IVM was chosen as the discriminating dose to allow approximately 10% of the L_3s of the MOTRI isolate to migrate through a Baermann chamber in the presence of the drug. However, on scaling up the method (Section 2.2.2) and with the larvae also being subsequently exsheathed, this concentration was found to be too high as the percentage migration was less than 1%. Thus, a lower concentration of 1µg/mL IVM was used. Three pools of xL_3 were obtained from this experiment; one pool (unexposed) comprised the xL_3s which had successfully migrated through the control (no IVM) Baermann chamber, the second pool (migrators) comprised the xL_3s which had successfully migrated through the Baermann chamber in the presence of 1µg/mL IVM and the third pool (non-migrators) comprised the xL_3s which failed to migrate through the Baermann chamber in the presence of 1µg/mL IVM. The percentage of xL_3s which migrated in the presence of IVM was 38.75% whilst in the control sample the proportion of xL_3s which migrated was 96.25%.

Real-time PCR was carried out to quantify the expression of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 in these three samples and also against the biological replicate MOTRI xL_3 sample. For *TeciPgp-2* NBD2, there was no statistically significant difference in expression between the migrators and non-migrators or between the non-migrators and unexposed samples. There was a statistically significant 1.88-fold increase in *TeciPgp-2* NBD2 expression in the unexposed sample compared to the migrators sample ($P<0.05$). For *TeciPgp-9* NBD2, there was no statistically significant difference in expression between the unexposed, migrators or non-migrators. When compared to the biological replicate MOTRI sample, there was no statistically significant difference in expression of *TeciPgp-2* NBD2 in any of the three LMIT samples and for *TeciPgp-9* NBD2, the only statistically significant difference was a 1.69-fold increase in expression in the unexposed sample compared to the MOTRI xL_3 biological replicate sample.

4.5: Relative quantification of expression of Cytochrome P450s in *T. circumcincta*

4.5.1: *In vivo* constitutive Cytochrome P450 expression: CVL vs. MOTRI

As with the Pgps, the relative constitutive expression of CYP1, 2 and 3 was investigated in eggs, L_1, xL_3, L_4 and adults of the CVL and MOTRI *T. circumcincta* isolates using real-time PCR. The specific primers and probes for the three CYP genes are shown in Table 4.1. Following triplicate real-time runs, the data were statistically analysed to determine the average fold change in expression of the genes in the MOTRI isolate compared to the CVL isolate, the statistical significance of each result and the range in relative expression. As before, a value greater than 1 for MOTRI indicated that the gene had increased expression compared to the CVL isolate, whilst a value between 0 and 1 for MOTRI indicated that the gene had reduced expression in MOTRI. The closer the relative fold change in expression was to 1, the more equal the level of expression between the MOTRI and CVL isolate. The results of this analysis are shown in Table 4.8. Statistically significant increases in expression of 11.73-fold for CYP1 ($P<0.01$) and 5.25-fold for CYP2 ($P<0.01$) were observed when comparing CVL and MOTRI eggs. Statistically significant reductions in expression of 1.56-fold for CYP1 ($P<0.05$), 1.98-fold for CYP2 ($P<0.05$) and 2.32-fold for CYP3 ($P<0.05$) were observed in the L_1 stage. In MOTRI adults there was a statistically significant 2.82-fold increase in expression of CYP1 ($P<0.01$) and a 2.60-fold increase in expression of CYP2 compared to CVL adults.

4.5.2: *In vivo* inducible Cytochrome P450 expression: MOTRI vs. Post IVM MOTRI

The relative inducible expression of the three CYP genes in *T. circumcincta* Post IVM MOTRI was determined compared to the MOTRI isolate. No comparison was made between the L_4 stage of these isolates as the average actin control gene Ct values, used to normalise the CYP gene data, was not comparable between the isolates as shown in Figure 4.3. As shown in Table 4.9, there were statistically significant reduced relative expression levels of 1.88-fold for CYP1 ($P<0.05$) and 1.86-fold for CYP2 ($P<0.05$) in Post IVM eggs compared to MOTRI eggs. For the same two CYP genes, in the L_1 stage there was an increase in relative expression of 2.91-fold ($P<0.05$) and 3.23-fold ($P<0.01$), respectively. The only statistically significant change in expression in the xL_3 stage was a 1.38-fold ($P<0.05$) increase in relative expression of CYP3 in Post IVM MOTRI compared to

MOTRI. Statistically significant reduced fold changes in expression of CYP1, 2 and 3 when comparing the adult Post IVM MOTRI to adult MOTRI of 1.23-fold ($P<0.05$), 3.08-fold ($P<0.01$) and 2.07-fold ($P<0.05$) were also observed.

4.6: Discussion

Using the relative quantification, or ΔΔCt, real-time PCR method, the expression of 11 Pgp and 3 CYP genes in three *T. circumcincta* isolates was investigated. Real-time PCR has advantages over standard PCR methods, being quicker to perform and also being considerably more sensitive, allowing the detection of smaller changes in expression, particularly in genes expressed at low levels (Pfaffl, 2001; Bustin *et al.*, 2009). The use of a reference gene (in this case actin) in the experimental procedure allows the results of the test gene to be normalised to take into account any potential variation in cDNA level, thus eliminating sample variation error (Van Zeveren *et al.*, 2007b). For the ΔΔCt method to be valid, the efficiencies of the target and control gene need to be approximately equal (Livak & Schmittgen, 2001). As shown above, this was not the case for *TeciPgp-9* NBD1 which was, therefore, excluded from further analysis. At the time of carrying out this experiment, actin was chosen as the control gene as it appeared to be more stably expressed between different *T. circumcincta* isolates compared to β-tubulin, as shown in the semi-quantitative PCR results in Chapter 3. However, since carrying out these experiments, the preferred method for performing relative quantification of gene expression has recommended that a panel of reference genes is investigated and that at least two of the most stably expressed genes are then used to normalise the test genes results (Van Zeveren *et al.*, 2007b; Strube *et al.*, 2008; Bustin *et al.*, 2009). One such method is the GeNorm approach which assumes that minimally regulated, stably transcribed genes stay in a constant ratio to each other and, therefore, provide a better baseline for normalising relative expression data to (Van Zeveren *et al.*, 2007b). Due to stage-specific variation in the control gene chosen for this experiment, no comparison of Pgp and CYP expression could be made across the respective life-cycle stages, however, this did not preclude the investigation of the relative expression between isolates. Such variability between life-cycle stages is not surprising given the extreme physiological changes observed in parasites as they develop from free-living to parasitic stages (Nisbet *et al.*, 2008).

These real-time data have shown that differences in expression of the Pgps and CYPs occur both between IVM-susceptible and -resistant *T. circumcincta* isolates (constitutive expression differences) and between pre- and post-IVM treated *T. circumcincta* isolates (inducible expression differences) as revealed through the range of *in vivo* and *in vitro* drug exposure experiments described above. Importantly, the constitutive expression differences found for *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 were (with the exception of *TeciPgp-2* NBD2 in the adult stage) statistically significant across all life-cycle stages. The expression of *TeciPgp-2* NBD2 is between 6.02 and 8.50-fold lower in the MOTRI isolate compared to the CVL isolate, whilst the expression of *TeciPgp-9* NBD2 is between 5.06 and 55.27-fold greater in the MOTRI isolate compared to the CVL isolate. Adult, L_4 and xL_3 parasites are exposed to IVM in the gut, eggs will be transiently exposed as they are excreted; whilst in the faecal pat, eggs and larval stages will remain exposed to IVM as 90% of the administered dose of IVM is excreted in faeces (Canga *et al.*, 2009). Therefore, altered expression of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 across all life-cycle stages could allow resistant parasites to survive continued IVM exposure.

Pgps are membrane transporters and IVM is known to be a potential substrate (Sangster *et al.*, 1999a; Prichard & Roulet, 2007). As such, an increase in the expression of *TeciPgp-9* NBD2 in the triple drug-resistant MOTRI isolate could represent the mechanism by which resistant parasites remove the drug, either from inside their bodies or away from the IVM target sites, such as the somatic and pharyngeal muscles (Gilleard & Beech, 2007; Prichard & Roulet, 2007; Blackhall, Prichard, & Beech, 2008). The reduced expression of *TeciPgp-2* NBD2, like the increased expression of *TeciPgp-9* NBD2 associated with the resistant phenotype of MOTRI, could also be related to the parasites preventing IVM reaching its target sites. Pgps are involved in transport of a range of endogenous and exogenous hydrophobic molecules into and out of cells, and are therefore not specifically designed to transport IVM (Kerboeuf, Guegnard, & Le Vern, 2003; Lespine *et al.*, 2008). Potentially, different Pgps could have different affinities for IVM and a parasite could reduce the expression of a particular Pgp with a high affinity to IVM, thus preventing the IVM from reaching its target site within the parasite's cells. Recent work, using fluorescent protein reporters in transgenic *C. elegans*, has shown that both *Pgp-9* and *Pgp-2* are expressed in the first and second bulbs of the pharynx and in the intestine (Zhao *et al.*, 2004) (http://wormbase.sanger.ac.uk/). Further work by Schroeder *et*

al (2007) and Nunes *et al* (2005), showed that in RNAi *C. elegans Pgp-2*(-) mutants, there was a dramatically reduced ability of the worms to store lipids; *Pgp-2* was shown to localise to the gut granules in the intestines of *C. elegans*, which are the site of fat storage. In this present study, *TeciPgp-2* NBD2 shows reduced expression in the resistant *T. circumcincta* isolate; it would be interesting to investigate whether this is linked to reduced storage of lipids in these parasites. Interestingly, experiments in *H. contortus* eggs have shown a link between the degree of anthelmintic resistance and cholesterol content; the lipid content of membranes appears to affect Pgp activity and transport of lipophilic molecules such as IVM (Riou *et al.*, 2003; Riou, Koch, & Kerboeuf, 2005; Riou *et al.*, 2007). Cholesterol depletion in *H. contortus* eggs resulted in an increase in TBZ resistance as measured by EHT (Riou *et al.*, 2003). Conversely free cholesterol concentration in *H. contortus* eggs was significantly higher in resistant isolates compared to susceptible isolates, and the presence of Pgps, as measured with monoclonal antibodies, was significantly correlated with the amount of free cholesterol (Riou *et al.*, 2007). It appears from these studies that Pgp function is dependent on the lipid environment of the membrane within which they are localised, yet the functional studies in *C. elegans* implied that, in the case of *Pgp-2*, at least, changes in expression affected the lipid environment (Riou, Koch, & Kerboeuf, 2005; Nunes *et al.*, 2005; Schroeder *et al.*, 2007). The 55.27-fold increase in expression of *TeciPgp-9* NBD2 in eggs in MOTRI compared to CVL is very interesting, especially as an increase in fluorescent staining with monoclonal antibodies for human and mouse multidrug resistance -1 (*mdr1*) was observed in anthelmintic resistant *H. contortus* eggs compared to susceptible eggs (Kerboeuf, Guegnard, & Le Vern, 2003). In a similar experiment in *T. circumcincta*, fluorescence of MOTRI eggs was three times greater than in susceptible (Moredun Ovine Susceptible Isolate, MOSI) eggs (P. Skuce. Pers. Comm.). In *C. elegans*, *mdr1* is orthologous to *Pgp-2* (http://wormbase.sanger.ac.uk/). However, none of the studies on *H. contortus*, mentioned above, investigated individual Pgp genes, and it was not possible to ascertain the Pgp-specificity of the immunofluorescence observed. It would be interesting to investigate, in the *T. circumcincta* isolates used in this thesis, the localisation of *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2, possibly through transgenic fluorescent protein reporter studies or the generation of specific monoclonal antibodies, and the effect that changes in expression of these genes could have on the lipid content of membranes, using RNAi, ultimately to elucidate further the role they play in IVM resistance.

An increase in PgpA (orthologous to *C. elegans Pgp-2*) mRNA level has been demonstrated using Northern blotting in IVM-selected *H. contortus* and, in adult *O. ostertagi*, a 3.4-fold increase in expression of a *Pgp2* orthologue was found using real-time PCR from a laboratory selected IVM resistant isolate (Xu *et al.*, 1998; Van Zeveren, 2009). More importantly, *TeciPgp-9* has been identified as playing a major role in the IVM resistant phenotype in a *T. circumcincta* near-isogenic anthelmintic-resistant line developed in New Zealand (Bisset, 2007). The near-isogenic line was developed by backcrossing an inbred multi-drug resistant isolate with an inbred anthelmintic susceptible isolate, with IVM selection at each generation. This research was conducted independently of the present study, and the results only became available at the time of writing. As in the present study, Bisset (2007) found increased polymorphisms in *TeciPgp-9* in the IVM resistant line. For example, individual resistant worms possessed three or four *TeciPgp-9* haplotypes compared to no more than two haplotypes in susceptible worms. Moreover, using real-time PCR, resistant worms exhibited a higher expression level of *TeciPgp-9* than susceptible worms, indicative of an increased gene copy number. Alongside this, amino acid substitutions were found when comparing *TeciPgp-9* coding sequences from resistant and susceptible worms. Unfortunately, we are not in a position to determine whether the same coding SNPs are implicated in our IVM-resistant MOTRI isolate at this stage because the regions of the gene where these amino acid substitutions are found in the New Zealand *T. circumcincta* lines are not covered by the sequence generated in this thesis. These amino acid substitutions may also be present in UK IVM-resistant *T. circumcincta* isolates. The generation of a full-length coding sequence for *TeciPgp-9* would be an obvious starting point for further work to investigate this. If found, they could form the basis of a molecular marker for IVM resistance.

Finding significant fold changes in expression in all life-cycle stages of *TeciPgp-9* NBD2 when comparing two, non-related UK isolates; one multi-drug resistant and the other drug-susceptible, becomes far more important in the light of the work carried out in New Zealand. Changes in the expression of *TeciPgp-9* could be caused by an increase in copy number, as found in the New Zealand study, or as a result of a mutation or change in the gene promoter or upstream regulatory element controlling the transcription of that gene. The cause of the change in expression of *TeciPgp-9* remains to be determined. Regulation of expression is dependent on regions of DNA either upstream of the gene in

question or in intronic regions; areas which would need to be investigated to determine how the fold change in expression of *TeciPgp-9* NBD2 between CVL and MOTRI occurs; very little is currently known about these regulatory sequences in nematodes (Prichard & Roulet, 2007). Even though the change in expression of *TeciPgp-9* NBD2 associated with IVM resistance has been found in both the UK and New Zealand, the next step in developing these changes into a marker for IVM resistance is to confirm the presence of the expression differences, or the amino acid substitutions, in other UK field isolates of *T. circumcincta* resistant to IVM and potentially in other parasites species too.

Another important step in fully investigating the role *TeciPgp-9* NBD2 plays in IVM resistance in the UK would be determining whether *TeciPgp-9* NBD1 (formerly PGP4) exhibits the same expression profile as *TeciPgp-9* NBD2, this was not possible in this investigation as the primers and probes for *TeciPgp-9* NBD1 were not sufficiently efficient and so the gene was excluded from the real-time PCR experiments described in this Chapter. More importantly, the generation of a full-length sequence for *TeciPgp-9* would confirm that they are genuinely two halves of the same gene. Ten NBD1 haplotypes and four NBD2 haplotypes for *TeciPgp-9* were identified in the New Zealand near-isogenic *T. circumcincta* lines, some of which were found exclusively or predominantly in the resistant lines. Some NBD1 haplotypes appeared to correspond to particular NBD2 haplotypes but have yet to be shown to be parts of the same haplotype (Bisset, 2007). This could be the same with *TeciPgp-9* NBD1 and *TeciPgp-9* NBD2 if they are found to have a different expression profile; the two parts, although potentially from the same gene may, in fact, have developed through a gene duplication event and therefore not have been under the same selection pressure. According to the alignments of *T. circumcincta* sequences against the *C. elegans* Pgps in Chapter 3, *TeciPgp-2* NBD1 (formerly PGP5) should represent the 5' end of the same Pgp molecule as *TeciPgp-2* NBD2, yet the gene expression profile of these two genes following triplicate real-time PCR runs was not identical. Indeed, for the xL_3, L_4 and adult stages there were non-statistically significant increased expression of *TeciPgp-2* NBD1 whilst for *TeciPgp-2* NBD2 in these stages there was decreased expression in the MOTRI isolate compared to the CVL isolate. As described for *TeciPgp-9* from New Zealand, there could be different haplotypes of the two NBDs, and so, although both *TeciPgp-2* NBD1 and *TeciPgp-2* NBD2 aligned most closely to the *C. elegans Pgp-2*; they could be fragments of two copies of the *TeciPgp-2* gene.

Changes in expression of other genes including the CYPs were also observed, these changes were not evident across all life-cycle stages and generally represented smaller fold changes than those found for *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2. As such, their importance in the role of expressing an IVM-resistant phenotype may not be as great. Interestingly, compared to the general pattern of expression of the panel of Pgp genes, the CYPs seemed to follow a similar pattern of expression to each other, as shown in Tables 4.8 and 4.9. For example, the expression of the three CYPs when comparing CVL and MOTRI L_1 all showed a statistically significant reduction in expression.

The first comparison described in this Chapter, investigating constitutive expression differences between the IVM-susceptible CVL and IVM-resistant MOTRI *T. circumcincta* isolates, was the same comparison as that carried out in Chapter 3 using semi-quantitative PCR. Comparing the semi-quantitative and real-time PCR data shows that the real-time PCR confirms the majority of the changes in gene expression identified whilst also identifying changes not visible in the semi-quantitative PCR, particularly where the expression of the genes was too low to be identified using semi-quantitative PCR. *TeciPgp-2* NBD2 (originally called PGP2 in the semi-quantitative experiment) was not amplified using semi-quantitative PCR but in the real-time PCR experiment, reduced expression was found in all life-cycle stages of the MOTRI isolate compared to the CVL isolate. PGP3 and PGP7 were the most stably expressed between CVL and MOTRI in the semi-quantitative PCR (Figure 3.3) and showed no statistically significant differences in expression in the real-time experiment (Table 4.4).

It was hypothesised that a stepwise change in expression going from the IVM-susceptible (CVL) isolate, through the IVM-resistant (MOTRI) isolate to the IVM survivors (Post IVM MOTRI) isolate might be observed. With the exception of *TeciPgp-9* NBD2 in the xL_3 stage, where the change in expression between CVL and MOTRI was a 17.49-fold increase and the change in expression between MOTRI and Post IVM MOTRI was a 1.54-fold increase, this was found not to be the case. This suggests that any increases in constitutive expression in resistant isolates were not enhanced when a resistant isolate was further selected with IVM treatment to remove any parasites within that isolate exhibiting a drug-susceptible phenotype. In fact, with the three CYP genes, it was

observed that when statistically significant increases in fold change in expression were observed between the CVL and MOTRI isolates, a corresponding statistically significant reduction in fold change in expression was observed between the MOTRI and Post IVM MOTRI (Tables 4.8 and 4.9).

To fully investigate the changes in expression found in isolates of *T. circumcincta* resistant to IVM, various experiments utilising both *in vivo* and *in vitro* approaches to IVM exposure were carried out. In all of these comparisons, parasites from the MOTRI isolate were compared to parasites of the same isolate which had undergone some form of selection with, or exposure to, IVM. As described above, the comparison between the Post IVM MOTRI (*in vivo* IVM treatment survivors) and MOTRI isolates did not show that parasites surviving IVM treatment exhibited any further change in expression of any given gene following on from the CVL-MOTRI constitutive gene expression comparison. Therefore, it was decided to determine whether any transient inducible changes in the expression of *TeciPgp-9* NBD2 and *TeciPgp-2* NBD2 could be identified. Utilising a modified LMIT method, a four way comparison between (i) MOTRI xL_3 able to migrate in the presence of IVM, (ii) those unable to migrate in the presence of IVM, (iii) those that migrated through a Baermann chamber whilst not exposed to IVM and (iv) the biological replicate xL_3 MOTRI isolate, was made. The only statistically significant differences in expression were a 1.88-fold increase in *TeciPgp-2* NBD2 expression in the unexposed sample compared to the migrators sample and a 1.69-fold increase in *TeciPgp-9* NBD2 expression in the unexposed sample compared to the MOTRI xL_3 biological replicate sample. Potentially, the LMIT method and the IVM dose chosen for this experiment (1µg/mL IVM) may not have been sufficient to induce any changes in expression of the two genes investigated. Alternatively, the non-migrator xL_3 could have also been killed by the exposure to IVM and, as such, would not have exhibited altered *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 expression profiles. However, the actin Ct values between the three LMIA samples, and when compared to the biological replicate MOTRI xL_3 sample, were similar (data not shown); dead or dying parasites would have been expected to have altered levels of mRNA.

As a final experiment, donor sheep were infected with MOTRI parasites which were allowed to develop to the adult stage prior to treatment with a full therapeutic dose of IVM. Three days after drug administration, adult parasites were collected from the abomasa, as described previously. This time course was chosen to give any adults affected by the IVM treatment a chance to be expelled by the host whilst those able to resist IVM treatment would still be exposed to IVM. Gill & Lacey (1998) found that 25% of *T. circumcincta* were expelled 8-10 hours after exposure with the rest remaining in the abomasum for at least 14 hours post treatment whilst the half-life of IVM is between 61 and 178 hours (Prichard *et al.*, 1985a; Marriner, Mckinnon, & Bogan, 1987; Canga *et al.*, 2009). The statistically significant relative fold change in expression of *TeciPgp-2* NBD2 was 13.69-fold lower in the IVM survivors compared to MOTRI whilst the 1.82-fold increase in expression of *TeciPgp-9* NBD2 in the IVM survivors compared to MOTRI was not statistically significant. This experiment gives an indication that inducible gene expression changes may occur in response to IVM treatment; however, the biggest challenge when carrying out these experiments is determining whether any changes observed are due to the IVM exposure or a generalised response to the experimental procedure. Roulet and Prichard (2006) showed that over-expression of five *H. contortus* Pgps occurred 24 hours after *in vivo* IVM exposure. As such, the experimental protocol chosen to investigate the inducible expression of Pgps in *T. circumcincta* may need to be altered to identify changes in expression which occur rapidly following IVM exposure. For example, the incubation and processing periods when carrying out the LMIT may have allowed altered gene expression patterns to return to baseline levels.

This work has shown that, in a panel of candidate resistant genes comprising of Pgps and CYPs, significant differences in gene expression were evident when comparing *T. circumcincta* isolates resistant and susceptible to IVM, as well as following IVM exposure. Significantly, the constitutive expression differences in *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 were found across all life-cycle stages and the differences found in expression levels of *TeciPgp-9* NBD2 were reinforced by similar findings in the same gene in laboratory-derived *T. circumcincta* isogenic lines in New Zealand. These differences in expression, alongside SNPs identified in *TeciPgp-9* NBD2 (Chapter 3) need further investigation to determine whether they occur in other isolates of *T. circumcincta*, in other parasitic species exhibiting IVM resistance, to determine what is the cause of the

differences in expression observed and what is their functional significance. If found in other isolates exhibiting IVM resistance, these differences in expression could form the basis of a molecular test for IVM resistance in parasitic nematodes.

Table 4.1: Table showing primers and probes designed for real-time PCR. All probes were fluorescently labelled with FAM.

Primer name	Primer direction	Sequence	Melting temp (^{o}C)	Target
ActinQTcFor	Sense	ACG ACG AGG TTG CTG CTC TT	59	*T. circumcincta* actin
ActinQTcProbe	NA	TGG TTG ACA ATG GAT CC	69	
ActinQTcRev	Antisense	GAA TCC GGC TTT GCA CAT TC	59	
P1QTcFor	Sense	GTG TGG AAA GAG CAC CGT GAT	59	*T. circumcincta* PGP1
P1QTcProbe	NA	CAG CTC GTC GAG AGG T	69	
P1QTcRev	Antisense	GCC ACA CAA GGC GTC GTA A	59	
P2QTcFor	Sense	TGC AAA TGC AGT TCG GTA AGA	58	*T. circumcincta* PGP2
P2QTcProbe	NA	AAT GCG GGA CAC AGA A	68	
P2QTcRev	Antisense	TGC CAG CCT CCT CCA GAA	59	
P3QTcFor	Sense	CGA GTT GCA ATT GAG AGA TTT CA	58	*T. circumcincta* PGP3
P3QTcProbe	NA	TGG CTG AAT GGC G	69	
P3QTcRev	Antisense	AGC CAG GGA TGC ATTGGA	59	
P4QTcFor	Sense	GAA CGT TGG TGA CCG AGG AA	60	*T. circumcincta* PGP4
P4QTcProbe	NA	TGT CTG GTG GCC AAA A	70	
P4QTcRev	Antisense	GCG GGC AAT GGC TAT ACG	59	
P5QTcFor	Sense	TCA TCA GCG ACC CAA CAT TG	59	*T. circumcincta* PGP5
P5QTcProbe	NA	ACC GTG GCC GGA TT	69	
P5QTcRev	Antisense	CAC CGC GAA GAA CAC AGT GA	59	
P6QTcFor	Sense	GAC GGT TCT GTG GAA GTT GAT G	58	*T. circumcincta* PGP6
P6QTcProbe	NA	CTC ATC TAC GTG CCC ATA T	70	
P6QTcRev	Antisense	TCG GCT CTT GCG ATA CCA A	59	
P7QTcFor	Sense	TCC GGC TCT CTG CCA AGT	58	*T. circumcincta* PGP7
P7QTcProbe	NA	ATG TTT GTT GTG ATT CCG	69	
P7QTcRev	Antisense	CGG CGA AAG TTC CAA CCA TA	60	
P8QTcFor	Sense	AAG TGG ATG CTC GTG AAC TGA A	58	*T. circumcincta* PGP8
P8QTcProbe	NA	TTG AGG CAC CTT CG	68	

Primer name	Primer direction	Sequence	Melting temp (^{o}C)	Target
P8QTcRev	Antisense	CCC GCT AGG GAG ATT TGT GA	59	
P9QTcFor	Sense	TCA GGA GCC GGC GTT GT	60	*T. circumcincta* PGP9
P9QTcProbe	NA	TGC TGA TAC AGT TGA AAA T	70	
P9QTcRev	Antisense	CCC GTC CCA AGC GGA TAT	59	
P10QTcFor	Sense	CCG CTA ACG GAA GGT GAG ATT	59	*T. circumcincta* PGP10
P10QTcProbe	NA	CCT TGA TGG TCG TCC AT	69	
P10QTcRev	Antisense	GCG CTG TGA CTT AGC GAA CTT	58	
P11QTcFor	Sense	CCG CTG GAG CTG GAC AAC	59	*T. circumcincta* PGP11
P11QTcProbe	NA	CAC CGA AAG ACG AGG TC	70	
P11QTcRev	Antisense	TTC CGT CAC GCG ACG TT	58	
C1QTcFor	Sense	AAC TGC TCA CGC TCA CAA AGG	59	*T. circumcincta* P450 1
C1QTcProbe	NA	ACA ACG CCT ACT CTC	69	
C1QTcRev	Antisense	AGT GTT CGG TCT GTC CAC GAT	58	
C2QTcFor	Sense	CGT TGT CGG TGG ACA TTC TG	59	*T. circumcincta* P450 2
C2QTcProbe	NA	CCC AAA AGG CAC TCC A	69	
C2QTcRev	Antisense	TTA TTA CCG ACA ATT CAG CTG CAA	60	
C3QTcFor	Sense	TTT GGA AAG AAC AAC GAC GTG TA	58	*T. circumcincta* P450 3
C3QTcProbe	NA	CAC TCC AAA TAC TTC G	68	

Table 4.2: Table showing the plasmids used to validate the efficiency of the gene specific real-time primer and probe sets.

Gene	Plasmid name	Plasmid size (bp)	Concentration (ng/µL)
Actin	Actin race a	3436	401.19
PGP1	P1 race 40 1b	3640	308.07
PGP2	29 Nov 11c	3799	320.18
PGP3	P3_2a	3752	285.31
PGP4	(17 Oct) P5b	3368	349.22
PGP5	25 Mar P5 3Ar	3772	295.69
PGP7	25 Mar P7 1a	4013	311.81
PGP6	2S2A1	3402	263.59
PGP8	1S2A1	3398	328.03
PGP9	1S1A1	3398	224.39
PGP10	Pgp1sm1	3395	333.7
PGP11	Pgp2sm1	3395	616.26
CYP1	Jan16C3a	3606	332.87
CYP2	C2L3a	3323	219.99
CYP3	C3L3a	3468	276.81

Table 4.3: Table showing the efficiencies of the reactions for each primer and probe set, as calculated from the standard curves. The percentage efficiency was calculated from the slope of the standard curve using the formula described in 2.4.3 whilst the R2 value gives an indication of how close the data points fit to the standard curve and the intercept is the point at which the standard curve crosses the y axis (number of real-time cycles).

Gene	Slope	Intercept	R2 value	Efficiency of reaction (%)
Actin	-3.4381	43.1010	0.9939	95.37
Actin	-3.4206	44.1108	0.9977	96.04
Actin	-3.1195	40.6668	0.9981	109.2
Actin	-3.1008	39.9113	0.9989	110.13
PGP1	-3.2210	41.6619	0.9930	104.39
TeciPgp-2 NBD2	-3.2791	41.7425	0.9991	101.81
PGP3	-2.9697	39.8586	0.9964	117.13
TeciPgp-9 NBD1	-3.6202	51.7543	0.9112	89.5
TeciPgp-9 NBD1	-4.4088	58.1673	0.9931	68.58
TeciPgp-2 NBD1	-3.2707	41.2686	0.9859	102.18
TeciPgp-9 NBD2	-3.3026	39.6582	0.9975	100.81
PGP7	-3.3170	42.2188	0.9993	100.21
PGP8	-3.2373	38.3497	0.9979	103.65
PGP9	-3.3567	39.4773	0.9964	98.568
PGP10	-3.5396	43.4254	0.9769	91.65
PGP11	-3.3911	42.1324	0.9982	97.19
CYP1	-3.3458	39.6981	0.9977	99.01
CYP2	-3.3274	42.8249	0.9973	99.77
CYP3	-3.2538	39.5755	0.9963	102.92

Table 4.4: Relative constitutive expression of the Pgp genes in *T. circumcincta* MOTRI compared to CVL. The fold change in expression in CVL equals 1 and the expression in MOTRI is expressed relative to this. The statistical significance (at either P<0.05 or P<0.01) of each result is shown in brackets below the fold change in expression. n.s. means the result is not statistically significant.

	PGP1	***TeciPgp-2*** **NBD2**	**PGP3**	***TeciPgp-2*** **NBD1**	***TeciPgp-9*** **NBD2**	**PGP7**	**PGP8**	**PGP9**	**PGP10**	**PGP11**
Eggs	3.90 (P<0.05)	0.17 (P<0.05)	2.00 (P<0.05)	0.96 (n.s.)	55.27 (P<0.01)	1.07 (n.s.)	3.50 (n.s.)	1.58 (n.s.)	1.07 (n.s.)	0.39 (n.s.)
L_1	0.73 (n.s.)	0.12 (P<0.01)	0.70 (n.s.)	0.73 (n.s.)	5.06 (P<0.05)	0.64 (n.s.)	1.45 (n.s.)	1.36 (P<0.05)	1.00 (n.s.)	0.87 (n.s.)
xL_3	0.82 (n.s.)	0.13 (P<0.05)	1.17 (n.s.)	1.07 (n.s.)	17.49 (P<0.01)	1.11 (n.s.)	0.99 (n.s.)	0.65 (n.s.)	1.88 (P<0.05)	0.82 (n.s.)
L_4	0.68 (P<0.01)	0.14 (P<0.01)	0.61 (n.s.)	1.37 (n.s.)	14.04 (P<0.05)	0.90 (n.s.)	3.39 (n.s.)	2.00 (P<0.01)	0.25 (P<0.01)	0.37 (P<0.05)
Adults	1.35 (n.s.)	0.37 (n.s.)	1.07 (n.s.)	1.78 (n.s.)	6.75 (P<0.01)	1.11 (n.s.)	1.12 (n.s.)	0.86 (n.s.)	1.51 (n.s.)	1.60 (P<0.05)

Table 4.5: Table showing the results of sequencing to confirm the specificity of the *T. circumcincta TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 real-time primers. The primers and probes used to determine the expression of *TeciPgp-2* NBD2 were P2QTcFor, P2QTcProbe and P2QTcRev and for *TeciPgp-9* NBD2 were P6QTcFor, P6QTcProbe and P6QTcRev as shown in Table 4.1. The percentage identity of each clone to the consensus sequence is shown in brackets.

Clone Name	**Sequence Result**
P2A1a	Aligned to *TeciPgp-2* NBD2 (95% identity)
P2A1b	No insert
P2C9a	Aligned to *TeciPgp-2* NBD2 (95% identity)
P2C9b	Aligned to *TeciPgp-2* NBD2 (96% identity)
P6A7a	Aligned to *TeciPgp-9* NBD2 (94% identity)
P6A7b	Aligned to *TeciPgp-9* NBD2 (100% identity)
P6B3a	Aligned to *TeciPgp-9* NBD2 (100% identity)

Table 4.6: Relative constitutive expression of the *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 in biological replicates of *T. circumcincta* MOTRI compared to CVL. The fold change in expression in CVL equals 1 and the expression in MOTRI is expressed relative to this. The statistical significance (at either P<0.05 or P<0.01) of each result is shown in brackets below the fold change in expression. n.s. means the result was not statistically significant.

	***TeciPgp-2* NBD2**	***TeciPgp-9* NBD2**
xL_3	0.07 (P<0.05)	14.93 (P<0.01)
Adults	0.08 (P<0.01)	1.00 (n.s)

Table 4.7: Relative inducible expression of the Pgp genes in *T. circumcincta* Post IVM MOTRI compared to MOTRI. The fold change in expression in MOTRI equals 1 and the expression in Post IVM MOTRI is expressed relative to this. The statistical significance (at either P<0.05 or P<0.01) of each result is shown in brackets below the fold change in expression. n.s. means the result is not statistically significant.

	PGP1	***TeciPgp-2* NBD2**	**PGP3**	***TeciPgp-2* NBD1**	***TeciPgp-9* NBD2**	**PGP7**	**PGP8**	**PGP9**	**PGP10**	**PGP11**
Eggs	0.53 (n.s.)	1.67 (n.s.)	1.34 (n.s.)	1.53 (n.s.)	1.10 (n.s.)	1.38 (n.s.)	0.76 (n.s.)	1.93 (n.s.)	1.17 (n.s.)	1.42 (n.s.)
L_1	1.24 (n.s.)	3.42 (P<0.05)	1.58 (P<0.05)	3.16 (P<0.01)	1.71 (n.s.)	1.33 (n.s.)	1.17 (n.s.)	0.81 (n.s.)	1.60 (n.s.)	1.79 (n.s.)
xL_3	1.67 (P<0.01)	1.25 (n.s.)	0.94 (n.s.)	0.93 (n.s.)	1.54 (P<0.01)	0.99 (n.s.)	3.84 (P<0.01)	0.73 (n.s.)	0.84 (n.s.)	0.80 (n.s.)
Adults	0.63 (n.s.)	0.76 (n.s.)	0.49 (P<0.01)	0.74 (n.s.)	0.69 (n.s.)	0.56 (P<0.05)	0.47 (P<0.05)	0.94 (n.s.)	0.65 (n.s.)	0.44 (P<0.05)

Table 4.8: Relative constitutive expression of the CYP genes in *T. circumcincta* MOTRI compared to CVL. The fold change in expression in CVL equals 1 and the expression in MOTRI is expressed relative to this. The statistical significance (at either P<0.05 or P<0.01) of each result is shown in brackets below the fold change in expression. n.s. means the result is not statistically significant.

	CYP1	**CYP2**	**CYP3**
Eggs	11.73 (P<0.01)	5.25 (P<0.01)	0.95 n.s.
L_1	0.64 (P<0.05)	0.51 (P<0.05)	0.43 (P<0.05)
xL_3	1.09 n.s.	1.19 n.s.	1.09 n.s.
L_4	1.16 n.s.	1.89 n.s.	1.10 n.s.
Adults	2.82 (P<0.01)	2.60 (P<0.05)	0.78 n.s.

Table 4.9: Relative inducible expression of the CYP genes in *T. circumcincta* Post IVM MOTRI compared to MOTRI. The fold change in expression in MOTRI equals 1 and the expression in Post IVM MOTRI is expressed relative to this. The statistical significance (at either P<0.05 or P<0.01) of each result is shown in brackets below the fold change in expression. n.s. means the result is not statistically significant.

	CYP1	**CYP2**	**CYP3**
Eggs	0.53 (P<0.05)	0.54 (P<0.05)	0.92 n.s.
L_1	2.91 (P<0.05)	3.23 (P<0.01)	1.92 n.s
xL_3	1.56 n.s.	1.33 n.s.	1.38 (P<0.05)
Adults	0.81 (P<0.05)	0.32 (P<0.01)	0.48 (P<0.05)

A

B

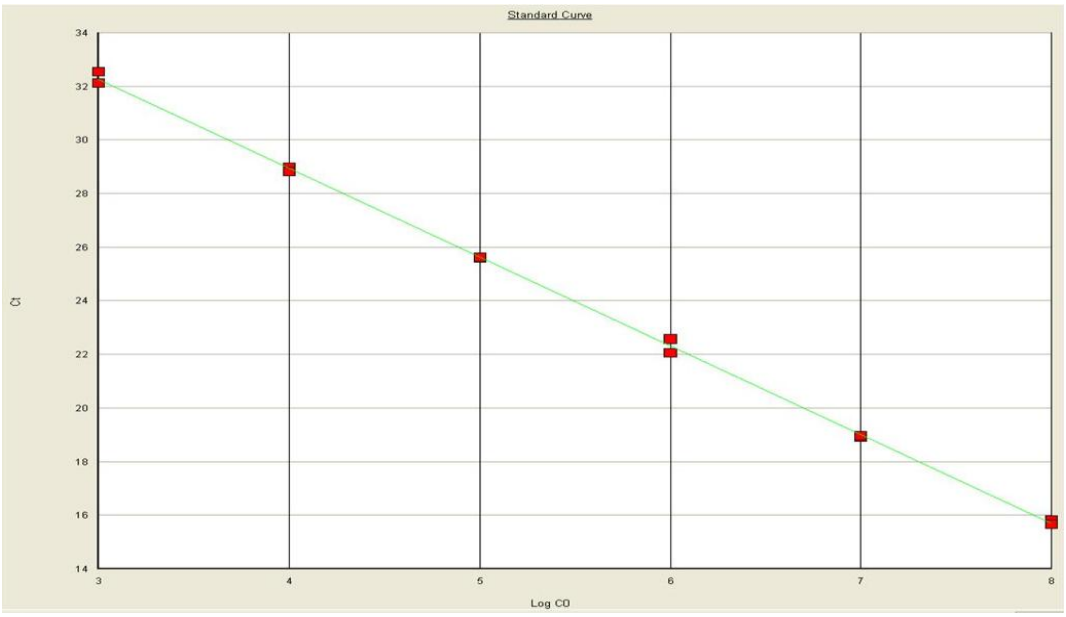


Figure 4.1: Graphs showing typical results from the validation of the real-time primer and probe sets. **A**: Representative graph of the results for PGP7 showing the duplicate results for each of the dilutions in the calibration curve. The point where the amplification curves cross the green line is the Ct value. **B**: The standard curve for PGP7 generated using the Ct values from the graph in part A. The standard curve is used to calculate the efficiency of the reaction. Graphs generated using the ABI 7000 sequence detection system.

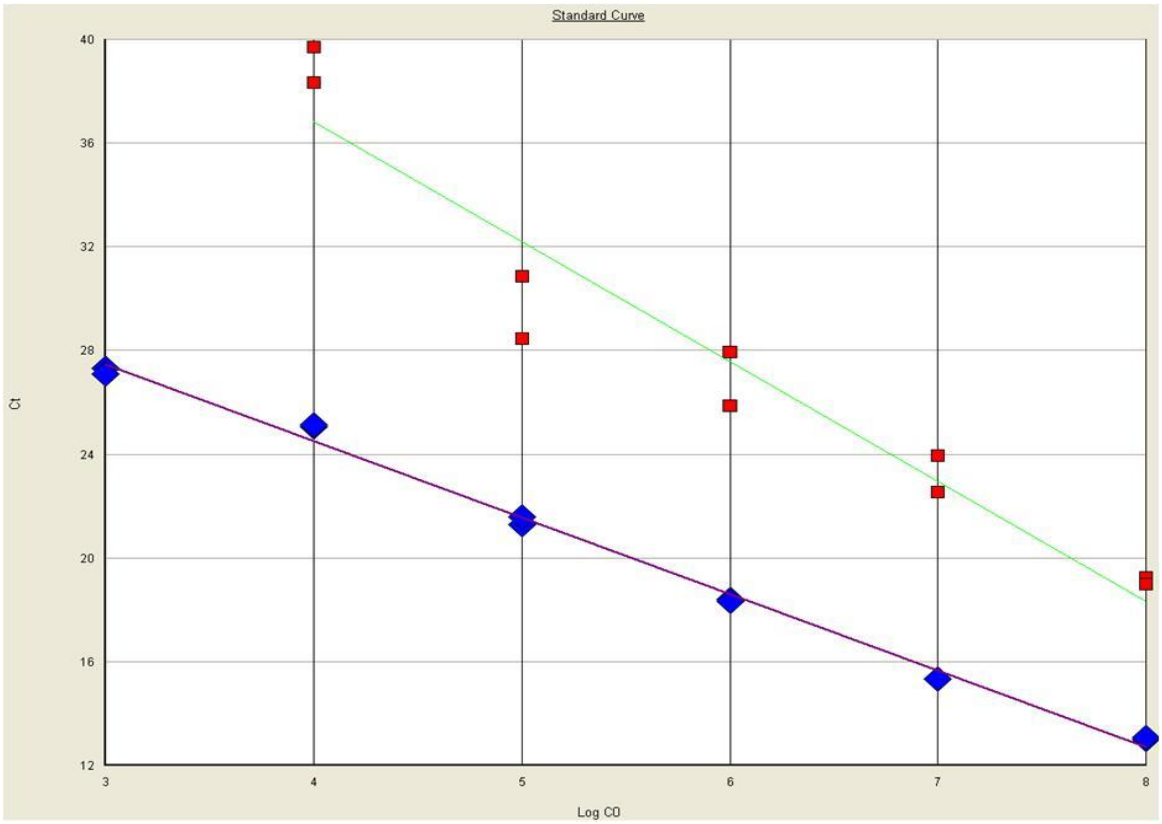


Figure 4.2: Standard curves for *TeciPgp-9* NBD1 and actin. These show that the *TeciPgp-9* NBD1 primers and probes did not efficiently amplify the plasmid containing the *TeciPgp-9* NBD1 gene, meaning accurate comparisons between the actin control gene and the *TeciPgp-9* NBD1could not be made. Graphs generated using the ABI 7000 sequence detection system.

