The role of microRNAs in the host-parasite relationship in the veterinary nematode *Haemonchus contortus*

Volume I of I

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Henry Yuekun Gu

BSc, MRes

UNIVERSITY of GLASGOW

Institute of Biodiversity, Animal Health and Comparative Medicine

School of Life Sciences

College of Medical, Veterinary and Life Sciences

University of Glasgow

2016

Abstract

Haemonchus contortus is one of the most pathogenic nematodes of small ruminants, particularly sheep, and has a major impact on welfare as well as causing significant economic losses to farmers worldwide. In this project, the possible interaction of parasite microRNAs with the host's immune response was investigated with a view to determining whether microRNAs may enhance parasite survival.

microRNAs are 22 nucleotides long, non-coding RNA molecules that bind to target sites with complementary sequences on mRNAs, usually in the 3' UTR. This interaction causes degradation of the mRNA or suppression of protein synthesis (Bartel, 2009; Selbach *et al.*, 2008). A previous study identified 192 microRNAs in *H. contortus*, a large proportion of which were unique to parasitic nematodes (Winter *et al.*, 2012). One particular microRNA, Hco-miR-5352 is of particular interest and is the major focus of this study. This microRNA is one of a cluster of four microRNAs (the miR-5352 cluster) which is conserved predominantly in nematodes that reside within the gastro-intestinal tract. Microarray and qRT-PCR data show that the miR-5352 cluster was most highly expressed in parasitic stages of the *H. contortus* life cycle. Some members of the cluster were detected in the excretory-secretory products of *H. contortus* as well as in abomasal and lymph node tissues taken from sheep infected with *H. contortus*. Transmission electron microscopy of the excretory-secretory products showed the presence of small vesicle-like structures.

The data presented here showed that *H. contortus* releases a range of microRNAs in the excretory-secretory products, some of which were present within extracellular vesicles. Bioinformatic target prediction methods identified CD69 as a likely target of Hco-miR-5352 and this interaction was demonstrated experimentally using a dual luciferase assay. However the interaction was not confirmed using an inducible CD69 system in Jurkat T cells. The impact of Hco-miR-5352 on global gene expression of an intestinal epithelial cell line identified several interesting targets with important roles in the host immune response against *H. contortus,* the regulation of which may be important in parasite survival within the host.

Table of Contents

List of Tables

List of Figures

Acknowledgment

This PhD would not have been possible without the support of my supervisors, Professor Eileen Devaney and Dr Collette Britton. I would like to thank them from the bottom of my heart for all their guidance, patience and encouragement.

I would also like to acknowledge Agriculture and Horticulture Development Board (AHDB) Beef & Lamb, Quality Meat Scotland (QMS) and Hybu Cig Cymru (HCC) for funding my studentship.

In addition, I would like to thank Dr. Alan Winter, Dr. Brett Roberts, Dr. Vicky Gillan, Dr Jane Kinnaird and Ms. Kirsty Maitland for their invaluable assistance in the lab, Dr Willie Weir for all his help with the bioinformatics, Dr. Roz Laing and Dr. Axel Martinelli for all the help with the *H. contortus* genome and RNA-Seq. Special thanks to Dr. Neil Marks for his support, especially with the qRT-PCR.

I would like to thank Kerry and Seamus O'Neill for being so welcoming, and my friends Kirsten Landsgaard, Rose Rislan, Jack Barton and Nibras Safdar for telling me to hurry up and finish so we can go to the Isle of Skye.

Finally, I have to thank my parents for their endless love and encouragement, without whom, this day would not be possible.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature: _____

Printed name: Henry Gu

Abbreviations

AAM	Alternatively activated macrophages
ago	argonaute
APC	antigen presenting cells
BLAST	Basic local alignment search tool
bp	Base Pair
BZ	Benzimidazole
CAM	Classically activated macrophages
DC	Dendritic cells
DNA	Deoxyribonucleic acid
dsRNA	double stranded RNA
ES	Excretory/secretory products
EV	Extracellular vesicles
FACS	Fluorescence activated cell sorting
FSC-A	forward scatter
GABA	g-aminobutyric acid
GFP	Green fluorescent protein
GI	Gastro-intestinal
GluCl	Glutamate-gated chloride channel
HCV