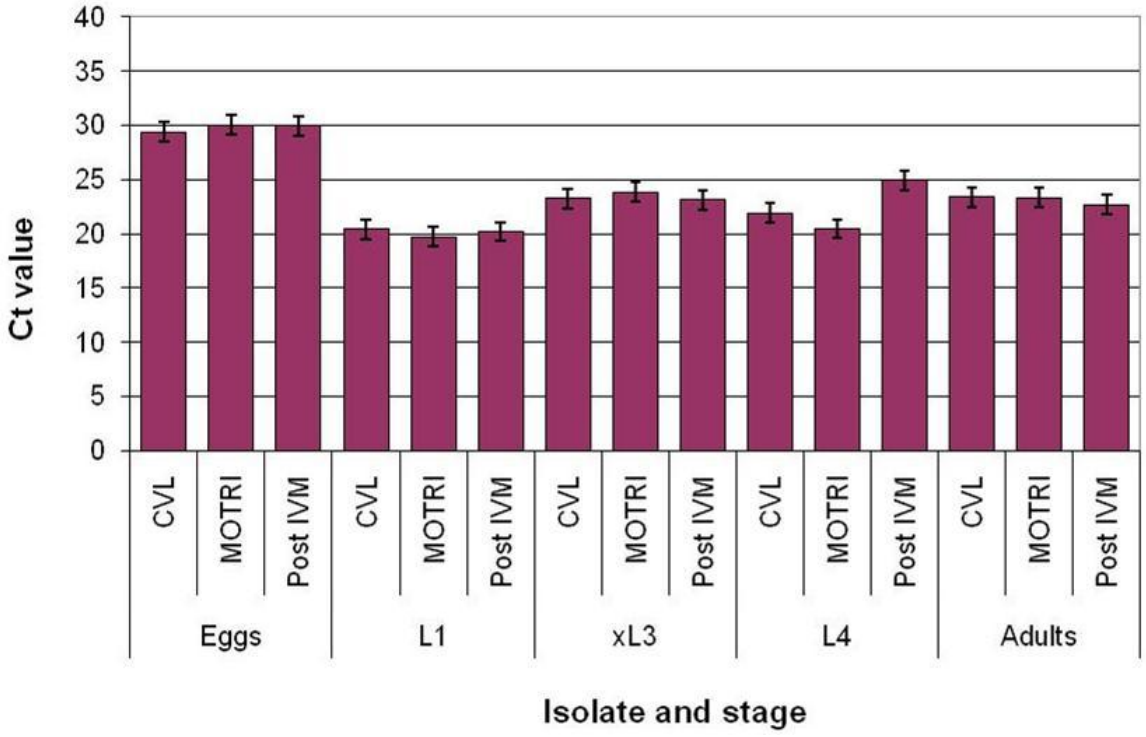


Figure 4.3: Graph showing the average actin Ct values for each isolate and life-cycle combination. The results are a combination of triplicate real-time run carried out to quantify the expression of the ten Pgp and three CYP genes in *T. circumcincta*. The error bars on the graph show the standard deviation of the actin Ct values

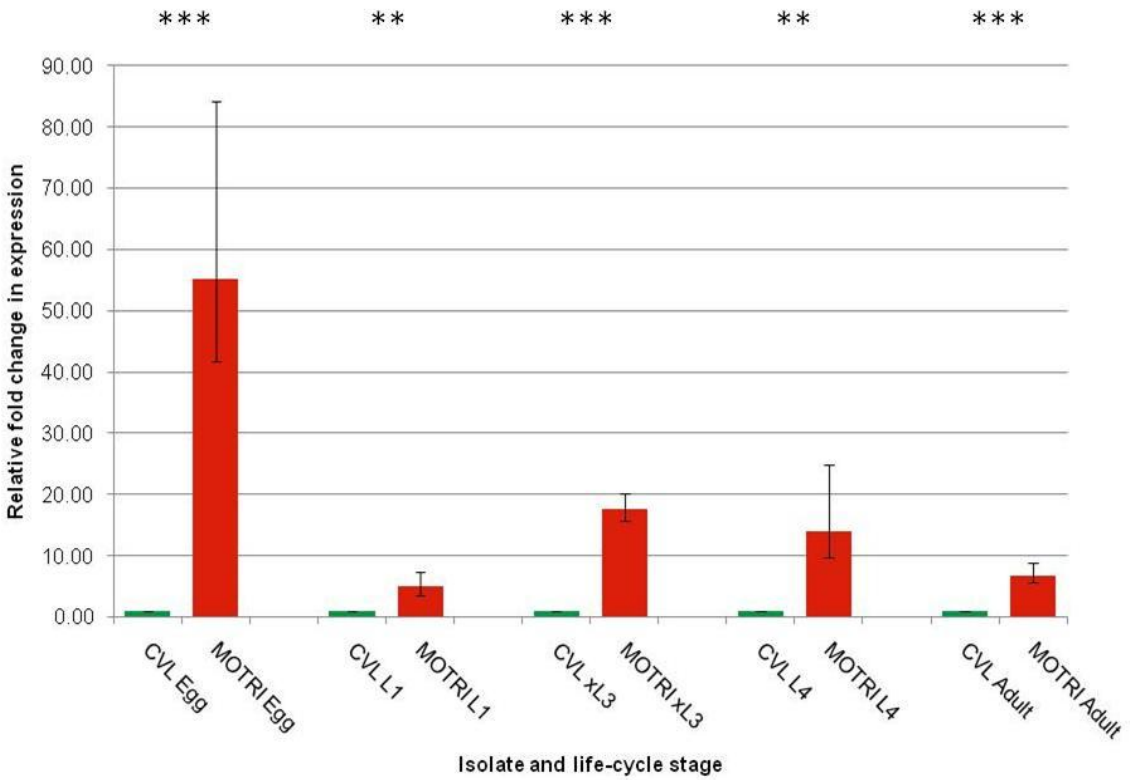
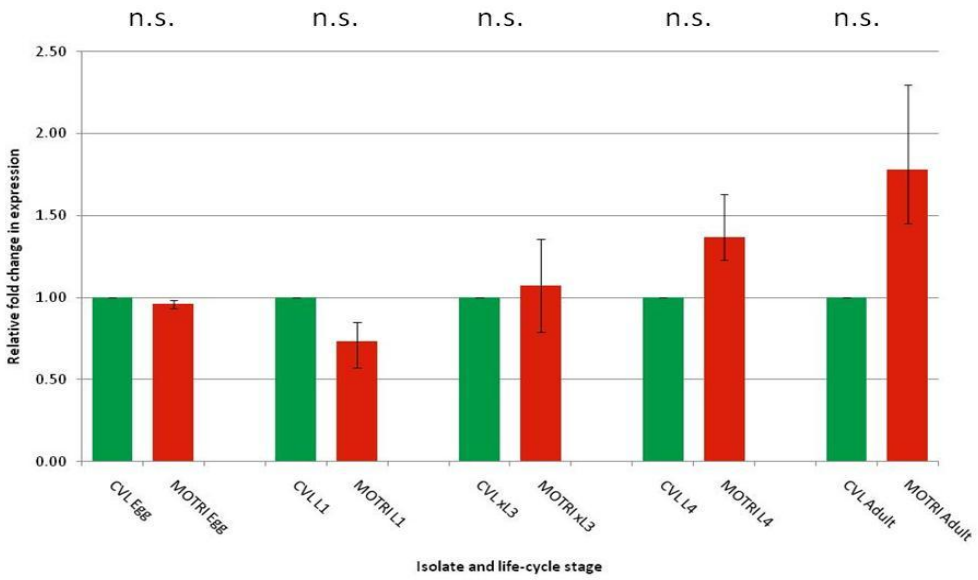
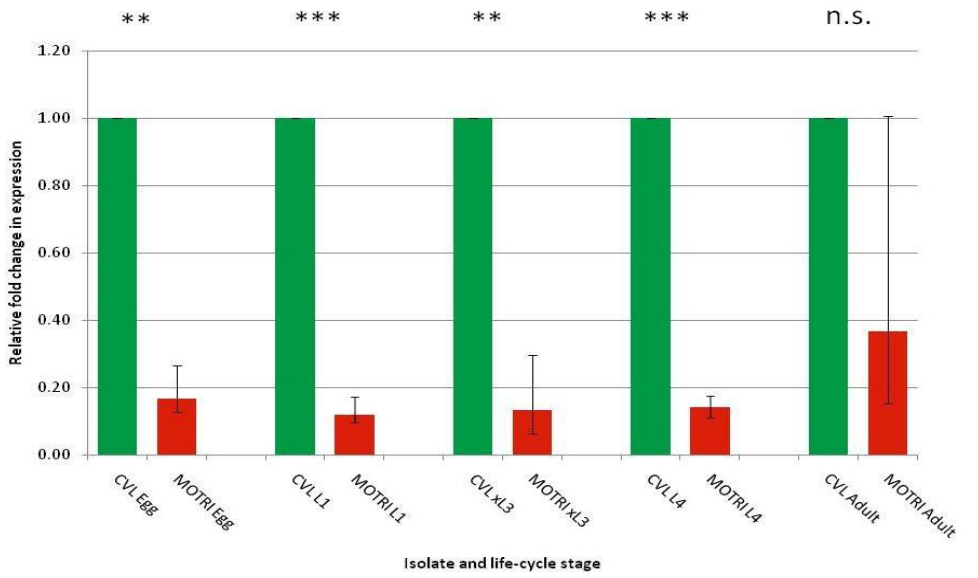


Figure 4.4: Graph showing the relative constitutive expression of *TeciPgp-9* NBD2 in *T. circumcincta.* The stars above the graph indicate the statistical significance of the relative fold change in expression of *TeciPgp-9* NBD2 in MOTRI compared to CVL. $P<0.01$ is represented by *** and $P<0.05$ is represented by **. The lines for each bar on the graph represent the range in relative expression.



A



B

Figure 4.5: Graphs showing the relative constitutive expression of *TeciPgp-2* NBD 1 and 2 in *T. circumcincta*. **A**: Expression of *TeciPgp-2* NBD1. **B**: Expression of *TeciPgp-2* NBD2. The stars above the graph indicate the statistical significance of the relative fold change in expression of *TeciPgp-2* in MOTRI compared to CVL. $P<0.01$ is represented by *** and $P<0.05$ is represented by **, statistically non-significant results are indicated with n.s. The lines for each bar on the graph represent the range in relative expression.

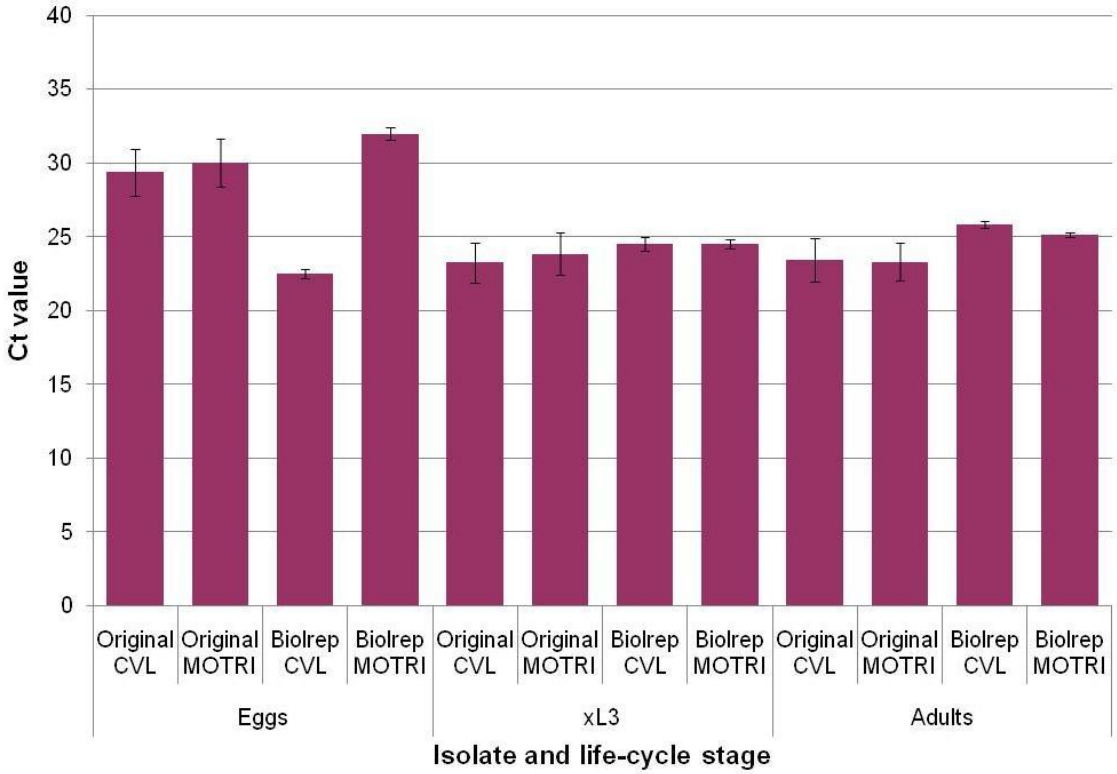


Figure 4.6: Graph showing the average actin Ct values for the biological replicates of the constitutive gene expression comparison between MOTRI and CVL *T. circumcincta* compared to the corresponding (original) isolate and life-cycle average actin Ct values. The error bars on the graph show the standard deviation of the actin Ct values.

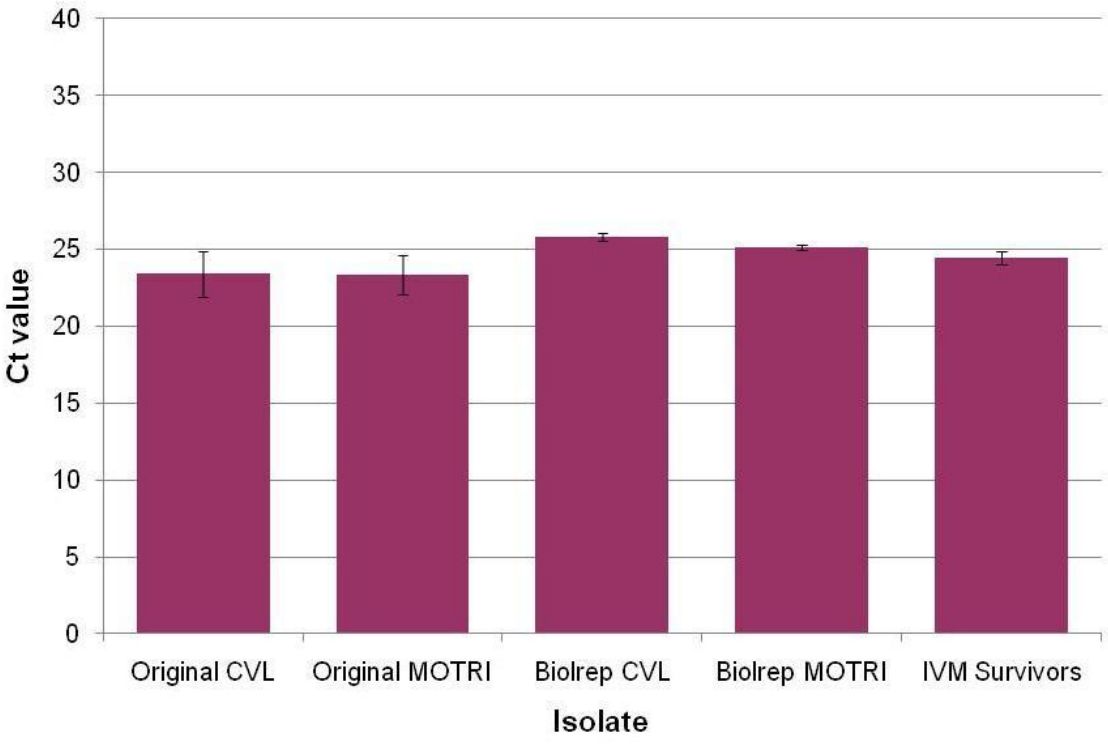
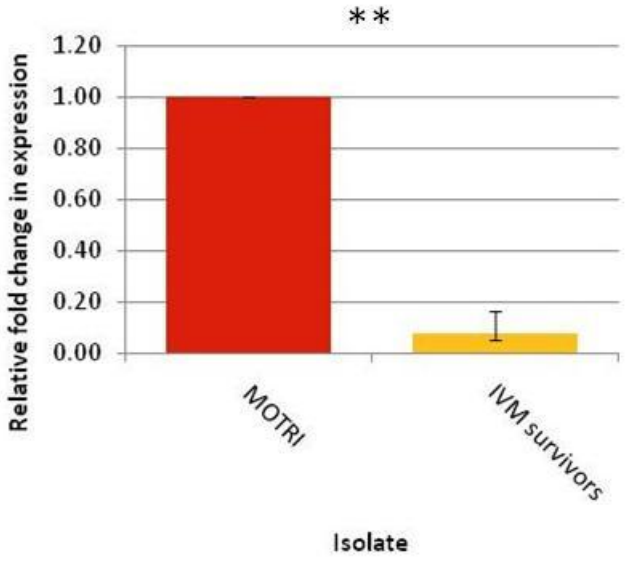
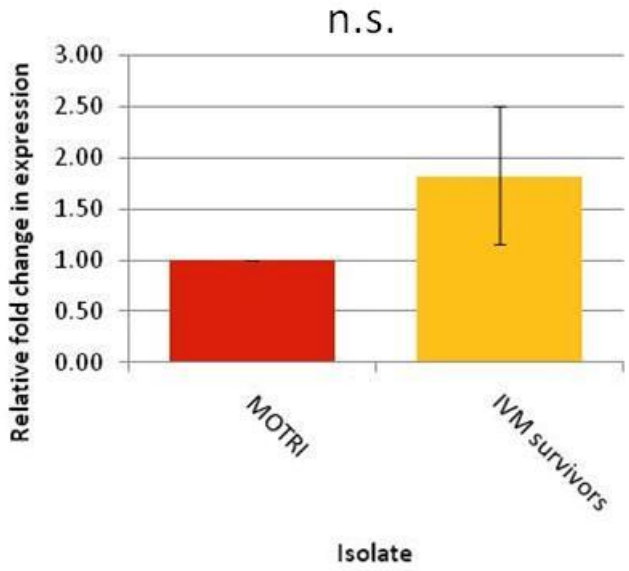


Figure 4.7: Graph showing the average actin Ct values for the adult *T. circumcincta* IVM survivors compared to the Ct values for the adult CVL and MOTRI *T. circumcincta* biological replicates and original adult CVL and MOTRI *T. circumcincta*. The error bars on the graph show the standard deviation of the actin Ct values.



A



B

Figure 4.8: Graphs showing the relative fold change in expression between MOTRI adults and IVM survivor MOTRI adults. **A**: Relative expression of *TeciPgp-2* NBD2. **B**: Relative expression of *TeciPgp-9* NBD2. P<0.05 is represented by **, statistically non-significant results are indicated with n.s. The lines for each bar on the graph represent the range in relative expression.

Chapter 5: Roche 454 Sequencing Analysis Comparing *in vitro* Ivermectin-Exposed and Ivermectin-Unexposed *Teladorsagia circumcincta* Adults*

5.1: Introduction

Anthelmintics are the principal method of control for PGE, yet their mode of action and the underlying genetic determinants of resistance exhibited by the parasites are still not adequately understood. Most of the studies in this area to date have used a "candidate gene approach" to look for key genetic changes, such as target site mutations, in related parasitic species or in the model nematode, *C. elegans* (Wolstenholme *et al.*, 2004). However, this has largely not been a fruitful exercise (Gilleard, 2006; von Samson-Himmelstjerna *et al.*, 2007). Classical genetic linkage studies to identify 'resistance genes' are challenging, particularly in nematode parasite populations made up of individuals with differing levels of resistance, where there are no parental anthelmintic-susceptible isolates to compare to the resistant populations, and especially when the mode of action of the anthelmintic is still not clearly understood (Gilleard, 2006). This makes it difficult to propose candidate resistance genes for further study. An overview of the current understanding of the modes of action of the anthelmintics and the mechanisms of resistance are given below, more detail is provided in Chapter 1.

The BZs bind to β-tubulin within the parasite, disrupting various essential cellular processes and resulting in cell lysis (McKellar & Jackson, 2004; von Samson-Himmelstjerna *et al.*, 2007; Mitreva *et al.*, 2007); resistance to the BZs is associated with point mutations at codon 200 (Phe200Tyr), codon 167 (Phe167Tyr) and codon 198 (Glu198Ala) in the isotype 1 β-tubulin gene (Silvestre & Cabaret, 2002; Wolstenholme *et al.*, 2004; Ghisi, Kaminsky, & Maser, 2007; de Lourdes Mottier & Prichard, 2008). However, some *T. circumcincta* isolates have been shown to survive BZ treatment *in vivo* and *in vitro* when they have susceptible genotypes at codon 200, suggesting other, as yet unidentified mechanisms, are enabling the parasites to exhibit resistance (Stenhouse, 2007).

LEV, as a cholinergic agonist at the nicotinic neuromuscular junctions, opens and blocks the acetylcholine receptor-mediated cation channels resulting in the rapid tonic paralysis of the somatic muscles (McKellar & Jackson, 2004; Rayes *et al.,* 2004; Prichard & Geary, 2008). Three nAChR receptors subunits, *unc*-38, *unc*-29 and *lev*-1, identified in *C. elegans* appear to be involved in LEV resistance (Kopp *et al.*, 2008). Mutations in *unc*-29 or *unc*-38 caused a complete loss of sensitivity to LEV (Fleming *et al.*, 1997). A mutation from glutamic acid (Glu) to glycine (Gly) at amino acid 153 (Glu153Gly) in the *unc*-38 gene of *C. elegans* was necessary for LEV resistance (Rayes *et al.*, 2004; Martin & Robertson, 2007).

IVM binds irreversibly to GABA and GluCl channels causing the hyperpolarisation and flaccid paralysis of pharyngeal and somatic muscle cells leading to starvation and immobility of the worms (Martin *et al.*, 1998; Blackhall *et al.*, 1998a; Blackhall, Prichard, & Beech, 2003; Prichard & Roulet, 2007; Stenhouse, 2007; James & Davey, 2008). In *C. elegans*, concurrent mutations in several GluCl α-type subunit genes, *avr*-14, *avr*-15 and *glc*-1, are required to confer high levels of IVM resistance; mutations of any two of the three genes only confers low or no resistance (Dent *et al.*, 2000; Cook *et al.*, 2006; McCavera, Walsh, & Wolstenholme, 2007). The molecular basis of resistance to IVM in trichostrongyle parasites, such as *T. circumcincta*, remains to be elucidated (Geary, 2005; Prichard *et al.*, 2007).

Alternative mechanisms of resistance to those described above, such as changes in gene expression levels, remain largely unexplored. As previously stated in Chapter 1, differences in expression levels of genes may be **constitutive**, where a gene is always expressed differentially between anthelmintic-susceptible and -resistant isolates of parasites at rest, or **inducible**, where a change in gene expression is observed between parasites that have been exposed to anthelmintic and those that remain unexposed. Again, most research in this area has adopted a "candidate gene approach", concentrating on genes that are suspected of being involved in drug efflux or metabolism, for example the Pgps and CYPs (Kotze, 1997; Prichard & Roulet, 2007). However other, as yet unidentified, genes may also be involved in resistance; this provides an opportunity to use a more global approach to identify changes in gene expression, which has potentially not been prejudiced

by selectively looking for particular genes. "Next generation" sequencing approaches, such as Roche 454, Illumina Solexa and ABI SOLiD technologies, offer an alternative to the candidate gene approach as they can offer an unbiased and global analysis of quantitative and qualitative genetic changes associated with a particular phenotype or in response to treatment (Marguiles *et al.*, 2005). The digital readout (Expressed Sequence Tag; EST) from the sequenced cDNA reflects the mRNA expression level of that particular gene or contiguous sequence (Nagaraj, Gasser, & Ranganathan, 2007). All of these next generation sequencing technologies offer advantages over the traditional Sanger sequencing method (Sanger, Nicklen, & Coulson, 1977), in terms of cost, throughput and amount of data generated, however, the length of the reads produced may be significantly shorter (Pop & Salzberg, 2008; Shin *et al.*, 2008; Shendure & Ji, 2008).

Roche 454 sequencing amplifies cDNA strands which have been attached to agarose beads in a picolitre sized emulsion-based pyrosequencing reaction; at each sequencing cycle a single type of nucleotide is added, followed by the substrate luciferin which causes light to be emitted from the beads where a nucleotide has been incorporated. Roche 454 sequencing generates hundreds of thousands of sequences of between 200 and 400 nucleotides in length in a single run with an average accuracy of 96% (Marguiles *et al.*, 2005; Graveley, 2008; Mardis, 2008; Rothberg & Leamon, 2008). Like the 454 platform, the ABI SOLiD (Sequencing by Oligo Ligation and Detection) also uses the emulsion PCR method, utilising ligase instead of polymerase. The cDNA is captured on paramagnetic beads, immobilized in a solid substrate for the sequencing reaction stage which utilises a fluorescently labelled octamer that generates a signal corresponding to base 5 of the octamer. The octamer is then cleaved between bases 5 and 6, removing the fluorescent part of the octamer. Further cycles enable the sequencing of every 5^{th} base. The rest of the sequence is generated using other octamers which target different bases (Shendure & Ji, 2008). SOLiD sequencing generates 3-4 gigabases of sequences in approximately five days with an average read length of 25-50 bases (Graveley, 2008; Mardis, 2008). Illumina Solexa attaches pieces of cDNA, which are flanked with adaptors, to a solid substrate. These are clonally amplified by bridge PCR to generate a cluster of identical amplicons which are then linearized to generate single stranded cDNA. Sequencing by synthesis is carried out utilising fluorescently labelled dNTPs with a terminating moiety that only allows a single nucleotide to attach at each cycle. The

fluorescence is measured and then the terminating moiety removed, allowing the next dNTP to attach (Mardis, 2008; Shendure & Ji, 2008). This process generates large numbers of sequences (1300Mb sequence/ run) of a similar length to those generated using the SOLiD method (Shendure & Ji, 2008; Graveley, 2008).

Utilizing these next generation sequencing approaches allows the rapid generation of databases containing novel sequences. This is particularly important for genomic resource-poor organisms such as *T. circumcincta*, where the amount of sequence data is limited; currently the *T. circumcincta* genome project has generated 14Mb of sequence (http://www.sanger.ac.uk/Projects/T_circumcincta/;) and only ~6,000 ESTs are in the public domain (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html). Although the cost of Roche 454 sequencing is higher compared to SOLiD or Solexa sequencing, using this method to investigate the transcriptome of *T. circumcincta* is preferable at this point in time as it increases the chance of generating meaningful data with sufficient read length to enable correct assembly and identification of ESTs when a genome sequence is not available to use as a scaffold (Shendure & Ji, 2008). The inducible transcriptomic response of the MOTRI multi-drug resistant isolate of *T. circumcincta* (MTci5) following *in vitro* IVM exposure was therefore investigated using Roche 454 sequencing. *In vitro* IVM exposure was chosen as very little is known about how drug-resistant *T. circumcincta* responds to IVM exposure (Prichard *et al.*, 2007). As the market leader in terms of anthelmintic sales in the livestock market, it is especially important to determine how parasites are able to survive IVM treatment (Van Zeveren, 2009). Using the methods described below, two pools of adult parasites were incubated for 5 hours, one pool exposed to IVM and the other pool as an unexposed control. RNA was extracted from each pool and subjected to Roche 454 sequencing and bioinformatic analysis of the subsequent EST dataset.

5.2: Materials and methods

5.2.1: *In vitro* exposure to ivermectin

Adult MOTRI *T. circumcincta* were collected from the abomasum of a sheep previously infected with 15000 L_3 21 days P.I using a modification to the method

described previously in Section 2.1.6. After removal of the abomasum and collection of the abomasal contents and washings, individual worms were identified under a compound microscope using the x10 objective. These were removed from the abomasal contents using a hypodermic needle bent into an L shape. Parasites were inspected to confirm they were not damaged and sexed under a microscope before being placed in Petri dishes, containing 9850µL RPMI and 100µL penicillin/streptomycin (final concentration 1%), with an approximate 50:50 sex ratio, a total of approximately 100 parasites in each of 4 dishes. 50µL DMSO (final concentration 0.05%) was added to two of the petri dishes to act as the unexposed controls and 50µL DMSO/ IVM (final concentration DMSO = 0.05%, final concentration IVM = 5µg/mL) was added to the other two dishes. The petri dishes were placed in an incubator at 37°C in 5% CO_2 for 5 hours before the parasites were removed from the culture medium, placed in 1x PBS, and snap frozen in liquid nitrogen.

5.2.2: 454 sequencing reaction

Total RNA was extracted from one IVM-exposed and one -unexposed pool of parasites using Trizol reagent (Invitrogen) according to the manufacturer's recommendations, as described in Section 2.3.1. Confirmation of RNA integrity was obtained by agarose gel electrophoresis and by determining the RNA concentration using a NanoDrop® ND-1000 spectrophotometer. 1µg total RNA from the IVM-exposed and -unexposed samples was supplied to Cogenics (www.cogenics.com) for cDNA synthesis and Roche 454 sequencing. Adaptors were ligated onto cDNA transcribed from the respective samples to aid subsequent transcript identification (exposed vs unexposed) and the cDNA subjected to a quarter plate Roche 454 sequencing run on a GS-FLX 454 genome sequencer. The basic technique is as described in Marguiles *et al* (2005). Briefly, individual fragments of the cDNA library were captured onto beads inside an oil emulsion containing the PCR reagents. Following PCR amplification, the emulsion was broken and a picolitre- sized pyrosequencing reaction set up for each bead containing amplified DNA on a fibre-optic slide. The light emitted by the individual pyrosequencing reactions in each well is converted into a digital read of sequence data.

5.2.3: Bioinformatics

Whole genome sequencing assembly was carried out using the proprietary Roche 454 Newbler Assembler software. The identity of contiguous sequences (contigs) was sought using BLAST [Basic Local Alignment Search Tool (Altschul *et al.*, 1990)], searching against three databases; Swissprot, Wormpep and the "est_others" nucleotide database curated by NCBI. Functional classification was also carried out on the contigs on the basis of homology with the KEGG genomes orthology database using KAAS (KEGG Automatic Annotation Server: http://www.genome.jp/tools/kaas/) (Kanehisa & Goto, 2000; Moriya *et al.*, 2007) with the single direction hit option. Contigs could be represented in more than one functional group.

Comparison of EST abundance was carried out by tagging contigs with identifiers to distinguish their origin (IVM-exposed or -unexposed) then pooling and assembling these sequences using SeqMan (DNASTAR Lasergene version 8) with the ProAssembler option. Sequences with a better than 95% match were assembled into clusters, manually curated and/or trimmed to resolve any conflicts arising from sequencing artefacts. The consensus sequences of the clusters were BLAST searched in the NCBInr database to give a putative identity.

5.2.4: Statistical analysis of 454 dataset

A 2×2 Fisher's Exact test was conducted to test for differences between the proportions of exposed and unexposed contigs for each functional classification as identified by KEGG analysis. The estimated p-values for all functional groups were then adjusted using the False Discovery Rate (FDR) approach (Benjamini & Hochberg, 1995) to take into account multiple comparisons of treatments.

The methodology used to identify clusters indicated that there was likely to be excess clustering or overdispersion in the data. However, with a single replicate for each combination of cluster and treatment, it was not possible to quantify this effect. Therefore

an approach was adopted which combined information from two statistical models that together spanned the range of results which would arise from the more appropriate model.

The baseline statistical model was a saturated model that assumed that differences observed between treatments for all clusters were fully explained as arising from the differences in the treatment means. This involved fitting a generalised linear model (GLM) with a binomial error distribution and logit link function (with cluster, treatment and a cluster by treatment interaction as fixed effects) or conducting a Fisher's Exact test for the number of reads in each cluster where a GLM was inappropriate. The p-values from this model were liberal. The estimated p-value indicating the level of statistical significance for differences between the exposed and unexposed groups in the numbers in each cluster was recorded. These p-values were adjusted using the FDR approach as discussed earlier. Subsequently, an estimate of the overdispersion in the data was obtained by refitting the GLM with only a treatment effect, assuming that all other observed differences were caused by excess clustering. This gave an overestimate of the true overdispersion in the data.

Finally, an alternative model was fitted, with over-dispersion fixed at the value obtained previously. This involved refitting the GLM with cluster, treatment and the cluster by treatment interaction as fixed effects, and overdispersion fixed at the previous estimate or conducting a Fisher's Exact test for the number of reads in each cluster and adjusting the resulting p-value for the previously estimated overdispersion. These p-values were adjusted using the FDR approach as mentioned earlier. The p-values from this model were conservative.

The FDR adjusted p-value for a cluster is the minimum FDR for which the observed difference between the mean expression level of exposed and unexposed worms would be accepted as statistically significant. It is possible to interpret these adjusted p-values as the expected probabilities of the clusters being assessed as showing no statistically significant differences, given the unadjusted p-values and a random choice of FDR threshold. The adjusted p-values from the two models were used to categorise each

cluster into one of three classes: clusters with statistically significant p-values under the conservative model were of prime scientific interest; clusters with non-significant p-values under the liberal model were unlikely to be of interest; other clusters were of indeterminate status, but could be prioritised for further work with respect to either set of p-values.

All statistical analyses were carried out using the 12th Edition of the GenStat statistical package (Payne *et al.*, 2009) except the Fisher's Exact test which was performed in the R software system version 2.9.2 (R Development Core Team, 2009: www.R-project.org)

5.3: Results

From the two samples (IVM –exposed and –unexposed) submitted for Roche 454 sequencing, a total of 98,685 reads (= ESTs) were obtained. A summary of these results is shown in Table 5.1. For the IVM-exposed sample, 55,341 individual sequence reads were obtained, of which 68.3% were assembled into 2,049 contigs with the remaining 15,488 reads not aligning to any other sequence (singletons). For the unexposed sample, a total of 43,344 individual sequence reads was obtained, of which 65% were assembled into 1,659 contigs with the remaining 13,857 reads as singletons. The average read length for the exposed and unexposed samples was 251 base pairs (Figure 5.1). Singletons which failed to form a contig were found to be either short sequences of poor quality (but with good homology to the sequencing primer); to have a high occurrence of repeats, or to contain sequencing errors and so were excluded from any further analyses.

Candidate resistance genes were absent from either dataset with the exception of one contig (contig 1062, containing 2 reads and 241bp long) in the IVM-exposed dataset with BLAST homology to an ABC transporter; subfamily B (MDR/TAP); member 1. The nucleotide sequence for this contig was aligned against the 11 available *T. circumcincta* Pgp sequences described in Chapter 3 and also against the 15 known *C. elegans* Pgp sequences. Phylogenetic trees were generated from these alignments using both the NJ and UPGMA distance matrices. Figure 5.2 shows the relationships between contig 1062 and

the 11 partial *T. circumcincta* Pgps previously identified, whilst Figure 5.3 shows the relationship between contig 1062 and the 15 *C. elegans* Pgps. Contig 1062 does not align exclusively to any individual sequence suggesting it may be a novel member of the *T. circumcincta* Pgp family or possibly a different part of one of the genes previously identified.

5.3.1: Comparison of transcript levels based on KEGG functional groups

Figure 5.4 shows the functional classification of each contig as identified using the KEGG genomes orthology database. 56.2% of the exposed and 55.8% of the unexposed samples were shown as having no orthologous hit. In the exposed sample 12.6% of the contigs were in the group 'energy metabolism', 8.6% of the contigs were in the group 'circulatory system' and 13% of the contigs in the group 'neurodegenerative disease'. In the unexposed sample, the percentages for these same groups were 12.8%, 8.9% and 13.1%, respectively. There was no statistically significant difference between the mean proportion of reads in exposed and unexposed samples for these four largest groups so they were removed from the data presented (Figure 5.4) to allow for easier visualisation of the smaller functional groups. However, all groups were included in the statistical analysis. The KEGG orthologous group for 'translation' showed a statistically significant increase in expression in the IVM-exposed sample compared to the IVM-unexposed sample ($P<0.001$), whereas the orthologous groups for 'amino acid metabolism', 'carbohydrate metabolism' and 'xenobiotic degradation and metabolism' showed statistically significant ($P=0.001$, $P=0.022$ and $P=0.042$ respectively) decreases in mean proportion of reads in the IVM-exposed sample compared to the -unexposed sample. All other groupings, including those with no orthologous hits, were found to have no statistically significant difference in the mean proportion of reads between the IVM-exposed and IVM-unexposed samples.

5.3.2: Comparison of transcript levels within clusters

2,570 clusters were obtained when the exposed and unexposed contigs were combined; individual clusters varied in size from 2 to 5,823 reads per cluster. Sixteen clusters were found to have statistically significant differences in expression level ($P<0.0001$) between the mean proportion of exposed and unexposed reads in the cluster

under the conservative statistical model. Under the liberal statistical model, a further 355 clusters were found to have statistically significant differences ($P<0.05$) between the mean proportion of exposed and unexposed reads in the cluster. The remaining 2,199 clusters were not statistically significant under either model.

Of the sixteen statistically significant clusters under the conservative model, half of these clusters had an increased mean proportion of exposed reads compared to unexposed reads. The change in read numbers and putative identities of these sixteen clusters is shown in Table 5.2. Four of the clusters had no BLAST homology assigned and two were un-named or hypothetical proteins. A further four clusters were identified as vitellogenin, however, there was no consistency in expression as one cluster showed an increased mean proportion of exposed reads whilst the three other vitellogenin clusters showed a reduced mean proportion of exposed reads, compared to unexposed reads. Cytochrome C oxidase subunits were also represented four times, three were of subunit type I and one was a subunit type III. Of the four cytochrome C oxidase subunits, half showed a reduced number of exposed reads compared to unexposed reads. The remaining two clusters were identified as a major sperm protein, showing a reduced mean proportion of exposed reads, and a NADH dehydrogenase subunit 4, showing an increased mean proportion of exposed reads.

Of the 355 statistically significant clusters under the liberal model, 228 (64%) of these clusters showed a reduced mean proportion of exposed reads compared to unexposed reads. Of the 355 clusters, 89 (25%) had no BLAST identity whilst 138 (39%) were hypothetical, predicted or putative proteins. Some cluster identities featured more commonly than other clusters; the majority of cluster identities were only represented once. Nine of the clusters were identified as heat shock proteins (HSPs), five showing an increased number of exposed reads and four a reduced number of exposed reads. The next most common identities were six clusters representing C-type single domain activation associated secreted proteins (ASPs). Apart from one cluster, all showed a reduced mean proportion of exposed reads compared to unexposed reads. Another five clusters were identified as secreted protein precursors; four had a reduced mean proportion of exposed reads compared to unexposed reads and were secreted protein 5 precursors. One showed

an increased mean proportion of exposed reads compared to unexposed reads but, unlike the previous four secreted protein precursors, was identified as a secreted protein 6 precursor. Like the clusters with the statistically significant results under the conservative model, vitellogenin also featured in the significant results in the liberal model, with four clusters again being assigned this identity and following the same expression pattern. Three more clusters were also identified as cytochrome C oxidase subunits, two were subunit type I and one subunit type III. However, compared to the statistically significant clusters with the same identity in the conservative model, all of these clusters showed an increase in the mean proportion of exposed reads. The identities of the 355 clusters are provided in Appendix 5.

5.4: Discussion

Limited sequence data are available for *T. circumcincta* and, although a genome sequencing project is underway (http://www.sanger.ac.uk/Projects/Helminths/), it is unlikely that an annotated genome sequence will be available in the near future. Currently, the *T. circumcincta* genomic resource represents ~6,000 ESTs (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html). Currently, it is estimated that parasitic nematodes have approximately 20,000 genes; in *C. elegans* the genome is approximately 100 megabases in size whilst in *H. contortus*, due to bigger and more numerous introns, the genome is estimated to be 200-350 megabases long, although the exact genome size has not yet been determined. The *T. circumcincta* genome is more likely to mirror that of *H. contortus* as both are parasitic species, although Leroy, Duperray & Morand (2003) estimated the genome size of *T. circumcincta* to be 59 megabases. Currently, the *T. circumcincta* genome sequencing stands at 14 megabases (http://www.sanger.ac.uk/Projects/T_circumcincta/) (Parkinson *et al.*, 2004b; Sutherland & Scott, 2010). Although not providing full coverage of the transcriptome, the generation of a total of 98,685 novel reads (i.e. ESTs) provides a valuable resource to allow the investigation of whether inducible gene expression changes have a role to play in resistance to anthelmintics.

The majority of the KEGG genome functional classification groups, including the largest 'no hits' group were equally represented in both exposed and unexposed datasets. The scale of changes observed following IVM exposure is very different to those observed when comparing the gene expression profile of *T. circumcincta* during profound changes in the worms' physiology, e.g. transition to a different life cycle stage (Nisbet *et al.*, 2008). A statistically significant increase in the expression of genes linked to translation in the exposed sample suggests that one of the responses to IVM exposure is a generalised increase in protein production. As expected, translation was one of the larger functional groups shown in Figure 5.4; most of the contigs in this group were of ribosomal origin. All of the groups showing a statistically significant reduction in expression appear to be linked to metabolism, including 'amino acid metabolism', 'carbohydrate metabolism' and 'xenobiotic degradation and metabolism'. One way the parasites could be trying to survive exposure to IVM could be to stop any non-essential cellular processes; however the general reduction in metabolic processes could also indicate the parasites being paralysed or killed by IVM.

The functional group 'xenobiotic degradation and metabolism' showed a reduction in expression in the exposed group. This was where any candidate resistance genes such as cytochrome P450s would be expected to sit, and an increase in expression of genes potentially involved in generalised drug handling mechanisms was anticipated. Genes of this class are generally expressed at low levels so the majority are likely to be missing from this transcriptomic analysis (Kotze, 1997; Barrett, 1998). This is because an EST dataset is subject to sampling bias which leads to under-representation of rare transcript (Nagaraj, Gasser, & Ranganathan, 2007). Interestingly, the only contig identified as a candidate resistance gene (IVM-exposed contig 1062) was identified as an ABCB1 which according to the KEGG functional classifications used was classified in the 'membrane transport' group not in the 'xenobiotic degradation and metabolism' group. This contig was aligned against Pgp sequences from both *C. elegans* and *T. circumcincta* as Pgps are also part of the B subfamily of ABC transporters (Sheps *et al.*, 2004). The alignment trees (Figures 5.2 and 5.3) did not indicate that contig 1062 was the same or closely related to any of the *T. circumcincta* previously identified, nor did it align to any particular *C. elegans* Pgp.

None of the putative target genes implicated in IVM resistance, such as the GABA and GluCl channels described in the Chapter introduction, were present in the cluster analysis. A more targeted approach such as real-time PCR, or more in-depth sequencing, would be required to assay all genes in this class and to provide statistically meaningful data on differential expression of these genes. It is also important to remember that the MOTRI isolate represents a spectrum of resistant phenotypes and will include a proportion of susceptible individuals that will undoubtedly be killed by the IVM exposure *in vitro*. It is possible that the level of IVM exposure chosen caused the death of the relatively susceptible parasites and induced stress responses in the more resistant parasites. The contribution of dead or dying individuals may have masked the expression of generalised drug handling mechanisms such as Pgps and CYPs.

Analysis of the functional classification of the contigs in the exposed and unexposed datasets revealed only slight differences between the two datasets. This suggested that any differences in expression levels were subtle and were more likely to be found within clusters rather than within functional groups. HSPs show increased expression when cells are exposed to heat or other environmental stresses (Lindquist, 1986). However four of the nine clusters identified as HSPs showed a reduced mean proportion of exposed reads when exposed to IVM. Six of the clusters were identified as C-type single domain ASPs and the majority were shown to have reduced expression in the exposed sample compared to the unexposed sample. ASPs are nematode-specific proteins and are believed to have roles in parasite establishment and maintenance and may also act as allergens (Visser *et al.*, 2008). A reduced expression of these proteins in response to IVM could be the adult parasites down-regulating activity associated with their parasitic behaviour.

More than one cluster was identified as vitellogenin or cytochrome C oxidase subunits; however, no clear pattern of increased or decreased expression of these cluster identities could be identified. Cytochrome C oxidases are part of the mitochondrial electron transport chain; changes in expression may be due to changes in the parasite's respiration rate in response to drug exposure. The majority of clusters identified as vitellogenin (six out of the eight identified as having statistically significant differences in

read numbers between the two samples) showed a decreased expression in the exposed sample, suggesting a reduction in egg production in response to IVM exposure.

Unfortunately, we do not have access to the parental susceptible population of *T. circumcincta* from which MOTRI was selected in the field. This makes meaningful genetic comparisons of susceptible and resistant isolates difficult as they do not share a common genetic background. However, in this study we have explored global changes in gene expression within one isolate following exposure to IVM *in vitro*. This experiment has shown that next generation sequencing approaches can be used to investigate gene expression changes in pools of complex nematode parasites such as *T. circumcincta*, for which genomic resources are limited. This small-scale study has facilitated the generation of two new sets of ESTs from a multiple drug-resistant *T. circumcincta* isolate, and allowed a comparative investigation into the transcriptomic response of this isolate to IVM exposure. Results clearly show that not all the observed expression changes reflect up-regulation, many of the changes involved down-regulation. Results also show that the MOTRI isolate, which exhibits an IVM efficacy of 60%, is still affected by IVM treatment.

In terms of understanding the genetic basis of resistance, next generation sequencing approaches have great potential to contribute to this area of research. These high-throughput sequencing technologies now routinely require very small amounts of starting material, making them ideal for studies of parasitic helminths for which biological material can be limiting. However, care also needs to be taken when using EST datasets to estimate gene numbers; the fragmented nature of the sequence generated (particularly those generated using Solexa and SOLiD) can cause over-representation of a gene when two non-overlapping fragments of the same gene are taken to be two separate genes (Abubucker *et al.*, 2009). This may be the case with the dataset described above, especially in the case of vitellogenin which is a very long transcript (Nisbet & Gasser, 2004). Until a full genome is available for *T. circumcincta*, it will not be possible to definitively say whether the gene expression changes identified in this Chapter are real or a result of erroneous gene fragmentation.

Although not exploited in this study, the assembly of large sequence datasets into contigs allows a precise analysis of qualitative genetic changes, for example SNPs and indels, which may be linked to resistance. Furthermore, detailed network analysis of large EST datasets such as those generated by next generation sequencing may contribute to making biological sense of complicated inter-related gene expression changes. The biggest challenge when generating large amounts of EST data is the correct application of downstream bioinformatic and statistical analysis to answer biologically relevant questions (Pop & Salzberg, 2008; Shendure & Ji, 2008). In this research we have developed a two-model based statistical method to allow the categorization of a large dataset into more refined groups, aiding the identification of the most significant changes. The application of next generation sequencing approaches when applied to the most appropriate biological resource should ultimately help identify the genuine genetic determinants underlying anthelmintic resistance.

Table 5.1: Summary of the results of the Roche 454 sequencing carried out on two pools of MOTRI mixed-sex *T. circumcincta* adults, one exposed to ivermectin *in vitro* and one as an unexposed control.

Sample	**Number of reads**	**% reads assembled**	**Number of contigs**	**Largest contig length (BP)**	**Number of singletons**
IVM exposed	55,341	68.3	2,049	3,735	15,488
IVM unexposed	43.344	65.0	1,659	2,872	13,857

Table 5.2: The sixteen clusters with statistically significant differences in gene expression between IVM-exposed and –unexposed worms under the conservative model. The columns show the cluster identity used in the analysis (**cluster**), the number of reads within that cluster (**U reads** = unexposed and **T reads** = exposed), whether gene expression up (+) or down (-) regulated compared to unexposed reads (**change**) and what the putative cluster identity was (**Best BLAST identity**).

Cluster	**U reads**	**T reads**	**Change**	**Best BLAST identity**
1782	0	863	+	Cytochrome c oxidase subunit I [*Ancylostoma caninum*] YP_002725710.1
210	508	1439	+	Cytochrome c oxidase subunit I [*Steinernema rarum*] AAY22431.1
538	1893	3930	+	Hypothetical protein Bm1_17870 [*Brugia malayi*] XP_001895031.1
1621	0	176	+	NADH dehydrogenase subunit 4 [*Cooperia oncophora*] NP_851330.1
378	0	62	+	No hits found
1747	0	253	+	No hits found
175	271	836	+	Unnamed protein product [*Macaca fascicularis*] BAE89779.1
1746	0	291	+	Vitellogenin-6; Flags: Precursor [*Oscheius brevesophaga*] Q94637.1
2557	516	0	-	Cytochrome oxidase subunit I [*Cooperia oncophora*] BAE72507.1
409	347	12	-	Cytochrome oxidase subunit III [*Ancylostoma duodenale*] CAD10435.1
672	348	199	-	Major sperm protein [*Dictyocaulus viviparus*] ACC64398.1
2550	283	0	-	No hits found
2558	46	0	-	No hits found
574	182	0	-	Vitellogenin [*Haemonchus contortus*] AF305957_1
102	418	241	-	Vitellogenin structural genes (yolk protein genes) family member (vit-2) [*Caenorhabditis elegans*] NP_001123117.1
653	783	299	-	Vitellogenin-6; Flags: Precursor [*Oscheius brevesophaga*] Q94637.1

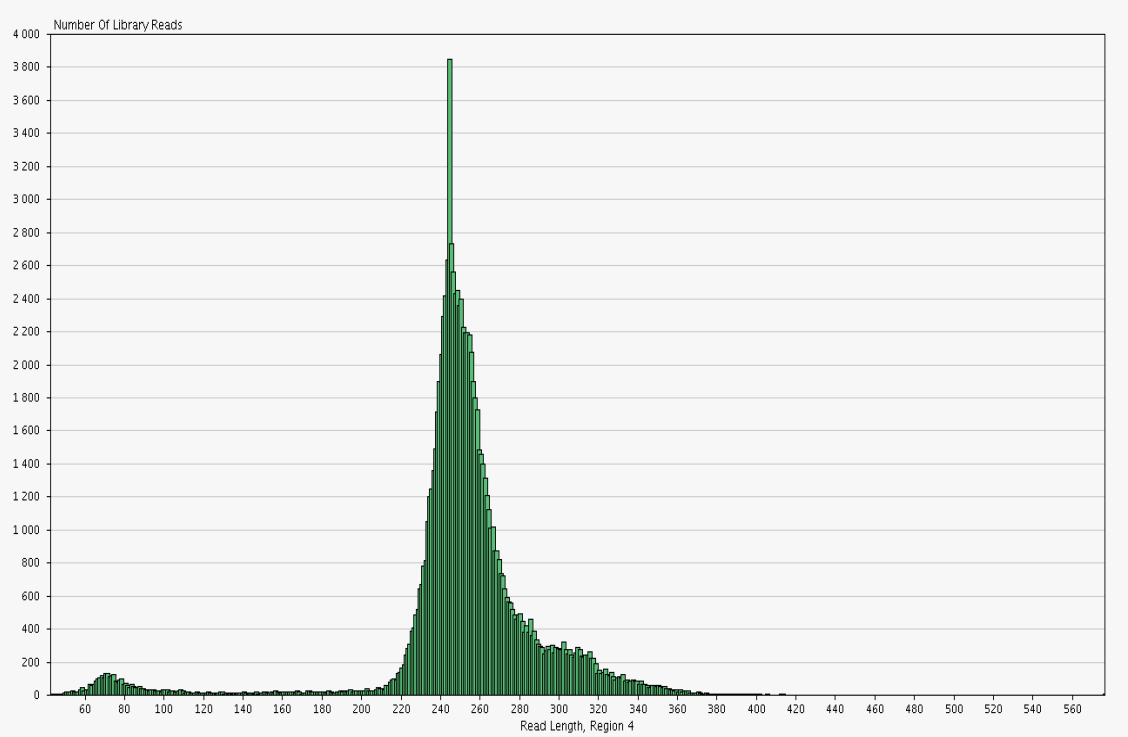
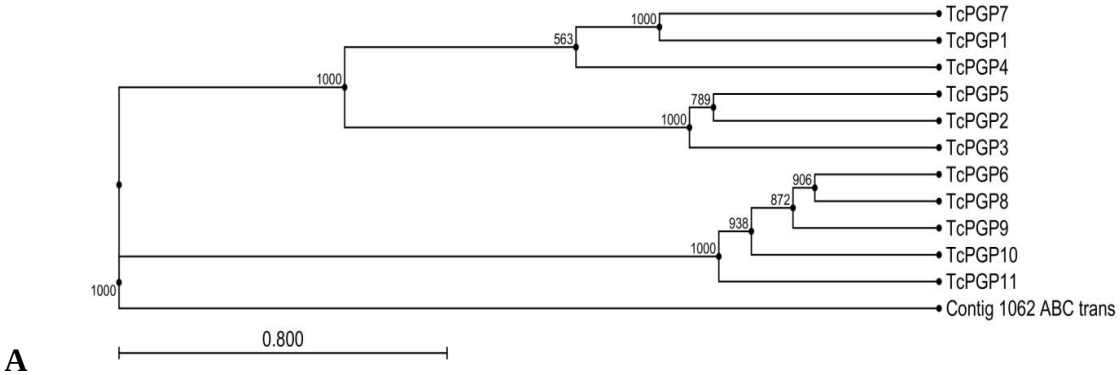
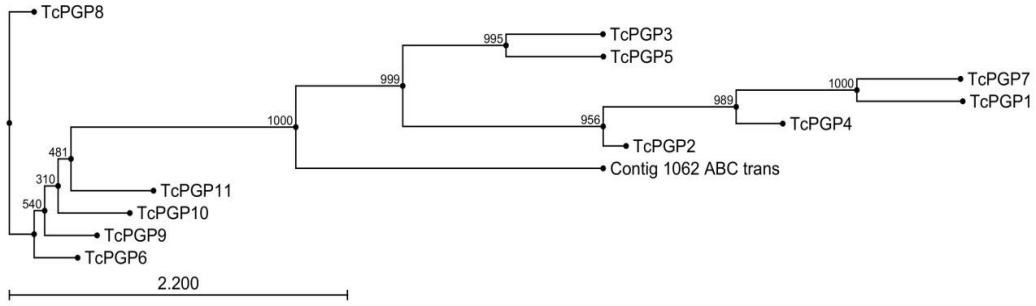


Figure 5.1: Graph showing the average read length of the sequences generated by Cogenics using the two pools of RNA generated from MOTRI *T. circumcincta* adults, one exposed to ivermectin and one unexposed *in vitro*. Reproduced from 454 run report from Cogenics



A



B

Figure 5.2: Relationship of contig 1062 from the IVM-exposed dataset to the 11 partial Pgp gene sequences identified in Chapter 3. **A**: Tree generated using the UPGMA distance matrix with a bootstrap value of 1000. **B**: Tree generated using the NJ distance matrix with a bootstrap value of 1000.

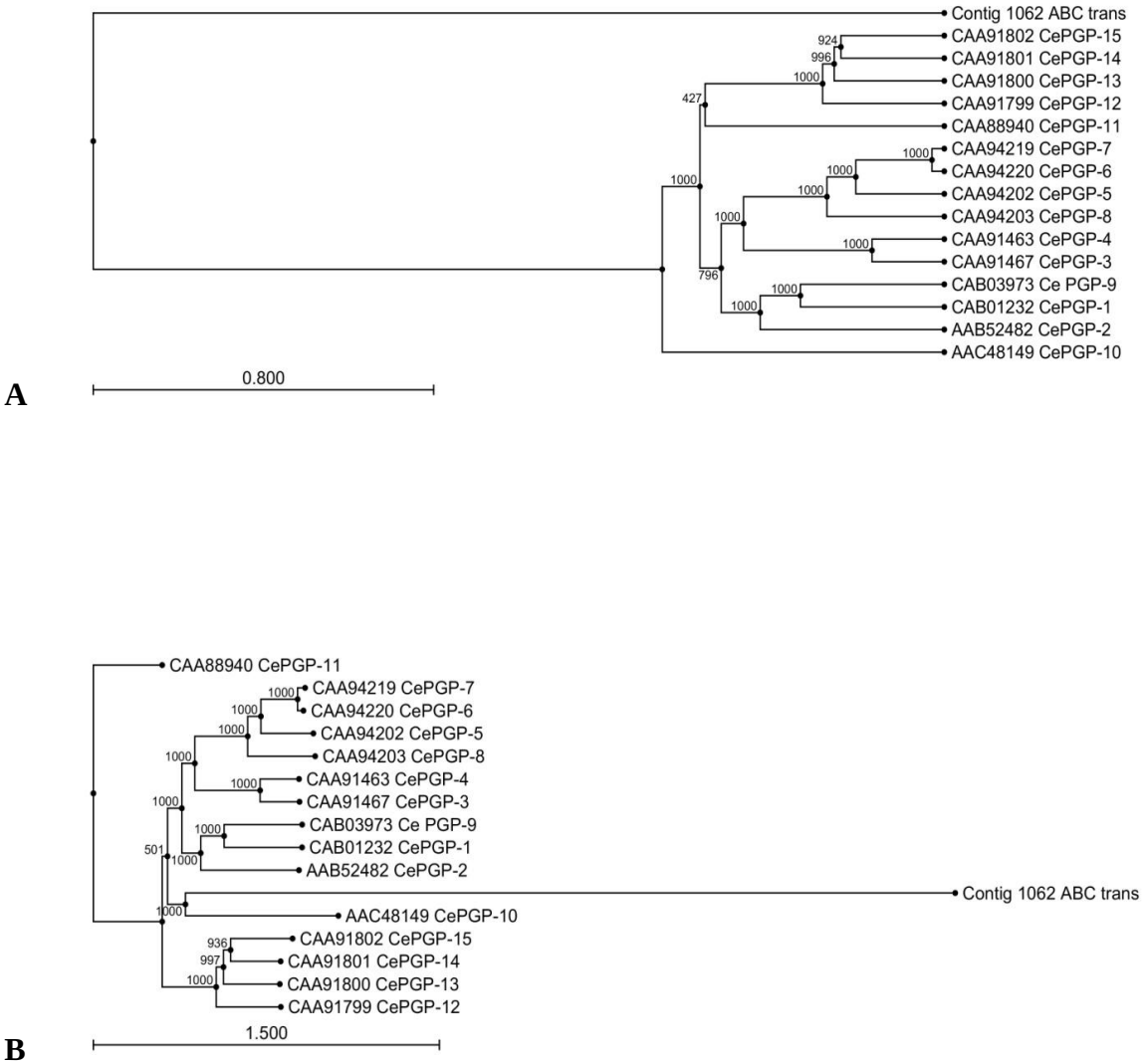


Figure 5.3: Relationship of contig 1062 from the IVM-exposed dataset to the 15 *C. elegans* Pgp gene sequences. **A**: Tree generated using the UPGMA distance matrix with a bootstrap value of 1000. **B**: Tree generated using the NJ distance matrix with a bootstrap value of 1000.

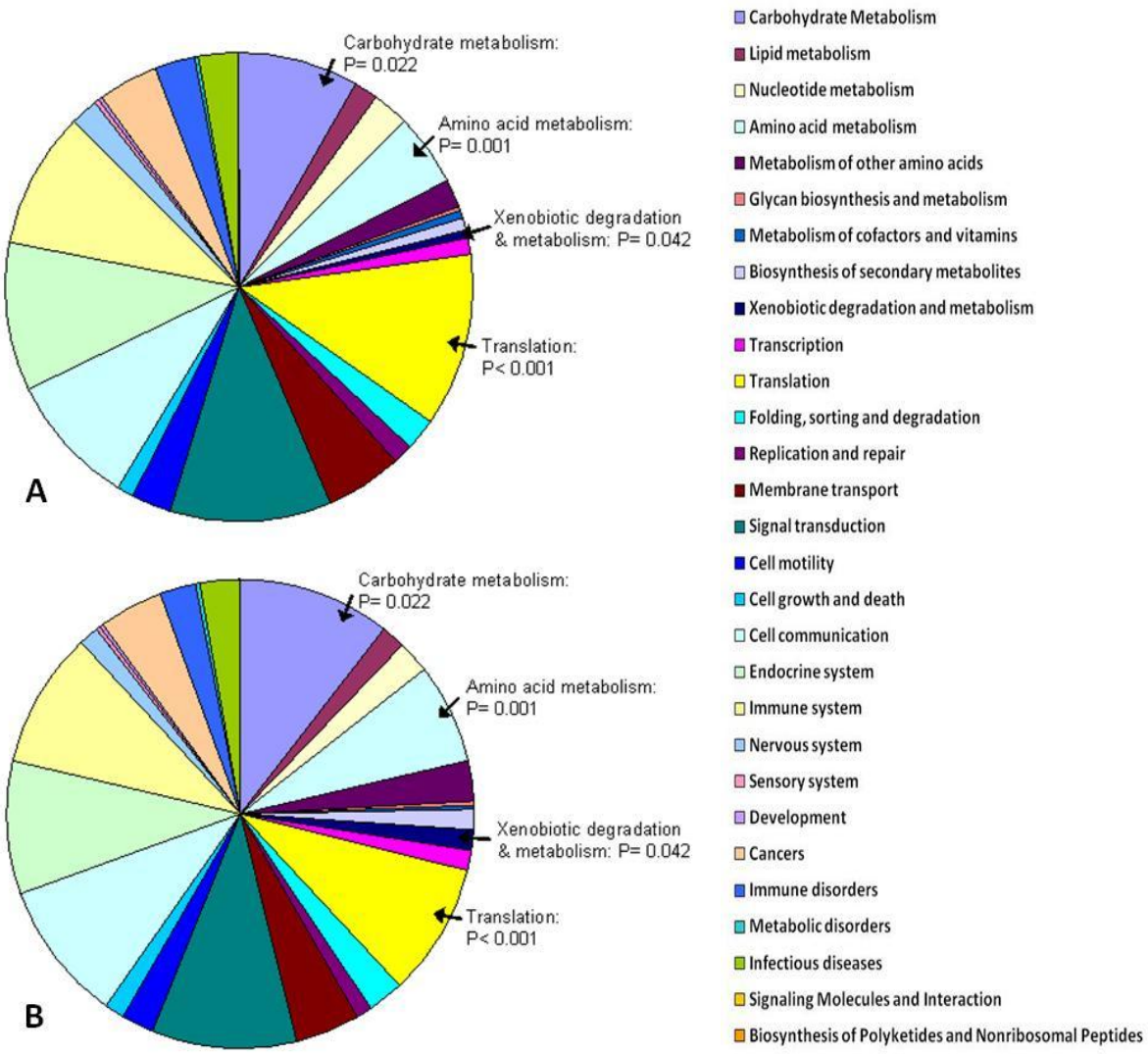


Figure 5.4: Functional classification of reads from 454 sequencing of (**A**) *in vitro* ivermectin exposed MOTRI *T. circumcincta* adults and (**B**) unexposed MOTRI *T. circumcincta* adults identified using KAAS. Not shown in the Figure are the 56.2% of exposed reads and 55.8% of unexposed reads with no known hits, the 12.6% of exposed reads and 12.8% of unexposed reads in the functional group 'energy metabolism', the 8.6% of exposed reads and 8.9% of unexposed reads in the functional group 'circulatory system' and the 13% of exposed reads and 13.1% of unexposed reads in the functional group 'neurodegenerative disease'. Individual reads may be present in more than one functional group.

Chapter 6: Suppression Subtractive Hybridisation Comparison of Gene Expression in CVL and MOTRI *Teladorsagia circumcincta* Isolates

6.1: Introduction

Candidate resistance gene studies aimed at identifying the genes associated with IVM resistance in parasitic nematodes such as *T. circumcincta* have not definitively identified either the genes involved nor markers associated with resistance (Gilleard, 2006; von Samson-Himmelstjerna *et al.*, 2007). Most of the studies to date have looked for mutations associated with IVM resistance, yet other mechanisms of resistance such as changes in expression levels of genes potentially involved in drug handling or metabolism have largely been ignored. The work described in Chapters 3 and 4 has investigated the role differential gene expression of Pgps and CYPs could play in the IVM resistance phenotype exhibited by certain isolates of *T. circumcincta*. However, as discussed in Chapter 5, changes in the expression of other genes, not classified as candidate resistance genes, could potentially enable *T. circumcincta* to exhibit IVM resistance. The use of a more global "non-hypothesis" driven approach to identify genes exhibiting altered expression profiles in response to IVM exposure (see Chapter 5) or in a IVM-resistant isolate could enhance the knowledge of how IVM-resistant *T. circumcincta* is able to handle the drug and also survive drug exposure. One such technique, which can be used to investigate changes in gene expression, is described below.

Suppression subtractive hybridisation (SSH) was first described by Diatchenko *et al* (1996) as a method to allow the identification and isolation of differentially expressed genes between two pools of cDNA. Firstly, normalisation of the tester cDNA sample equalises the abundance of the cDNAs within that pool, then sequences common to both samples are removed by hybridising the driver cDNA to the tester cDNA. A final PCR step selectively amplifies only those cDNAs which are unique to the tester cDNA population whilst preventing the amplification of non-unique transcripts. Identification of the unique transcripts in the tester population is then achieved by cloning and sequencing the resultant PCR products. The method is especially useful for identifying transcripts with low abundance as the first, normalisation, step equalises the abundance of the transcripts, whilst the suppressive PCR step means single-stranded, non-differentially expressed double-stranded or driver specific cDNAs fail to amplify. SSH has advantages

over previous methods, such as differential display, as it is less-time consuming and requires less starting material (Gasser & Newton, 2000; Cottee *et al.*, 2006). Compared to standard EST library construction, SSH also produces fewer redundant sequences, made up of common, abundantly expressed genes and enables the identification of genes expressed at low levels by selectively amplifying them (Diatchenko *et al.*, 1996; Nisbet & Gasser, 2004; Cottee *et al.*, 2006).

Few studies to date have utilised SSH to investigate differences in gene expression in parasitic nematodes. Where it has been employed, SSH has been used to investigate stage-specific changes in gene expression, particularly those changes associated with the transition between the free-living and parasitic stages, whilst other studies have investigated gender differences in expression. Two-way (identifying increased and reduced expression of genes between the tester and driver cDNA) stage-specific expression of genes has been investigated in *Trichinella spiralis*, where comparisons were made between newly hatched larvae, day 3 adults and day 5 adults; in *An. caninum*, where genes with differential expression between free-living and parasitic larval stages were identified and, finally, in *T. circumcincta*, where gene expression was compared between xL_3 and early L_4 parasites (Liu *et al.*, 2007; Nisbet *et al.*, 2008; Datu *et al.*, 2008). Comparisons of gene expression between male and female adult parasites, using SSH, have been carried out in *O. dentatum* and *T. vitrinus* (Nisbet & Gasser, 2004; Cottee *et al.*, 2006). Uniquely, one study in *T. spiralis* has used SSH to investigate differential expression in life-cycle stages and between the parasite sexes. This study investigated the genes induced in *T. spiralis* following *in vitro* heat shock and identified 12 genes induced in infective larvae following exposure to 43ºC, as compared to 37ºC (Mak, Sun, & Ko, 2001). No studies, to date have employed SSH, in parasitic nematodes to investigate the gene expression changes associated with anthelmintic resistance.

The investigation of constitutive expression differences has, until now, focused on the panel of candidate resistance genes as described in Chapters 3 and 4. The final comparison in this thesis was to use a ‘non-hypothesis’ driven one-way SSH approach to identify genes with increased expression in the resistant MOTRI *T. circumcincta* isolate compared to the susceptible CVL isolate. This was to complement the Roche 454 dataset

generated in Chapter 5 comparing the inducible expression of genes in the MOTRI isolate which had been exposed to IVM *in vitro*. The SSH method was chosen because it amplifies unique transcripts with low abundance whereas the methods used prior to SSH were inefficient at amplifying transcripts of low abundance (Diatchenko *et al.*, 1996). To identify genes with increased expression in, or unique to MOTRI, MOTRI cDNA was chosen as the tester sample while the CVL cDNA was the driver sample.

6.2: Materials and methods

6.2.1: Extraction of RNA

RNA was extracted from two pools of adult worms, one from the known IVM-susceptible *T. circumcincta* CVL (MTci2) isolate and the other from the triple BZ, LEV and IVM-resistant *T. circumcincta* isolate (MOTRI or MTci5) using the Trizol® method, as described previously (Section 2.3.1). Estimations of the concentration of the RNA were made using a NanoDrop® ND-1000 spectrophotometer. From these readings it was decided that the subtraction experiment would utilise cDNA made using total RNA as there were insufficient quantities of total RNA to purify sufficient poly A^+ RNA for the subtraction experiment.

6.2.2: Synthesis of cDNA

cDNA was synthesised from the above total RNA using the SMARTer™ PCR cDNA Synthesis kit (Clontech). First strand synthesis was carried out by combining, for each of the two isolates, 1µg total RNA with 1µL 3' SMART™ CDS Primer II A and nuclease free water to give a total volume of 4.5µL. This was incubated at 72ºC for 3 mins followed by 42ºC for 2 mins in an Applied Biosystems 2720 thermal cycler. While this reaction was incubating, a master mix for the two reactions was set up by combining; 4µL 5X First-Strand Buffer, 0.5µL 100mM DTT, 2µL 10mM dNTP mix, 2µL 12µM SMARTer II A Oligonucleotide, 0.5µL RNase Inhibitor and 2µL 100U SMARTScribe™ Reverse Transcriptase. 5.5µL of the master mix was transferred by pipette into each of the reaction tubes as soon as the thermal cycling was completed and the tubes incubated for 90 mins at 42ºC. The incubation time was extended to 90 mins to ensure full-length cDNAs were

obtained. The reaction was terminated by heating the tubes to 70ºC for 10 mins and the first strand reaction product diluted in 40µL TE buffer.

cDNA amplification, using long distance PCR, utilised 30µL of each of the first strand reaction mixes. A master mix comprising of 666µL sterile deionised water, 90µL 10X Advantage 2 PCR Buffer, 18µL 50X 10mM dNTP mix, 18µL 12µM 5' PCR Primer II A and 18µL 50X Advantage 2 Polymerase Mix was set up with 270µL aliquots of this master mix placed into each of the two tubes containing the first strand reaction mix. The resulting PCR reaction mix was split into three aliquots of 100µL, labelled A, B, and C, for each of the two samples and thermal cycling commenced in an Applied Biosystems 2720 thermal cycler pre-heated to 95ºC. Cycling parameters were 95ºC for 1min followed by 15 cycles of 95ºC for 15 secs, 65ºC for 30 secs and 68ºC for 6 mins. After the 15 cycles had finished, 30µL from each of the tubes labelled C was removed for PCR optimisation, 5µL of which was kept for gel analysis, whilst the remainder of tubes C and all of tubes A and B were stored at 4ºC. Further PCR cycles were carried out on the 30µL aliquots from tubes C, using the conditions described above. After 18, 21, 24 and 27 cycles, 5µl aliquots were taken for gel analysis. All of the 5µL aliquots were run on a 1.2% agarose gel and the gel visualised under UV. The results of this gel determined that the samples A, B and the remainder of C did not require further PCR cycles; 15 cycles was optimal for amplification of the PCR products.

6.2.3: *Rsa* I digestion of cDNA

The three tubes for each of the samples were combined and the PCR product cleaned up using the QIAquick® PCR purification kit (Qiagen), eluting into 50µL nuclease free water. The DNA was diluted to give a total volume of 322.5µL with 10µL set aside for later analysis (pre-digestion sample). To each sample, 36µL 10X *Rsa* I restriction buffer and 1.5µL *Rsa* I were added and the samples incubated at 37ºC for three hours. Another 10µL aliquot (post-digestion sample) was removed and the rest of the sample stored at 4ºC while the pre- and post-digestion samples were run on a 1.2% agarose gel. The gel confirmed the *Rsa* I digestion had worked so the reactions were terminated by the addition of 8µL 0.5M EDTA. The digested cDNA was cleaned up using the QIAquick®

PCR purification kit (Qiagen), eluting into 30µL, and the concentration determined using a NanoDrop® ND-1000 spectrophotometer. The more concentrated CVL sample was diluted using nuclease free water to the same concentration as the MOTRI sample, namely 276ng/µL, which was deemed suitable for the subtraction experiment.

6.2.4: Subtraction reaction

A one-way subtraction was carried out with the MOTRI *Rsa* I digested cDNA as the tester sample and the CVL *Rsa* I digested cDNA as the driver sample. The aim of this subtraction was to identify sequences which exhibited increased expression in the MOTRI sample as compared to the CVL sample. The kit used for this experiment was the BD PCR-Select cDNA Subtraction Kit (Clontech), following the manufacturer's protocol. A diagram of the subtraction experiment is shown in Figure 6.1.

Firstly, the tester cDNA underwent adaptor ligation; 1µL of the tester cDNA was diluted in 5µL sterile water and a master mix set up by combining 9µL sterile water, 6µL 5X Ligation Buffer and 3µL T4 DNA Ligase. Two reactions were set up as follows; for tester1-1, 2µL diluted tester cDNA were mixed with 2µL 10mM Adaptor 1 and 6µL master mix whilst for tester 1-2, 2µL diluted tester cDNA were mixed with 2µL 10mM Adaptor 2R and 6µL master mix. For the unsubtracted tester control, 1-c, 2µL of tester 1-1 and 2µL of tester 1-2 were combined. The three tubes (1-1, 1-2 and 1-c) were incubated overnight at 16ºC. The samples were stored at -20ºC whilst ligation efficiency analysis was carried out. To stop and inactivate the ligation step, 1µL 0.2M EDTA/ 1mg/mL glycogen mix was added to each tube, mixed and the samples heated at 72ºC for 5 mins.

Ligation efficiency was evaluated by PCR amplification of the *T. circumcincta* β-tubulin housekeeping gene. This gene was chosen as the amplicons obtained using the β-tubulin forward and reverse primers (Table 2.1) did not contain any *Rsa* I digestion sites. The master mix for the reactions was set up by combining 92.5µL sterile water, 12.5µL 10X PCR reaction buffer, 2.5µL 10mM dNTP mix and 2.5µL 50X BD Advantage cDNA Polymerase Mix. 1µL of each of tester 1-1 and tester 1-2 was diluted in 200µl nuclease

free water and four PCR reactions set up. In reaction tube 1, 1µl each of diluted Tester 1-1, β-tubulin reverse primer and PCR Primer 1 was mixed with 22µL master mix. In reaction tube 2, 1µl each of diluted Tester 1-1, β-tubulin reverse primer and β-tubulin forward primer was mixed with 22µL master mix. In reaction tube 3, 1µl each of diluted Tester 1-2, β-tubulin reverse primer and PCR Primer 1 was mixed with 22µL master mix. Finally, in reaction tube 4, 1µl each of diluted Tester 1-2, β-tubulin reverse primer and β-tubulin forward primer was mixed with 22µL master mix. The four reaction tubes were subjected to PCR on an Applied Biosystems 2720 thermal cycler with the following reaction conditions; 94ºC for 30 secs followed by 25 cycles of 94ºC for 10 secs, 59ºC for 30 secs and 68ºC for 2 mins 30 secs. 5µL of each of the four PCR reactions was analysed on a 2% agarose gel.

The first hybridisation between the tester samples (1-1 and 1-2) and the driver (CVL) cDNA was carried out by combining the two pools of DNA. In hybridisation tube 1, 1.5µL of the *Rsa* I-digested driver (CVL) cDNA from Section 6.2.3 above was mixed with 1.5µL tester 1-1 and 1µL 4X hybridisation buffers which had previously been incubated at room temperature for 15 mins. In hybridisation tube 2, 1.5µL of the *Rsa* I-digested driver (CVL) cDNA from Section 6.2.3 above was mixed with 1.5µL tester 1-2 and 1µL 4X hybridisation buffer which had previously been incubated at room temperature for 15 mins. Both samples were overlaid with a drop of mineral oil and incubated on an Applied Biosystems 2720 thermal cycler at 98ºC for 1 min 30 secs followed by 8 hours at 68ºC.

The second hybridisation involved the addition of fresh denatured driver cDNA to the combined reaction mixes of tubes 1 and 2. To do this, fresh driver cDNA (CVL from Section 6.2.3) was firstly denatured by combining 1µL of the driver cDNA with 1µL of room temperature 4X Hybridisation buffer and 2µL of sterile water. This was overlaid with a drop of mineral oil and incubated on a thermal cycler at 98ºC for 1 min 30 secs. A pipette and pipette tips were warmed to 45ºC and the sample in hybridisation tube 2 drawn up into the pipette tip followed by a small amount of air and then the freshly denatured driver cDNA. Finally, the contents of the pipette tip were transferred into hybridisation tube 1, mixed and incubated in an Applied Biosystems 2720 thermal cycler set at 68ºC

overnight. These steps were carried out as quickly as possible, removing the tubes only briefly from the thermal cyclers set at 68ºC to prevent the denatured cDNA from re-annealing. After overnight incubation, 200µL pre-warmed dilution buffer was added to the sample and mixed; the sample was then heated at 68ºC for 7 mins and stored at -20ºC.

Amplification of cDNA was carried out by a two-step nested PCR. In the first PCR reaction, a 1µL aliquot of the subtracted cDNA and a 1µL aliquot of the 1-c sample (from the adaptor ligation step, which had been diluted in 1mL of water) were placed into separate tubes. 24µL of the master mix, made by mixing 58.5µL sterile water with 7.5µL of 10X PCR reaction buffer, 1.5µL 10mM dNTP mix, 3µL 10µM PCR primer 1 and 1.5µL 50X Advantage cDNA Polymerase mix, was added to each reaction tube. The reaction conditions for the PCR were; 75ºC for 5 mins, 94ºC for 25 secs followed by 27 cycles of 94ºC for 10 secs, 66ºC for 30 secs and 72ºC for 1 min 30 secs. On completion of the PCR, 8µL aliquots were taken from each reaction tube for analysis, and a further 3µL of each of the amplified cDNAs diluted in 27µL sterile water. 1µL of the resulting diluted cDNAs were transferred to a fresh PCR tube ready for the second PCR reaction and the rest of the cDNA stored at -20ºC. The second master mix was made by mixing 55.5µL sterile water, 7.5µL 10X PCR reaction buffer, 1.5µL 10mM dNTP mix, 3µL 10µM Nested PCR primer 1, 3µL 10µM Nested PCR primer 2R and 1.5µL 50X Advantage cDNA Polymerase mix. 24µL of the master mix was aliquoted into the two PCR reaction tubes which were subjected to 11 cycles of 94ºC for 10 secs, 68ºC for 30 secs and 72ºC for 1 min 30 secs in an Applied Biosystems 2720 thermal cycler. 8µL of the amplified cDNA was removed and analysed on a 2% agarose gel alongside the earlier samples. The final PCR product was cleaned up using the QIAquick® PCR purification kit (Qiagen), eluting in a final volume of 30µL and stored at -20ºC.

6.2.5: Cloning and sequencing

Cloning of the SSH products from the previous section was carried out using JM109 competent cells (Stratagene) and the pGEM®-T vector system (Promega), as described in Sections 2.3.7 and 2.3.8. Initially, 10 white colonies were selected from each of two plates of transformed cells, set up from the ligation mix containing the purified

hybridised cDNA. These were grown up in glass universals containing 10mL LB broth and 20µL 25mg/mL AMP, the plasmids purified and sent for sequencing using the T7 plasmid primer, as described in Section 2.3.9. Subsequently, a further 10 white colonies from each of the two plates were cloned and sequenced.

6.2.6: Analysis of results

Analysis of the sequence data was carried out using SeqMan (DNASTAR Lasergene version 8) to enable the accuracy of base calls to be checked, to align sequences into contigs and to remove plasmid and adaptor sequence. Consensus sequences were used to perform a tBLASTx search on NCBI (www.ncbi.nlm.nih.gov/blast/Blast.cgi) and a BLASTn search on EMBL-EBI (www.ebi.ac.uk/blast2/parasites.html). Consensus sequences were BLAST searched, using the local BLAST server, against the Roche 454 sequencing dataset (Chapter 5) to determine if any of the sequences with constitutive expression differences between MOTRI and CVL also exhibited inducible expression differences under IVM exposure *in vitro*.

6.2.7: Confirmation of results using semi-quantitative PCR

Specific primers for six of the sequences obtained were designed to allow semi-quantitative PCR to be carried out to confirm the validity of the SSH results. The protocol was as follows: 32µL of nuclease free water was combined with 5µL 10X PCR buffer, 1.5µL 50mM $MgCl_2$, 2µL dNTP, 0·5µL Platinum® *Taq* DNA Polymerase, 2µL each of the forward and reverse primers, and 5µL template cDNA to give a total reaction volume of 50µL. The template for the PCR reaction was either CVL and MOTRI cDNA, synthesised from the original RNA generated for the SSH experiment, using random primers as described in Section 2.3.2, diluted to 10ng/µL, or the purified *Rsa* I digested CVL and MOTRI cDNA from Section 6.2.3 which had been diluted to 10ng/µL. The PCR reaction conditions were 5 mins at 94ºC followed by 30 cycles of 94ºC for 30 secs, 58ºC for 30 secs and 72ºC for 30 secs, followed by 10 mins at 72ºC and a hold at 4ºC. After 15, 20, 25 cycles and on completion of the PCR, 5µL of the reaction mix was removed from each tube and stored at 4ºC prior to all the samples being visualised on a 1% agarose gel as described previously. B tubulin primers (Table 2.1) were used as a positive control and the

primers used to amplify the SSH products are shown in Table 6.1. The protocols described above required the following reagents which were not supplied as parts of the kits used:

Agarose gel

2% agarose gel was prepared by dissolving 8g agarose and 8mL 50X TAE in 392mL of dH_2O using a microwave. 20µL 10000X Gel Red (Biotium Inc) was added to enable visualisation of DNA under UV light.

1.2% agarose gel was prepared by dissolving 4.8g agarose and 8mL 50X TAE in 392mL of dH_2O using a microwave. 20µL 10000X Gel Red (Biotium Inc) was added to enable visualisation of DNA under UV light.

0.5M Ethylenediaminetetraacetic acid (EDTA)

18.61g EDTA was diluted in water to a final volume of 100mL.

1M TrisHCl, pH8.0

12.11g TrisHCl was diluted in water to a final volume of 100mL. The pH was adjusted to pH8.0 using sodium hydroxide.

TE buffer (10mM Tris [pH8.0], 0.1mM EDTA)

1mL of the 1M TrisHCl and 20µl of the 0.5M EDTA was diluted to give a total volume of 100mL.

Other reagents required during the SSH experiment, such as TAE buffer and cloning reagents have been previously described in Chapter 2.

6.3: Results

6.3.1: Genes identified using SSH

On completion of the subtraction experiment and the cloning of the resultant PCR product, a total of 40 colonies were sequenced. One colony (2g) failed the sequencing quality control checks and was excluded from any further analysis. The results of the sequence analysis, as described in Section 6.2.6, are shown in Table 6.2 below. Of the 39 remaining colonies, one (colony 2k) was found, upon sequence analysis, to not contain an insert in the plasmid whilst the remaining thirty-eight aligned into 28 contigs ranging in size from 215 to 742 bp in length. The contigs were named SSH1 to 28, sequentially as they were identified. The majority of contigs contained a single sequence, with the maximum number of sequences within a contig being 4. The EBI and NCBI best BLAST hits for 12 of the 28 contigs gave similar predicted sequence identities while a further 5 contigs did not have a definitive BLAST identity, being labelled as predicted or hypothetical proteins or the BLAST identities were not sufficiently detailed (e.g. SSH11 in Table 6.2). Of 28 contigs, 3 were identified as heat shock protein (HSP) homologues, 2 of these were HSP-16 and the third HSP-20. Three of the contigs had putative BLAST identities of ribosomal proteins whilst a further 2 contigs had putative BLAST identities of *T. circumcincta* mitochondrial sequences.

6.3.2: Confirmation of results using semi-quantitative PCR

Using the primers in Table 6.1 and the semi-quantitative PCR protocol, as described in Section 6.2.7 above, the SSH results for a selection of the contigs were confirmed. The contigs were chosen on the basis of their putative BLAST identities and were as follows: SSH3 (axonemal dynein), SSH4 (HSP-16), SSH6 (SXC1), SSH7 (HSP-16), SSH12 (HSP-20) and SSH17 (Cathepsin B). The expression of these genes was compared to the control gene, β-tubulin (Table 2.1). The amplification of SSH 6, 7 and 12 was very weak and for SSH17 not visible at all after completion of the first semi-quantitative PCR, so a second semi-quantitative PCR was set up, with the number of cycles extended to 35 cycles. The results of the semi-quantitative PCR experiments are shown in Figure 6.2. For SSH 3 and 4, the MOTRI sample showed a greater band intensity compared to the CVL sample at 20 cycles whilst even after 35 cycles and using *Rsa* I-

digested cDNA as template, no band could be detected in either the MOTRI or CVL sample for SSH17. The first bands visible for SSH 6 and 12 both appeared in the MOTRI sample lane after 30 cycles, whilst the difference in intensity of the two bands present at 30 cycles for SSH7 indicates that this gene was more highly expressed in the MOTRI sample compared to the CVL sample. Therefore, with the exception of SSH17, where no expression of the gene could be demonstrated, the semi-quantitative PCR results confirm the SSH results, indicating that for SSH 3, 4, 6, 7, and 12, there is increased expression in the MOTRI *T. circumcincta* isolate compared to the CVL isolate.

6.3.3: Comparison of SSH results with Roche 454 sequencing results (Chapter 5).

BLAST searches of the nucleotide sequences obtained from the SSH analysis were made against the contigs generated using the Roche 454 sequencing platform (Chapter 5) to determine whether any of the sequences exhibiting increased expression in MOTRI also exhibited an altered expression profile following *in vitro* IVM exposure of the MOTRI isolate. Once the top BLAST hit (or hits if more than 1 shared the top E score) was determined, the identity of the cluster containing that contig was made and the number of exposed and unexposed reads in that cluster determined by searching the cluster database generated in Chapter 5. Table 6.3 shows the results of this analysis; from this it is possible to see that the E scores for the BLAST searches of SSH 3, 8, 15, 16, 21, 22, 24 and 27 against the 454 dataset were very poor and so these SSH sequences are not discussed further. The identities for the top BLAST cluster hit in the 454 dataset for the remaining SSH contigs was generally very similar to the BLAST hits generated when searching the SSH nucleotide sequences against the EBI and NCBI databases. The exception to this was SSH7, which did not align to a HSP as the top 454 cluster BLAST hit; however, other clusters with similar E scores to the top cluster BLAST hit were identified as HSPs. SSH14, when BLAST searched against the 454 dataset, aligned most closely to cluster 538, which was one of the 16 clusters in Chapter 5 with a statistically significant difference in gene expression in exposed and unexposed reads in the conservative statistical model. The results for this gene indicate that it exhibits both inducibly and constitutively increased expression in response to IVM. SSH 5, 6, 10 and 23 aligned most closely to clusters in the 454 dataset which were statistically significant under the liberal statistical model (clusters 464, 674, 2274 and 494, respectively), indicating that the genes represented by these SSH sequences could exhibit constitutive and, potentially, inducible differences in expression in

response to IVM. However, despite all SSH sequences showing increased expression in MOTRI compared to CVL, 2 (SSH 6 and 10) showed decreased numbers of exposed reads compared to unexposed reads in the Roche 454 dataset. Of the remaining SSH sequences, 13 (SSH 1, 2, 4, 7, 11, 12, 17, 18, 19, 20, 25, 26 and 28) had top BLAST identities of clusters exhibiting non statistically significant increases in number of exposed compared to unexposed reads whilst 1 top cluster BLAST identity (SSH13) showed no change in number of reads and a further 1 top cluster BLAST identity (SSH9) showed a non statistically significant decrease in number of exposed compared to unexposed reads.

6.4: Discussion

The SSH method has identified a panel of 28 sequences exhibiting constitutively increased expression in the MOTRI *T. circumcincta* isolate compared to the CVL isolate. BLAST searches in both EBI and NCBI have given putative identities for these sequences. In some cases, more than one sequence had the same BLAST identity (e.g. SSH 4 and 7), this could be an artefact of the SSH process; cDNA is digested using *Rsa* I so the two sequences may be non-overlapping parts of the same gene. However, *Rsa* I digestion is recommended as longer cDNA fragments may not hybridise as effectively and smaller cDNA fragments also provide a better representation of the genes (Diatchenko *et al.*, 1996). The sequence generated was BLAST searched against the 454 datasets also generated in this thesis (Chapter 5) to determine whether any constitutive changes in gene expression between MOTRI and CVL *T. circumcincta* adults were mirrored by inducible changes in expression as determined by the statistical analysis of the 454 dataset. Encouragingly, many of the SSH sequences aligned very closely to sequences generated using the 454 sequencing platform, suggesting (particularly in the cases of SSH 5, 6, 10, 14 and 23, which aligned most closely to clusters identified as exhibiting statistically significant differences in the 454 dataset) that both inducible and constitutive changes in gene expression could play a role in the IVM-resistance phenotype in MOTRI *T. circumcincta*. Unfortunately, the function of SSH14 (and cluster 538) is not known and, due to the significance in the change observed in the 454 study, alongside the current observation, this gene (and others identified in this SSH work) could merit further investigation to determine the role its expression plays in the IVM-resistant phenotype in *T. circumcincta*.

The accuracy of the SSH results was confirmed using semi-quantitative PCR, this is an important step because, if the tester and driver samples are very similar (as is potentially the case when comparing two adult stage *T. circumcincta* isolates), the SSH technique is more prone to generating false positives (where a sequence is selectively subtracted despite being present in equal proportions in both the tester and driver samples). The results for SSH 3, 4, 6, 7 and 12 were all confirmed by semi-quantitative PCR. Unfortunately, amplification of SSH17 failed despite using the *Rsa* I-digested cDNA, which had already been amplified as part of the SSH protocol and, therefore, should have been enriched for rare transcripts. Confirmation of the differential expression of SSH17 is still required. The confirmatory results for SSH 3, 4, 6, 7 and 12 suggests that the other gene changes here may also be reliable, however, these results would have to be confirmed by real-time or semi-quantitative PCR before any further conclusions are made. These genes were chosen for semi-quantitative PCR partially based on their BLAST identities and merit further discussion.

SSH 4, 7 and 12 had BLAST identities of HSPs; these are a large gene family present in a range of species, including parasitic nematodes. HSPs are believed to act as molecular chaperones, binding and stabilising proteins during folding, assembly, transport across membranes and degradation (Vercauteren *et al.*, 2006). HSPs are both constitutively expressed in a range of life-cycle stages and induced in response to heat shock and exposure to heavy metals, organic compounds and oxidants (Hartman *et al.*, 2003; Vercauteren *et al.*, 2006). Constitutively increased expression of these genes in a *T. circumcincta* isolate resistant to IVM could enable these parasites to more efficiently manufacture and transport other proteins required to metabolise and handle the anthelmintics. Alternatively, it could be indicative of a generalised up-regulation of genes associated with the parasite's stress response, a preconditioning of the resistant parasites enabling them to cope with the anthelmintics more easily and rapidly.

SSH3 was also chosen for confirmation of the SSH result with semi-quantitative PCR because of its apparent homology with an intermediate chain 1 axonemal dynein; however, this identity needs to be used cautiously due to the low E score of the NCBI BLAST hit and also because the best EBI and 454 dataset BLAST hits did not support this

BLAST identity. However, dynein may have a role to play in the IVM-resistant phenotype and, as such, merits further discussion. Dynein has been shown to be required for the correct structure and function of amphidial neurones which are located on either side of the pharynx in nematodes (Wicks *et al.*, 2000; Freeman *et al.*, 2003). Mutations in dynein genes, such as CHE-3, have caused disruption in the structure of these amphidial neurones leading to disrupted chemotaxis and reduced uptake of lipophilic dyes (Wicks *et al.*, 2000; Freeman *et al.*, 2003; Bisset, 2007). IVM is a lipophilic drug and disrupted amphidial neurones have been associated with IVM-resistant isolates in *C. elegans* and *H. contortus*, however, this phenomenon appears to be more associated with cytoplasmic dynein rather than axonemal dynein (the BLAST identity of SSH3), which appears to have a role in the beating of flagella and cilia (Wicks *et al.*, 2000; Freeman *et al.*, 2003; Lespine *et al.*, 2007; Bisset, 2007).

Six cysteine (SXC) motif-containing genes (the putative BLAST identity of SSH6) are short amino acid sequences with six conserved potentially disulphide-bonded cysteines. SXCs are suggested to act as signalling ligands or are involved in protein-protein interactions (Blaxter, 1998; De Maere *et al.*, 2002). They can exist in several forms, as genes containing only the SXC motif, as mucin-like genes or with an enzymatic domain flanked by SXC motifs. Those with enzymatic domains have been identified as tyrosinases, myeloperoxidases and astacin-like zinc metalloproteases (Blaxter, 1998; De Maere *et al.*, 2002). SXC genes were first identified in *Toxocara canis* in surface coat proteins and are part of the small secreted protein class; as such they are seen as a potential vaccine candidate (Daub *et al.*, 2000; De Maere *et al.*, 2002). Why IVM-resistant *T. circumcincta* adults exhibit constitutively increased expression of an *O. ostertagi* SXC1 homologue is not clear but, like the HSPs, this observation could be linked to the role of SXC1 in protein interactions and signalling pathways, allowing the parasites to respond more rapidly to IVM exposure.

Finally, cathepsin B (the putative identity of SSH17) is a peptidase and part of the C1 family of the CA clan of cysteine proteases (Ranjit *et al.*, 2008; Atkinson, Babbitt, & Sajid, 2009). The functions of cysteine proteases include protein processing and turnover, degradation of host protein (for example, extracellular dermal matrix during skin

penetration by *Necator americanus*), hydrolysis of haemoglobin (particularly in blood feeders like *H. contortus* and *N. americanus*) and inhibition of the host protective immune responses (Ranjit *et al.*, 2008). Nisbet *et al* (2008) identified a cathepsin B-like sequence which was highly represented in the L_4-specific SSH dataset, whilst immunoscreening of a *H. contortus* cDNA expression library with the protective S3 TSBP fraction identified three cathepsin B homologues which localised to the gut microvilli (Skuce *et al.*, 1999). Like the SXC genes, cathepsin B molecules are seen as good potential vaccine targets (Knox & Smith, 2001; Ranjit *et al.*, 2008). Again, why genes with potentially immunogenic and protein interaction roles exhibit increased expression in the IVM-resistant MOTRI isolate is not clear but potentially could be a result of a generalised increase in the turnover of all proteins enabling the parasites to cope better with IVM exposure.

Greater depth, and potential accuracy of the sequences generated could have been achieved through re-sequencing of the forty selected clones, whilst selecting further colonies from the plates or repeated cloning of the SSH PCR product could have enabled the identification of further sequences exhibiting increased expression in the MOTRI isolate compared to the CVL isolate. However, the amount of novel sequence generated would have become increasingly small as more sequencing was carried out. As a result, the number of sequences selected was capped at forty. In comparison, a two-way SSH experiment looking at differential gene expression between xL_3 and L_4 *T. circumcincta* generated 361 and 472 unique sequences, respectively (Nisbet *et al.*, 2008). This result illustrates the scale of gene expression changes associated with the transition from the free-living xL_3 stage to the parasitic L_4 stage (Nisbet *et al.*, 2008), in comparison with the more subtle gene expression changes associated with an IVM resistance phenotype in the adult stage.

One of the challenges in making comparisons between different isolates of the same species of parasite which have been identified from different geographical regions is proving whether any differences observed are due to the different genetic backgrounds of the isolates or are due to the differences in their drug resistance status. The effective size of a parasite population and the rate at which genes flow among sub-populations

determines the population genetic structure (Blouin *et al.*, 1995). Population genetic studies of *T. circumcincta*, using mitochondrial sequence data and microsatellite markers, have indicated that the majority of genetic diversity in *T. circumcincta* is found within the populations rather than between populations (Blouin *et al.*, 1995; Grillo *et al.*, 2007). Animal movement determines, to a certain extent, how much gene flow occurs in the parasite population they harbour and is seen as a major risk factor in the development and spread of anthelmintic resistance in the UK (Skuce *et al.*, 2010). The population genetic structure of *T. circumcincta* suggests that the comparison of different isolates, not sharing a common genetic background, is scientifically valid. As such, identifying constitutively expressed differences between the MOTRI and CVL isolates, utilising the "non-hypothesis" driven SSH approach, has enabled the identification of 28 gene fragments which could form the basis of further investigations into how IVM resistant *T. circumcincta* is able to survive anthelmintic exposure.

These results could be investigated further, through increasing the amount of sequence generated from this SSH experiment or by carrying out the SSH experiment in the opposite direction (CVL as the tester and MOTRI as the driver) to identify genes exhibiting constitutive decreased expression in the IVM-resistant MOTRI isolate. Utilising the real-time PCR approach, as used for the candidate resistance genes, the Pgps and CYPs, in Chapter 4, could provide relative quantification of the expression of the panel of 28 sequences identified in this chapter in MOTRI. The close homology of some of the sequences identified in this SSH dataset with sequences generated using the 454 sequencing approach also indicates that the "non-hypothesis" driven approach is a worthwhile addition to understanding the role of gene expression in IVM resistance. The SSH genes also exhibiting statistically significant changes in inducible gene expression (SSH 5, 6, 10, 14 and 23) would be the priority for any future studies.

Table 6.1: Primers used to amplify *T. circumcincta* SSH products from the *Rsa* I digested cDNA.

Primer name	**Primer direction**	**Sequence**	**Melting temp (°C)**	**Target**	**Expected size of PCR product (bp)**
SSH3 For	Sense	CTA CTC CGA TGC CGT TAG GA	59.4	SSH3	127
SSH3 Rev	Antisense	TCC AGA TCG GGA CAA TTG AG	57.3		
SSH4 For	Sense	TTT TCT GGT AGG GTC CAT CG	57.3	SSH4	202
SSH4 Rev	Antisense	AAG CGC TAA GCG GAC AAT TA	55.3		
SSH6 For	Sense	CAG TGG CAA CCG GTA ACA AT	57.3	SSH6	191
SSH6 Rev	Antisense	CAA TTA GCC TTC GCA AGA CC	57.3		
SSH7 For	Sense	TGG ATT GGG ATG GTT CTA GC	57.3	SSH7	200
SSH7 Rev	Antisense	GGC CAA ATC CTT ACC ATT GA	55.3		
SSH12 For	Sense	CAA GCA GGA GCA CAA AAC TG	57.3	SSH12	194
SSH12 Rev	Antisense	TCT GTG GAG CAG CCA CTA TG	59.4		
SSH17 For	Sense	CAA TGG CTC GAC CGT TAG TG	59.4	SSH17	191
SSH17 Rev	Antisense	GAC GGC AGC AGT CCT TTT TA	57.3		

Table 6.2: Results of cloning and sequencing of the PCR product generated from a one-way suppressive subtractive hybridisation of CVL from MOTRI *T. circumcincta* adults.

Contig name	**Contig length (bp)**	**No. seqs**	**Best EBI BLAST hit**	**EMBL acc no.**	**E score**	**Best NCBI tBLASTx hit**	**NCBI acc no.**	**E Score**
SSH1	322	1	*Ostertagia ostertagi* L4 pAMP1 v1 cDNA 5' similar to SW:KAG1_CAEEL Q10454 probable arginine kinase	BM896907	1.4e-60	*Caenorhabditis briggsae* Hypothetical protein CBG11000	XM_002644962.1	7E-63
SSH2	302	2	*Necator americanus* L3 cDNA clone Na_L3_40B03 5' similar to F57B10.3 CE11302 phosphoglycerate mutase status	BU666597	1.20E-36	*Brugia malayi* 2,3-bisphosphoglycerate-independent phosphoglycerate mutase partial mRNA	XM_001892376.1	2E-52
SSH3	215	3	*Steinernema feltiae* IS6 - desiccation stress related ESTs cDNA clone S24-3	BQ579856	1.00E-08	Predicted: *Taeniopygia guttata* similar to dynein, axonemal, intermediate chain 1	XM_002189680.1	1.1
SSH4	679	4	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_32D04 5' similar to P06582 Heat shock protein HSP16-2. *Caenorhabditis elegans*	CB037306	2.00E-89	*Angiostrongylus cantonensis* mRNA for putative Heat Shock Protein (hsp-16.1b)	FM207717.1	7E-24
SSH5	401	2	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_24G11 5' similar to O62337 R06C1.4 protein. *Caenorhabditis elegans*	CB036677	2.80E-45	*Caenorhabditis briggsae* Hypothetical protein CBG18692	XM_002640484.1	3E-22
SSH6	296	1	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_36B08 5' similar to Q9GNW3 SXC1 protein *Ostertagia ostertagi*	CB037620	4.00E-53	*Ostertagia ostertagi* partial mRNA for SXC1 protein	AJ302944.1	1E-13
SSH7	430	1	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_32F06 5' similar to P06582 Heat shock protein HSP16-2. *Caenorhabditis elegans*	CB037332	7.2e-44	*Angiostrongylus cantonensis* mRNA for putative Heat Shock Protein (hsp-16.1b gene)	FM207717.1	9E-11

Contig name	Contig length (bp)	No. seqs	Best EBI BLAST hit	EMBL acc no.	E score	Best NCBI tBLASTx hit	NCBI acc no.	E Score
SSH8	429	1	*Brugia malayi* L3 subtracted cDNA library clone SWBmL3SA163	AA933193	0.21	*Brugia malayi* Phosphatidylinositol 3- and 4-kinase family protein partial mRNA	XM_001901052.1	7E-54
SSH9	452	1	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_50B08 5'	CB038796	1.8e-91	*Drosophila mojavensis* GI14783 mRNA	XM_002010235.1	0.00001
SSH10	449	1	*Strongyloides ratti* L2 pAMP1 v1 Chiapelli McCarter cDNA 5' similar to 60S acidic ribosomal protein P1	BI073294	7.20E-09	*Streptococcus equi* subsp. equi 4047, complete genome	FM204883.1	0.8
SSH11	462	1	*Pristionchus pacificus* mixed stage SL2b TOPO cDNA 5'	CN656930	1.40E-07	*Homo sapiens* 3 BAC RP11-143P4 (Roswell Park Cancer Institute Human BAC Library) complete sequence	AC092966.6	0.09
SSH12	536	1	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_44E08 5' similar to Q07160 Heat shock protein homolog (HSP20) *Nippostrongylus*	CB038331	1.1e-82	*Ostertagia ostertagi* mRNA for heat shock protein 20	AJ310811.2	1E-42
SSH13	597	1	*Caenorhabditis briggsae* mRNA for cyclophilin 4 isoform	AJ004826	1.7e-46	*Caenorhabditis elegans* cyclophilin MOG-6	AF421146.1	9E-76
SSH14	485	2	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_02B09 5' similar to AAB81938.1 non-functional folate binding protein - *Homo sapiens*	BM052010	1.2e-97	*Trichostrongylus colubriformis* 18S rRNA gene	AJ920350.1	2E-85
SSH15	423	1	*Teladorsagia circumcincta* L3 library cDNA clone Tc_L3_18A02 5' similar to T27E9.2 CE14265 ubiquinol-cytochrome c reductase complex	CB035418	2.4e-52	*Caenorhabditis elegans* hypothetical protein (T27E9.2)	NM_067380.3	2E-23

Contig name	Contig length (bp)	No. seqs	Best EBI BLAST hit	EMBL acc no.	E score	Best NCBI tBLASTx hit	NCBI acc no.	E Score
SSH16	248	1	*Caenorhabditis briggsae* contig cb25.fpc2220	CAAC01000044	0.94	Pig DNA sequence from clone CH242-240L11 on chromosome 13	CU468995.15	2.1
SSH17	248	1	*Ostertagia ostertagi* partial mRNA for putative cathepsin B.4 (catB.4 gene)	AJ296147	6.10E-21	*Ostertagia ostertagi* partial mRNA for cathepsin B.2	AJ401373.2	1.00E-36
SSH18	300	1	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_29H04 5'	CB037096	1.80E-58	*Teladorsagia circumcincta* mitochondrion	GQ888720.1	2.00E-41
SSH19	303	1	*Teladorsagia circumcincta* adults library 2 cDNA clone Tc_ad2_40H03 5' similar to P53014 Myosin, essential light chain	CB038013	1.5e-60	*Caenorhabditis briggsae* CBR-MLC-3 protein	XM_002646977.1	4.00E-43
SSH20	263	1	*Nippostrongylus brasiliensis* uni-zap adult library cDNA clone Nb_ad1_06A11 5'	BM279355	9.90E-28	*Teladorsagia circumcincta* mitochondrion	GQ888720.1	7.00E-33
SSH21	521	1	*Caenorhabditis bri*ggsae contig cb25.fpc0091	CAAC01000022	1.6e-10	*Caenorhabditis briggsae* CBR-MIG-1 protein	XM_002638736.1	2.00E-57
SSH22	489	1	*Teladorsagia circumcincta* L4 library cDNA clone Tc_L4_18E09 5',	CB043721	7.0e-33	*Hirschia baltica* ATCC 49814	CP001678.1	0.26
SSH23	409	1	*Haemonchus contortus* cDNA clone Hc_ad_41B01 5',	CB012573	7.4e-12	*Dictyostelium discoideum* AX4 actin binding protein	XM_641035.1	8.00E-06
SSH24	742	1	*Ostertagia ostertagi* L4 SL1 TOPO v1 cDNA similar to keratins in a glycine-rich region	BQ100330	7.5e-78	*Ostertagia ostertagi* partial mRNA for keratin	AJ429146.1	5.00E-56
SSH25	656	1	*Ostertagia ostertagi* L3 SL1 TOPO v2 cDNA 5' similar to SW:RHOA_CAEEL Q22038 RAS-like GTP-binding protein RHOA	BQ625968	7.4e-50	*Caenorhabditis briggsae* CBR-RHO-1 protein	XM_002632397.1	3.00E-20

Contig name	**Contig length (bp)**	**No. seqs**	**Best EBI BLAST hit**	**EMBL acc no.**	**E score**	**Best NCBI tBLASTx hit**	**NCBI acc no.**	**E Score**
SSH26	241	2	*Haemonchus contortus* cDNA clone Hc_L4_24H05 5' similar to F28D1.7 CE05747 locus:rps-23 ribosomal protein S23	CA994729	2.20E-47	*Caenorhabditis elegans* Ribosomal Protein, Small subunit family	NM_069964.4	4.00E-32
SSH27	295	2	*Meloidogyne chitwoodi* female SL1 pGEM cDNA 5' contains PTR5 repetitive element	CF801743	3.50E-06	*Mus musculus* BAC clone RP23-377L13 from chromosome 17	AC134908.5	1.7
SSH28	231	1	*Haemonchus contortus* cDNA clone Hc_L4_02B08 5' similar to NP_001021.1 ribosomal protein S27 (metallopanstimulin 1)	BF186770	9.30E-44	*Ancylostoma duodenale* ribosomal protein S27e mRNA	EF490130.1	3.00E-27

Table 6.3: Table showing the top BLAST cluster hits when the nucleotide sequences for each of the SSH contigs was BLAST searched against the contigs generated in the Roche 454 sequence analysis (Chapter 5). ∞ represents an E score where the value was so small it registered as 0.00 (to 2 decimal places). The change in cluster reads indicates whether there were more unexposed (-) or exposed (+) or equal (=) reads in the cluster, whilst the statistical significance indicates whether the change in number of reads in that cluster was statistically significant under the conservative statistical model (1), statistically significant under the liberal statistical model (2) or not statistically significant under either statistical model (3). If more than one cluster has an identical E score as the best BLAST hit for a particular SSH sequence, then all the clusters with that E score are shown.

SSH sequence	Best BLAST hit (Cluster number)	E score	NCBI Cluster best BLAST hit	Change in cluster reads	Statistical significance
SSH1	215	∞	Hypothetical protein F46H5.3 *Caenorhabditis elegans*	+	3
SSH2	1007	e^-115	Independent phosphoglycerate mutase *Onchocerca volvulus*	+	3
SSH3	634	0.84	S60004 hypothetical protein - common roundworm retrotransposon R4	+	3
SSH4	680	e^-123	Heat shock protein 20 *Haemonchus contortus*	+	3
SSH5	464	e^-177	Hypothetical protein CBG18692 *Caenorhabditis briggsae* AF16	+	2
SSH6	674	e^-104	SXC1 protein *Ostertagia ostertagi*	-	2
SSH7	110	e^-128	Hypothetical protein CBG18371 *Caenorhabditis briggsae* AF16	+	3
SSH8	1756	1.70	No hits found	+	3
	1720	1.70	Hypothetical protein Bm1_02090 *Brugia malayi*	+	3
	918	1.70	eXPOrtin (nuclear export receptor) family member (xpo-1)	+	2
SSH9	575	∞	GH24252 *Drosophila grimshawi*	-	3
SSH10	2274	e^-109	No hits found	-	2
SSH11	136	e^-124	Hypothetical protein TGME49_048900 *Toxoplasma gondii*	+	3
SSH12	131	e^-148	Heat shock protein 20 Ostertagia ostertagi	+	3
SSH13	152	e^-130	CYclophyliN family member (cyn-4) *Caenorhabditis elegans*	=	3
SSH14	538	∞	Hypothetical protein Bm1_17870 *Brugia malayi*	+	1
SSH15	537	0.43	Pre-mRNA cleavage complex II protein Clp1 *Brugia malayi*	+	2
SSH16	369	0.25	Hypothetical protein F40A3.3 *Caenorhabditis elegans*	+	3
SSH17	201	8e^-60	Cathepsin B-like cysteine proteinase 1	+	3
SSH18	10	8e^-39	NADH dehydrogenase subu3t 4L Necator americanus	+	3

SSH sequence	Best BLAST hit (Cluster number)	E score	NCBI Cluster best BLAST hit	Change in cluster reads	Statistical significance
SSH19	88	e^-147	C. briggsae CBR-MLC-3 protein *Caenorhabditis briggsae*	+	3
SSH20	10	4e^-28	NADH dehydrogenase subunit 4L Necator americanus	+	3
SSH21	121	0.14	Putative zinc metallopeptidase precursor; MEP1b *Haemonchus*	-	3
	654	0.14	Hypothetical protein CBG20630 *Caenorhabditis briggsae* AF16	-	3
SSH22	2140	0.50	Integrin alpha cytoplasmic region family protein *Brugia malayi*	-	3
	349	0.50	No hits found	-	3
SSH23	494	∝	No hits found	+	2
SSH24	100	0.77	Hypothetical protein CBG03794 *Caenorhabditis briggsae* AF16	+	3
	114	0.77	Hypothetical protein Y66H1B.2 *Caenorhabditis elegans*	+	3
SSH25	1160	∝	RAS-like GTP-binding protein RhoA *Brugia malayi*	+	3
SSH26	1649	e^-117	Ribosomal Protein, Small subunit family member (rps-23)	+	3
SSH27	399	1.20	Hypothetical protein CBG22942 *Caenorhabditis briggsae* AF16	+	3
	1744	1.20	No hits found	+	3
	768	1.20	Hypothetical protein *Plasmodium yoelii yoelii* str. 17XNL	+	3
SSH28	391	e^-123	Ribosomal protein S27e *Ancylostoma duodenale*	+	3

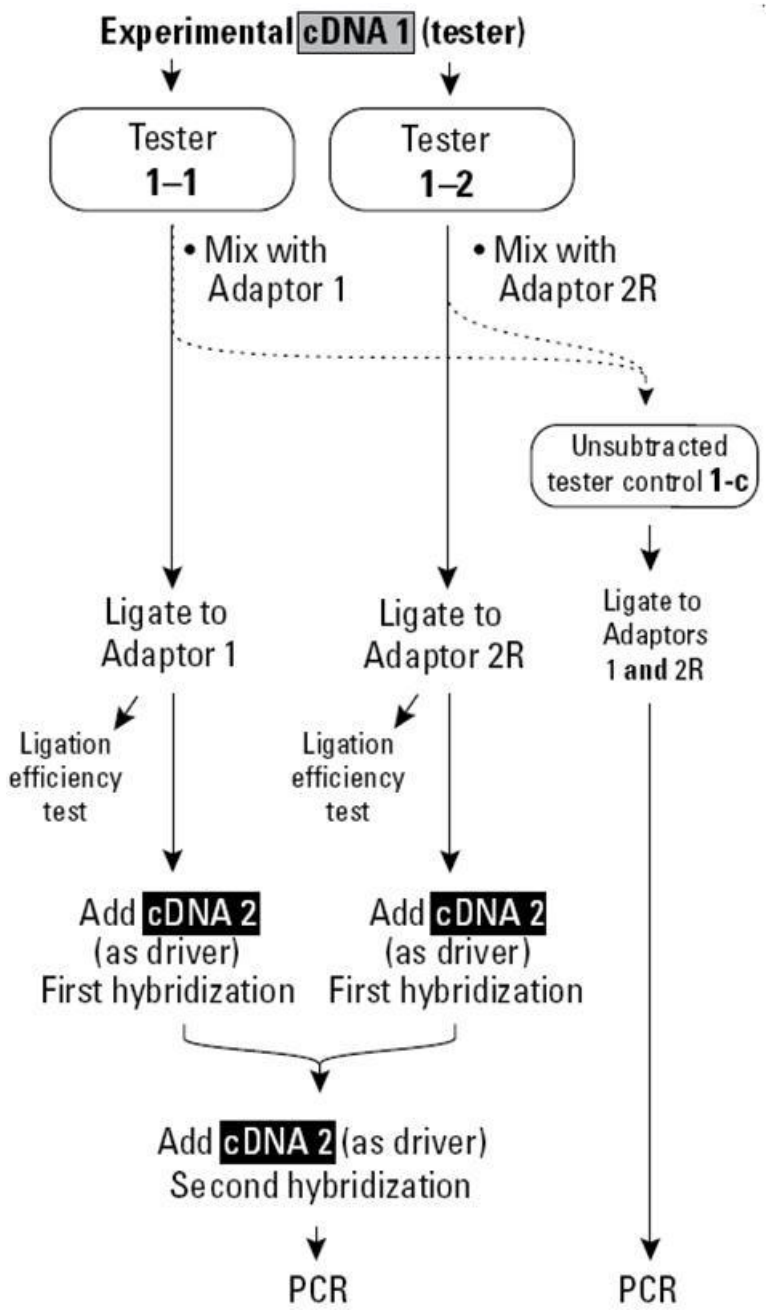


Figure 6.1: Diagram showing the steps involved in the subtraction reaction. cDNA 1 (tester) was MOTRI and cDNA 2 (driver) was *T. circumcincta* CVL. Both the driver and tester cDNA was *Rsa* I digested, as described in Section 6.2.3 prior to the subtraction experiment. Diagram modified from the BD PCR-select™ cDNA subtraction kit user manual (http://www.clontech.com/images/pt/PT1117-1.pdf).

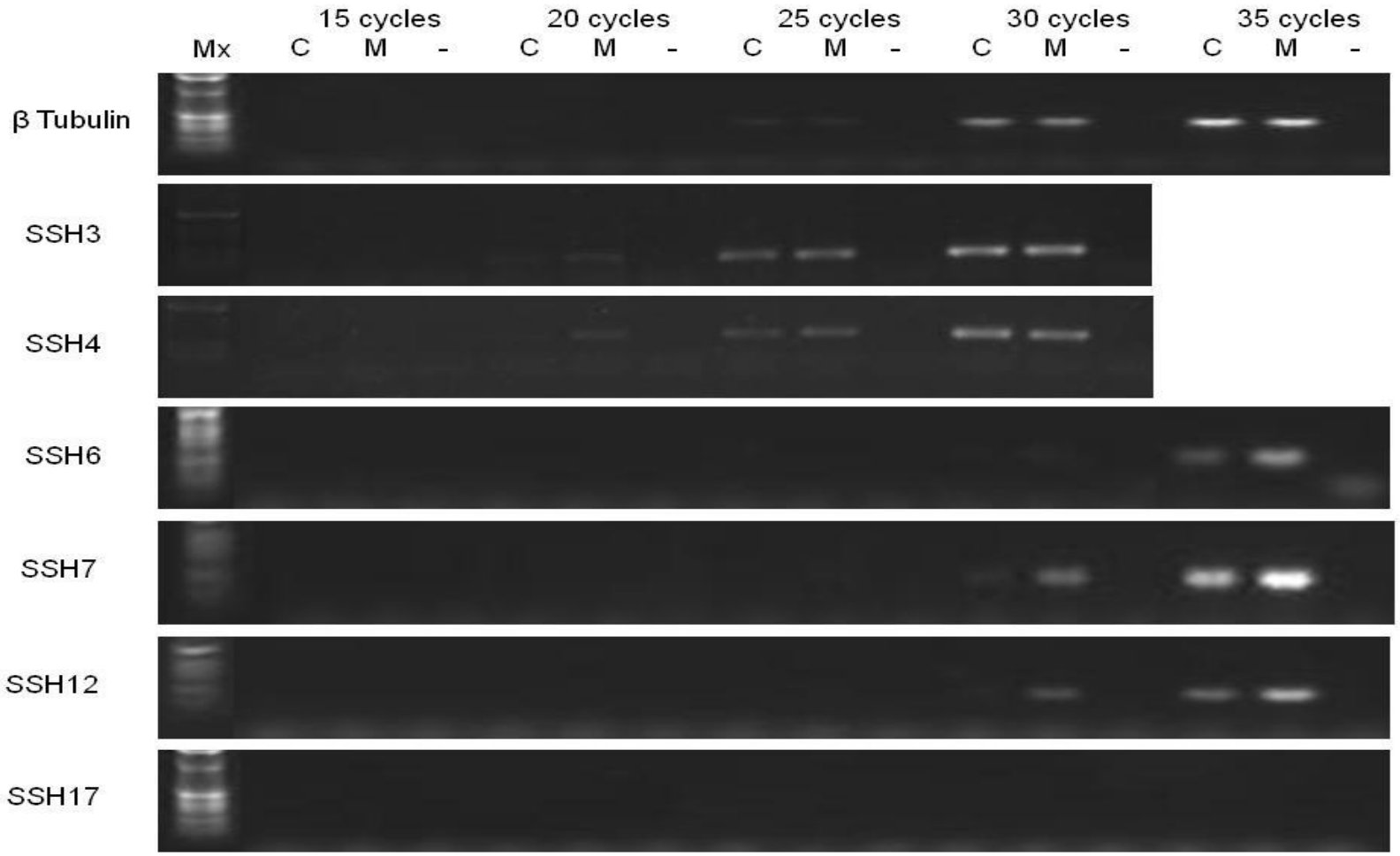


Figure 6.2: 1% agarose gel electrophoresis of semi-quantitative PCR products for SSHs 3, 4, 6, 7, 12 and 17 against the T. circumcincta CVL (C) or MOTRI (M) isolate. 5µL PCR product was removed after 15, 20, 25, 30 and (for SSHs 6, 7, 12 and 17) 35 cycles. A negative control (-) was included for each gene and the gel marker lane (Mx) containing TrackIt™ 1Kb Plus DNA Ladder (Invitrogen) is on the left.

Chapter 7: General Discussion and Conclusions

This study examined the potential role of changes in gene expression in the anthelmintic resistant phenotype of *T. circumcincta*, particularly in relation to IVM resistance. To understand the role of gene expression in the IVM-resistance phenotype of *T. circumcincta*, a range of experimental approaches were employed. Identification of the expression profile of "candidate resistance" genes potentially involved in drug handling and metabolism in different isolates of *T. circumcincta*, both IVM-resistant and –susceptible, and IVM-exposed and -unexposed, was carried out using semi-quantitative and real-time PCR as described in Chapters 3 and 4, respectively. Alongside this, two further experiments were carried out, as described in Chapters 5 and 6, utilising the Roche 454 sequencing platform to investigate the expression of genes following *in vitro* IVM exposure, and SSH to identify genes with constitutively increased expression in the MOTRI isolate compared to the CVL isolate. These two approaches were termed "non-hypothesis" driven as neither technique attempted to focus on particular genes prior to carrying out the experiments.

Anthelmintics are the mainstay of modern parasitic nematode control and, in some areas of the world, the levels of resistance to the BZs, LEVs and MLs exhibited by endoparasites of small ruminants is so high that economically viable farming is becoming impossible (Jackson & Coop, 2000; Sargison *et al.*, 2005; Blake & Coles, 2007). In the UK, BZ resistance is widespread, ML resistance is becoming increasingly common and MDR parasite isolates are being reported (Bartley *et al.*, 2003; Bartley *et al.*, 2006; Sargison *et al.*, 2007b). No routine surveillance is carried out, so the levels of resistance exhibited by *T. circumcincta*, the predominant parasitic nematode species on UK farms, to the different classes of anthelmintics is not known (Sargison, Scott, & Jackson, 2001; Bartley *et al.*, 2003). The detection of resistance is problematic; current diagnostic tests, using *in vivo* and *in vitro* methods such as the FECRT, EHT and LFIT, can be time-consuming, expensive and lacking in sensitivity (Coles *et al.*, 1992; McKellar & Jackson, 2004). Alongside this, the modes of action of the anthelmintics and mechanisms of resistance employed by the parasites they target are still not fully understood, as described in Chapter 1. Naturally, the genes believed to be the targets of the anthelmintics have been prioritised for candidate gene studies investigating the mechanisms of anthelmintic

resistance. However, to date, this approach has not been very successful and, in the case of IVM, has failed to identify any SNPs or mutations conclusively linked to resistance. As such, it was decided that alternative mechanisms of resistance were worthy of investigation. One alternative mechanism of resistance could be changes in the expression of genes, for example, generic drug handling mechanisms allowing *T. circumcincta* to survive exposure to IVM. These changes would still have a genetic basis, which, if identified, could be developed into a molecular-based test for diagnosis of IVM resistance in parasitic nematodes. Genetic tests have advantages over the current, commonly used, diagnostic tests, being quicker and easier to perform, more sensitive and able to be developed into high-throughput formats (von Samson-Himmelstjerna, 2006; von Samson-Himmelstjerna *et al.*, 2009a). Molecular markers for ML resistance, developed into a diagnostic test, would enable the spread of anthelmintic resistance to be monitored more easily and allow the development of management practices such as the TST approach, which could reduce the inevitable spread of anthelmintic resistance, to be investigated more thoroughly (Kenyon *et al.*, 2009b).

In the first series of experiments, described in Chapters 3 and 4, two families of genes in *T. circumcincta* were investigated; the Pgps and CYPs. These were chosen for investigation due to their possible roles in drug transport and metabolism: Pgps are large transmembrane proteins which transport a range of hydrophobic molecules into and out of cells (Valverde *et al.*, 1992; Kerboeuf *et al.*, 2003). Pgps have been implicated in drug resistance in human immunodeficiency viruses, malaria parasites and cancerous tumours whilst, in parasitic nematodes, a range of experiments have shown that changes in mRNA levels and allele frequencies occur in parasites in response to IVM selection and that co-administration of Pgp inhibitors with IVM can increase the efficacy of IVM against phenotypically IVM-resistant parasites isolates (Beugnet, Gauthey, & Kerboeuf, 1997; Xu *et al.*, 1998; Blackhall *et al.*, 1998b; Loo & Clarke, 1999; Jones & George, 2005; Ardelli, Guerriero, & Prichard, 2006b). However, the majority of work on Pgps in parasitic nematodes has focused on *H. contortus* and *O. volvulus*, no work on the role of Pgps in the anthelmintic resistant phenotype in *T. circumcincta* has been found in peer-reviewed journals to date. Investigations into the role of Pgps in anthelmintic resistance in *T. circumcincta* may be hampered by the lack of genomic sequence for this species in publically available databases. In Chapter 3 of this thesis, eleven novel partial Pgp gene

sequences from *T. circumcincta* were identified, eight of which appear to align to NBD1 indicating that there are at least eight Pgp genes in *T. circumcincta*. This is similar to what has been identified in other nematode species; for example, in *O. volvulus* at least 2 Pgp genes have been identified, at least 12 from *H. contortus* and 15 from *C. elegans* (Huang & Prichard, 1999; Sangster *et al.*, 1999a; Sheps *et al.*, 2004). Alignment of the *T. circumcincta* Pgp sequences against the *C. elegans* Pgp sequences was carried out to determine the correct nomenclature for these genes, as suggested in the guidelines on the *Caenorhabditis* Genetics Centre website (http://www.cbs.umn.edu/CGC/nomenclature/index.html).

Eukaryotic CYPs are single domain membrane-bound proteins that catalyse the metabolism of a range of predominantly hydrophobic molecules, such as IVM (Mansuy, 1998; Graham & Peterson, 1999). Like the Pgps, the CYPs appear to play a role in drug resistance, as over-expression of CYPs or increased CYP activity has been shown to correlate with insecticide resistance in *Drosophila* spp, *An. gambiae*, and *C. p. pallens* and chloroquine resistance in *P. falciparum* and *P. berghei* (Ndifor, Ward, & Howells, 1990; Daborn *et al.*, 2002; Vontas *et al.*, 2005; Gong *et al.*, 2005). In parasitic nematodes, there is some evidence that CYPs play a role in anthelmintic resistance, however, CYP activity in parasitic nematodes appears to be difficult to measure (Barrett, 1998). The use of CYP inhibitors, such as piperonyl butoxide, ketoconazole and metyrapone has been shown to improve the efficacy of the BZs (Kotze, 1997; McKellar & Jackson, 2004; Bartley *et al.*, 2009; Virkel *et al.*, 2009). Changes in "non-specific" drug transport affecting one drug class could also have an effect with another drug class. In Chapter 3, three novel CYPs were identified from *T. circumcincta*, using both degenerate PCR and bioinformatic approaches. Assigning the correct nomenclature to these partial genes was not attempted; the CYP superfamily is very large and the naming of any newly discovered CYPs is carried out through the P450 Nomenclature Committee, following the recommendations in Nelson *et al.* (1996), and also requires full length gene sequences.

The first step in determining whether any of the genes identified in Chapter 3 could play a role in an IVM-resistant phenotype in *T. circumcincta* involved the design of semi-quantitative PCR assays. Validation of the assays for PGPs 2 (subsequently named

TeciPgp-2 NBD2), 3, 5 (subsequently named *TeciPgp-2* NBD1), 6 (subsequently named *TeciPgp-9* NBD2), 7 and 9 and CYPs 1, 2 and 3 were successful. The results of the PCRs, comparing the expression of these genes in the egg, L_1, xL_3, L_4 and adult stages of the CVL and MOTRI *T. circumcincta* isolates, indicated that there were differences in the expression levels of some of these genes between the isolates (Figures 3.3 and 3.13). As a result, it was decided that quantifying the expression of the 11 Pgp and 3 CYP genes using real-time PCR would give a more accurate result compared to the semi-quantitative PCR approach; analysis of the real-time results would determine what changes in expression were statistically significant, something that was not possible using semi-quantitative PCR. Smaller changes in expression are more easily quantified using real-time PCR compared to identifying changes in the intensity of bands on agarose gels; real-time PCR is also faster to perform than standard PCR (Pfaffl, 2001; Bustin *et al.*, 2009).

Using the relative quantitative, or ΔΔCt, real-time PCR method, the relative expression of PGP1, *TeciPgp-2* NBD2, PGP3, *TeciPgp-2* NBD1, *TeciPgp-9* NBD2, PGP7, PGP8, PGP9, PGP10, PGP11, CYP1, CYP2 and CYP3 was determined in several isolates and following exposure of some of these isolates to IVM *in vitro* and *in vivo*. One concern when comparing isolates which do not share a common genetic background (e.g. CVL and MOTRI) was that any differences in expression found are a result of the inherent genetic variability between the isolates. However, population genetic studies of *T. circumcincta* have indicated that the majority of genetic diversity occurs within subpopulations rather than between subpopulations (Blouin *et al.*, 1995; Grillo *et al.*, 2007). Comparing the actin control gene Ct values for the different isolate and life-cycle stages (Figure 4.3) indicated that, with the exception of the MOTRI L_4 and Post IVM MOTRI L_4 comparison, the actin Ct values were sufficiently similar between isolates to allow valid comparisons of the test genes. Likewise, comparing the actin Ct values between the different life-cycle stages also indicated that comparing expression of the test genes between life-cycle stages was not feasible; normalisation of the test gene Ct values with the actin Ct values would not give accurate results. However, although academically interesting, determining the relative expression of the 3 CYP and 10 validated Pgp genes between life-cycle stages was not part of this work. In future experiments, a different approach to normalising the Pgp and CYP expression levels would be recommended. The GeNorm approach starts with a panel of reference genes from which the two most stably expressed genes are selected. This

approach, which has only relatively recently been adopted as the "gold-standard" method, assumes that stably transcribed genes stay in a constant ratio to each other and provide a better baseline to normalise the test gene expression with (Van Zeveren *et al.*, 2007b; Strube *et al.*, 2008; Bustin *et al.*, 2009).

In the first real-time experiment, constitutive relative gene expression was investigated by comparing the CVL and MOTRI isolates at rest; the same comparison as was carried out using semi-quantitative PCR. When comparing the real-time and semi-quantitative data, the majority of the results were in agreement, with the exception of *TeciPgp-2* NBD2 where no amplification across all life-cycle stages (Chapter 3) was observed using semi-quantitative PCR but amplification of this gene was observed using real-time PCR. This reflects the relative sensitivities of the respective methods. As shown in Chapter 4, no statistically significant changes in expression level were found for *TeciPgp-2* NBD1, PGP7 and PGP8 when comparing CVL and MOTRI; statistically significant differences in expression level of PGP1, PGP3, PGP9, PGP10, PGP11, CYP1, CYP2 and CYP3 in individual life-cycle stages were observed when comparing CVL and MOTRI; but, more importantly, changes in expression across all life-cycle stages were observed for *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2. For *TeciPgp*-2 NBD2 there was a statistically significant reduction in expression in the IVM-resistant MOTRI isolate compared to the IVM-susceptible CVL isolate, whereas, for *TeciPgp*-9 NBD2, there was a statistically significant increase in expression in the MOTRI isolate compared to the CVL isolate. Biological replicates for xL_3 and adult stages of the CVL and MOTRI isolates were analysed and, with the exception of a non-statistically significant result for *TeciPgp-9* NBD2 in the adult stage, confirmed the findings of the original triplicate real-time experiments. Changes across all life-cycle stages could indicate that genetic changes such as altered copy number underlie the expression differences observed, as opposed to alterations in the regulation of the gene which would be anticipated to only occur in the life-cycle stages exposed to IVM.

Interestingly, despite the phylogenetic analysis of the partial Pgp sequences indicating that two halves of the same, *TeciPgp-2*, gene had been identified, the real-time PCR results for *TeciPgp-2* NBD1 and 2 did not follow the same pattern, as shown in

Figure 4.5. Further work is required to investigate this; potentially the phylogentic analysis and subsequent re-naming of the Pgp sequences could have erroneously classified these partial sequences as the two separate NBDs of the same gene. Unfortunately, the relative expression levels of *TeciPgp-9* NBD1 could not be determined as the probe and primers designed did not amplify *TeciPgp-9* NBD1 efficiently. Therefore, it was not possible to determine whether the two partial *TeciPgp-9* gene fragments (NBD1 and NBD2) exhibited the same expression pattern. The sequencing of the full-length coding sequences of *TeciPgp-2* and *TeciPgp*-9 would conclusively determine whether the nomenclature for these genes is correct and would be a priority now that these genes have been shown to exhibit statistically significant changes in expression when comparing an IVM-resistant and -susceptible *T. circumcincta* isolate.

It has become apparent, by comparing nucleotide sequences, that *TeciPgp-9* NBD2 is the same gene as identified by Bisset (2007) in New Zealand from laboratory derived near-isogenic inbred *T. circumcincta* lines. The pattern of allelic, or haplotype, variation of *TeciPgp -1*, *-2* and *-9* was compared between the drug-susceptible and -resistant lines. Analysis of genomic sequence from individual adult male worms showed that five IBDA (equivalent to NBD1) and two IBDB (equivalent to NBD2) haplotypes in the New Zealand *TeciPgp-9* were under positive selection pressure in the MDR line and of these 7 haplotypes, four IBDA and one IBDB were exclusive to the MDR line. Alongside this, Bisset (2007) showed a large number of SNPs in the New Zealand *TeciPgp-9* gene, which, with the exception of four, were non-coding or silent. The finding of changes in expression levels of *TeciPgp-9* NBD2 in UK field-derived isolates becomes even more significant in the light of these independent findings from laboratory-derived isolates from the other side of the world. In this present study, we did not attempt to identify whether haplotypes of *TeciPgp-9* were present as the full-length genomic sequence was not available. However, as shown in Chapter 3, the presence of at least four silent SNPs (A138G, C177T, A201T and A216C) in *TeciPgp-9* NBD2 in the UK field-derived, unrelated isolates has been shown; unfortunately none of the sequence data obtained from the UK isolates for *TeciPgp-9* NBD2 covers the regions identified in the New Zealand near-isogenic lines containing the four coding SNPs. In work currently on-going, the sequence of *TeciPgp-9* NBD2 is being extended through 5’ and 3’ RACE PCR and has

shown that at least one of the coding SNPs, identified from the New Zealand near-isogenic lines, has been identified in the UK CVL and MOTRI isolates (F. Turball, Pers. Comm.).

In the New Zealand study, the expression level of *TeciPgp-9* IBDA in individual adult male worms was determined using SYBR® green real-time PCR. Using the ΔΔCt method, there was an average of a 3.4 fold increase in *TeciPgp-9* in the resistant worms compared to the susceptible worms (Bisset, 2007). In the present study, using pools of mixed sex adult *T. circumcincta* to generate cDNA, a 6.75 fold increase in expression of *TeciPgp-9* NBD2 ($P<0.01$) was observed when comparing CVL to MOTRI, whilst the increase in expression for this gene in the other life-cycle stages was 55.27 fold ($P<0.01$), 5.06 fold ($P<0.05$), 17.49 fold ($P<0.01$) and 14.04 fold ($P<0.05$) for eggs, L_1, xL_3 and L_4, respectively. As such, the change in expression of *TeciPgp-9* NBD2 in IVM-resistant UK isolates, compared to IVM-susceptible UK isolates is greater than the change in expression observed between the New Zealand isolates. By comparing the results of the SYBR® green real-time PCR with the haplotype data for the individual worms, Bisset (2007) was able to determine that the likely cause of the increase in abundance was an increase in copy number of *TeciPgp-9* as the MDR worms (exhibiting a higher expression level of *TeciPgp-9*) possessed 3 or 4 different haplotypes of *TeciPgp-9* IBDA compared to at most 2 haplotypes in the susceptible worms. In fact, the presence of different 'resistance-associated' haplotypes altered the expression of *TeciPgp-9*. For example, the presence of IBDA haplotype 10 in resistant worms was related to a 10 fold increase in abundance of *TeciPgp-9* template in the resistant worms compared to susceptible worms (Bisset, 2007). Genes believed to be orthologous to *TeciPgp-9* have been studied previously; the gene designated as *Hcpgp-1* is implicated in ML resistance in *H. contortus* and ML selection was shown to increase the frequency of *Hcpgp-1* in a ML-resistant *H. contortus* isolate (Kwa *et al.*, 1998; Le Jambre, Lenane, & Wardrop, 1999; Bisset, 2007). In *C. elegans*, *Pgp-9* has been shown to be expressed in the intestine and the first and second bulbs of the pharynx (Zhao *et al.*, 2004) (http://wormbase.sanger.ac.uk/).

The identification of reduced *TeciPgp-2* NBD2 expression across all life-cycle stages, when comparing CVL and MOTRI, was unexpected; the role of Pgps as generic detoxification pumps suggested that increased expression of these genes would be

expected to produce an IVM-resistant phenotype in *T. circumcincta* (Sangster, 1994; Kerboeuf, Guegnard, & Le Vern, 2003). However, drug resistance could also be mediated by reduced uptake of the drug, drug sequestration and/ or by preventing the drug from reaching its target site (Wolstenholme *et al.*, 2004; Jones & George, 2005). *TeciPgp-2* NBD2 could have a high affinity for IVM and so, by reducing expression of this gene (and its gene products) parasites are able to reduce the amount of IVM transported into or around their tissues. Although identified in the New Zealand study, no real-time quantification of this gene (*Tecipgp-2* IBDB) was carried out so it is not possible to determine whether reduced expression of this gene is also found in New Zealand. Haplotype data for *Tecipgp-2* IBDB is available, however, indicating that seven haplotypes are present in the New Zealand near-isogenic lines of *T. circumcincta*, one of which show reduced frequencies in the resistant line whilst another was found to be exclusively present in the resistant line (Bisset, 2007). It could be possible that the *TeciPgp-2* NBD2 sequence obtained in this project is specific for a haplotype which is found more commonly in the IVM-susceptible (CVL) isolate than the IVM-resistant (MOTRI) isolate thus, *TeciPgp-2* NBD2 could represent a potential marker for IVM susceptibility. The PgpA gene from *H. contortus* is believed to be orthologous to *C. elegans* and *T. circumcincta Pgp-2*; changes in *H. contortus* PgpA allele frequencies are associated with IVM and MOX resistance and this gene is over-expressed in IVM-selected *H. contortus* (Xu *et al.*, 1998; Blackhall *et al.*, 1998b; Bisset, 2007). Interestingly, reduced or disrupted expression of *C. elegans Pgp*-2, which is localised to the gut granule membrane and the first and second bulbs of the pharynx, is associated with impaired fat storage (Zhao *et al.*, 2004; Nunes *et al.*, 2005; Schroeder *et al.*, 2007); in eggs of *H. contortus* cholesterol depletion increases TBZ resistance (Riou *et al.*, 2003). It appears that Pgp function, and hence drug resistance, is affected by the concentration of lipids (such as cholesterol) in the membrane within which it is localised (Riou, Koch, & Kerboeuf, 2005; Nunes *et al.*, 2005; Schroeder *et al.*, 2007). As such, like *TeciPgp-9* NBD2, the *TeciPgp-2* NBD2 gene warrants further investigation, particularly to determine why, in *T. circumcincta*, a reduced expression of this gene is associated with IVM-resistance whilst in *H. contortus*, this gene is over-expressed in IVM-selected strains.

The expression of a panel of "candidate resistance" genes was also investigated in *T. circumcincta* isolates which had been exposed to IVM treatment. This was to determine

what changes in gene expression were evident in an IVM-resistant isolate (MOTRI) upon exposure to IVM. A range of experiments was carried out; firstly the IVM-resistant MOTRI isolate was compared to the Post IVM MOTRI isolate, generated by collecting the survivors of a full therapeutic dose of IVM. The rationale behind this experiment was that MOTRI is known to comprise individuals displaying a range of sensitivities to IVM; treatment would select the most resistant sub-population within MOTRI for further analysis. Due to previously described differences in actin Ct values, no comparison between the MOTRI and Post IVM MOTRI L_4 stages was made. Fewer statistically significant changes in expression between the isolates were observed and no changes in expression of any single gene across all life-cycle stages were observed. Potentially, the MOTRI isolate is so resistant to IVM that exposing it to IVM failed to further select resistant parasites, although this is counter-intuitive as the efficacy of IVM against MOTRI is 60% (Bartley *et al.*, 2004). Alternatively, if parasites do alter expression of resistance-associated genes, such as the Pgps and CYPs, this change could occur more rapidly and only be detectable whilst the parasites are still exposed to the drug. Thus, the timing of IVM exposure and collection of the surviving parasites may be crucial to being able to identify inducible expression differences. To answer this, two further experiments were carried out to investigate inducible changes, as described below.

The LMIT can be used to determine the phenotypic IVM-resistance status of parasite populations (Wagland *et al.*, 1992; Rabel, McGregor, & Douch, 1994). By scaling up the assay, as described in Chapter 2, it was possible to separate IVM-exposed MOTRI *T. circumcincta* into those able to migrate, or not, in the presence of a discriminating dose of IVM. The expression of *TeciPgp-9* NBD2 and *TeciPgp-2* NBD2 were compared in these selected pools of xL_3 and a pool of IVM-unexposed xL_3. Only one significant change in expression of these genes was found; a statistically significant 1.88-fold increase in *TeciPgp-2* NBD2 expression in the unexposed sample compared to the migrators sample ($P<0.05$). Potentially, as described earlier, this could be because the MOTRI isolate is constitutively expressing these genes at different levels and does not alter expression further in response to IVM exposure. This could be the case with *TeciPgp-9* NBD2, if the changes in expression are found to be due to an increase in copy number rather than due to increased transcription levels. Alternatively, the design of the experiment may not have been optimal; for example, the dose of IVM chosen could have been insufficient, the

length of time taken to collect the xL_3 following exposure may have been too long or it might have been better to expose the IVM-susceptible CVL isolate.

As a final attempt to elucidate whether changes in expression of *TeciPgp-9* NBD2 and *TeciPgp-2* NBD2 were inducible, adult *T. circumcincta* MOTRI in a donor sheep were exposed to a full therapeutic dose of IVM and subsequently collected at necropsy three days post-treatment. This time scale was chosen as the half-life of IVM is between 61 and 102 hours, so the adult worms would still be exposed to IVM but any susceptible worms should have been killed and expelled from the abomasum (Canga *et al.*, 2009). Using this approach, it was determined that the expression of *TeciPgp-2* NBD2 was reduced in the MOTRI survivors of IVM treatment and, although not statistically significant, the expression of *TeciPgp-9* NBD2 was increased in the survivors of IVM treatment. As stated previously, this suggests that the correct design of experiments is crucial to identifying inducible gene expression differences. Alongside this, the majority of genes, particularly the three CYPs, did not exhibit significant changes in expression, potentially subtle changes in gene expression could still have an effect on the ability of parasites to survive anthelmintic treatment. For example, in a study of IVM-resistance in *O. ostertagi*, a 3.4 fold increase in expression of PGP2a was found in laboratory-derived IVM-resistant adults compared to IVM-susceptible adults (Van Zeveren, 2009). It is possible that, as the genes were identified in susceptible parasites, and appear to be expressed at low levels, resistance-associated genes may be present in susceptible parasites below the threshold of detection (Barrett, 1998). As such other, as yet unidentified Pgps and CYPs (or previously unknown genes) could also be involved in the anthelmintic resistance phenotype in *T. circumcincta*.

The fact that similar changes in *TeciPgp-9* NBD2 have been observed in *T. circumcincta* isolates from opposite sides of the globe is potentially significant, especially given the fact that the IVM-resistant isolates were generated by two completely different methods. In New Zealand, the multi-drug resistant isolate was derived in the laboratory by inbreeding and back-crossing, with IVM selection at each generation whereas in the UK, MOTRI was naturally selected over many generations by IVM treatment in the field. This suggests that there is some universal commonality in the IVM-resistance mechanisms and

that the polymorphisms and changes in gene expression observed in *TeciPgp-9* NBD2 are potentially significant in the IVM-resistance phenotype in *T. circumcincta*. This study provides further evidence that Pgps, especially *TeciPgp-9* NBD2, are involved in IVM-resistance; increased expression of this Pgp is associated with resistance; anthelmintic resistant *H. contortus* eggs were shown to have increased amounts of Pgps, as measured through the level of fluorescent monoclonal antibody, measured using flow cytometry (Kerboeuf, Guegnard, & Le Vern, 2003). A priority for further studies on *TeciPgp-9* and its role in the IVM resistant phenotype in *T. circumcincta* would be the generation of the full-length coding and genomic sequence from the CVL, MOTRI and Post IVM MOTRI field-derived isolates, respectively. This would enable the determination of the presence of specific resistance-associated alleles, haplotypes or coding SNPs in this gene; as have already been identified in the New Zealand laboratory derived near-isogenic lines. Alongside this, the changes in gene expression found in *TeciPgp-9* NBD2 between the CVL and MOTRI isolates need to be investigated in other isolates with differing levels of IVM-resistance and from different geographic locations. If this was found to be the case, then the changes in *TeciPgp-9* NBD2 could form the basis of a universal molecular test (like those available for the BZs) for IVM resistance in *T. circumcincta*. The generation of a full-length coding sequence for this gene would also help to determine whether the observed changes in gene expression are a result of increased transcription, mediated through a mutation in an upstream regulatory element, for example, or, like in the New Zealand studies, due to an increase in gene copy number. As has been shown with some *in vitro* inhibitor studies, the use of Pgp inhibitors causes phenotypically IVM-resistant *T. circumcincta* and *H. contortus* to shift towards an IVM-susceptible phenotype; the further knowledge gained during this study, suggesting that the Pgps may play a role in IVM resistance in *T. circumcincta*, would indicate that studying the co-administration of these inhibitors with anthelmintics as a method to overcome MDR parasites could be a viable option (Virkel *et al.*, 2009; Bartley *et al.*, 2009). However, the recently released AAD anthelmintic, monepantel, has been shown to be effective against MDR parasites (Kaminsky *et al.*, 2008) and may delay the inevitable loss of ability to control parasites, so expensive animal trials involving Pgp inhibitors may not, at this present time, be commercially viable.

As an alternative method to investigate more global gene expression events in relation to IVM resistance, Roche 454 sequencing was used to generate nearly 100,000 novel sequence reads from two pools of adult *T. circumcincta*, one of which had been exposed to IVM *in vitro* whilst the other was left unexposed as a control. The subsequent objective bioinformatic and statistical analysis of the *in silico* sequence data generated, enabled the sequences to be classified according to the significance of the change in number of reads (reflecting the mRNA level of that sequence in the two samples) and, for many sequences, provided a putative BLAST identity. Sixteen sequence clusters were shown to have the most statistically significant changes in expression between the exposed and unexposed samples and a further 355 were shown to have statistically significant changes in expression under a more liberal statistical model. Further details of these sequences can be found in Chapter 5 and Appendix 5. To provide more information on the putative function of individual contigs, functional classification was assigned using KAAS; from this any statistically significant changes in the proportion of functional groupings was identified, indicating that for the orthologous groups 'translation', 'amino acid metabolism', 'carbohydrate metabolism' and 'xenobiotic degradation and metabolism', statistically significant changes in the mean proportion of reads within these groupings between the two samples was observed. This work is particularly important for organisms like *T. circumcincta* where the amount of sequence data available in the public domain is limited; at present there are only ~6,000 ESTs available on the NCBI dbEST website. There are a further 14 megabases of genomic sequence available, generated through the genome sequencing project, on the Sanger website (http://www.sanger.ac.uk/Projects/Helminths/), but candidate resistance genes are conspicuously absent to date. The available sequence does not provide full coverage of the *T. circumcincta* genome; which has been estimated to be in the order of 59 megabases in size (Leroy, Duperray, & Morand, 2003). The generation of nearly 100,000 novel sequences in this project is a valuable genetic resource for studying changes that occur in *T. circumcincta* in response to IVM exposure.

One criticism of the approach used to generate this 454 sequencing EST dataset is the lack of candidate resistance genes identified, however this could, like the current size of the *T. circumcincta* genome sequencing project be purely due to the depth of coverage achieved. Even with this relatively limited sequencing dataset, the approach would still be

anticipated to identify the major changes in gene expression associated with IVM exposure. Only one contig, containing 2 reads, was identified as a Pgp in the IVM-exposed dataset and the change in expression of this gene was not found to be statistically significant. However, as the approach was designed to be 'non-hypothesis' driven, focusing on finding candidate resistance genes such as the Pgps and CYPs could result in ignoring other, more significant, changes in gene expression which do enable the parasites to cope with IVM exposure. As such, a statistically robust objective analysis of the changes in gene expression enabled a list of potentially interesting genes to be drawn up without any preconceived ideas as to their function in the IVM-resistance phenotype. Mainly for financial reasons, there was only one opportunity to perform this experiment, so the experimental design was selected to provide the simplest comparison and the best possible chance of identifying gene expression changes within an IVM-resistant *T. circumcincta* isolate in response to IVM exposure. The standard therapeutic dose of IVM in sheep is 0.25mg/kg body weight, with a half-life of between 61 and 102 hours (Canga *et al.*, 2009). In one experiment, following intra-ruminal dosing with 0.2mg/kg body weight IVM, the peak plasma concentration was approximately 17.6ng/mL (Prichard *et al.*, 1985b) but how this relates to the concentration *T. circumcincta* is exposed to *in vivo* in the abomasum is not known. At least a proportion of the MOTRI isolate is able to survive a dose of 0.25mg/kg body weight IVM *in vivo*, so an *in vitro* IVM exposure concentration of 5μg/mL was chosen to ensure that the IVM exposed parasites were stressed. This is reflected in the BLAST identies of the most significant changes in read number which included HSPs, which have been shown to increase expression in response to environmental stresses; cytochrome oxidases, which are part of the mitochondrial electron transport chain; and vitellogenin, which is linked to egg production (Lindquist, 1986). However, it is possible that some of the changes in gene expression observed could be due to the presence of dead or dying adult worms within the IVM-exposed sample due to MOTRI comprising individuals with a range of IVM sensitivities, as previously stated. Alternatively, other resistance mechanisms utilised by surviving worms in the IVM-exposed sample could have been masked by the contribution from dead or dying worms. If this experiment was to be carried out again, an alternative approach could be to select the IVM-exposed survivors by either putting the parasites through a phenotypic assay, such as the scaled-up LMIT (Chapter 4), or selecting only adult worms which still appeared to be motile or alive following exposure to IVM. This approach could, however, increase the time that the parasites are exposed to IVM, and potentially cause degradation in the RNA extracted. Selection of parasites in this experiment was not carried out as only a limited

number of worms were exposed to IVM and the priority was to ensure there was sufficient, good quality RNA extracted as rapidly as possible. This was also to ensure that any transient, inducible gene expression changes could be identified.

Roche 454 sequencing is one of the "next-generation" sequencing technologies which have the potential to revolutionise genome sequencing. Compared to traditional Sanger sequencing, these techniques are cheaper, faster and generate considerably more data, although often at the expense of read length, however technical improvements are rapidly beginning to overcome this (Marguiles *et al.*, 2005; Pop & Salzberg, 2008; Shendure & Ji, 2008). A novel sequencing method called real-time DNA sequencing, has recently been described; this could potentially be an improvement on the "next-generation" approaches and could be used to investigate gene expression in genome-poor parasitic nematodes, such as *T. circumcincta*, as the sequence generated is longer compared to the current "next generation" approaches (Eid *et al.*, 2009). The main challenge, however, of producing large EST datasets, such as that described in Chapter 5, is the subsequent bioinformatic and statistical analysis required to generate meaningful results, in order to answer biologically relevant questions. A novel statistical approach has been applied to this dataset and has allowed the identification of a panel of genes showing altered expression profiles in MOTRI in response to *in vitro* IVM exposure; these genes could form the basis of further studies to identify how IVM-resistant parasites survive anthelmintic treatment.

As a second "non-hypothesis" driven approach to investigate IVM-resistance in *T. circumcincta*, a SSH experiment was undertaken (Chapter 6) specifically designed to identify genes exhibiting increased expression in the MOTRI isolate compared to CVL. Cloning of the subsequent products identified a panel of 28 contiguous sequences exhibiting increased expression in the MOTRI isolate. Confirmation of the SSH results for five out of six selected sequences was obtained by carrying out semi-quantitative PCR, the sixth result could not be confirmed as no PCR amplification occurred. These SSH sequences were chosen for further investigation based on their BLAST identities, which included HSPs, SXC1, dynein and cathepsin B. Sequences with HSP homology were also prominent in the Roche 454 sequencing dataset, and HSPs are known to be constitutively

expressed in a range of parasite life-cycle stages and can be induced in response to various physical and chemical challenges (Hartman et al., 2003; Vercauteren et al., 2006). The fact that HSPs have also been identified as exhibiting statistically significant changes in expression in response to IVM exposure, as well as in a comparison between IVM-resistant and -susceptible *T. circumcincta* isolates, suggests that this large gene family is worthy of further investigations. Both SXC1 and Cathepsin B are seen as potential vaccine candidates in parasitic nematodes; if an effective vaccine was to be developed against these genes, it could be even more effective against resistant parasites which have a greater expression level of these genes compared to susceptible parasites (Daub et al., 2000; De Maere et al., 2002; Ranjit et al., 2008). The number of sequences identified through this subtraction experiment is less than those generated by Nisbet *et al* (2008) in a study investigating gene expression associated with the transition from the free-living to parasitic stage in *T. circumcincta*. However, a second round of cloning and sequencing of the SSH products in this experiment indicated that the yield, in terms of novel sequences obtained, was diminishing. Potentially, further rounds of sequencing of the SSH product could identify other sequences exhibiting increased expression in the IVM-resistant isolate compared to the IVM-susceptible but, as a method to identify genes for further study, the SSH experiment has already generated a panel of 28 potential novel IVM-resistance associated genes.

In order to determine whether the two non-candidate approaches had identified any of the same, or similar, genes, the sequences identified using SSH were BLAST searched against the EST dataset generated using Roche 454 sequencing. High BLAST E scores for some of these searches indicated that the same molecules may be present in the SSH dataset as those identified in the Roche 454 dataset. One SSH sequence aligned to a contig which was deemed to be significant in the conservative statistical model of the 454 sequencing analysis, whilst a further 5 SSH sequences aligned to contigs deemed to be significant in the liberal statistical model of the 454 sequencing analysis. The fact that these gene sequences appear to exhibit constitutive expression differences between the CVL and MOTRI *T. circumcincta* isolates, and inducible expression differences following IVM expression suggests that these genes need to be investigated further to determine whether they enable the expression of an IVM resistance phenotype in *T. circumcincta* through some as yet unidentified resistance mechanism.

One concern whilst undertaking this comparison was that the two isolates used (CVL and MOTRI) were isolated independently from different regions of the UK. This could mean that background genetic differences between the two isolates could be important and lead to the generation of spurious results, linked to the different genetic background of the parasites, in the subtraction experiment. However, using mitochondrial sequence data and neutral microsatellite markers, Blouin *et al* (1995) and Grillo *et al* (2007) were able to determine that the majority of genetic diversity in *T. circumcincta* is found within rather than between populations. As such, although it is possible to compare field-derived isolates, the best possible comparison would still be parental versus field-derived isolates, or failing that, through the generation of back-crossed near-isogenic lines, such as those described in Bisset (2007). This would enable any non-resistance associated gene expression changes to be excluded or, at least, minimised.

The SSH method has proven to be a useful tool to identify a panel of putative IVM-resistance associated genes for further study. One criticism of this method is the potential to falsely identify differences in gene expression if the tester and driver cDNA are very similar to each other (A. Nisbet, Pers. Comm.). Confirmation of the results using semi-quantitative PCR is an important control in removing these false positives but whether this could be an issue with the comparison described in Chapter 6 is uncertain as the genetic relationship of MOTRI and CVL is not known. If an alternative comparison was to be made within an isolate, such as that between the migrators and non-migrators in the LMIT experiment, then this tendency of the SSH technique to give false positives could be more of a problem. The panel of genes identified in this SSH experiment provides a starting point for further investigations into the role of altered constitutive expression of genes in IVM-resistant and -susceptible *T. circumcincta* isolates. Following the success of the real-time PCR assays to determine the relative constitutive expression of a panel of Pgps and CYPs in a range of *T. circumcincta* isolates, this would be the preferred method to fully investigate the genes identified through SSH. It would also be interesting to carry out this subtraction experiment in the opposite direction, to identify genes with reduced expression in the MOTRI isolate compared to the CVL isolate. As has been shown using real-time PCR, a statistically significant constitutive reduction in expression across all life-cycle stages of *TeciPgp-2* NBD2 was observed; as such constitutive decreased expression of

other, as yet identified, genes could also play a part in the expression of an IVM resistant phenotype in *T. circumcincta*.

Both the "non-hypothesis" driven Roche 454 sequencing and SSH approaches attempted to identify other genes which could be associated with the IVM-resistant phenotype either through a change in expression in response to IVM exposure or constitutively. The Roche 454 sequencing of IVM-exposed MOTRI adults showed that even a drug resistant isolate is affected by exposure to IVM. A novel method of statistically analysing the results has drawn up a panel of genes which warrant further investigation; caution has to be applied however as the relatively short sequences generated could mean some of the changes observed are false positives due to gene fragmentation. This is also a problem which can also occur following the *Rsa* I digestion used in the SSH experiment. The SSH experiment has identified other genes, not previously investigated in the context of anthelmintic resistance, showing increased expression in MOTRI compared to CVL. Both panels of genes identified using these approaches could form the basis of a novel panel of "candidate resistance" genes; their role in expression of a IVM-resistant phenotype in *T. circumcincta* investigated further using real-time PCR, as carried out for the panel of Pgp and CYP genes in this thesis.

The work described in this thesis has shown that changes in expression, both in genes putatively associated with IVM resistance and other, non-candidate resistance genes, do occur as identified in IVM-resistant and -susceptible *T. circumcincta* isolates and following IVM exposure *in vitro* and *in vivo*. The MDR MOTRI isolate is affected by exposure to IVM and also shows constitutive differences in expression when compared to an unrelated drug susceptible isolate (CVL). Identification of changes in expression associated with IVM resistance could form the basis of a molecular marker for resistance; something which is currently impeding the development of targeted control strategies aimed at minimising the spread of IVM resistance in *T. circumcincta* (von Samson-Himmelstjerna *et al.*, 2009a). Maintenance of the efficacy of anthelmintics can only be achieved if accurate diagnosis of the resistance status of a parasite population can be made (von Samson-Himmelstjerna, 2006). Ultimately, a sensitive and specific molecular diagnostic test for IVM resistance needs to be developed, which can discriminate between

nematode species in natural infections, and identify the IVM-resistance status of each species. The identification of the mechanisms of IVM-resistance, and genetic changes which control it, would be the first step towards this. Significantly, in this project, two partial Pgp genes, namely *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2, have been shown to exhibit a constitutively altered expression profile between IVM-resistant and -susceptible isolates and non-coding SNPs in *TeciPgp-9* NBD2 have been identified when comparing MOTRI and CVL sequences. The results for *TeciPgp-9* NBD2 are even more significant in the light of recent findings from New Zealand, as discussed above.

Important questions remain to be answered before these genes are heralded as potential molecular markers for IVM resistance in *T. circumcincta*: The full-length sequence of both these genes needs to be identified so that any further SNPs (silent or coding) and alleles or haplotypes, associated with IVM-resistance, can be identified. These genetic changes could form the basis of future molecular markers for resistance and would also mean that the cause (either a transcriptional change or a change in gene copy number) of the altered gene expression profiles of these genes could be determined. It would be interesting to see whether the numerous silent SNPs identified in *TeciPgp-9* NBD2 cause an altered protein conformation which affects its ability to bind to substrates, as hypothesised in Kimchi-Sarfaty *et al* (2007); this would require the codon usage in *T. circumcincta* to be ascertained, if it is not already known. At the very least, a silent SNP is still a useful molecular marker for IVM-resistance, provided it is shown to accurately distinguish between resistant and susceptible parasites. Lastly, and most importantly, the changes found in *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2, when comparing CVL and MOTRI, need to be confirmed in other isolates of *T. circumcincta*. There is no point in developing a molecular marker for resistance which is subsequently found to only be present in one or a few isolates. This was the case with the leucine to phenylalanine substitution (Leu256Phe) in the GluClα3 subunit of *C. oncophora* which was subsequently not found in any other parasite species or *C. oncophora* isolates (Njue *et al.*, 2004; Van Zeveren, 2009).

IVM is still the mainstay of many treatment regimes for reducing the burden of PGE in livestock worldwide; it is also used in mass drug treatment programmes for human

parasitic nematodes such as *O. volvulus* (Ardelli, Guerriero, & Prichard, 2006a; Ardelli, Guerriero, & Prichard, 2006b; Prichard & Roulet, 2007; Omura, 2008). As such, the rapid development of IVM resistance, coupled with climate change potentially causing an increase in disease incidence is a worrying development. This is compounded by the lack of rapid, sensitive and cheap diagnostic tests for IVM resistance; the anthelmintic resistance status of parasite populations is generally unknown (Kaplan, 2004; Prichard *et al.*, 2007; Kenyon *et al.*, 2009a). Knowledge of the anthelmintic resistance status of a parasite population and the mechanisms by which the parasites exhibit resistance is a prerequisite for developing strategies for controlling resistant parasites. The change in gene expression observed in this project could be a step forward in determining how nematode parasites, such as *T. circumcincta*, are affected by IVM exposure, and how resistant parasites are able to survive IVM treatment. The changes in *TeciPgp-2* NBD2 and *TeciPgp-9* NBD2 are particularly noteworthy; further work on these genes is required but the changes in expression of these genes could form the basis of a much needed molecular marker for IVM resistance. This would be particularly beneficial in being able to understand the implications, at the genetic level, of control and management practices, such as targeted selective treatments, aimed at slowing down the inevitable spread of IVM resistance (Kenyon *et al.*, 2009b).

References

Abbott, K. A., Taylor, M. A., & Stubbings, L. A. (2004) Sustainable worm control strategies for sheep. SCOPS, www.nationalsheep.org.uk.

Abubucker, S., Zarlenga, D. S., Martin, J., Yin, Y., Wang, Z., McCarter, J. P., Gasbarree, L., Wilson, R. K., & Mitreva, M. (2009) The transcriptomes of the cattle parasitic nematode *Ostertagia ostartagi*. *Veterinary Parasitology*, **162**, (1-2), 89-99.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology*, **215**, (3), 403-410.

Alvarez-Sanchez, M. A., Garcia, J. P., Bartley, D., Jackson, F., & Rojo-Vazquez, F. A. (2005) The larval feeding inhibition assay for the diagnosis of nematode anthelmintic resistance. *Experimental Parasitology*, **110**, (1), 56-61.

Alvinerie, M., Dupuy, J., Kiki-Mvouaka, S., Sutra, J. F., & Lespine, A. (2008) Ketoconazole increases the plasma levels of ivermectin in sheep. *Veterinary Parasitology*, **157**, (1-2), 117-122.

Amichot, M., Brun, A., Cuany, A., De Souza, G., Le Mouel, T., Bride, J. M., Babault, M., Salaun, J. P., Rahmani, R., & Berge, J. B. (1998) Induction of cytochrome P450 activities in *Drosophila melanogaster* strains susceptible or resistant to insecticides. *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology*, **121**, (1-3), 311-319.

Ardelli, B. F., Guerriero, S. B., & Prichard, R. K. (2005) Genomic organization and effects of ivermectin selection on *Onchocerca volvulus* P-glycoprotein. *Molecular and Biochemical Parasitology*, **143**, (1), 58-66.

Ardelli, B. F., Guerriero, S. B., & Prichard, R. K. (2006a) Characterization of a half-size ATP-binding cassette transporter gene which may be a useful marker for ivermectin selection in *Onchocerca volvulus*. *Molecular and Biochemical Parasitology*, **145**, (1), 94-100.

Ardelli, B. F., Guerriero, S. B., & Prichard, R. K. (2006b) Ivermectin imposes selection pressure on P-glycoprotein from *Onchocerca volvulus*: linkage disequilibrium and genotype diversity. *Parasitology*, **132**, 375-386.

Ardelli, B. F., Stitt, L. E., Tompkins, J. B., & Prichard, R. K. (2009) A comparison of the effects of ivermectin and moxidectin on the nematode Caenorhabditis elegans. *Veterinary Parasitology*, **165**, (1-2), 96-108.

Atkinson, H. J., Babbitt, P. C., & Sajid, M. (2009) The global cysteine peptidase landscape in parasites. *Trends in Parasitology*, **25**, (12), 573-581.

Baliharova, V., Skalova, L., Maas, R. F. M., De Vrieze, G., Bull, S., & Fink-Gremmels, J. (2003) The effects of benzimidazole anthelmintics on P4501A in rat hepatocytes and HepG2 cells. *Research in Veterinary Science*, **75**, (1), 61-69.

Barnes, E. H., Dobson, R. J., & Barger, I. A. (1995) Worm control and anthelmintic resistance - Adventures with a model. *Parasitology Today*, **11**, (2), 56-63.

Barnett, C., Hossell, J., Perry, M., Procter, C., & Hughes, G. 2006, *A handbook of climate trends across Scotland. SNIFFER project CC03, Scotland & Northern Ireland Forum for Environmental Research, 62pp*.

Barrett, J. (1998) Cytochrome P450 in parasitic protozoa and helminths. *Comparative Biochemistry and Physiology C-Pharmacology Toxicology & Endocrinology*, **121**, (1-3), 181-183.

Bartley, D. J., Jackson, E., Johnston, K., Coop, R. L., Mitchell, G. B. B., Sales, J., & Jackson, F. (2003) A survey of anthelmintic resistant nematode parasites in Scottish sheep flocks. *Veterinary Parasitology*, **117**, (1-2), 61-71.

Bartley, D. J., Jackson, F., Jackson, E., & Sargison, N. (2004) Characterisation of two triple resistant field isolates of *Teladorsagia* from Scottish lowland sheep farms. *Veterinary Parasitology*, **123**, (3-4), 189-199.

Bartley, D. J., Donnan, A. A., Jackson, E., Sargison, N., Mitchell, G. B. B., & Jackson, F. (2006) A small scale survey of ivermectin resistance in sheep nematodes using the faecal egg count reduction test on samples collected from Scottish sheep. *Veterinary Parasitology*, **137**, (1-2), 112-118.

Bartley, D. J., McAllister, H., Bartley, Y., Dupuy, J., Nez, C., Alvinerie, M., Jackson, F., & Lespine, A. (2009) P-glycoprotein interfering agents potentiate ivermectin susceptibility in ivermectin sensitive and resistant isolates of *Teladorsagia circumcincta* and *Haemonchus contortus*. *Parasitology*, **136**, (09), 1081-1088.

Barton, N. J. (1983) Development of anthelmintic resistance in nematodes from sheep in Australia subjected to different treatment frequencies. *International Journal for Parasitology*, **13**, (2), 125-132.

Bartos, M., Rayes, D., & Bouzat, C. (2006) Molecular determinants of pyrantel selectivity in nicotinic receptors. *Molecular Pharmacology*, **70**, (4), 1307-1318.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the False Discovery Rate - A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological*, **57**, (1), 289-300.

Beraldi, D., Craig, B. H., Bishop, S. C., Hopkins, J., & Pemberton, J. M. (2008) Phenotypic analysis of host-parasite interactions in lambs infected with *Teladorsagia circumcincta*. *International Journal for Parasitology*, **38**, (13), 1567-1577.

Berge, J. B., Feyereisen, R., & Amichot, M. (1998) Cytochrome P450 monooxygenases and insecticide resistance in insects. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **353**, (1376), 1701-1705.

Besier, B. (2009) A novel anthelmintic group: Welcome relief for sheep farmers. *New Zealand Veterinary Journal*, **57**, (1), 1-2.

Beugnet, F., Gauthey, M., & Kerboeuf, D. (1997) Partial in vitro reversal of benzimidazole resistance by the free-living stages of *Haemonchus contortus* with verapamil. *Veterinary Record*, **141**, (22), 575-576.

Bisset, S. A. & Morris, C. A. (1996) Feasibility and implications of breeding sheep for resilience to nematode challenge. *International Journal for Parasitology*, **26**, (8-9), 857-868.

Bisset, S. A. (2007). Genetic basis of multiple-anthelmintic resisitance in *Teladorsagia circumcincta.* Doctor of Philosophy, Flinders University of South Australia.

Blackhall, W. J., Pouliot, J. F., Prichard, R. K., & Beech, R. N. (1998a) *Haemonchus contortus*: Selection at a glutamate-gated chloride channel gene in ivermectin- and moxidectin-selected strains. *Experimental Parasitology*, **90**, (1), 42-48.

Blackhall, W. J., Liu, H. Y., Xu, M., Prichard, R. K., & Beech, R. N. (1998b) Selection at a P-glycoprotein gene in ivermectin- and moxidectin-selected strains of *Haemonchus contortus*. *Molecular and Biochemical Parasitology*, **95**, (2), 193-201.

Blackhall, W. J., Prichard, R. K., & Beech, R. N. (2003) Selection at a γ-aminobutyric acid receptor gene in *Haemonchus contortus* resistant to avermectins/milbemycins. *Molecular and Biochemical Parasitology*, **131**, (2), 137-145.

Blackhall, W. J., Prichard, R. K., & Beech, R. N. (2008) P-glycoprotein selection in strains of *Haemonchus contortus* resistant to benzimidazoles. *Veterinary Parasitology*, **152**, (1-2), 101-107.

Blake, N. & Coles, G. (2007) Flock cull due to anthelmintic-resistant nematodes. *Veterinary Record*, **161**, (1), 36.

Blaxter, M. (1998) *Caenorhabditis elegans* is a nematode. *Science*, **282**, (5396), 2041-2046.

Blouin, M. S., Yowell, C. A., Courtney, C. H., & Dame, J. B. (1995) Host Movement and the Genetic-Structure of Populations of Parasitic Nematodes. *Genetics*, **141**, (3), 1007-1014.

Boag, B. & Thomas, R. J. (1973) Epidemiological studies on gastrointestinal nematode parasites of sheep - Control of infection in lambs on clean pasture. *Research in Veterinary Science*, **14**, (1), 11-20.

Boyne, s., Stott, A. W., & Gunn, G. J. 2006, *Disease priority survey to establish priorities for cattle and sheep disease modelling and cost-benefit studies*.

Britt, D. P. (1982) Benzimidazole resistant nematodes in Britain. *Veterinary Record*, **110**, (14), 343-344.

Broeks, A., Janssen, H. W. R. M., Calafat, J., & Plasterk, R. H. A. (1995) A P-glycoprotein protects *Caenorhabditis elegans* against natural toxins. *Embo Journal*, **14**, (9), 1858-1866.

Broughan, J. M. & Wall, R. (2007) Faecal soiling and gastrointestinal helminth infection in lambs. *International Journal for Parasitology*, **37**, (11), 1255-1268.

Burgess, C., Bennett, M., & Kenyon, F. (2009) Promoting sustainable use of anthelmintics. *The Veterinary Record*, **164**, (14), 412-413.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009) The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, **55**, (4), 611-622.

Canga, A. G., Prieto, A. M. S., Liebana, M. J. D., Martinez, N. F., Vega, M. S., & Vieitez, J. J. G. (2009) The pharmacokinetics and metabolism of ivermectin in domestic animal species. *Veterinary Journal*, **179**, (1), 25-37.

Cawthorne, R. J. G. & Whitehead, J. D. (1983) Isolation of benzimidazole resistant strains of *Ostertagia circumcincta* from British sheep. *Veterinary Record*, **112**, (12), 274-277.

Chen, C. D. & Kemper, B. (1996) Different structural requirements at specific proline residue positions in the conserved proline-rich region of cytochrome P450 2C2. *Journal of Biological Chemistry*, **271**, (45), 28607-28611.

Colditz, I. G., Watson, D. L., Gray, G. D., & Eady, S. J. (1996) Some relationships between age, immune responsiveness and resistance to parasites in ruminants. *International Journal for Parasitology*, **26**, (8-9), 869-877.

Coles, G. C., Bauer, C., Borgsteede, F. H. M., Geerts, S., Klei, T. R., Taylor, M. A., & Waller, P. J. (1992) World Association for the Advancement of Veterinary Parasitology (WAAVP) Methods for the detection of anthelmintic resistance in nematodes of veterinary importance. *Veterinary Parasitology*, **44**, (1-2), 35-44.

Coles, G. C. (2005) Anthelmintic resistance - looking to the future: a UK perspective. *Research in Veterinary Science*, **78**, (2), 99-108.

Coles, G. C., Jackson, F., Pomroy, W. E., Prichard, R. K., von Samson-Himmelstjerna, G., Silvestre, A., Taylor, M. A., & Vercruysse, J. (2006) The detection of anthelmintic resistance in nematodes of veterinary importance. *Veterinary Parasitology*, **136**, (3-4), 167-185.

Cook, A., Aptel, N., Portillo, V., Siney, E., Sihota, R., Holden-Dye, L., & Wolstenholme, A. (2006) *Caenorhabditis elegans* ivermectin receptors regulate locomotor behaviour and are functional orthologues of *Haemonchus contortus* receptors. *Molecular and Biochemical Parasitology*, **147**, (1), 118-125.

Cottee, P. A., Nisbet, A. J., El-Osta, Y. G. A., Webster, T. L., & Gasser, R. B. (2006) Construction of gender-enriched cDNA archives for adult *Oesophagostomum dentatum* by suppressive-subtractive hybridization and a microarray analysis of expressed sequence tags. *Parasitology*, **132**, 691-708.

Daborn, P. J., Yen, J. L., Bogwitz, M. R., Le Goff, G., Feil, E., Jeffers, S., Tijet, N., Perry, T., Heckel, D., Batterham, P., Feyereisen, R., Wilson, T. G., & Ffrench-Constant, R. H. (2002) A single P450 allele associated with insecticide resistance in *Drosophila*. *Science*, **297**, (5590), 2253-2256.

Dalton, J. P. & Mulcahy, G. (2001) Parasite vaccines - a reality? *Veterinary Parasitology*, **98**, (1-3), 149-167.

Datu, B., Gasser, R., Nagaraj, S., Ong, E., O'Donoghue, P., McInnes, R., Ranganathan, S., & Loukas, A. (2008) Transcriptional changes in the hookworm, *Ancylostoma caninum*, during the transition from a free-living to a parasitic larva. *PLoS Neglected Tropical Diseases*, **2**, (1), e130.

Daub, J., Loukas, A., Pritchard, D. I., & Blaxter, M. (2000) A survey of genes expressed in adults of the human hookworm, *Necator americanus*. *Parasitology*, **120**, 171-184.

de Lourdes Mottier, M. & Prichard, R. K. (2008) Genetic analysis of a relationship between macrocyclic lactone and benzimidazole anthelmintic selection on *Haemonchus contortus*. *Pharmacogenet Genomics*, **18**, (2), 129-140.

De Maere, V., Vercauteren, I., Saverwyns, H., Claerebout, E., Berx, G., & Vercruysse, J. (2002) Identification of potential protective antigens of *Ostertagia ostertagi* with local antibody probes. *Parasitology*, **125**, 383-391.

Denham, D. A. (1969) The development of *Ostertagia circumcincta* in lambs. *Journal of Helminthology*, **43**, (3-4), 299-310.

Dent, J. A., Smith, M. M., Vassilatis, D. K., & Avery, L. (2000) The genetics of ivermectin resistance in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, (6), 2674-2679.

Diatchenko, L., Lau, Y. F. C., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D., & Siebert, P. D. (1996) Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, (12), 6025-6030.

Dobson, R. J., Donald, A. D., Waller, P. J., & Snowdon, K. L. (1986) An egg hatch assay for resistance to levamisole in trichostrongyloid nematode parasites. *Veterinary Parasitology*, **19**, (1-2), 77-84.

Dobson, R. J., Le Jambre, L. F., & Gill, J. H. (1996) Management of anthelmintic resistance: Inheritance of resistance and selection with persistent drugs. *International Journal for Parasitology*, **26**, (8-9), 993-1000.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X. X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C. C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J., & Turner, S. (2009) Real-Time DNA Sequencing from Single Polymerase Molecules. *Science*, **323**, (5910), 133-138.

Ekong, R. M., Robson, K. J. H., Baker, D. A., & Warhurst, D. C. (1993) Transcripts of the multidrug resistance genes in chloroquine-sensitive and chloroquine-resistant *Plasmodium falciparum*. *Parasitology*, **106**, 107-115.

Elard, L., Sauve, C., & Humbert, J. F. (1998) Fitness of benzimidazole-resistant and -susceptible worms of *Teladorsagia circumcincta*, a nematode parasite of small ruminants. *Parasitology*, **117**, 571-578.

Elard, L., Cabaret, J., & Humbert, J. F. (1999) PCR diagnosis of benzimidazole-susceptibility or -resistance in natural populations of the small ruminant parasite, *Teladorsagia circumcincta*. *Veterinary Parasitology*, **80**, (3), 231-237.

Faessler, H., Torgerson, P. R., & Hertzberg, H. (2007) Failure of *Duddingtonia flagrans* to reduce gastrointestinal nematode infections in dairy ewes. *Veterinary Parasitology*, **147**, 96-102.

Ffrench-Constant, R. H., Daborn, P. J., & Le Goff, G. (2004) The genetics and genomics of insecticide resistance. *Trends in Genetics*, **20**, (3), 163-170.

Fleming, J. T., Squire, M. D., Barnes, T. M., Tornoe, C., Matsuda, K., Ahnn, J., Fire, A., Sulston, J. E., Barnard, E. A., Sattelle, D. B., & Lewis, J. A. (1997) *Caenorhabditis elegans* levamisole resistance genes *lev*-1, *unc*-29, and *unc*-38 encode functional nicotinic acetylcholine receptor subunits. *Journal of Neuroscience*, **17**, (15), 5843-5857.

Freeman, A. S., Nghiem, C., Li, J., Ashton, F. T., Guerrero, J., Shoop, W. L., & Schad, G. A. (2003) Amphidial structure of ivermectin-resistant and susceptible laboratory and field strains of *Haemonchus contortus*. *Veterinary Parasitology*, **110**, (3-4), 217-226.

Gasser, R. B. & Newton, S. E. (2000) Genomic and genetic research on bursate nematodes: significance, implications and prospects. *International Journal for Parasitology*, **30**, (4), 509-534.

Geary, T. G. (2005) Ivermectin 20 years on: maturation of a wonder drug. *Trends in Parasitology*, **21**, (11), 530-532.

Ghisi, M., Kaminsky, R., & Maser, P. (2007) Phenotyping and genotyping of *Haemonchus contortus* isolates reveals a new putative candidate mutation for benzimidazole resistance in nematodes. *Veterinary Parasitology*, **144**, (3-4), 313-320.

Gibbs, H. C. (1986) Hypobiosis in parasitic nematodes - An update. *Advances in Parasitology*, **25**, 129-174.

Gibson, T. E. (1973) Recent advances in epidemiology and control of parasitic gastroenteritis in sheep. *Veterinary Record*, **92**, (18), 469-473.

Gill, J. H. & Lacey, E. (1998) Avermectin/milbemycin resistance in trichostrongyloid nematodes. *International Journal for Parasitology*, **28**, (6), 863-877.

Gilleard, J. S. (2006) Understanding anthelmintic resistance: The need for genomics and genetics. *International Journal for Parasitology*, **36**, 1227-1239.

Gilleard, J. S. & Beech, R. N. (2007) Population genetics of anthelmintic resistance in parasitic nematodes. *Parasitology*, **134**, 1133-1147.

Gong, M. Q., Gu, Y., Hu, X. B., Sun, Y., Ma, L., Li, X. L., Sun, L. X., Sun, J., Qian, J., & Zhu, C. L. (2005) Cloning and overexpression of CYP6F1, a cytochrome p450 gene, from

deltamethrin-resistant *Culex pipiens pallens*. *Acta Biochimica et Biophysica Sinica*, **37**, (5), 317-326.

Gotoh, O. (1998) Divergent structures of *Caenorhabditis elegans* cytochrome P450 genes suggest the frequent loss and gain of introns during the evolution of nematodes. *Molecular Biology and Evolution*, **15**, (11), 1447-1459.

Grade, J. T., Arble, B. L., Weladji, R. B., & Van Damme, P. (2008) Anthelmintic efficacy and dose determination of *Albizia anthelmintica* against gastrointestinal nematodes in naturally infected Ugandan sheep. *Veterinary Parasitology*, **157**, (3-4), 267-274.

Graham, S. E. & Peterson, J. A. (1999) How similar are P450s and what can their differences teach us? *Archives of Biochemistry and Biophysics*, **369**, (1), 24-29.

Graveley, B. R. (2008) Power Sequencing. *Nature*, **453**, 1197-1198.

Greer, A. W., Kenyon, F., Bartley, D. J., Jackson, E. B., Gordon, Y., Donnan, A. A., McBean, D. W., & Jackson, F. (2009) Development and field evaluation of a decision support model for anthelmintic treatments as part of a targeted selective treatment (TST) regime in lambs. *Veterinary Parasitology*, **164**, (1), 12-20.

Grillo, V., Jackson, F., Cabaret, J., & Gilleard, J. S. (2007) Population genetic analysis of the ovine parasitic nematode *Teladorsagia circumcincta* and evidence for a cryptic species. *International Journal for Parasitology*, **37**, (3-4), 435-447.

Gruner, L. & Suryahadi, S. (1993) Irrigation, faecal water content and development rate of free-living stages of sheep trichostrongyles. *Veterinary Research*, **24**, (4), 327-334.

Guengerich, F. P. (1991) Reactions and significance of cytochrome P450 enzymes. *Journal of Biological Chemistry*, **266**, (16), 10019-10022.

Hartman, D., Cottee, P. A., Savin, K. W., Bhave, M., Presidente, P. J. A., Fulton, L., Walkiewicz, M., & Newton, S. E. (2003) *Haemonchus contortus*: molecular characterisation of a small heat shock protein. *Experimental Parasitology*, **104**, (3-4), 96-103.

Heath, G. B. S. & Michel, J. F. (1969) A contribution to epidemiology of parasitic gastroenteritis in lambs. *Veterinary Record*, **85**, (11), 305-&.

Heckendorn, F., Haring, D. A., Maurer, V., Senn, M., & Hertzberg, H. (2007) Individual administration of three tanniferous forage plants to lambs artificially infected with *Haemonchus contortus* and *Cooperia curticei*. *Veterinary Parasitology*, **146**, (1-2), 123-134.

Higgins, C. F., Callaghan, R., Linton, K. J., Rosenberg, M. F., & Ford, R. C. (1997) Structure of the multidrug resistance P-glycoprotein. *Seminars in Cancer Biology*, **8**, (3), 135-142.

Hodgkinson, J., Clark, H., Kaplan, R., Lake, S., & Matthews, J. (2008) The role of polymorphisms at beta tubulin isotype 1 codons 167 and 200 in benzimidazole resistance in cyathostomins. *International Journal for Parasitology*, **38**, (10), 1149-1160.

Huang, Y. J. & Prichard, R. K. (1999) Identification and stage-specific expression of two putative P-glycoprotein coding genes in *Onchocerca volvulus*. *Molecular and Biochemical Parasitology*, **102**, (2), 273-281.

Huntley, J. F., Schallig, H. D. F. H., Kooyman, F. N. J., Mackellar, A., Jackson, F., & Smith, W. D. (1998a) IgE antibody during infection with the ovine abomasal nematode, *Teladorsagia circumcincta*: primary and secondary responses in serum and gastric lymph of sheep. *Parasite Immunology*, **20**, (11), 565-571.

Huntley, J. F., Schallig, H. D. F. H., Kooyman, F. N. J., Mackellar, A., Millership, J., & Smith, W. D. (1998b) IgE responses in the serum and gastric lymph of sheep infected with *Teladorsagia circumcincta*. *Parasite Immunology*, **20**, (4), 163-168.

Jackson, F., Jackson, E., & Coop, R. L. (1992) Evidence of multiple anthelmintic resistance in a strain of *Teladorsagia circumcincta* (*Ostertagia circumcincta*) isolated from goats in Scotland. *Research in Veterinary Science*, **53**, (3), 371-374.

Jackson, F. & Coop, R. L. (2000) The development of anthelmintic resistance in sheep nematodes. *Parasitology*, **120**, S95-S107.

James, C. E. & Davey, M. W. (2008) Increased expression of ABC transport proteins is associated with ivermectin resistance in the model nematode *Caenorhabditis elegans*. *International Journal for Parasitology*, **39**, (2), 213-220..

James, C. E., Hudson, A. L., & Davey, M. W. (2009) Drug resistance mechanisms in helminths: is it survival of the fittest? *Trends in Parasitology*, **25**, (7), 328-335.

Jones, P. M. & George, A. M. (2005) Multidrug resistance in parasites: ABC transporters, P-glycoproteins and molecular modelling. *International Journal for Parasitology*, **35**, (5), 555-566.

Kaminsky, R., Ducray, P., Jung, M., Clover, R., Rufener, L., Bouvier, J., Weber, S. S., Wenger, A., Wieland-Berghausen, S., Goebel, T., Gauvry, N., Pautrat, F., Skripsky, T., Froelich, O., Komoin-Oka, C., Westlund, B., Sluder, A., & Maser, P. (2008) A new class of anthelmintics effective against drug-resistant nematodes. *Nature*, **452**, (7184), 176-U19.

Kanehisa, M. & Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, **28**, (1), 27-30.

Kaplan, R. M. (2004) Drug resistance in nematodes of veterinary importance: a status report. *Trends in Parasitology*, **20**, (10), 477-481.

Kaplan, R. M., Vidyashankar, A. N., Howell, S. B., Neiss, J. M., Williamson, L. H., & Terrill, T. H. (2007) A novel approach for combining the use of in vitro and in vivo data to measure and detect emerging moxidectin resistance in gastrointestinal nematodes of goats. *International Journal for Parasitology*, **37**, (7), 795-804.

Kenyon, F., Sargison, N. D., Skuce, P. J., & Jackson, F. (2009a) Sheep helminth parasitic disease in south eastern Scotland arising as a possible consequence of climate change. *Veterinary Parasitology*, **163**, (4), 293-297.

Kenyon, F., Greer, A. W., Coles, G. C., Cringoli, G., Papadopoulos, E., Cabaret, J., Berrag, B., Varady, M., Van Wyk, J. A., Thomas, E., Vercruysse, J., & Jackson, F. (2009b) The role of targeted selective treatments in the development of refugia-based approaches to the control of gastrointestinal nematodes of small ruminants. *Veterinary Parasitology*, **164**, (1), 3-11.

Kerboeuf, D. (1985) Winter survival of trichostrongyle larvae - A study using tracer lambs. *Research in Veterinary Science*, **38**, (3), 364-367.

Kerboeuf, D., Guegnard, F., & Le Vern, Y. (2002) Analysis and partial reversal of multidrug resistance to anthelmintics due to P-glycoprotein in *Haemonchus contortus* eggs using *Lens culinaris* lectin. *Parasitology Research*, **88**, (9), 816-821.

Kerboeuf, D., Guegnard, F., & Le Vern, Y. (2003) Detection of P-glycoprotein-mediated multidrug resistance against anthelmintics in *Haemonchus contortus* using anti-human mdr1 monoclonal antibodies. *Parasitology Research*, **91**, (1), 79-85.

Kerboeuf, D., Blackhall, W., Kaminsky, R., & von Samson-Himmelstjerna, G. (2003) P-glycoprotein in helminths: function and perspectives for anthelmintic treatment and reversal of resistance. *International Journal of Antimicrobial Agents*, **22**, (3), 332-346.

Kidane, A., Houdijk, J. G. M., Athanasiadou, S., Tolkamp, B. J., & Kyriazakis, I. (2010) Effects of maternal protein nutrition and subsequent grazing on chicory (Cichorium intybus) on parasitism and performance of lambs. *Journal of Animal Science*, **88**, (4), 1513-1521.

Kimchi-Sarfaty, C., Oh, J. M., Kim, I. W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V., & Gottesman, M. M. (2007) A "Silent" Polymorphism in the MDR1 Gene Changes Substrate Specificity. *Science*, **315**, (5811), 525-528.

Knox, D. P. & Smith, W. D. (2001) Vaccination against gastrointestinal nematode parasites of ruminants using gut-expressed antigens. *Veterinary Parasitology*, **100**, (1-2), 21-32.

Komar, A. A. (2007) Silent SNPs: impact on gene function and phenotype. *Pharmacogenomics*, **8**, (8), 1075-1080.

Kopp, S. R., Coleman, G. T., Traub, R. J., McCarthy, J. S., & Kotze, A. C. (2008) Acetylcholine receptor subunit genes from *Ancylostoma caninum*: Altered transcription patterns associated with pyrantel resistance. *International Journal for Parasitology*, **39**, (4), 435-441.

Kotze, A. C. (1997) Cytochrome P450 monooxygenase activity in *Haemonchus contortus* (Nematoda). *International Journal for Parasitology*, **27**, (1), 33-40.

Kwa, M. S. G., Veenstra, J. G., & Roos, M. H. (1994) Benzimidazole resistance in *Haemonchus contortus* is correlated with a conserved mutation at amino acid 200 in β-tubulin isotype 1. *Molecular and Biochemical Parasitology*, **63**, (2), 299-303.

Kwa, M. S. G., Okoli, M. N., Schulz-Key, H., Okongkwo, P. O., & Roos, M. H. (1998) Use of P-glycoprotein gene probes to investigate anthelmintic resistance in *Haemonchus*

contortus and comparison with *Onchocerca volvulus*. *International Journal for Parasitology*, **28**, (8), 1235-1240.

Larsen, M. (1999) Biological control of helminths. *International Journal for Parasitology*, **29**, (1), 139-146.

Le Goff, G., Boundy, S., Daborn, P. J., Yen, J. L., Sofer, L., Lind, R., Sabourault, C., Madi-Ravazzi, L., & Ffrench-Constant, R. H. (2003) Microarray analysis of cytochrome P450 mediated insecticide resistance in *Drosophila*. *Insect Biochemistry and Molecular Biology*, **33**, (7), 701-708.

Le Jambre, L. F. (1976) Egg hatch as an *in vitro* assay of thiabendazole resistance in nematodes. *Veterinary Parasitology*, **2**, (4), 385-391.

Le Jambre, L. F., Lenane, I. J., & Wardrop, A. J. (1999) A hybridisation technique to identify anthelmintic resistance genes in *Haemonchus*. *International Journal for Parasitology*, **29**, (12), 1979-1985.

Le Jambre, L. F., Dobson, R. J., Lenane, I. J., & Barnes, E. H. (1999) Selection for anthelmintic resistance by macrocyclic lactones in *Haemonchus contortus*. *International Journal for Parasitology*, **29**, (7), 1101-1111.

Le Jambre, L. F., Gill, J. H., Lenane, I. J., & Baker, P. (2000) Inheritance of avermectin resistance in *Haemonchus contortus*. *International Journal for Parasitology*, **30**, (1), 105-111.

Leroy, S., Duperray, C., & Morand, S. (2003) Flow cytometry for parasite nematode genome size measurement. *Molecular and Biochemical Parasitology*, **128**, (1), 91-93.

Lespine, A., Martin, S., Dupuy, J., Roulet, A., Pineau, T., Orlowski, S., & Alvinerie, M. (2007) Interaction of macrocyclic lactones with P-glycoprotein: Structure-affinity relationship. *European Journal of Pharmaceutical Sciences*, **30**, (1), 84-94.

Lespine, A., Alvinerie, M., Vercruysse, J., Prichard, R. K., & Geldhof, P. (2008) ABC transporter modulation: a strategy to enhance the activity of macrocyclic lactone anthelmintics. *Trends in Parasitology*, **24**, (7), 293-298.

Lichtenfels, J. R. & Hoberg, E. P. (1993) The systematics of nematodes that cause Ostertagiasis in domestic and wild ruminants in North America - An update and a key to species. *Veterinary Parasitology*, **46**, (1-4), 33-53.

Lifschitz, A., Suarez, V. H., Sallovitz, J., Cristel, S. L., Imperiale, F., Ahoussou, S., Schiavi, C., & Lanusse, C. (2010) Cattle nematodes resistant to macrocyclic lactones: comparative effects of P-glycoprotein modulation on the efficacy and disposition kinetics of ivermectin and moxidectin. *Experimental Parasitology*, **125**, (2), 172-178.

Lincke, C. R., Broeks, A., The, I., Plasterk, R. H. A., & Borst, P. (1993) The expression of two P-glycoprotein (Pgp) genes in transgenic *Caenorhabditis elegans* is confined to intestinal cells. *Embo Journal*, **12**, (4), 1615-1620.

Lindquist, S. (1986) The Heat-Shock Response. *Annual Review of Biochemistry*, **55**, 1151-1191.

Liu, M. Y., Wang, X. L., Fu, B. Q., Li, C. Y., Wu, X. P., Le Rhun, D., Chen, Q. J., & Boireau, P. (2007) Identification of stage-specifically expressed genes of *Trichinella spiralis* by suppression subtractive hybridization. *Parasitology*, **134**, 1443-1455.

Livak, K. J. & Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods*, **25**, (4), 402-408.

Loo, T. W. & Clarke, D. M. (1999) Molecular dissection of the human multidrug resistance P-glycoprotein. *Biochemistry and Cell Biology-Biochimie et Biologie Cellulaire*, **77**, (1), 11-23.

Mak, C. H., Sun, K. W., & Ko, R. C. (2001) Identification of some heat-induced genes of *Trichinella spiralis*. *Parasitology*, **123**, 293-300.

Mansuy, D. (1998) The great diversity of reactions catalyzed by cytochromes P450. *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology*, **121**, (1-3), 5-14.

Mardis, E. R. (2008) The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, **24**, (3), 133-141.

Marguiles, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J.-B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, k. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F., & Rothberg, J. M. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **437**, 376-380.

Marriner, S. E., Mckinnon, I., & Bogan, J. A. (1987) The Pharmacokinetics of Ivermectin After Oral and Subcutaneous Administration to Sheep and Horses. *Journal of Veterinary Pharmacology and Therapeutics*, **10**, (2), 175-179.

Martin, P. J., Anderson, N., & Jarrett, R. G. (1989) Detecting Benzimidazole Resistance with Fecal Egg Count Reduction Tests and Invitro Assays. *Australian Veterinary Journal*, **66**, (8), 236-240.

Martin, R. J., Murray, I., Robertson, A. P., Bjorn, H., & Sangster, N. (1998) Anthelmintics and ion-channels: after a puncture, use a patch. *International Journal for Parasitology*, **28**, (6), 849-862.

Martin, R. J. & Robertson, A. P. (2007) Mode of action of levamisole and pyrantel, Anthelmintic Resistance, E153 and Q57. *Parasitology*, **134**, 1093-1104.

Martin, S. K., Oduola, A. M. J., & Milhous, W. K. (1987) Reversal of chloroquine resistance in *Plasmodium falciparum* by verapamil. *Science*, **235**, (4791), 899-901.

McCavera, S., Walsh, T. K., & Wolstenholme, A. J. (2007) Nematode ligand-gated chloride channels: an appraisal of their involvement in macrocyclic lactone resistance and prospects for developing molecular markers. *Parasitology*, **134**, 1111-1121.

McKellar, Q. A. (1993) Interactions of *Ostertagia* species with their bovine and ovine hosts. *International Journal for Parasitology*, **23**, (4), 451-462.

McKellar, Q. A. & Jackson, F. (2004) Veterinary anthelmintics: old and new. *Trends in Parasitology*, **20**, (10), 456-461.

McLeod, R. S. (1995) Costs of major parasites to the Australian livestock industries. *International Journal for Parasitology*, **25**, (11), 1363-1367.

McNeilly, T., Devaney, E., & Matthews, J. (2009) Teladorsagia circumcincta in the sheep abomasum: defining the role of dendritic cells in T cell regulation and protective immunity. *Parasite Immunology*, **31**, (7), 347-356.

Menzel, R., Bogaert, T., & Achazi, R. (2001) A systematic gene expression screen of *Caenorhabditis elegans* cytochrome P450 genes reveals CYP35 as strongly xenobiotic inducible. *Archives of Biochemistry and Biophysics*, **395**, (2), 158-168.

Menzel, R., Rodel, M., Kulas, J., & Steinberg, C. E. W. (2005) CYP35: Xenobiotically induced gene expression in the nematode *Caenorhabditis elegans*. *Archives of Biochemistry and Biophysics*, **438**, (1), 93-102.

Michel, J. (1974) Arrested development of nematodes and some related phenomena. *Advances in parasitology*, **12**. 279-366.

Miller, H. R. P. (1996) Prospects for the immunological control of ruminant gastrointestinal nematodes: Natural immunity, can it be harnessed? *International Journal for Parasitology*, **26**, (8-9), 801-811.

Mitreva, M., Zarlenga, D. S., McCarter, J. P., & Jasmer, D. P. (2007) Parasitic nematodes - From genomes to control. *Veterinary Parasitology*, **148**, (1), 31-42.

Molento, M. (2009) Parasite control in the age of drug resistance and changing agricultural practices. *Veterinary Parasitology*, **163**, 229-234.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, **35**, W182-W185.

Morley, F. H. W. & Donald, A. D. (1980) Farm management and systems of helminth control. *Veterinary Parasitology*, **6**, (1-3), 105-134.

Nagaraj, S. H., Gasser, R. B., & Ranganathan, S. (2007) A hitchhiker's guide to expressed sequence tag (EST) analysis. *Briefings in Bioinformatics*, **8**, (1), 6-21.

Ndifor, A. M., Ward, S. A., & Howells, R. E. (1990) Cytochrome P450 activity in malarial parasites and its possible relationship to chloroquine resistance. *Molecular and Biochemical Parasitology*, **41**, (2), 251-257.

Nelson, D. R., Koymans, L., Kamataki, T., Stegeman, J. J., Feyereisen, R., Waxman, D. J., Waterman, M. R., Gotoh, O., Coon, M. J., Estabrook, R. W., Gunsalus, I. C., & Nebert, D. W. (1996) P450 superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics*, **6**, (1), 1-42.

Nelson, D. R. (1998) Metazoan cytochrome P450 evolution. *Comparative Biochemistry and Physiology C-Pharmacology Toxicology & Endocrinology*, **121**, (1-3), 15-22.

Nelson, D. R. (1999) Cytochrome P450 and the individuality of species. *Archives of Biochemistry and Biophysics*, **369**, (1), 1-10.

Neveu, C., Charvet, C., Fauvin, A., Cortet, J., Castagnone-Sereno, P., & Cabaret, J. (2007) Identification of levamisole resistance markers in the parasitic nematode *Haemonchus contortus* using a cDNA-AFLP approach. *Parasitology*, **134**, 1105-1110.

Nieuwhof, G. J. & Bishop, S. C. (2005) Costs of the major endemic diseases of sheep in Great Britain and the potential benefits of reduction in disease impact. *Animal Science*, **81**, 23-29.

Niezen, J. H., Robertson, H. A., Waghorn, G. C., & Charleston, W. A. G. (1998) Production, faecal egg counts and worm burdens of ewe lambs which grazed six contrasting forages. *Veterinary Parasitology*, **80**, (1), 15-27.

Nikou, D., Ranson, H., & Hemingway, J. (2003) An adult-specific CYP6 P450 gene is overexpressed in a pyrethroid-resistant strain of the malaria vector, *Anopheles gambiae*. *Gene*, **318**, 91-102.

Nisbet, A. J. & Gasser, R. B. (2004) Profiling of gender-specific gene expression for *Trichostrongylus vitrinus* (Nematoda : Strongylida) by microarray analysis of expressed sequence tag libraries constructed by suppressive-subtractive hybridisation. *International Journal for Parasitology*, **34**, (5), 633-643.

Nisbet, A. J., Redmond, D. L., Matthews, J. B., Watkins, C., Yaga, R., Jones, J. T., Nath, M., & Knox, D. P. (2008) Stage-specific gene expression in *Teladorsagia circumcincta* (Nematoda : Strongylida) infective larvae and early parasitic stages. *International Journal for Parasitology*, **38**, (7), 829-838.

Njue, A. I., Hayashi, J., Kinne, L., Feng, X. P., & Prichard, R. K. (2004) Mutations in the extracellular domains of glutamate-gated chloride channel α-3 and β subunits from ivermectin-resistant *Cooperia oncophora* affect agonist sensitivity. *Journal of Neurochemistry*, **89**, (5), 1137-1147.

Nunes, F., Wolf, M., Hartmann, J., & Paul, R. J. (2005) The ABC transporter PGP-2 from *Caenorhabditis elegans* is expressed in the sensory neuron pair AWA and contributes to lysosome formation and lipid storage within the intestine. *Biochemical and Biophysical Research Communications*, **338**, (2), 862-871.

O'Connor, L. J., Walkden-Brown, S. W., & Kahn, L. P. (2006) Ecology of the free-living stages of major trichostrongylid parasites of sheep. *Veterinary Parasitology*, **142**, (1-2), 1-15.

Omura, S. (2008) Ivermectin: 25 years and still going strong. *International Journal of Antimicrobial Agents*, **31**, (2), 91-98.

Omura, T. (1999) Forty years of cytochrome P450. *Biochemical and Biophysical Research Communications*, **266**, (3), 690-698.

Parkinson, J., Whitton, C., Schmid, R., Thomson, M., & Blaxter, M. (2004a) NEMBASE: a resource for parasitic nematode ESTs. *Nucleic Acids Research*, **32**, D427-D430.

Parkinson, J., Mitreva, M., Whitton, C., Thomson, M., Daub, J., Martin, J., Schmid, R., Hall, N., Barrell, B., Waterston, R. H., McCarter, J. P., & Blaxter, M. L. (2004b) A transcriptomic analysis of the phylum Nematoda. *Nature Genetics*, **36**, (12), 1259-1267.

Payne, R. W., Murray, D. A., Harding, S. A., Baird, D. B., & Souter, D. M. (2009) GenStat for Windows. 12th, VSN International.

Pfaffl, M. W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, **29**, (9).

Pop, M. & Salzberg, S. L. (2008) Bioinformatics challenges of new sequencing technology. *Trends in Genetics*, **24**, (3), 142-149.

Pouliot, J. F., L'Heureux, F., Liu, Z., Prichard, R. K., & Georges, E. (1997) Reversal of P-glycoprotein associated multidrug resistance by ivermectin. *Biochemical Pharmacology*, **53**, (1), 17-25.

Prichard, R. K., Steel, J. W., Lacey, E., & Hennessy, D. R. (1985a) Pharmacokinetics of Ivermectin in Sheep Following Intravenous, Intra-Abomasal Or Intraruminal Administration. *Journal of Veterinary Pharmacology and Therapeutics*, **8**, (1), 88-94.

Prichard, R. K., Steel, J. W., Lacey, E., & Hennessy, D. R. (1985b) Pharmacokinetics of ivermectin in sheep following intravenous, intra-abomasal or intraruminal administration. *Journal of Veterinary Pharmacology and Therapeutics*, **8**, (1), 88-94.

Prichard, R. K. (2001) Genetic variability following selection of *Haemonchus contortus* with anthelmintics. *Trends in Parasitology*, **17**, (9), 445-453.

Prichard, R. K. & Roulet, A. (2007) ABC transporters and β-tubulin in macrocyclic lactone resistance: prospects for marker development. *Parasitology*, **134**, 1123-1132.

Prichard, R. K., von Samson-Himmelstjerna, G., Blackhall, W. J., & Geary, T. G. (2007) Foreword: Towards markers for anthelmintic resistance in helminths of importance in animal and human health. *Parasitology*, **134**, 1073-1076.

Prichard, R. K. & Geary, T. G. (2008) Fresh hope to can the worms. *Nature*, **452**, (7184), 157-158.

Rabel, B., McGregor, R., & Douch, P. G. C. (1994) Improved bioassay for estimation of inhibitory effects of ovine gastrointestinal mucus and anthelmintics on nematode larval migration. *International Journal for Parasitology*, **24**, (5), 671-676.

Ranjit, N., Zhan, B., Stenzel, D. J., Mulvenna, J., Fujiwara, R., Hotez, P. J., & Loukas, A. (2008) A family of cathepsin B cysteine proteases expressed in the gut of the human hookworm, *Necator americanus*. *Molecular and Biochemical Parasitology*, **160**, (2), 90-99.

Rayes, D., De Rosa, M. A. J., Bartos, M., & Bouzat, C. (2004) Molecular basis of the differential sensitivity of nematode and mammalian muscle to the anthelmintic agent levamisole. *Journal of Biological Chemistry*, **279**, (35), 36372-36381.

Redman, E., Packard, E., Grillo, V., Smith, J., Jackson, F., & Gilleard, J. S. (2008) Microsatellite analysis reveals marked genetic differentiation between *Haemonchus contortus* laboratory isolates and provides a rapid system of genetic fingerprinting. *International Journal for Parasitology*, **38**, (1), 111-122.

Riou, M., Guegnard, F., Le Vern, Y., & Kerboeuf, D. (2003) Modulation of the multidrug resistance (MDR) system in the nematode *Haemonchus contortus* by changing cholesterol content: effects on resistance to anthelmintics. *Journal of Antimicrobial Chemotherapy*, **52**, (2), 180-187.

Riou, M., Koch, C., & Kerboeuf, D. (2005) Increased resistance to anthelmintics of *Haemonchus contortus* eggs associated with changes in membrane fluidity of eggshells during embryonation. *Parasitology Research*, **95**, (4), 266-272.

Riou, M., Koch, C., Delaleu, B., Berthon, P., & Kerboeuf, D. (2005) Immunolocalisation of an ABC transporter, P-glycoprotein, in the eggshells and cuticles of free-living and parasitic stages of *Haemonchus contortus*. *Parasitology Research*, **96**, (3), 142-148.

Riou, M., Grasseau, I., Blesbois, E., & Kerboeuf, D. (2007) Relationships between sterol/phospholipid composition and xenobiotic transport in nematodes. *Parasitology Research*, **100**, (5), 1125-1134.

Rodpradit, P., Boonsuepsakul, S., Chareonviriyaphap, T., Bangs, M. J., & Rongnoparut, P. (2005) Cytochrome P450 genes: Molecular cloning and overexpression in a pyrethroid-resistant strain of *Anopheles minimus* mosquito. *Journal of the American Mosquito Control Association*, **21**, (1), 71-79.

Rothberg, J. M. & Leamon, J. H. (2008) The development and impact of 454 sequencing. *Nature biotechnology*, **26**, (10), 1117-1124.

Roulet, A. & Prichard, R. K. "Ivermectin and moxidectin cause constitutive and induce over expression of different P-glycoproteins in resistant *Haemonchus contortus*", in *Annual Meeting of the American Association of Veterinary Parasitologists*.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) DNA Sequencing with Chain-Terminating Inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, **74**, (12), 5463-5467.

Sangster, N. C. (1994) P-glycoproteins in nematodes. *Parasitology Today*, **10**, (8), 319-322.

Sangster, N. C., Redwin, J. M., & Bjorn, H. (1998) Inheritance of levamisole and benzimidazole resistance in an isolate of *Haemonchus contortus*. *International Journal for Parasitology*, **28**, (3), 503-510.

Sangster, N. C., Bannan, S. C., Weiss, A. S., Nulf, S. C., Klein, R. D., & Geary, T. G. (1999a) *Haemonchus contortus*: Sequence heterogeneity of internucleotide binding domains from P-glycoproteins and an association with avermectin/milbemycin resistance. *Experimental Parasitology*, **91**, (3), 250-257.

Sangster, N. C., Bannan, S. C., Weiss, A. S., Nulf, S. C., Klein, R. D., & Geary, T. G. (1999b) *Haemonchus contortus*: Sequence heterogeneity of internucleotide binding domains from P-glycoproteins and an association with avermectin/milbemycin resistance. *Experimental Parasitology*, **91**, (3), 250-257.

Sargison, N., Scott, P., & Jackson, F. (2001) Multiple anthelmintic resistance in sheep. *Veterinary Record*, **149**, (25), 778-779.

Sargison, N. D., Jackson, F., & Scott, P. R. (2002) Teladorsagiosis in young lambs and extended postparturient susceptibility in moxidectin-treated ewes grazing heavily contaminated pastures. *Veterinary Record*, **151**, (12), 353-355.

Sargison, N. D., Jackson, F., Bartley, D. J., & Moir, A. C. P. (2005) Failure of moxidectin to control benzimidazole-, levamisole- and ivermectin-resistant *Teladorsagia circumcincta* in a sheep flock. *Veterinary Record*, **156**, (4), 105-109.

Sargison, N. D., Wilson, D. J., Bartley, D. J., & Penny, C. D. (2007a) Haemonchosis and teladorsagiosis in a Scottish sheep flock putatively associated with the overwintering of hypobiotic fourth stage larvae. *Veterinary Parasitology*, **147**, (3-4), 326-331.

Sargison, N. D., Jackson, F., Bartley, D. J., Wilson, D. J., Stenhouse, L. J., & Penny, C. D. (2007b) Observations on the emergence of multiple anthelmintic resistance in sheep flocks in the south-east of Scotland. *Veterinary Parasitology*, **145**, (1-2), 65-76.

Schroeder, L. K., Kremer, S., Kramer, M. J., Currie, E., Kwan, E., Watts, J. L., Lawrenson, A. L., & Hermann, G. J. (2007) Function of the *Caenorhabditis elegans* ABC transporter PGP-2 in the biogenesis of a lysosome-related fat storage organelle. *Molecular Biology of the Cell*, **18**, (3), 995-1008.

Scott, I., Hodgkinson, S. M., Khalaf, S., Lawton, D. E. B., Collett, M. G., Reynolds, G. W., Pomroy, W. E., & Simpson, H. V. (1998) Infection of sheep with adult and larval *Ostertagia circumcincta*: abomasal morphology. *International Journal for Parasitology*, **28**, (9), 1383-1392.

Shendure, J. & Ji, H. (2008) Next-generation DNA sequencing. *Nature biotechnology*, **26**, (10), 1135-1145.

Sheps, J. A., Ralph, S., Zhao, Z. Y., Baillie, D. L., & Ling, V. (2004) The ABC transporter gene family of *Caenorhabditis elegans* has implications for the evolutionary dynamics of multidrug resistance in eukaryotes. *Genome Biology*, **5**, (3).

Shin, H., Hirst, M., Bainbridge, M. N., Magrini, V., Mardis, E., Moerman, D. G., Marra, M. A., Baillie, D. L., & Jones, S. J. (2008) Transcriptome analysis for *Caenorhabditis elegans* based on novel expressed sequence tags. *BMC.Biol.*, **6**, (30).

Silvestre, A. & Humbert, J. F. (2000) A molecular tool for species identification and benzimidazole resistance diagnosis in larval communities of small ruminant parasites. *Experimental Parasitology*, **95**, (4), 271-276.

Silvestre, A., Cabaret, J., & Humbert, J. F. (2001) Effect of benzimidazole under dosing on the resistant allele frequency in *Teladorsagia circumcincta* (Nematoda). *Parasitology*, **123**, 103-111.

Silvestre, A. & Humbert, J. F. (2002) Diversity of benzimidazole-resistance alleles in populations of small ruminant parasites. *International Journal for Parasitology*, **32**, (7), II.

Silvestre, A. & Cabaret, J. (2002) Mutation in position 167 of isotype 1 β-tubulin gene of Trichostrongylid nematodes: role in benzimidazole resistance? *Molecular and Biochemical Parasitology*, **120**, (2), II.

Silvestre, A., Sauve, C., Cortet, J., & Cabaret, J. (2009) Contrasting genetic structures of two parasitic nematodes, determined on the basis of neutral microsatellite markers and selected anthelmintic resistance markers. *Molecular Ecology*, **18**, (24), 5086-5100.

Skuce, P. J., Redmond, D. L., Liddell, S., Stewart, E. M., Newlands, G. F. J., Smith, W. D., & Knox, D. P. (1999) Molecular cloning and characterization of gut-derived cysteine proteinases associated with a host protective extract from Haemonchus contortus. *Parasitology*, **119**, 405-412.

Skuce, P. J., Stenhouse, L. J., Jackson, F., Hypsa, V., & Gilleard, J. (2010) Benzimidazole resistance allele haplotype diversity in United Kingdom isolates of *Teladorsagia circumcincta* supports a hypothesis of multiple origins of resistance by recurrent mutation. *International Journal for Parasitology*, **40**, (11), 1247-1255.

Smith, J. M. & Prichard, R. K. (2002) Localization of P-glycoprotein mRNA in the tissues of *Haemonchus contortus* adult worms and its relative abundance in drug-selected and susceptible strains. *Journal of Parasitology*, **88**, (3), 612-620.

Smith, S., Nisbet, A., Meikle, L., I, Inglis, N., Sales, J., Beynon, R., & Matthews, J. (2009) Proteomic analysis of excretory/secretory products released by Teladorsagia circumcincta larvae early post-infection. *Parasite Immunology*, **31**, (1), 10-19.

Smith, W. D., Van Wyk, J. A., & van Strijp, M. F. (2001) Preliminary observations on the potential of gut membrane proteins of *Haemonchus contortus* as candidate vaccine antigens in sheep on naturally infected pasture. *Veterinary Parasitology*, **98**, (4), 285-297.

Smith, W. D. & Zarlenga, D. S. (2006) Developments and hurdles in generating vaccines for controlling helminth parasites of grazing ruminants. *Veterinary Parasitology*, **139**, (4), 347-359.

Soulsby (2007) New concepts in strongyle control and anthelmintic resistance: The role of refugia. *Veterinary Journal*, **174**, (1), 6-7.

Stafford, K. A., Morgan, E. R., & Coles, G. C. (2009) Weight-based targeted selective treatment of gastrointestinal nematodes in a commercial sheep flock. *Veterinary Parasitology*, **164**, (1), 59-65.

Stear, M. J., Bishop, S. C., Doligalska, M., Duncan, J. L., Holmes, P. H., Irvine, J., McCririe, L., McKellar, Q. A., Sinski, E., & Murray, M. (1995) Regulation of egg production, worm burden, worm length and worm fecundity by host responses in sheep infected with *Ostertagia circumcincta*. *Parasite Immunology*, **17**, (12), 643-652.

Stear, M. J., Bairden, K., Duncan, J. L., Eckersall, P. D., Fishwick, G., Graham, P. A., Holmes, P. H., McKellar, Q. A., Mitchell, S., Murray, M., Parkins, J. J., & Wallace, D. S. (2000) The influence of relative resistance and urea-supplementation on deliberate infection with *Teladorsagia circumcincta* during winter. *Veterinary Parasitology*, **94**, (1-2), 45-54.

Stear, M. J., Abuagob, O., Benothman, M., Bishop, S. C., Innocent, G., Kerr, A., & Mitchell, S. (2006) Variation among faecal egg counts following natural nematode infection in Scottish Blackface lambs. *Parasitology*, **132**, 275-280.

Stear, M. J., Doligalska, M., & Donskow-Schmelter, K. (2007) Alternatives to anthelmintics for the control of nematodes in livestock. *Parasitology*, **134**, 139-151.

Stear, M. J., Boag, B., Cattadori, I., & Murray, I. (2009) Genetic variation in resistance to mixed, predominantly *Teladorsagia circumcincta* nematode infections of sheep: from heritabilities to gene identification. *Parasite Immunology*, **31**, (5), 274-282.

Stenhouse, L. J. (2007). Characterisation of anthelmintic resistance in a multiple drug resistant *Teladorsagia circumcincta* isolate. Doctor of Philosophy, University of Glasgow.

Strube, C., Buschbaum, S., Wolken, S., & Schnieder, T. (2008) Evaluation of reference genes for quantitative real-time PCR to investigate protein disulfide isomerase transcription pattern in the bovine lungworm *Dictyocaulus viviparus*. *Gene*, **425**, (1-2), 36-43.

Stubbings, L. A. 2003, "Internal parasite control in sheep", in *Short Term Strategies to Slow the Development of Anthelmintic Resistance in Internal Parasites of Sheep in the United Kingdom*.

Sutherland, I. A., Leathwick, D. M., Moen, I. C., & Bisset, S. A. (2002) Resistance to therapeutic treatment with macrocyclic lactone anthelmintics in O*stertagia circumcincta*. *Veterinary Parasitology*, **109**, (1-2), II.

Sutherland, I. A., Brown, A. E., Leathwick, D. M., & Bisset, S. A. (2003) Resistance to prophylactic treatment with macrocyclic lactone anthelmintics in *Teladorsagia circumcincta*. *Veterinary Parasitology*, **115**, (4), 301-309.

Sutherland, I. A. & Scott, I. 2010, *Gastrointestinal Nematodes of Sheep and Cattle, Biology and Control*, First edn, Wiley-Blackwell.

Sutherland, I. A., Shaw, J., & Shaw, R. J. (2010) The production costs of anthelmintic resistance in sheep managed within a monthly preventive drench program. *Veterinary Parasitology*, **171**, (3-4), 300-304.

Sykes, A. R. & Coop, R. L. (2001) Interaction between nutrition and gastrointestinal parasitism in sheep. *New Zealand Veterinary Journal*, **49**, (6), 222-226.

Taylor, M. A., Hunt, K. R., & Goodyear, K. L. (2002) Anthelmintic resistance detection methods. *Veterinary Parasitology*, **103**, (3), 183-194.

Taylor, M. A., Coop, R. L., & Wall, R. L. 2007, *Veterinary Parasitology*, 3rd edn, Blackwell Publishing Ltd.

Urquhart, G. M., Armour, J., Duncan, J. L., Dunn, A. M., & Jennings, F. W. 1996, *Veterinary Parasitology*, 2nd edn, Blackwell Publishing Ltd.

Valverde, M. A., Diaz, M., Sepulveda, F. V., Gill, D. R., Hyde, S. C., & Higgins, C. F. (1992) Volume regulated chloride channels associated with the human multidrug resistance P-glycoprotein. *Nature*, **355**, (6363), 830-833.

van Dijk, J., Sargison, N. D., Kenyon, F., & Skuce, P. J. (2010) Climate change and infectious disease: helminthological challenges to farmed ruminants in temperate regions. *Animal*, **4**, (3), 377-392.

Van Wyk, J. A. & Malan, F. S. (1988) Resistance of field strains of *Haemonchus contortus* to ivermectin, closantel, rafoxanide and the benzimidazoles in South-Africa. *Veterinary Record*, **123**, (9), 226-228.

Van Zeveren, A. M., Casaert, S., Alvinerie, M., Geldhof, P., Claerebout, E., & Vercruysse, J. (2007a) Experimental selection for ivermectin resistance in *Ostertagia ostertagi* in cattle. *Veterinary Parasitology*, **150**, (1-2), 104-110.

Van Zeveren, A. M., Visser, A., Hoorens, P. R., Vercruysse, J., Claerebout, E., & Geldhof, P. (2007b) Evaluation of reference genes for quantitative real-time PCR in *Ostertagia ostertagi* by the coefficient of variation and geNorm approach. *Molecular and Biochemical Parasitology*, **153**, (2), 224-227.

Van Zeveren, A. M. (2009). Ivermectin resistance in the bovine nematode *Ostertagia ostertagi*. Doctor of Philosphy, Ghent University.

Vercauteren, I., Maere, V., Vercruysse, J., Stevens, M., Gevaert, K., & Claerebout, E. (2006) A small heat shock protein of *Ostertagia ostertagi*: Stage-specific expression, heat inducibility, and protection trial. *Journal of Parasitology*, **92**, (6), 1244-1250.

Vercruysse, J., Knox, D. P., Schetters, T. P. M., & Willadsen, P. (2004) Veterinary parasitic vaccines: pitfalls and future directions. *Trends in Parasitology*, **20**, (10), 488-492.

Virkel, G., Lifschitz, A., Sallovitz, J., Ballent, M., Scarcella, S., & Lanusse, C. (2009) Inhibition of cytochrome P450 activity enhances the systemic availability of triclabendazole metabolites in sheep. *Journal of Veterinary Pharmacology and Therapeutics*, **32**, (1), 79-86.

Visser, A., Van Zeveren, A. M., Meyvis, Y., Peelaers, I., Van den Broeck, W., Gevaert, K., Vercruysse, J., Claerebout, E., & Geldhof, P. (2008) Gender-enriched transcription of activation associated secreted proteins in *Ostertagia ostertagi*. *International Journal for Parasitology*, **38**, (3-4), 455-465.

von Samson-Himmelstjerna, G. (2006) Molecular diagnosis of anthelmintic resistance. *Veterinary Parasitology*, **136**, (2), 99-107.

von Samson-Himmelstjerna, G., Blackhall, W. J., McCarthy, J. S., & Skuce, P. J. (2007) Single nucleotide polymorphism (SNP) markers for benzimidazole resistance in veterinary nematodes. *Parasitology*, **134**, 1077-1086.

von Samson-Himmelstjerna, G., Walsh, T. K., Donnan, A. A., Carriere, S., Jackson, F., Skuce, P. J., Rohn, K., & Wolstenholme, A. J. (2009a) Molecular detection of benzimidazole resistance in *Haemonchus contortus* using real-time PCR and pyrosequencing. *Parasitology*, **136**, (3), 349-358.

von Samson-Himmelstjerna, G., Coles, G. C., Jackson, F., Bauer, C., Borgsteede, F., Cirak, V. Y., Demeler, J., Donnan, A., Dorny, P., Epe, C., Harder, A., Hoglund, J., Kaminsky, R., Kerboeuf, D., Kuttler, U., Papadopoulos, E., Posedi, J., Small, J., Varady, M., Vercruysse, J., & Wirtherle, N. (2009b) Standardization of the egg hatch test for the detection of benzimidazole resistance in parasitic nematodes. *Parasitology Research*, **105**, (3), 825-834.

Vontas, J., Blass, C., Koutsos, A. C., David, J. P., Kafatos, F. C., Louis, C., Hemingway, J., Christophides, G. K., & Ranson, H. (2005) Gene expression in insecticide resistant and susceptible *Anopheles gambiae* strains constitutively or after insecticide exposure. *Insect Molecular Biology*, **14**, (5), 509-521.

Waghorn, T. S., Miller, C. M., Oliver, A. M., & Leathwick, D. M. (2009) Drench-and-shift is a high-risk practice in the absence of refugia. *N Z Vet J*, **57**, (6), 359-363.

Wagland, B. M., Jones, W. O., Hribar, L., Bendixsen, T., & Emery, D. L. (1992) A new simplified assay for larval migration inhibition. *International Journal for Parasitology*, **22**, (8), 1183-1185.

Waller, P. J. & Thamsborg, S. M. (2004) Nematode control in 'green' ruminant production systems. *Trends in Parasitology*, **20**, (10), 493-497.

Walsh, T. K., Donnan, A. A., Jackson, F., Skuce, P., & Wolstenholme, A. J. (2007) Detection and measurement of benzimidazole resistance alleles in *Haemonchus contortus* using real-time PCR with locked nucleic acid Taqman probes. *Veterinary Parasitology*, **144**, (3-4), 304-312.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Chuch, D. M., DiCuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L. Y., Helmberg, W., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D. J., Madden, T. L., Maglott, D. R., Miller, V., Ostell, J., Pruitt, K. D., Schuler, G. D., Shumway, M., Sequeira, E., Sherry, S. T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R. L., Tatusova, T. A., Wagner, L., & Yaschenko, E. (2008) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, **36**, D13-D21.

Wicks, S. R., de Vries, C. J., van Luenen, H. G. A. M., & Plasterk, R. H. A. (2000) CHE-3, a cytosolic dynein heavy chain, is required for sensory cilia structure and function in Caenorhabditis elegans. *Developmental Biology*, **221**, (2), 295-307.

Willadsen, P., Bird, P., Cobon, G. S., & Hungerford, J. (1995) Commercialization of A Recombinant Vaccine Against Boophilus-Microplus. *Parasitology*, **110**, S43-S50.

Wilson, D. & Sargison, N. (2007) Anthelmintic resistance in *Teladorsagia circumcincta* in sheep in the UK. *Veterinary Record*, **161**, 535-536.

Wolstenholme, A. J., Fairweather, I., Prichard, R., von Samson-Himmelstjerna, G., & Sangster, N. C. (2004) Drug resistance in veterinary helminths. *Trends in Parasitology*, **20**, (10), 469-476.

Wood, I. B., Amaral, N. K., Bairden, K., Duncan, J. L., Kassai, T., Malone, J. B., Pankavich, J. A., Reinecke, R. K., Slocombe, O., Taylor, S. M., & Vercruysse, J. (1995) World Association for the Advancement of Veterinary Parasitology (WAAVP) 2nd edition of guidelines for evaluating the efficacy of anthelmintics in ruminants (bovine, ovine, caprine). *Veterinary Parasitology*, **58**, (3), 181-213.

Woolaston, R. R. & Baker, R. L. (1996) Prospects of breeding small ruminants for resistance to internal parasites. *International Journal for Parasitology*, **26**, (8-9), 845-855.

Xu, M., Molento, M., Blackhall, W., Ribeiro, P., Beech, R., & Prichard, R. (1998) Ivermectin resistance in nematodes may be caused by alteration of P-glycoprotein homolog. *Molecular and Biochemical Parasitology*, **91**, (2), 327-335.

Zarlenga, D. S. & Gasbarre, L. C. (2009) From parasite genomes to one healthy world: Are we having fun yet? *Veterinary Parasitology*, **163**, (3), 235-249.

Zhang, Y. C., Guo, X. H., Lin, E. T., & Benet, L. Z. (1998) Overlapping substrate specificities of cytochrome P450 3A and P-glycoprotein for a novel cysteine protease inhibitor. *Drug Metabolism and Disposition*, **26**, (4), 360-366.

Zhao, Z. Y., Sheps, J. A., Ling, V., Fang, L. L., & Baillie, D. L. (2004) Expression analysis of ABC transporters reveals differential functions of tandemly duplicated genes in *Caenorhabditis elegans*. *Journal of Molecular Biology*, **344**, (2), 409-417.

Appendices

Appendix 1: Clustal W alignment of the nucleotide sequences of the 11 partial Pgp genes identified from *T. circumcincta*

```
Pgp1            CTAATACGACTCACTATAGGGCAAGAGTGGTATCAACGCAGAGTACGCGGGGATTCCATT 60
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            ------------------------------------------------------------
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            ACGAGAATATAATGTGAGGTGGTGGCGTGGAGTAGTCGGCGTTGTTCAGCAGGAACCAGT 120
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            ------------------------------------------------------------
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            RATATTTCGTGCCACGGTCGCAGAAAATGTTCGAATGGGAGACGACAGCCTAAGTGATGC 180
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            ------------------------------------------------------------
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            AGATGTTGAAGAAGCGTGTAAAGTTGCGAATGCTTTGGGATTCATCAGGAACCTCAGTGA 240
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            ------------------------------------------------------------
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            GGGCTTCAACACTGYGATCGGTGAGGGTGCCGTGCAGCTGTCAGGTGGTCAGAAACAACG 300
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            ------------------------------------------------------------
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
```

```
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            AATCGCTATCGCGAGAGCTCTCGTTAGAAACCCACAGATTTTGCTTCTGGACGAGGCGAC 360
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            ------------------------------------------------------------
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            GAGTGCCCTGGACACAGAAAGCGAACATGCAGTGCAGAAGGCGCTTGACAAGGCACGAGA 420
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            ------------------------------------------------------------
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            AAATCGAACTACTATATGCATAGCGCATCGGCTTTCTACCATCAGAGATGCTGATAAAAT 480
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGGATTGAG 60
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            CATTGTTTTCGACGAGGGACACGTTGTTGAGCAAGGGACACATGATGAACTGATGGCAGT 540
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            TAGACTTCTTGAGCCTGTTTTCGTCTGCTAAGTGTGTCCACTCCTCATTTCGGTGGAGAT 120
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            TGAAGGTGGCATGTACCGGAGTATGGTAAAAGCACAGGCGATTGATAAAGGCGAAGAGGA 600
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            TTGTTCTACGATGTTCACTTCATTGCTAGTCCGGCGAGCTATTTGTACTCGTTGCTATTC 180
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            CACCACACTAGATGATGTGAATCCCACGGAACTCAAACGAGGTGTGTCGCGAGTGGGATC 660
```

```
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            ACTTCCGTCTAGTTCAAAAAGGTTATTCTGTGCTCTACCTCGACAGTCGATTACGCCATC 240
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            GGAGGAAGACACGCATCGTGTCGAAATGGCTAGAGAATCGGCACGACTTCGTCAAAGTAT 720
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            TTTGAAGTTATTAAAAGTCGGTCGAACTCTGTTGAGAAGTCCCTTTGTAAGAAATGTTAC 300
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            GATCAGCACTTCGACACAGGAGCCGGATTGGGAGATAGAAAGTGCCCGTGACAACATGGT 780
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ------------------------------------------------------------
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            GACTGTGGAGACGAAACGTGCCTCGGTTAAACTTAAAAATGCAACATTCAAAGATCTGAA 360
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            TGAAGAAGGCGCTATGGAAGCCTCTCTCCTCGACATTTTTGGCTACGCCAAGCCAGAACT 840
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ----------------------------------------CTAATACGACTCACTATAGGGC 22
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            GGAGATATTCTCTTTAGCTGGACCGTATAAGTGGAGAATTCTTCTAGGC---CTTTCATT 417
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            TCCAATGGCTGGCATTGCCTTGATCTTCACTTTATTRCGTGGACTCACATGGCCGCTGTT 900
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            AAGCAGTGGTATCAACGCAGAGTACGCGGGAGGACTGCCTATGATGTCTATCATTATGGG 82
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            CCTCGGCGTAAGCAGCTCGATTTTCCTGATAACACCTAGAGTGCTGGGAAAACTAATAGA 477
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            CTCTATAATCTACGGAAAACTGTTTCTGTTGCTCTCAAGTCCGGATCCGGAAACTCTCGC 960
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            CAATATTTCACAAAACTTCATGAGTATCACTGGTAACACAACCTCTATCCAACAATTCGA 142
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
```

```
Pgp7            TGAGTTCGACGAAACGAAACGAGCCGCTCCAGGATATGAAGATCTATCGCTCAAGATTGC 537
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            TAGTGGTAACATCCTCAATTCGATCTTCTTTCWATTGCTCGCCATCGCTTCCGGTATCAC 1020
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            ACATGATGTGATCCAGAACTGCCTTAAATATGTCTACCTYGGYTGCGGAGTATTCACGGC 202
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            TAGGTATTTAAAGGACAATCCCATTGCACTGGTTGGGTTGCTGTTTCTTGGAGCTGCTGC 597
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            AACTTTCGCTTCTGGATCACTGTTTGGTATCACCGGCGAAAAGATGGCAATGCGATTACG 1080
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            GGCRACGATTCAGGCAATGTGTTTTCTAACGGTMTGCGARAATCTTGTKAATCAACTCAG 262
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            TATAACTGCAAGAGTGTATTGTATGCACACAGCCGGACAACTCATTATAAATGATCTCCG 657
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            AATGGATGTCTTCAAGAATATAATGCGCCAGGATGCTTCCTATTTTGATAATCCGAATCA 1140
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            AAGACAGTTCTTCAAGTCRATTCTTCGTCAAGACATCACGTGGTTCGACAA------RAA 316
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            TAAGAAAGTTTTCAATTCGGTTCTACGCCAAGACATCGCCTTCTTCGATAA------AAA 711
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            CAACACAGGCAACTTGACAGCCCATTTGGCTTCGGACACACCGAATGTGCAAGCTGCGAT 1200
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            CAATTCAGGAACTCTYGCCACRAAACTRTTCGACAATCTGGAACGAGTCAAAGAGGGWAC 376
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            CAAAGTGGGCGAAATTGTCTCTCGTCTATCAACTGATGCCCTCATCGTGGGATACTCGGT 771
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            CGACCAACGACTCGCAGAGGTGTTGCAAGGTGTTTGCGCTTTGGTGGCTGGAATTGTCGT 1260
Pgp2            ------------------------------------------------------------
Pgp3            ------------------------------------------------------------
Pgp4            MGGYGACAAACTTGGCCTTATGATCCAATTYGTGGCGCAGTTYTTCGGCGGTTTCATCGT 436
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            -GTCCATGAATTTGAGCGATGGTGCAAGAGCTCTCTSACATGCTTAGGCTCAGGCGGTCT 830
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------
```

```
Pgp1            TGCATTTTCCTACGGTTGGAATGTCGCTCCTATTGGCCTAGCCACCGCTCTTCTATTAGT 1320
Pgp2            ---------------------------TCTAATACGACTCACTATAGGGCAAGCAGTGGT 33
Pgp3            ------------------------------------------------------------
Pgp4            GGCGTTCACTTACGACTGGAAACTYACTCTGATCATGATGTCATTGGCYCCATTCATGAT 496
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            AATGGTTTACACATCTCCGGCTCTCTGCCAAGTGATGTTTGTTGTGATTCCGATTATGGT 890
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            TTTCGCTCAATCGTCTGTCGCTCAATATCTGAAGTTCAGAGGACAGAGGGATATGGAGTC 1380
Pgp2            ATCAACGCAGAGTACGCGGGTGCAAATGCAGTTCGGTAAGAAAATGCGGGACACAGAACT 93
Pgp3            ------------------------------------------------------------CT 2
Pgp4            YATTTGTGGAGCWTTCATTGCYAAGTTGATGGCCAGTGCAGCTACTCGGGAAGCCAARAA 556
Pgp5            ------------------------------------------------------------
Pgp6            ------------------------------------------------------------
Pgp7            TGGAACTTTCGCCGTTTTCGGTAAGCTGCAACGGAAGTACACTCTCCAAATGCAGGAGGC 950
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            TGCAGTGGAAGCCAGTCAGATAGTGACGGAATCAATCTCGAATACACGA----ACGATTC 1436
Pgp2            TCTGGAGGAGGCTGGCAAGGTTGCATCGCAAGCTGTAGAGAACATCCGC----ACCGTGC 149
Pgp3            AATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGAGACAGTAA 62
Pgp4            GTACGCCGTGGCGGGAGGAATTGCTGARGAAGTGCTCACHTCWATGAGA----ACTGTKA 612
Pgp5            -------------GCGGCTATAGCCGAGGAAACGTTCTCGTCCATACGT----ACTGTGC 43
Pgp6            ------------------------------------------------------------
Pgp7            TGTAGCTGGGGCTAATCAGGTGGCAACAGAACGATTRTCGAATGTTCGA----ACGGTTA 1006
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            AAGCACTCTGTAAGGAGGGCTACATGTACGAAGCATACTGTGCTGCAGCACAGGAACCTC 1496
Pgp2            ATGCCCTGAATCGGCAAGAGCAATTCCATTTCATGTACTGTGAATATCTGAAAGAGCCTC 209
Pgp3            AGTACTACGGTAATGAGATGTATGAAGTGGATCGTTTCCGAGTTGCAATTGAGAGATTTC 122
Pgp4            TCGCATTCAACGGACAGCCTTACGAATGCGAGAGGTACGAAAAAGCAYTGGAAGACGGCA 672
Pgp5            ATGCATTATGTGGTCATAGAAGAGAGCTTACAAGGTTTGAGGCAGCGTTGGAGAAAGGAC 103
Pgp6            ------------------------------------------------------------
Pgp7            GAATGTTGGTTGCCGAAAAGAAGGAACTAGGTGCGTACGCCGACAAGATTTACGACATCT 1066
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            ATAAACGGGCTCTTGTACGAGGCC-TATGGCAGGCGCTGTCCTTAGCGCTGTCAAATAGT 1555
Pgp2            ATCGGGAGAATCTTTGCCAAACCCACACTTACGGTGGTGTGTTT-GCATTCTCACAATCG 268
Pgp3            AGTTGGC-TGAATGGCGMTCCAATGCATCCCTGGCTCTGCTAAACTTTCTGCAAAACGGC 181
Pgp4            AATCGACMGGAATCAAGAAATCCTTGTACATYGGCATTGGYCTCGGGATCACTTTTCTCA 732
Pgp5            GTAAGACAGGACTTGTCAAATATTTCTATATGGGCGTAGGTGTCGGATTCGGTCAGATGT 163
Pgp6            ------------------------------------------------------------
Pgp7            GGATGATATCGAGAAAAGAAGGTTTGGCTCGTGGTGCCATGTATGGCTCGTTCCAATTCA 1126
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            TTCGTTGTGGTGAACTTCGCTATCGCCTACGCTTTTGGACTATGGCTCATACGAAACGAA 1615
Pgp2            TTACTGTTCTTCATGTATGCGATAGCGTTTTGGATTGGTGCAATCTTCGTGGATAACCAT 328
Pgp3            ACCATTGGTCTCGGRATGACAGCCGGTTCTATACTTGTCGCTTACTTRGTYACAGTTGAC 241
```

```
Pgp4            TCATGTTCTCKTCGTACTGYCTGGCTTTCTGGGTTGGCACGGATTTTGTCTTCAARAATC 792
Pgp5            GTACCTACGTGTCGTACGCCTTGGCTTTTTGGTACGGTAGCACACTGATCATCAGCGACC 223
Pgp6            ------------------------------------------------------------
Pgp7            CGGGGTAT-------------ATCGCCCTTTCAACCGTACTGTTCTACGGAAGTAACCTT 1173
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            TGGAGTACACCATTTATAGTTTTCCAGGTGATAGAAGCGT-TAA-ACATGGCGTCAATGA 1673
Pgp2            AGTATGCAACCGATTGATGTCTWCCGAGTGTTTTTCGCGT-TCATGTTCTGCGGTCAAAT 387
Pgp3            CACGARCTCACGGTAGGAGATTATGTTCTGTTCACCACTTATATTCTACAGCTGTATTCT 301
Pgp4            AAATGCAAG---GAGGAACTGTKATGACGGTATTCTTCTC-CGTGATGATGGGYTCWATG 848
Pgp5            CAACATTGGACCGTGGCCGGATTTTCACTGTGTTCTTCGC-GGTGATGTCCGGCTCATCA 282
Pgp6            ------------------------------------------------------------
Pgp7            ATCAGTCAGGGMTTGCTCACTTACGGTGATTTGTCATCGT-TCTGCCTTTATGCCGTTCT 1232
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            GCGTGAYGATGG-CCGCCTCATATTTTCCGGAGTATATTCGTGCCCGGATATCAGCAGGG 1732
Pgp2            GGTCGGCAACA--TCTCCTCATTCATTCCGGACGTTGTGAAAGCTCGCCTAGCTGCATCG 445
Pgp3            CCATTGAACTTTTTTGGGACAGTGTATAGAACTATTCAAAAGTCTTTCATTGATATGGAG 361
Pgp4            GCGCTCGGWCAGGCYGGACCRCAATTTGCTGTCCTTGGYACAGCTATGGGTGCYGCTGGG 908
Pgp5            GCTCYAGGTACCTGCCTCCCATATCTTAACACCATATCAATCGCCCAAGGTGCTGTTCGA 342
Pgp6            ------------------------------------------------------------
Pgp7            CTGTGCAGCGAGTTTGTCAAACATATCTGGTTTTTTCATCGAAATCATGAAGGGYTTGGG 1292
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            GTCATGTTCACGATGATGCGACAGCGGCCAAAAATCGACA-ATATGAGTCATCAAGGR-- 1789
Pgp2            CTCCTGTTTTACCTTATTGAACACCCTACGGAGATCGATA-ATCTATCGGAGGACGGC-- 502
Pgp3            AATATGTTYGACCTAATGAACGAGGAGGTCGATGTAAGGG---ATGCTCCAAACGCAATT 418
Pgp4            TCTCTYTATCAGATTATCGMKSGGGAACCAGAAATAGACTCCTACTCCTCCGAAGGAGTT 968
Pgp5            AGTGTGCTGAAAGTGATCAATAGTCGTCCAAAGATCGATCCATATTCATTGGATGGGATT 402
Pgp6            ------------------------------------------------------------
Pgp7            CGCCAGTTCTAGACTGTTCGAACTTCGCAATACAGTGCCGCGCATCCCTCTTGAGGGA-- 1350
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            GATAAACCGGCGCTCAAAGGCGACATAGCGCTGAGAAACGTCTATTTCTCGTATCCGGTA 1849
Pgp2            ATTAAGAAGAAACTCTCCGGTCATGTTGTATTCCGCAATGTCTATTTCAATTATCCAACC 562
Pgp3            GAATACCATCCAACGAATGGGCARATTGTAGTGAAAGATCTTACATTCGCCTAC------ 472
Pgp4            AGGCCATCGAATCTCAAAGGAAAAATCACTGTCTCAAATCTGAAGTTCACTTATCCAACA 1028
Pgp5            GTGCTGAACAATATGAGAGGATCTATTCGCTTGAAGAACGTACATTTCTCTTACCCATCT 462
Pgp6            ------------------------------------------------------------
Pgp7            GGTATCAARAAGGGAAATGTCGAGGATRCAATCAGATTTGACCGTGTGGCGTTTGCATAT 1410
Pgp8            ------------------------------------------------------------
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------

Pgp1            CGACGGCGGCAGCTTGTCTTTCAAGGCATGAACCTGAGCGTCCGACATGGACAGACTGTG 1909
Pgp2            AGAAAACAGATCAGAGTACTTCGTGGATTAAGCTTAGAGATTCAACCCGGTACTACGGTA 622
Pgp3            AACGAGAACCGTATAGTRCTGAATGACATCTCGTTYGAAGTCGGCAGTGGACAGTCYATC 532
Pgp4            CGACCAGATGTCCCGATTCTTAAGGGTGTTTCATTTGAAGCGAAACCCGGTGAGACGATA 1088
Pgp5            AGAAGTACACTCCCGATATTGAGAGGCGTGTCCCTCCAAGTGGCAGCCGGCCAAAAGATC 522
Pgp6            ------------------------------------------------------------
Pgp7            CCGGGACGAGATCCGCTGTTCCACAATATCTCCTTCACAGTWCCGGCTGGTCAAATCACT 1470
Pgp8            ------------------------------------------------------------
```

```
Pgp9            ------------------------------------------------------------
Pgp10           ------------------------------------------------------------
Pgp11           ------------------------------------------------------------


Pgp1            GCACTGGTAKGCGCCAGCGGGTGTGGAAAGAGCACCGTGATACAGCTCGTCGAGAGGTAT 1969
Pgp2            GCACTGGTTGGTCGATGTGGGTGTGGRAARAGCACCGTGATGGCRTTRTTGGAAAGGTTT 682
Pgp3            GCATTCGTYGGCCCATCRGGTGGTGGGAARAGTACGTTGATTCGTTTGTTGTTCCGTCTG 592
Pgp4            GCACTGGTAGGTTCTAGTGGRTGTGGAAARAGYACCATAATTCAGCTGCTACTACGGTAC 1148
Pgp5            GCCCTGGTGGGTTCAAGYGGGTGTGGAAAGTCGACAATCGTTAATTTGCTGCTGAGATTT 582
Pgp6            ---------------AGTGGGTGTGGGAAGTCWACAGTGATATCGTTACTTGAAMGGCTT 45
Pgp7            GCAGTGGTCGGATCGTCYGGTAGCGGWAARTCWACCATCGCAAATCTTYTGTTGAGGCTT 1530
Pgp8            ---------------AGCGGGTGTGGGAAAAGCACCTCCATTCAGTTAATCGAACGATTC 45
Pgp9            ---------------AGTGGGTGTGGGAAGAGTACCCTTGTTGGACTACTTCTTCGCTTT 45
Pgp10           ---------------AGYGGGTGYGGGAAGAGYACCCTCGCCAGTTTATTGATGGGCTAT 45
Pgp11           ---------------AGCGGGTGTGGTAAGAGTACGCTTATACGTTGTGTAAACCTGCTG 45
                                    **  * ** **    **            *

Pgp1            TACGACGCCTTGTGTGGCAACGTGAGTATCGACACATACGACATACGTGATCTCTCCATT 2029
Pgp2            TACAATCAAAGCAGGGGTGTCATCACGGTCGATGGCGAAAACATCAGAAACATGAATATC 742
Pgp3            TTCGAGTGTCCTGAGGGAACTATTTTCYTCGACGGGAAAGATGTACGACATCTGAAGTTG 652
Pgp4            TACAATCCTGCAGATGGAAAGATTACAATAGACGGTGTGGAAATCGACAAGATTAATATC 1208
Pgp5            TACGATCCGACCAGRGGSAAGGTAACCATAGATGACATTGACGTCTGCGATCTCAACGTT 642
Pgp6            TACGATGCYTTGGACGGTTCTGTGGAAGTTGATGGYAATGATCTWCGCSAAGTRAAYCCY 105
Pgp7            TACGATCCGGACCGTGGTCATATCATGATTGATGATGTCGATTTGAAACAAACAGATCCT 1590
Pgp8            TACGATCCTGTGGCTGGTGCTGTGCTGTTCGATGAAGTGGATGCTCGTGAACTGAATTTG 105
Pgp9            TACGAACAGAAATCGGGAAAGATCTCTATCGATGACGTTCCCATATCCGAACTGAATATC 105
Pgp10           TACCCGCTAACGGAAGGTGAGATTCGCCTTGATGGTCGTCCATTAAGTTCGCTAAGTCAC 105
Pgp11           GAGCGCCCAACCGAGGGTAGCGTGCTGGCCGATGGCCAGGA-ACTGACCACGCTGTCAGA 104
                              **     *       **

Pgp1            CGTTACGTAAGGGACAACATGGCATTAGTTGGACAGGAGCCGACACTATTCAATGTGACG 2089
Pgp2            CGTGATCTTCGTGAACAAGTGTGCATCGTCAGTCAGGAACCGACGCTGTTYGAYTGTACG 802
Pgp3            GTATCGCTGCGTAAACAAATTGGAATCGTACCGCAGGATACGGTACTGTTCAACGATACG 712
Pgp4            GAATTTCTCCGAAATTACGTTGGAGTAGTGTCACAAGAGCCTATGCTGTTTAACACAACG 1268
Pgp5            CACAAACTTCGTGAACAAATCGGTGTTGTCAGTCAGGAGCCAGTACTATTTGATGGGACA 702
Pgp6            ACTCATCTACGTGCCCATATAGCYTTGGTATCGCAAGAGCCGATTCTTTTCGACAGATCM 165
Pgp7            TCTTACTGGCGACGGCARATTGGAACTGTAGGGCAGGAACCAATATTGTTCTCAACCACG 1650
Pgp8            AGGCACCTTCGATCACAAATCTCCCTAGCGGGCCAGGAGCCTATACTTTTCAACTATTCC 165
Pgp9            GAACATCTGAGGAACATCGTAGGAGTCGTGTCTCAGGAGCCGGCGTTGTTTGCTGATACA 165
Pgp10           AGCGCGCTGCGCCAGGGCGTGGCAATGGTGCAGCAAGATCCGGTGGTGCTGGCGGATACC 165
Pgp11           ATCCGAGTTGACCAAAGCTCGCCGCCAGATTGGTATGATTTTCCAGCATTTTAACCTGCT 164
                                        *      * **          *        *

Pgp1            ATACGCGAGAACATCATGTATGGTCTTG---AC------AAATGTTCTCAGGAGGAAATT 2140
Pgp2            ATCCTTGAGAACATYTGCTATGGACTGG---ACGACCCTAAACCTTCCTACGAGAAGGTG 859
Pgp3            ATACGGTACAATATCCGCTTTGGTCGCC---------CGTCAGCCACAGACGAGGAGGTT 763
Pgp4            ATTGAACAGAATATCCGTTATGGACGTG---------AAAAAGTCACAGATGCTGAAATC 1319
Pgp5            TTATTTGAGAATATTCGAATGGGATATG---------AACACGCTACTATGGAAGAGGTC 753
Pgp6            ATYCGMGACAACATCCTCTACGGYCTY---CCWCCAGGWTCCGTTAGTGAYGCACAAGTK 222
Pgp7            ATTCGGGAGAATATYCTCTATGGTGCTGAGTACCCCGATAAAATCACGGAGGCTAAGATA 1710
Pgp8            ATAAGAGAGAATATTGCGTACGGTTTCG---AG------GAAGCAACTACAGCTCAAATC 216
Pgp9            GTTGAAAATAATATCCGCTTGGGACGGG---------TGGACATT-TCGCGGCAGGAAAT 215
Pgp10           TTCCTCGCCAACGTGACGCTGGGGCGGG---------ATATC--T-CCGAAGAACGCGTC 213
Pgp11           CTCTTCGCG--TACTGTTTTTGGCAACG----------TGGCTCTGCCGCTG--GAGCTG 210
                 *                 **                           *

Pgp1            GAGCACGCCGCACGTCTCGCCAATATTCATGATTTTATTAATAGCCTTCCT---GAGCGT 2197
Pgp2            GTCGCTGCGGCAAAAATGGCMAATATYCACAACTTTGTGCTRGGGCTGCCA---GAGGGT 916
Pgp3            TATGAAGCCGCAAAAGCTGCAATGATTCACGATAAGATCATGACCCTTCCT---GAAGGC 820
Pgp4            ACGGCGGCACTCCGTAAAGCAAACGCCTACAATTTTGTGCAGTCATTCCCT---GACGGA 1376
Pgp5            CAGGAAGCTTGTCGTATTGCAAATGCRGCTGATTTTATCAAACGACTTCCG---GACGGT 810
Pgp6            CACGAAGTCGCTCAACGTGCYAACATTCACAGCTTCATCATAGGMCTGCCT---GATGGM 279
Pgp7            GAAGAGGCTGCAGAGCAGTCAAACGCTCTCGATTTCATCCAAGCGTTCTCT---GAAAAA 1767
Pgp8            GAGGATGCCGCTAGACTGGCCAACGCCCACAACTTCATTGTCAAATTGCCA---GCAGGC 273
Pgp9            GGAAGATTGTTGCAAGATGGCAAATGCGCATGATTTCATCATGAATTTAAGT--CAGGGC 273
Pgp10           TGGCAGGCGCTGGAAACCGTGCAACTGGCGGAGCTGGCGCGT-AGCATGAGC--GACGGT 270
Pgp11           GACAACACACCGAAAGACGAGGTCAAACGTCGCGTGACGGAATTGCTGTCATTGGTTGGT 270
```

```
Pgp1            TACAATACGGTGGTTGGAGCAAAAGGTGCACTGTTGTCGGGCGGTCAAAAGCAGCGCATC 2257
Pgp2            TAYGACACACGTGTGGGCGAGAAAGGCACACARTTATCGGGCGGGCAGAAGCAGCGAATC 976
Pgp3            TATGAAACGATGGTTGGTGAGCGGGGATTGAAGCTATCTGGTGGTGAGAAACAACGTGTC 880
Pgp4            ATTTACACGAACGTTGGTGACCGAGGAACCCAGATGTCTGGTGGCCAAAAGCAACGTATA 1436
Pgp5            TACGGTACYCGCGTTGGSGAGCGTGGTGTGCAGCTGAGTGGTGGACAAAAACAGCGAATA 870
Pgp6            TAYAACACGCGTGCAGGAGAAAAGGGRGCGCAGCTRTCTGGKGGGCAGAAACAACGGATC 339
Pgp7            TTCGATACTATGGTCGGTGAACAGGGATCAATGCTCAGTGGCGGCCAGAAACAGCGAATA 1827
Pgp8            TACGACACGGTCGTCGGTGAGCGAGGTAGCATGCTGTCCGGTGGCCAAAAGCAGCGAATC 333
Pgp9            TACCAGACAAGGATAGGAGACGGTGGAGTTCAGCTGTCTGGTGGTCGAAAGCAGCGAGTT 333
Pgp10           ATTTACACGCCGCTGGGCGAGCAGGGGAATAATCTCTCAGTTGGGCAAAAGCAACTGCTG 330
Pgp11           CTTGGCGATAAGCATGATGGCTACCCGTCGAATCTTTCCGGTGGGCAGAAACAACGTGTG 330
                              *  *              *    *  **    ** ** *   *

Pgp1            GCCATCGCTCGTGCTATCGTCAGAGATCCAAAGATACTACTACTTGACGAAGCCACCTCC 2317
Pgp2            GCGATCGCTCGGGCGTTGATYCGAGATCCACCTATACTTCTTCTTGACGAGGCAACCWSC 1036
Pgp3            GCGATTGCACGAACCATTCTGAAGAAACCACAGTTTATATTTTTGGACGAAGCAACCAGC 940
Pgp4            GCCATTGCCCGCGCTTTGGTCAGAGACCCAAAAATTCTTCTACTCGACGAAGCCACATCC 1496
Pgp5            GCCATTGCTCGTGCRATYATCAARAAYCCACGTATACTYTTGCTCGAYGAAGCAACCWSC 930
Pgp6            GCCATYGCACRTGCACTTGTTCGGAATCCAAAAATCCTACTACTTGACGARGCCACAWSY 399
Pgp7            GCAATTGCGAGAGCACTGGTCACGAATCCAAGGATTCTAATCCTCGATGAAGCCACTAGC 1887
Pgp8            GCAATAGCTCGAGCTGTGATCCGTAACCCGAAAGTTCTTCTATTGGACGAGGCAACCAGT 393
Pgp9            GCTATAGCTCGAGCTCTGGTGAGAAATCCGCGAATTCTGCTACTGGACGAGGCCACATCC 393
Pgp10           GCACTGGCGCGCGTGCTGGTCGAGACGCCGCAAATCCTGATCCTTGATGAAGCCACATCC 390
Pgp11           GCAATTGCCCGTGCGTTAGCCAGCAATCCCAAAGTATTGCTGTGTGACGAAGCAACCAGT 390
                **  * **        *          **     *  *  *    ** ** ** **

Pgp1            GCCCT 2322
Pgp2            GCCCT 1041
Pgp3            GCCCT 945
Pgp4            GCCCT 1501
Pgp5            GCCCT 935
Pgp6            GCCCT 404
Pgp7            GCCCT 1892
Pgp8            GCCCT 398
Pgp9            GCCCT 398
Pgp10           GCCCT 395
Pgp11           GCCCT 395
                *****
```

Appendix 2: Clustal W alignment of the translated protein sequences of the 11 partial Pgp genes identified from *T. circumcincta*

PGP7 has two alignments due to the frame shift that occurs in the nucleotide sequence. Rf after the Pgp name stands for reading frame, with the end number denoting whether the reading frame was 1, 2 or 3.

```
Pgp1Rf2         YDSLGKSGINAEYAGIPLREYNVRWWRGVVGVVQQEPVIFRATVAENVRMGDDSLSDADV 60
Pgp2Rf2         ------------------------------------------------------------
Pgp3Rf2         ------------------------------------------------------------
Pgp4Rf3         ------------------------------------------------------------
Pgp5Rf1         ------------------------------------------------------------
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         ------------------------------------------------------------
Pgp7Rf1         ------------------------------------------------------------
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         EEACKVANALGFIRNLSEGFNTXIGEGAVQLSGGQKQRIAIARALVRNPQILLLDEATSA 120
Pgp2Rf2         ------------------------------------------------------------
Pgp3Rf2         ------------------------------------------------------------
Pgp4Rf3         ------------------------------------------------------------
Pgp5Rf1         ------------------------------------------------------------
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         ------------------------------------------------------------
Pgp7Rf1         ------------------------------------------------------------
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         LDTESEHAVQKALDKARENRTTICIAHRLSTIRDADKIIVFDEGHVVEQGTHDELMAVEG 180
Pgp2Rf2         ------------------------------------------------------------
Pgp3Rf2         ------------------------------------------------------------
Pgp4Rf3         ------------------------------------------------------------
Pgp5Rf1         ------------------------------------------------------------
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         ------------------------------------------------------YDSLGK 6
Pgp7Rf1         ----------------------------LIRLTIGQAVVSTQSTRGIETSACFRLLSVST 32
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         GMYRSMVKAQAIDKGEEDTTLDDVNPTELKRGVSRVGSEEDTHRVEMARESARLRQSMIS 240
Pgp2Rf2         ------------------------------------------------------------
Pgp3Rf2         ------------------------------------------------------------
Pgp4Rf3         ------------------------------------------------------------
Pgp5Rf1         ------------------------------------------------------------
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         QWYQRRVRGGLSRLLEPVFVCVCPLLISVEICSTMFTSLLVRRAICTR------------ 54
Pgp7Rf1         PHFGGDLFYDVHFIASPASYLYSLLFTSVFKKVILCSTSTVDYAIFEVIKSR-------- 84
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         TSTQEPDWEIESARDNMVEEGAMEASLLDIFGYAKPELPMAGIALIFTLLRGLTWPLFSI 300
Pgp2Rf2         ------------------------------------------------------------
Pgp3Rf2         ------------------------------------------------------------
Pgp4Rf3         ------------------------NTTHYRASSGINAEYAGGLPMMSIIMGNISQNFMSI 36
```

```
Pgp5Rf1         ------------------------------------------------------------
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         ------------------------CYSLPSSSKRLFCALPRQSITPSLKLLKVGRTLLRS  90
Pgp7Rf1         ---------SNSVEKSLCKKCYDCGDETCLGTKCNIQRSEGDILFSWTVVENSSRPFIPR 135
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         IYGKLFLLLSSPDPETLASGNILNSIFFXLLAIASGITTFASGSLFGITGEKMAMRLRMD 360
Pgp2Rf2         ------------------------------------------------------------
Pgp3Rf2         ------------------------------------------------------------
Pgp4Rf3         TG-------NTTSIQQFEHDVIQNCLKYVYLGCGVFTAATIQAMCFLTVCENLVNQLRRQ  89
Pgp5Rf1         ------------------------------------------------------------
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         PFVRNVTTVETKRASVKLKNATFKDLKEIFSLAGPYKWRILLGLSFLGVSSSIFLITPRV 150
Pgp7Rf1         RKQLDFPDNTSAGKTNRVRRNETSRSRIRSIAQDCVFKGQSHCTGWVAVSWSCCYNCKSV 195
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         VFKNIMRQDASYFDNPNHNTGNLTAHLASDTPNVQAAIDQRLAEVLQGVCALVAGIVVAF 420
Pgp2Rf2         ------------------------------------------------------------
Pgp3Rf2         ------------------------------------------------------------
Pgp4Rf3         FFKSILRQDITWFDKNNSGT--LATKLFDNLERVKEGTGDKLGLMIQFVAQFFGGFIVAF 147
Pgp5Rf1         ------------------------------------------------------------
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         LGKLIDEFDETKRAAPGYEDLSLKIARYLKDNPIALVGLLFLGAAAITARVYCMHTAGQL 210
Pgp7Rf1         LYAHSRTTHYKSPESFQFGSTPRHRLLRKQSGRNCLSSINCPHRGILGVHEFERWCKSSL 255
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         SYGWN----------VAPIGLATALLLVFAQSSVAQYLKFRGQRDMESAVEASQIVTESI 470
Pgp2Rf2         ----------------LIRLTIGQAVVSTQSTRVQMQFGKKMRDTELLEEAGKVASQAV  43
Pgp3Rf2         ------------------------------------------------YDSLGKQWYQRR  12
Pgp4Rf3         TYDWK----------LTLIMMSLAPFMIICGAFIAKLMASAATREAKKYAVAGGIAEEVL 197
Pgp5Rf1         ----------------------------------------------------AAIAEETF   8
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         IINDLRKKVFNSVLRQDIAFFDKNKVGEIVSRLSTDALIVGYSVSMNLSDGARALXHAAQ 270
Pgp7Rf1         TCLGSGGLMVYTSPALCQVMFVVIPIMVGTFAVFGKLQRKYTLQMQEAVAGANQVATERL 315
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         SNTRTIQALCKEGYMYEAYCAAAQEPHKRALVRGLWQALSLALSNSFVVVNFAIAYAFG- 529
Pgp2Rf2         ENIRTVHALNRQEQFHFMYCEYLKEPHRENLCQTHTYGGVFAFSQSLLFFMYAIAFWIG- 102
Pgp3Rf2         VRGETVKYYGNEMYEVDRFRVAIERFQLAEWRSNASLALLNFLQNGTIGLGMTAGSILVA  72
Pgp4Rf3         TSMRTVIAFNGQPYECERYEKALEDGKSTGIKKSLYIGIGLGITFLIMFSSYCLAFWVGT 257
Pgp5Rf1         SSIRTVHALCGHRRELTRFEAALEKGRKTGLVKYFYMGVGVGFGQMCTYVSYALAFWYGS  68
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         AVWFTHLRLSAKCLLFRLWLELSPFSVSCNGSTLSKCRRLLGLIRWQQNDXRMFERLECW 330
Pgp7Rf1         SNVRTVRMLVAEKKELGAYADKIYDIWMISRKEGLARGAMYGSFQFTGYIALSTVLFYG- 374
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------

Pgp1Rf2         LWLIRNEWSTPFIVFQVIEALNMASMSVXMAASYFPEYIRARISAGVMFTMMRQRPKIDN 589
Pgp2Rf2         AIFVDNHSMQPIDVXRVFFAFMFCGQMVGNISSFIPDVVKARLAASLLFYLIEHPTEIDN 162
Pgp3Rf2         YLVTVDHELTVGDYVLFTTYILQLYSPLNFFGTVYRTIQKSFIDMENMFDLMNEEVDVR- 131
Pgp4Rf3         DFVFKN-QMQGGTVMTVFFSVMMGSMALGQAGPQFAVLGTAMGAAGSLYQIIXXEPEIDS 316
```

```
Pgp5Rf1         TLIISDPTLDRGRIFTVFFAVMSGSSAXGTCLPYLNTISIAQGAVRSVLKVINSRPKIDP 128
Pgp6Rf1         ------------------------------------------------------------
Pgp7Rf2         LPKRRNVRTPTRFTTSGYREKKVWLVVPCMARSNSRGISPFQPYCSTEVTLSVRXCSLTV 390
Pgp7Rf1         SNLISQGLLTYGDLSSFCLYAVLCAASLSNISGFFIEIMKGLGASSRLFELRNTVPRIP- 433
Pgp8Rf1         ------------------------------------------------------------
Pgp9Rf1         ------------------------------------------------------------
Pgp10Rf1        ------------------------------------------------------------
Pgp11Rf1        ------------------------------------------------------------


Pgp1Rf2         MSHQGDKPA-LKGDIALRNVYFSYPVRRRQLVFQGMNLSVRHGQTVALVXA---SGCGKS 645
Pgp2Rf2         LSEDGIKKK-LSGHVVFRNVYFNYPTRKQIRVLRGLSLEIQPGTTVALVGR---CGCGKS 218
Pgp3Rf2         DAPNAIEYHPTNGQIVVKDLTFAYNENR--IVLNDISFEVGSGQSIAFVGP---SGGGKS 186
Pgp4Rf3         YSSEGVRPSNLKGKITVSNLKFTYPTRPDVPILKGVSFEAKPGETIALVGS---SGCGKX 373
Pgp5Rf1         YSLDGIVLNNMRGSIRLKNVHFSYPSRSTLPILRGVSLQVAAGQKIALVGS---XGCGKS 185
Pgp6Rf1         ------------------------------------------------------SGCGKS 6
Pgp7Rf2         ICHRSAFMPFSVQRVCQTYLVFSSKSRXWAPVLDCSNFAIQCRASLLREVSXREMSRXQS 450
Pgp7Rf1         -LEGGIKKGNVEDXIRFDRVAFAYPGRD--PLFHNISFTVPAGQITAVVGS---SGSGKS 487
Pgp8Rf1         ------------------------------------------------------SGCGKS 6
Pgp9Rf1         ------------------------------------------------------SGCGKS 6
Pgp10Rf1        ------------------------------------------------------XGCGKX 6
Pgp11Rf1        ------------------------------------------------------SGCGKS 6
                                                                       .  :

Pgp1Rf2         TVIQLVERYYDALCGNVSIDTYDIRDLSIRYVRDNMA----------LVGQEPTLFNVT 694
Pgp2Rf2         TVMALLERFYNQSRGVITVDGENIRNMNIRDLREQVC----------IVSQEPTLFDCT 267
Pgp3Rf2         TLIRLLFRLFECPEGTIFXDGKDVRHLKLVSLRKQIG----------IVPQDTVLFNDT 235
Pgp4Rf3         TIIQLLLRYYNPADGKITIDGVEIDKINIEFLRNYVG----------VVSQEPMLFNTT 422
Pgp5Rf1         TIVNLLLRFYDPTRGKVTIDDIDVCDLNVHKLREQIG----------VVSQEPVLFDGT 234
Pgp6Rf1         TVISLLERLYDALDGSVEVDGNDLRXVNPTHLRAHIA----------LVSQEPILFDRS 55
Pgp7Rf2         DLTVWRLHIRDEIRCSTISPSQXRLVKSLQWSDRXVAXXXPSQIFCGFTIRTVVISLMMS 510
Pgp7Rf1         TIANLLLRLYDPDRGHIMIDDVDLKQTDPSYWRRQIG----------TVGQEPILFSTT 536
Pgp8Rf1         TSIQLIERFYDPVAGAVLFDEVDARELNLRHLRSQIS----------LAGQEPILFNYS 55
Pgp9Rf1         TLVGLLLRFYEQKSGKISIDDVPISELNIEHLRNIVG----------VVSQEPALFADT 55
Pgp10Rf1        TLASLLMGYYPLTEGEIRLDGRPLSSLSHSALRQGVA----------MVQQDPVVLADT 55
Pgp11Rf1        TLIRCVNLLERPTEGSVLADGQELTTLSESELTKARR----------QIGMIFQHFNLL 55
                                        .                          :

Pgp1Rf2         IRENIMYGLD---KCSQEEIEHAARLANIHDFINSLPER----YNTVVGAKGALLSGGQK 747
Pgp2Rf2         ILENICYGLDD-PKPSYEKVVAAAKMANIHNFVLGLPEG----YDTRVGEKGTQLSGGQK 322
Pgp3Rf2         IRYNIRFGRP---SATDEEVYEAAKAAMIHDKIMTLPEG----YETMVGERGLKLSGGEK 288
Pgp4Rf3         IEQNIRYGRE---KVTDAEITAALRKANAYNFVQSFPDG----IYTNVGDRGTQMSGGQK 475
Pgp5Rf1         LFENIRMGYE---HATMEEVQEACRIANAADFIKRLPDG----YGTRVGERGVQLSGGQK 287
Pgp6Rf1         IRDNILYGLPP-GSVSDAQVHEVAQRANIHSFIIGLPDG----YNTRAGEKGAQLSGGQK 110
Pgp7Rf2         INKQILLTGDGXLELGRNQYCSQPRFGRIXSMVLSTPIKSRRLRKRLQSSQTLSISSKRS 570
Pgp7Rf1         IRENILYGAEYPDKITEAKIEEAAEQSNALDFIQAFSEK----FDTMVGEQGSMLSGGQK 592
Pgp8Rf1         IRENIAYGFE---EATTAQIEDAARLANAHNFIVKLPAG----YDTVVGERGSMLSGGQK 108
Pgp9Rf1         VENNIRLGRV---DISRQEMEDCCKMANAHDFIMNLSQG----YQTRIGDGGVQLSGGRK 108
Pgp10Rf1        FLANVTLGRD----ISEERVWQALETVQLAELARSMSDG----IYTPLGEQGNNLSVGQK 107
Pgp11Rf1        SSRTVFGNVALPLELDNTPKDEVKRRVTELLSLVGLGDK----HDGYPSN----LSGGQK 107
                    :                   .                       .     :*  ..

Pgp1Rf2         QRIAIARAIVRDPKILLLDEATSA------------------ 771
Pgp2Rf2         QRIAIARALIRDPPILLLDEATXA------------------ 346
Pgp3Rf2         QRVAIARTILKKPQFIFLDEATSA------------------ 312
Pgp4Rf3         QRIAIARALVRDPKILLLDEATSA------------------ 499
Pgp5Rf1         QRIAIARAIIKNPRILLLDEATXA------------------ 311
Pgp6Rf1         QRIAIAXALVRNPKILLLDEATXA------------------ 134
Pgp7Rf2         LKNSILWSVNRDQCSVAARNSEQLREHWSRIQGFSSMKPLAP 612
Pgp7Rf1         QRIAIARALVTNPRILILDEATSA------------------ 616
Pgp8Rf1         QRIAIARAVIRNPKVLLLDEATSA------------------ 132
Pgp9Rf1         QRVAIARALVRNPRILLLDEATSA------------------ 132
Pgp10Rf1        QLLALARVLVETPQILILDEATSA------------------ 131
Pgp11Rf1        QRVAIARALASNPKVLLCDEATSA------------------ 131
                   ::   :      :   ::
```

Appendix 3: Clustal W alignment of the nucleotide sequences of the 3 partial CYP genes identified from *T. circumcincta*

```
CYP1TC          ------------------------------------------------------------
CYP2TC          CAAAGAAAAGACGATATCGCATCTGGACAACACACGTTAGATGGTGATGGCAATGATTTT 60
CYP3TC          ------------------------------------------------------------


CYP1TC          ------------------------------------------------------------
CYP2TC          GTTGATGCCTTCTTTATCAAGATGGAGAAAGATAACAAAGAAGGCCGATCTCCCAACGAA 120
CYP3TC          ------------------------------GATAGCAATATGTTTGCCGTCGTGATTCT 29


CYP1TC          ----------------------------------TTTGATCTTTG-GATTGCAGGTCAGGA 26
CYP2TC          TCATACAAAGAAGAGGAGCTGCTCTATGATATATTTGATCTTTG-GATTGCTGGTCAGGA 179
CYP3TC          TTTCCTGTTATCGTTTGTCGTCTACTACGTATGGTTGTTCTATGCGAACGTGAAACGTTA 89
                                              *** *** ** **  *     *   *

CYP1TC          AACCACAATAACTACGCTGGAGTGGGCGTTTTCCTACTTACTACTTAATCCACAGGTAAC 86
CYP2TC          AACGACAACAATCACAATATTGTGGGGAATGATGCATCTCATCAAAAATCCTGAGGTCAT 239
CYP3TC          CCCCAAAGGACCAACACCTCTGCCAGTTGTCGGAAATCTTTTG-TCGATCAAC------T 142
                  * * *  *   **      *   *   *     *  *  *     ***

CYP1TC          AGCACGGGTAGAGGAGGAACTGCTCACGCTCACAAAGGGACAACGCCTACTCTCTATCGT 146
CYP2TC          GCATAAAATTCGAACAGAGTTGAACACAGTCACTGGTGGCAACAGATTGATCTCCTTATC 299
CYP3TC          TGCGGAAGCTCAATGATGACTTCGTGCACATATCAAAGGATTATGG-TGATATCTTTACT 201
                                *     *    *     **     *  *  * **  *

CYP1TC          GGACAGACCGAACACTCCATACTACAACGCCACTTTGAACGAAGTCCATAGATGCGCTAC 206
CYP2TC          CGACAGAGAACACACTCACTACCTGAACTGGACAATACTGGAAATACATCGGCTTGCGTC 359
CYP3TC          GTTTGGC-------TTCCAAAGCCGTATGTGGTGAT-CATGAACTACGACAAT--ATTAA 251
                     *         **   *     *        *    *** * *

CYP1TC          ATTAGTGCCCATGAATTTGTGGCGTGATACCTCTGAAGATACTGTTGTTGGATCTTATGT 266
CYP2TC          TATTCTGAATTTGAATCTGTTCCGTAAAACCAAGGAAAATCTCGTTGTCGGTGGACATTC 419
CYP3TC          GGAGGCGTTCGCGAAAAAAGGAGATGACTTTAACGGACGGTCTGGTCTCTTTCCTGACAC 311
                      *     ***         * *       * *      * * *        *

CYP1TC          GAT-TCCCAAAGGAACCGCAATCACAGCTCAAATCTCACTCATTATGACTGACGAGAAAT 325
CYP2TC          TGT-CCCAAAAGGCACTCCAATTGCAGCTGAATTGTCGGTAATAATGCAAGATGAACAAC 478
CYP3TC          ATTGTACCAAAGCATGCACAATGGAGGCATTGTTTTCTCCCAGGGAGACCTTTGGAAAGA 371
                  *   * ****      ****    **     * **    *    *      *   *

CYP1TC          ACTTCGAAAATTSCAAAGAGTTCAATCCTGACCGYTATTTCRATGGCAAYAA----ATTG 381
CYP2TC          GCTTCGAGGAGGCTAACAAGTTCAATCCAGAACGGTATAAACATGGCGGTAAG-ACACTG 537
CYP3TC          ACAACGACG--TGTATCACTCCAAATACTTCGCGATTTTGGAATGGGAAAGAGCGCAATG 429
                 *  ***       *        *** *    ** * *    ****     *    * **

CYP1TC          GAGCAAAAGGTGGTCGCTTTCGGTTTAGGAAAGCGATCCTGTTTGGGTGAATCACTTGCT 441
CYP2TC          GAACAGCAAGTGATTCCTTTTGGTCTTGGTAAACGATCGTGTATTGGCGAGAATTTGGCA 597
CYP3TC          GAGGAACAGGTT-TCTCTTTCTGCACAAGAATT---TCTCAATCATATGAACAGCAT-CA 484
                **  *  * **  *  ****  *     * *     **          **        *

CYP1TC          CARGCAGAACTSTATCTGATCAYYGCRAAC---TTCCTGCTTCKCTACAAAATCTCGGCA 498
CYP2TC          AGAGCGGAAATTTTCTTGATTTTGGCCAAC---TTGATTTCACGCTACGAAAT------- 647
CYP3TC          AAAACAAGGATGAAGTTGATCTACGAAGACCATTACAGGTTTTTGTCGGGAACGTTATCA 544
                    *     *     ****    *   **   *           *    **

CYP1TC          GATCCTCTACATATGCCAAGGACGAAAGCCACAAATGAAATGGGAACCATGAGAAAAC-- 556
CYP2TC          ------------------------------------------------------------
CYP3TC          ACCGAACATTATTCGGATACGGATACAGTTACGATAACAGCGATCGCCTTATGAAAGTGG 604


CYP1TC          CATACCCCTATCACATGCARTT---TGAACKTCGATAAAATGCCTTGACTTTTTTGCTTT 613
CYP2TC          ------------------------------------------------------------
CYP3TC          CGGATACTTTAGCCATATTATTCGATGAAACGAGGAGCAGCAAGATGGTATTTCTAGCTC 664
```

```
CYP1TC          TCGTTATGCCTTGCTCATTTTTTGTTGTTGTTTATTATGCCTGTATTGTCTCAGTTGTTG 673
CYP2TC          ------------------------------------------------------------
CYP3TC          AACTTATGCCTTTCATA---------------------------------------------- 681


CYP1TC          TCTATTTTTTCAACGTGAGCCACCTTCATTTTCGCTGAAAAAACCTGCGAG 724
CYP2TC          ---------------------------------------------------
CYP3TC          ---------------------------------------------------
```

Appendix 4: Clustal W alignment of the translated protein sequences of the 3 CYP genes identified from *T. circumcincta*.

Rf after the CYP name stands for reading frame, with the end number denoting whether the reading frame was 1, 2 or 3.

```
CYP1TcRf1       ------------------------------------------------------FDL 3
CYP2TcRf1       QRKDDIASGQHTLDGDGNDFVDAFFIKMEKDNKEGRSPNESY------KEEELLYDIFDL 54
CYP3TcRf1       -DSNMFAVVILFLLSFVVYYVWLFYANVKRYPK-GPTPLPVVGNLLSINLRKLNDDFVHI 58
                                                                    ..:

CYP1TcRf1       WIAGQETTITTLEWAFSYLLLNPQVTARVEEELLTLTKGQRLLS--IVDRPNTPYYNATL 61
CYP2TcRf1       WIAGQETTTITILWGMMHLIKNPEVMHKIRTELNTVTGGNRLIS--LSDREHTHYLNWTI 112
CYP3TcRf1       SKDYGDIFTVWLPKPYVVIMNYDNIKEAFAKKGDDFNGRSGLFPDTLYQSMHNGGIVFSQ 118
                     :     :      ::   ::   .  :   ..  . *:.  : :  :.     :

CYP1TcRf1       NEVHRCATLVPMNLWRDTSEDTVVGSYVIPKGTAITAQISLIMTDEKYFENXKEFNPDRY 121
CYP2TcRf1       LEIHRLASILNLNLFRKTKENLVVGGHSVPKGTPIAAELSVIMQDEQRFEEANKFNPERY 172
CYP3TcRf1       GDLWKEQRRVSLQILR----DFGMGKSAMEEQVSLSAQEFLNHMNSIKNKDEVDLRRPLQ 174
                 :: :    : ::: *    :  :*   : : ..::*:  :   :.   ::  .:.

CYP1TcRf1       FX-GNKLEQKVVAFGLGKRSCLGESLAQAELYLIXANFLLXYKISADPLHMPRTKATNEM 180
CYP2TcRf1       KHGGKTLEQQVIPFGLGKRSCIGENLARAEIFLILANLISRYE----------------- 215
CYP3TcRf1       VFVGNVINRTLFGYGYSYDNSD-RLMKVADTLAILFDETRSSKMVFLAQLMPFI------ 227
                   *: ::: :. :* .  ..  . :  *:   *  :     :

CYP1TcRf1       GTMRKPYPYHMQFEXRNALTFLLFVMPCSFFVVVYYACIVSVVVYFFNVSHLHFRKNLR 239
CYP2TcRf1       ------------------------------------------------------------
CYP3TcRf1       ------------------------------------------------------------
```

Appendix 5: Table showing the 355 clusters that were statistically significant under the liberal model following bioinformatic and statistical analysis of the sequences generated using Roche 454 sequencing

The columns show the cluster identity used in the analysis (**cluster**), the number of reads within that cluster (**U reads** = unexposed and **T reads** = exposed), whether gene expression up (+) or down (-) regulated compared to unexposed reads (**change**) and what the putative cluster identity was (**BLAST identity**).

Cluster	U reads	T reads	Change	BLAST Identity
803	0	15	+	14-3-3 protein [*Caenorhabditis brenneri*] ACE74683.1
552	13	40	+	24 kDa excretory/secretory protein [*Haemonchus contortus*] AAV83999.1
1661	0	19	+	60S ribosomal protein L3 [*Caenorhabditis briggsae*] Q9NBK4
691	0	19	+	Abnormal cell lineage family member (lin-2) [*Caenorhabditis elegans*] NP_001024588.1
798	0	9	+	Abnormal cell Lineage family member (lin-53) [*Caenorhabditis elegans*] NP_492552.1
900	0	11	+	ADAM (disintegrin plus metalloprotease) family member (adm-4) [*Caenorhabditis elegans*] NP_509318.1
2	77	157	+	ATP synthase F0 subunit 6 [*Cooperia oncophora*] NP_851326.1
1105	0	13	+	ATP synthase subunit family member (atp-2) [*Caenorhabditis elegans*] NP_498111.2
1729	0	13	+	Bifunctional glyoxylate cycle protein [*Caenorhabditis elegans*] AAA85857.1
840	0	9	+	cDNA sequence BC017158 [*Brugia malayi*] XP_001896227.1
1771	0	12	+	Chain A, Glutathione Transferase-2, Apo Form [*Heligmosomoides polygyrus*] 1TW9_A
1345	0	9	+	Chaperonin Containing TCP-1 family member (cct-8) [*Caenorhabditis elegans*] NP_500035.2
858	0	15	+	Chitin-Binding Domain protein family member (cbd-1) [*Caenorhabditis elegans*] NP_502145.2
455	5	39	+	Conserved Germline Helicase family member (cgh-1) [*Caenorhabditis elegans*] NP_498646.1
1653	0	10	+	C-type single domain activation associated secreted protein [*Ostertagia ostertagi*] CAO00416.1
410	982	1531	+	Cytochrome c oxidase subunit I [*Cooperia oncophora*] NP_851331.1
117	306	538	+	Cytochrome c oxidase subunit I [*Cooperia oncophora*] NP_851331.1

Cluster	U reads	T reads	Change	BLAST Identity
99	422	703	+	Cytochrome c oxidase subunit III [*Cooperia oncophora*] NP_851329.1
1260	0	9	+	Delta6-fatty-acid-desaturase [*Caenorhabditis elegans*] AAC15586.1
1133	0	9	+	DNaJ domain (prokaryotic heat shock protein) family member (dnj-13) [*Caenorhabditis elegans*] NP_496468.1
649	124	239	+	Elongation FacTor family member (eft-2) [*Caenorhabditis elegans*] NP_492457.1
1187	0	13	+	Enhancer of PolyComb-like family member (epc-1) [*Caenorhabditis elegans*] NP_499642.2
652	16	73	+	Eukaryotic translation elongation factor 1A [*Ancylostoma ceylanicum*] DAA05878.1
918	0	9	+	Exportin (nuclear export receptor) family member (xpo-1) [*Caenorhabditis elegans*] NP_741567.1
874	0	12	+	F-box/LRR-repeat protein, putative [*Pediculus humanus corporis*] EEB16685.1
1728	0	9	+	GF24664 [*Drosophila ananassae*] XP_001956296.1
1266	0	11	+	GJ11426 [*Drosophila virilis*] XP_002048294.1
1110	0	10	+	GJ11487 [*Drosophila virilis*] XP_002048175.1
1149	0	9	+	GK13711 [*Drosophila willistoni*] XP_002072638.1
934	0	31	+	Globin-like ES protein F6 [*Ostertagia ostertagi*] CAD20463.1
1082	0	9	+	Groucho/TLE N-terminal Q-rich domain containing protein [*Brugia malayi*] XP_001901168.1
969	0	9	+	GTP-binding nuclear protein RAN/TC4 [*Brugia malayi*] XP_001900408.1
1671	0	16	+	Heat shock protein 20 [*Haemonchus contortus*] AAN05752.1
1739	0	27	+	Heat shock protein 20 [*Ostertagia ostertagi*] CAG25499.1
358	0	25	+	Heat shock protein 20 [*Ostertagia ostertagi*] CAG25499.1
60	12	38	+	Heat shock protein 20 [*Ostertagia ostertagi*] CAG25499.1
630	106	256	+	Heat Shock Protein family member (hsp-16.2) [*Caenorhabditis elegans*] NP_503507.1
765	0	10	+	Histone family member (his-35) [*Caenorhabditis elegans*] NP_505463.1
925	0	9	+	Hypothetical protein [*Brugia malayi*] XP_001896639.1

Cluster	U reads	T reads	Change	BLAST Identity
1705	0	11	+	Hypothetical protein [*Leishmania infantum* JPCM5] XP_001464917.1
522	23	64	+	Hypothetical protein [*Tetrahymena thermophila* SB210] XP_001028745.1
773	0	9	+	Hypothetical protein B0365.1 [*Caenorhabditis elegans*] NP_506267.1
1636	0	12	+	Hypothetical protein BRAFLDRAFT_124962 [*Branchiostoma floridae*] XP_002225324.1
1011	0	9	+	Hypothetical protein C05C8.2 [*Caenorhabditis elegans*] NP_504837.1
893	0	11	+	Hypothetical protein C07E3.9 [*Caenorhabditis elegans*] NP_496228.1
1066	0	10	+	Hypothetical protein C28H8.3 [*Caenorhabditis elegans*] NP_498283.1
1740	0	13	+	Hypothetical protein C39D10.7 [*Caenorhabditis elegans*] NP_509334.2
1044	0	11	+	Hypothetical protein C50F4.14 [*Caenorhabditis elegans*] NP_505467.2
1703	0	14	+	Hypothetical protein CBG00021 [*Caenorhabditis briggsae* AF16] XP_001666234.1
1710	0	10	+	Hypothetical protein CBG01069 [*Caenorhabditis briggsae* AF16] XP_001672943.1
454	236	395	+	Hypothetical protein CBG01578 [*Caenorhabditis briggsae* AF16] XP_001671326.1
1002	0	16	+	Hypothetical protein CBG01670 [*Caenorhabditis briggsae* AF16] XP_001671249.1
266	2	32	+	Hypothetical protein CBG02470 [*Caenorhabditis briggsae* AF16] XP_001679726.1
623	0	12	+	Hypothetical protein CBG02815 [*Caenorhabditis briggsae* AF16] XP_001679448.1
457	162	348	+	Hypothetical protein CBG03358 [*Caenorhabditis briggsae* AF16] XP_001670763.1
1719	0	9	+	Hypothetical protein CBG07973 [*Caenorhabditis briggsae*] CAP27887.1
502	336	349	+	Hypothetical protein CBG09272 [*Caenorhabditis briggsae* AF16] XP_001674277.1
275	130	301	+	Hypothetical protein CBG12450 [*Caenorhabditis briggsae* AF16] XP_001668445.1
788	0	10	+	Hypothetical protein CBG12789 [*Caenorhabditis briggsae* AF16] XP_001668986.1
687	0	9	+	Hypothetical protein CBG13605 [*Caenorhabditis briggsae* AF16] XP_001669195.1
839	0	12	+	Hypothetical protein CBG17994 [*Caenorhabditis briggsae* AF16] XP_001665645.1
464	84	284	+	Hypothetical protein CBG18692 [*Caenorhabditis briggsae* AF16] XP_001669248.1

Cluster	U reads	T reads	Change	BLAST Identity
1648	0	14	+	Hypothetical protein CBG18692 [*Caenorhabditis briggsae* AF16] XP_001669248.1
1669	0	15	+	Hypothetical protein CBG19186 [*Caenorhabditis briggsae* AF16] XP_001674556.1
181	0	16	+	Hypothetical protein CBG21168 [*Caenorhabditis briggsae* AF16] XP_001666373.1
1094	0	9	+	Hypothetical protein Csp3_JD02.002 [*Caenorhabditis* sp. PS1010] ACI49174.1
1690	0	13	+	Hypothetical protein Csp3_JD03.008 [*Caenorhabditis* sp. PS1010] ACI49199.1
1633	0	9	+	Hypothetical protein Csp3_JD05.001 [*Caenorhabditis* sp. PS1010] ACI49216.1
910	0	11	+	Hypothetical protein F25B4.6 [*Caenorhabditis elegans*] NP_504496.1
571	0	10	+	Hypothetical protein F41C3.5 [*Caenorhabditis elegans*] NP_494846.1
1664	0	15	+	Hypothetical protein F42A8.3 [*Caenorhabditis elegans*] NP_495993.2
1763	0	30	+	Hypothetical protein L3ni51 [*Dictyocaulus viviparus*] AAT02162.1
933	0	9	+	Hypothetical protein Y46G5A.13 [*Caenorhabditis elegans*] NP_496718.1
1190	0	10	+	Hypothetical protein Y57G11C.9 [*Caenorhabditis elegans*] NP_502785.2
1229	0	9	+	Hypothetical protein Y69E1A.1 [*Caenorhabditis elegans*] NP_502038.1
1078	0	9	+	Hypothetical protein ZC416.2 [*Caenorhabditis elegans*] NP_500388.2
1344	0	9	+	Metalloprotease IV [*Ostertagia ostertagi*] AAS47831.1
1134	0	11	+	Myosin heavy chain [*Trypanosoma cruzi* strain CL Brener] XP_808187.1
78	197	449	+	NADH dehydrogenase subunit 1 [*Cooperia oncophora*] NP_851325.1
666	278	532	+	NADH dehydrogenase subunit 2 [*Bunostomum phlebotomum*] YP_002725706.1
412	373	677	+	No hits found
1686	0	28	+	No hits found
1783	0	21	+	No hits found
1008	0	17	+	No hits found
750	0	16	+	No hits found

Cluster	U reads	T reads	Change	BLAST Identity
647	338	540	+	No hits found
721	0	12	+	No hits found
1780	0	10	+	No hits found
1726	0	10	+	No hits found
1632	0	10	+	No hits found
822	0	10	+	No hits found
903	0	11	+	No hits found
310	0	11	+	No hits found
1775	0	11	+	No hits found
494	31	73	+	No hits found
1076	0	9	+	No hits found
772	0	9	+	No hits found
957	0	9	+	No hits found
749	0	9	+	No hits found
882	0	9	+	No hits found
889	0	9	+	No hits found
805	0	9	+	No hits found
1742	0	9	+	No hits found
1051	0	9	+	No hits found
37	11	41	+	Paralysed Arrest at Two-fold family member (pat-10) [*Caenorhabditis elegans*] NP_491501.1
1693	0	14	+	PCNA (Proliferating Cell Nuclear Antigen) homolog family member (pcn-1) [*Caenorhabditis elegans*] NP_500466.2
932	0	9	+	Peptidase family M1 containing protein [*Brugia malayi*] XP_001897028.1

Cluster	U reads	T reads	Change	BLAST Identity
1135	0	10	+	Predicted protein [*Nematostella vectensis*] XP_001633893.1
668	28	91	+	Predicted: hypothetical protein [*Monodelphis domestica*] XP_001375856.1
711	0	14	+	Predicted: similar to Aconitase CG9244-PB [*Apis mellifera*] XP_391994.1
537	3	24	+	Pre-mRNA cleavage complex II protein Clp1 [*Brugia malayi*] XP_001895648.1
1716	0	10	+	Putative hexose transporter 2 [*Babesia bovis*] ABY75261.1
1663	0	10	+	Putative secretory protein precursor; HC40 [*Haemonchus contortus*] AAC03562.1
753	0	9	+	Putative ubiquitin conjugating enzyme 2 [*Oesophagostomum dentatum*] CAJ57642.1
961	0	9	+	Receptor Mediated Endocytosis family member (rme-2) [*Caenorhabditis elegans*] NP_500815.2
1065	0	9	+	Ribosomal protein S9 [*Danio rerio*] NP_957146.1
1701	0	11	+	Ribosomal Protein, Large subunit family member (rpl-11.2) [*Caenorhabditis elegans*] NP_508413.1
1761	0	10	+	RNA recognition motif domain containing protein [*Brugia malayi*] XP_001894321.1
629	490	925	+	rRNA promoter binding protein [*Brugia malayi*] XP_001891902.1
1665	0	21	+	Secreted protein 6 precursor [*Ancylostoma caninum*] AAO63578.1
793	0	9	+	SideroFleXiN (mitochondrial iron transporter) family member (sfxn-1.4) [*Caenorhabditis elegans*] NP_496493.1
789	0	9	+	SUMO (ubiquitin-related) homolog family member (smo-1) [*Caenorhabditis elegans*] NP_490842.1
1170	0	10	+	Troponin T [*Brugia malayi*] XP_001902458.1
1631	0	11	+	Ubiquitin Conjugating enzyme family member (ubc-3) [*Caenorhabditis elegans*] NP_490882.3
700	0	10	+	Venom-Allergen-like Protein family member (vap-1) [*Caenorhabditis elegans*] NP_001024553.1
287	66	200	+	Vitellogenin [*Haemonchus contortus*] AF305957_1
591	65	173	+	Vitellogenin-6; Flags: Precursor [*Oscheius brevesophaga*] Q94637.1
2039	9	0	-	Alpha-1,6-mannosyltransferase subunit (Och1), putative [*Talaromyces stipitatus* ATCC 10500] EED17616.1
17	24	11	-	Amino acid Transporter GlycoProtein subunit family member (atgp-2) [*Caenorhabditis elegans*]

Cluster	U reads	T reads	Change	BLAST Identity
				NP_497043.2
1804	6	0	-	Anaphase-promoting complex subunit 2-like [*Brugia malayi*] XP_001894849.1
626	98	73	-	Ancyclostoma-secreted protein-like protein [*Ostertagia ostertagi*] CAD56659.1
2233	9	0	-	Argonaute protein [*Brugia malayi*] XP_001892201.1
2529	7	0	-	Aspartic protease [*Ostertagia ostertagi*] CAC14005.1
2098	7	0	-	BESS motif family protein [*Brugia malayi*] XP_001900724.1
1840	8	0	-	Bifunctional glyoxylate cycle protein [*Caenorhabditis elegans*] AAA85857.1
87	10	0	-	C. briggsae CBR-ATG-2 protein [*Caenorhabditis briggsae*] CAP30290.1
1919	9	0	-	C. briggsae CBR-PRX-11 protein [*Caenorhabditis briggsae*] CAP27856.1
447	37	17	-	C. elegans protein F53B2.5, partially confirmed by transcrip [*Caenorhbditis elegans*] CAA98127.3
2473	6	0	-	CD151 antigen, putative [*Ixodes scapularis*] EEC13496.1
2511	7	0	-	Chitin binding Peritrophin-A domain containing protein [*Brugia malayi*] XP_001895704.1
2498	10	0	-	Conserved hypothetical protein [*Brugia malayi*] XP_001898151.1
1903	6	0	-	Conserved hypothetical protein [*Brugia malayi*] XP_001900985.1
605	29	0	-	C-type single domain activation associated secreted protein [*Ostertagia ostertagi*] CAO00416.1
2547	6	0	-	C-type single domain activation associated secreted protein ASP3 precursor [*Ostertagia ostertagi*] CAO00416.1
456	172	116	-	C-type single domain activation associated secreted protein ASP3 precursor [*Ostertagia ostertagi*] CAO00416.1
646	166	113	-	C-type single domain activation associated secreted protein ASP3 precursor [*Ostertagia ostertagi*] CAO00416.1
616	61	31	-	C-type single domain activation associated secreted protein ASP3 precursor [*Ostertagia ostertagi*] CAO00416.1

Cluster	U reads	T reads	Change	BLAST Identity
2116	10	0	-	Cyclin, N-terminal domain containing protein [*Brugia malayi*] XP_001891912.1
1918	6	0	-	Cyclophylin family member (cyn-1) [*Caenorhabditis elegans*] NP_506561.1
1912	6	0	-	DIRP family protein [*Brugia malayi*] XP_001900735.1
2567	8	0	-	E5 [Human papillomavirus type 69] BAA90732.1
2518	24	0	-	Endochitinase [*Schizosaccharomyces japonicus* yFS275] XP_002171895.1
1983	11	0	-	Fatty Acid CoA Synthetase family member (acs-16) [*Caenorhabditis elegans*] NP_498568.1
1947	6	0	-	FHA domain-containing protein [*Myxococcus xanthus* DK 1622] YP_632168.1
2109	9	0	-	Fructose-bisphosphate aldolase 1 [*Brugia malayi*] XP_001894530.1
1936	6	0	-	GABA TransAminase family member (gta-1) [*Caenorhabditis elegans*] NP_501862.1
2541	15	0	-	Gag-pol polyprotein [*Schistosoma mansoni*] CAJ00230.1
435	30	10	-	GEX Interacting protein family member (gei-7) [*Caenorhabditis elegans*] NP_503306.1
1964	18	0	-	Globin-like host-protective antigen; Flags: Precursor [*Trichostrongylus colubriformis*] P27613.1
2488	10	0	-	Glycerol-3-phosphate dehydrogenase [*Ctenolepisma longicaudata*] AAR13229.1
2521	8	0	-	Heat shock protein 20 [*Haemonchus contortus*] AAN05752.1
1871	15	0	-	Heat shock protein 20 [*Ostertagia ostertagi*] CAG25499.1
651	85	49	-	Heat Shock Protein family member (hsp-16.1) [*Caenorhabditis elegans*] NP_505354.1
2060	8	0	-	Heat Shock Protein family member (hsp-16.2) [*Caenorhabditis elegans*] NP_503507.1
2212	6	0	-	Helicase conserved C-terminal domain containing protein [*Brugia malayi*] XP_001901253.1
2153	8	0	-	Hypothetical protein [*Monosiga brevicollis* MX1] XP_001743138.1
1985	9	0	-	Hypothetical protein [*Plasmodium falciparum* 3D7] XP_001350649.1
1920	6	0	-	Hypothetical protein Bm1_13905 [*Brugia malayi*] XP_001894245.1
2035	11	0	-	Hypothetical protein BRAFLDRAFT_233880 [*Branchiostoma floridae*] XP_002235185.1
664	87	72	-	Hypothetical protein C04G2.9 [*Caenorhabditis elegans*] NP_501838.2

Cluster	U reads	T reads	Change	BLAST Identity
94	50	34	-	Hypothetical protein C07D8.6 [*Caenorhabditis elegans*] NP_509242.1
2013	6	0	-	Hypothetical protein C08H9.2 [*Caenorhabditis elegans*] NP_001129807.1
1786	6	0	-	Hypothetical protein C10G11.8 [*Caenorhabditis elegans*] NP_491811.1
1905	7	0	-	Hypothetical protein C14B1.10 [*Caenorhabditis elegans*] NP_497751.1
1990	7	0	-	Hypothetical protein C16C8.18 [*Caenorhabditis elegans*] NP_494539.1
2234	6	0	-	Hypothetical protein CBG00644 [*Caenorhabditis briggsae* AF16] XP_001678643.1
2371	6	0	-	Hypothetical protein CBG00694 [*Caenorhabditis briggsae* AF16] XP_001678605.1
2475	7	0	-	Hypothetical protein CBG01426 [*Caenorhabditis briggsae* AF16] XP_001673254.1
543	354	260	-	Hypothetical protein CBG01578 [*Caenorhabditis briggsae* AF16] XP_001671326.1
2503	6	0	-	Hypothetical protein CBG02238 [*Caenorhabditis briggsae* AF16] XP_001668494.1
588	8	0	-	Hypothetical protein CBG02304 [*Caenorhabditis briggsae* AF16] XP_001679863.1
1927	6	0	-	Hypothetical protein CBG02746 [*Caenorhabditis briggsae* AF16] XP_001679501.1
2263	8	0	-	Hypothetical protein CBG03108 [*Caenorhabditis briggsae* AF16] XP_001679208.1
2524	11	0	-	Hypothetical protein CBG04059 [*Caenorhabditis briggsae* AF16] XP_001668188.1
2015	6	0	-	Hypothetical protein CBG04059 [*Caenorhabditis briggsae* AF16] XP_001668188.1
2120	9	0	-	Hypothetical protein CBG04320 [*Caenorhabditis briggsae* AF16] XP_001678833.1
2122	9	0	-	Hypothetical protein CBG04454 [*Caenorhabditis briggsae* AF16] XP_001671526.1
1948	6	0	-	Hypothetical protein CBG05408 [*Caenorhabditis briggsae* AF16] XP_001670625.1
2175	11	0	-	Hypothetical protein CBG08653 [*Caenorhabditis briggsae* AF16] XP_001673371.1
1806	13	0	-	Hypothetical protein CBG08658 [*Caenorhabditis briggsae* AF16] XP_001673367.1
361	82	67	-	Hypothetical protein CBG08717 [*Caenorhabditis briggsae* AF16] XP_001669329.1
2214	7	0	-	Hypothetical protein CBG08774 [*Caenorhabditis briggsae* AF16] XP_001672562.1
2520	7	0	-	Hypothetical protein CBG09451 [*Caenorhabditis briggsae* AF16] XP_001674131.1

Cluster	U reads	T reads	Change	BLAST Identity
2528	6	0	-	Hypothetical protein CBG09806 [*Caenorhabditis briggsae* AF16] XP_001665103.1
1908	10	0	-	Hypothetical protein CBG09896 [*Caenorhabditis briggsae* AF16] XP_001665184.1
2378	6	0	-	Hypothetical protein CBG11896 [*Caenorhabditis briggsae* AF16] XP_001667421.1
2133	9	0	-	Hypothetical protein CBG12251 [*Caenorhabditis briggsae* AF16] XP_001666259.1
1922	8	0	-	Hypothetical protein CBG12316 [*Caenorhabditis briggsae* AF16] XP_001668329.1
2157	8	0	-	Hypothetical protein CBG12597 [*Caenorhabditis briggsae* AF16] XP_001668827.1
1866	7	0	-	Hypothetical protein CBG13426 [*Caenorhabditis briggsae* AF16] XP_001678475.1
2526	8	0	-	Hypothetical protein CBG15925 [*Caenorhabditis briggsae* AF16] XP_001667359.1
2132	10	0	-	Hypothetical protein CBG16572 [*Caenorhabditis briggsae* AF16] XP_001665476.1
2089	6	0	-	Hypothetical protein CBG17431 [*Caenorhabditis briggsae* AF16] XP_001676058.1
2199	6	0	-	Hypothetical protein CBG17919 [*Caenorhabditis briggsae*] CAP35459.1
2048	18	0	-	Hypothetical protein CBG17992 [*Caenorhabditis briggsae* AF16] XP_001665643.1
1994	7	0	-	Hypothetical protein CBG18189 [*Caenorhabditis briggsae* AF16] XP_001665804.1
2522	8	0	-	Hypothetical protein CBG18268 [*Caenorhabditis briggsae* AF16] XP_001665868.1
2561	6	0	-	Hypothetical protein CBG19105 [*Caenorhabditis briggsae* AF16] XP_001674484.1
1813	7	0	-	Hypothetical protein CBG21198 [*Caenorhabditis briggsae* AF16] XP_001666345.1
1837	7	0	-	Hypothetical protein CBG22304 [*Caenorhabditis briggsae* AF16] XP_001667777.1
2381	6	0	-	Hypothetical protein CBG24043 [*Caenorhabditis briggsae*] CAP20747.1
1961	7	0	-	Hypothetical protein CBG24742 [*Caenorhabditis briggsae* AF16] XP_001678957.1
1896	8	0	-	Hypothetical protein D1037.1 [*Caenorhabditis elegans*] NP_491200.1
2069	14	0	-	Hypothetical protein F11C1.1 [*Caenorhabditis elegans*] NP_510159.1
1923	6	0	-	Hypothetical protein F13H8.7 [*Caenorhabditis elegans*] NP_495261.1
499	22	10	-	Hypothetical protein F23B12.4 [*Caenorhabditis elegans*] NP_001122926.2

Cluster	U reads	T reads	Change	BLAST Identity
1889	10	0	-	Hypothetical protein F52E1.2 [*Caenorhabditis elegans*] NP_505170.2
321	22	9	-	Hypothetical protein F53B6.7 [*Caenorhabditis elegans*] NP_492409.2
507	70	40	-	Hypothetical protein F54H12.6 [*Caenorhabditis elegans*] NP_498737.1
2523	9	0	-	Hypothetical protein HcanM9_03085 [*Helicobacter canadensis* MIT 98-5491] ZP_03656248.1
65	20	6	-	Hypothetical protein K03E5.2d [*Caenorhabditis elegans*] ACO15785.1
1930	15	0	-	Hypothetical protein K09E9.4 [*Caenorhabditis elegans*] NP_001024780.1
2554	9	0	-	Hypothetical protein K12G11.6 [*Caenorhabditis elegans*] CAF31481.2
1827	23	0	-	Hypothetical protein L3ni51 [*Dictyocaulus viviparus*] AAT02162.1
2462	8	0	-	Hypothetical protein M176.4 [*Caenorhabditis elegans*] NP_496013.2
2497	11	0	-	Hypothetical protein NEMVEDRAFT_v1g141408 [*Nematostella vectensis*] XP_001622397.1
419	24	2	-	Hypothetical protein PC101070.00.0 [*Plasmodium chabaudi chabaudi*] CAH88354.1
84	58	28	-	Hypothetical protein R09E10.6 [*Caenorhabditis elegans*] NP_501889.1
116	7	0	-	Hypothetical protein SS1G_07081 [*Sclerotinia sclerotiorum* 1980 UF-70] XP_001591635.1
1873	7	0	-	Hypothetical protein T10G3.3 [*Caenorhabditis elegans*] NP_506346.2
2224	8	0	-	Hypothetical protein Y110A7A.19 [*Caenorhabditis elegans*] NP_491536.1
1812	6	0	-	Hypothetical protein Y1A5A.1 [*Caenorhabditis elegans*] NP_497801.1
394	23	6	-	Hypothetical protein Y38F1A.2 [*Caenorhabditis elegans*] NP_496760.1
527	27	10	-	Hypothetical protein Y69A2AR.7 [*Caenorhabditis elegans*] NP_741338.1
2144	9	0	-	Hypothetical protein ZK1307.8 [*Caenorhabditis elegans*] NP_496073.1
2169	6	0	-	Hypothetical protein ZK795.4 [*Caenorhabditis elegans*] NP_001023622.1
2551	8	0	-	Hypothetical protein ZK970.8 [*Caenorhabditis elegans*] NP_001022541.1
1808	10	0	-	IntraMembrane Protease (IMPAS) family member (imp-2) [*Caenorhabditis elegans*] NP_502079.1
14	22	8	-	Kelch-like protein X [*Brugia malayi*] XP_001898633.1

Cluster	U reads	T reads	Change	BLAST Identity
1834	6	0	-	Lipid Binding Protein family member (lbp-9) [*Caenorhabditis elegans*] NP_001033511.1
1797	6	0	-	Lysozyme family member (lys-10) [*Caenorhabditis elegans*] NP_501405.1
2552	7	0	-	Maternal Effect Lethal family member (mel-11) [*Caenorhabditis elegans*] NP_001021928.1
2508	7	0	-	Metalloprotease I [*Ostertagia ostertagi*] CAD28559.2
2052	7	0	-	MSP domain containing protein [*Brugia malayi*] XP_001898260.1
2559	9	0	-	Multiple PDZ domain protein family member (mpz-1) [*Caenorhabditis elegans*] NP_001122601.1
2570	26	0	-	No hits found
243	75	24	-	No hits found
2531	21	0	-	No hits found
1799	18	0	-	No hits found
2516	15	0	-	No hits found
193	45	18	-	No hits found
2515	12	0	-	No hits found
178	33	11	-	No hits found
2500	11	0	-	No hits found
2501	10	0	-	No hits found
2509	10	0	-	No hits found
2539	10	0	-	No hits found
608	10	0	-	No hits found
204	77	52	-	No hits found
1819	9	0	-	No hits found
1789	9	0	-	No hits found
2544	9	0	-	No hits found

Cluster	U reads	T reads	Change	BLAST Identity
514	9	0	-	No hits found
1845	8	0	-	No hits found
2215	8	0	-	No hits found
2002	7	0	-	No hits found
569	7	0	-	No hits found
1921	7	0	-	No hits found
1976	7	0	-	No hits found
2560	7	0	-	No hits found
2198	7	0	-	No hits found
2077	7	0	-	No hits found
2171	7	0	-	No hits found
1928	7	0	-	No hits found
1856	7	0	-	No hits found
2566	7	0	-	No hits found
235	101	87	-	No hits found
2274	6	0	-	No hits found
2093	6	0	-	No hits found
2321	6	0	-	No hits found
1851	6	0	-	No hits found
2279	6	0	-	No hits found
2569	6	0	-	No hits found
2032	6	0	-	No hits found
2377	6	0	-	No hits found

Cluster	U reads	T reads	Change	BLAST Identity
64	6	0	-	No hits found
2242	6	0	-	No hits found
2055	6	0	-	No hits found
2064	6	0	-	No hits found
2080	6	0	-	No hits found
2513	6	0	-	No hits found
2165	6	0	-	No hits found
1906	6	0	-	No hits found
1830	6	0	-	No hits found
1917	6	0	-	No hits found
2328	6	0	-	No hits found
2148	6	0	-	No hits found
2176	6	0	-	No hits found
612	6	0	-	No hits found
2272	6	0	-	No hits found
2542	6	0	-	No hits found
2091	6	0	-	No hits found
1999	6	0	-	No hits found
2117	6	0	-	No hits found
2143	6	0	-	No hits found
1931	6	0	-	No hits found
1878	6	0	-	No hits found
1965	6	0	-	No hits found

Cluster	U reads	T reads	Change	BLAST Identity
2268	6	0	-	No hits found
2155	6	0	-	No hits found
1816	7	0	-	Paramyosin [*Dictyocaulus viviparus*] ABO07440.1
2555	6	0	-	Precursor transthyretin like protein 1 [*Ostertagia ostertagi*] CAP45649.1
2510	6	0	-	Predicted protein [*Nematostella vectensis*] XP_001627195.1
1824	10	0	-	Predicted: hypothetical protein [*Apis mellifera*] XP_001121281.1
353	24	4	-	Predicted: similar to AGRin (synaptic protein) homolog family member (agr-1) [*Hydra magnipapillata*] XP_002157711.1
49	101	44	-	Predicted: similar to cell division cycle 2-like 5 isoform 2 [*Rattus norvegicus*] XP_001053609.1
2266	6	0	-	Predicted: similar to collagen, type XXVIII [*Hydra magnipapillata*] XP_002166419.1
2273	6	0	-	Predicted: similar to conserved hypothetical protein [*Tribolium castaneum*] XP_971602.1
1795	7	0	-	Predicted: similar to DEAD (Asp-Glu-Ala-Asp) box polypeptide [*Strongylocentrotus purpuratus*] XP_785431.2
2478	13	0	-	Predicted: similar to egg bindin receptor 1 precursor, partial [*Strongylocentrotus purpuratus*] XP_001177184.1
290	113	64	-	Predicted: similar to hect (homologous to the E6-AP (UBE3A) [*Ornithorhynchus anatinus*] XP_001514757.1
2545	6	0	-	Predicted: similar to nahoda CG12781-PA [*Danio rerio*] XP_698035.3
620	21	3	-	Predicted: similar to polyubiquitin [*Nasonia vitripennis*] XP_001599434.1
2241	6	0	-	Predicted: similar to predicted protein [*Hydra magnipapillata*] XP_002160784.1
1933	7	0	-	Predicted: similar to predicted protein, partial [*Hydra magnipapillata*] XP_002164783.1
1978	9	0	-	Predicted: similar to Protein phosphatase 1 regulatory subunit 11 [*Ciona intestinalis*] XP_002120515.1
1833	7	0	-	Predicted: similar to spen homolog, transcriptional regulator (Drosophila) [*Equus caballus*]

Cluster	U reads	T reads	Change	BLAST Identity
				XP_001914854.1
1979	6	0	-	Protein kinase domain containing protein [*Brugia malayi*] XP_001900551.1
1988	7	0	-	Protein RIO1 kinase [*Trichostrongylus vitrinus*] CAR64255.1
1898	6	0	-	Putative cathepsin B.1 [*Ostertagia ostertagi*] CAB97364.2
1874	7	0	-	Putative L3 ES protein [*Ostertagia ostertagi*] CAH23216.1
2525	8	0	-	Putative secretory protein precursor; HC40 [*Haemonchus contortus*] AAC03562.1
673	47	28	-	Putative secretory protein precursor; HC40 [*Haemonchus contortus*] AAC03562.1
1916	16	0	-	Putative serine/threonine phosphatase [*Oesophagostomum dentatum*] AF496634_1
140	21	8	-	Regulator of Microtubule Dynamics family member (rmd-2) [*Caenorhabditis elegans*] NP_741608.1
1893	6	0	-	RiboNucleotide Reductase family member (rnr-1) [*Caenorhabditis elegans*] NP_499039.1
2472	16	0	-	RNA recognition motif domain containing protein [*Brugia malayi*] XP_001894321.1
2530	12	0	-	Secreted protein 5 precursor [*Ancylostoma caninum*] AAO63577.1
1910	9	0	-	Secreted protein 5 precursor [*Ancylostoma caninum*] AAO63577.1
2482	8	0	-	Secreted protein 5 precursor [*Ancylostoma caninum*] AAO63577.1
2477	14	0	-	Secreted protein 5 precursor [*Ancylostoma ceylanicum*] ABB53347.1
2262	6	0	-	Splicing factor, arginine/serine-rich 4 [*Brugia malayi*] XP_001902042.1
669	20	6	-	SXC1 protein [*Ostertagia ostertagi*] CAC17797.1
674	31	17	-	SXC1 protein [*Ostertagia ostertagi*] CAC17797.1
2070	11	0	-	TBC (Tre-2/Bub2/Cdc16) domain family member (tbc-3) [*Caenorhabditis elegans*] NP_001023165.1
1998	7	0	-	Temporarily Assigned Gene name family member (tag-192) [*Caenorhabditis elegans*] NP_491426.2
2496	9	0	-	Trypsin-like Protease family member (try-1) [*Caenorhabditis elegans*] NP_494910.2
2023	8	0	-	Tubulin alpha chain [*Haemonchus contortus*] P50719.1
2504	6	0	-	Ubiquitin [*Lepeophtheirus salmonis*] ACO11811.1

Cluster	U reads	T reads	Change	BLAST Identity
1890	6	0	-	Ubiquitin, putative [*Pediculus humanus corporis*] EEB18688.1
418	226	160	-	Vitellogenin [*Haemonchus contortus*] AF305957_1
583	64	24	-	Vitellogenin structural genes (yolk protein genes) family member (vit-6) [*Caenorhabditis elegans*]. NP_001023274.1
169	310	233	-	Vitellogenin-6; Flags: Precursor [*Oscheius brevesophaga*] Q94637.1
2033	6	0	-	Zinc finger, C2H2 type family protein [*Brugia malayi*] XP_001901497.1
1963	7	0	-	Zinc finger, C2H2 type family protein [*Brugia malayi*] XP_001901967.1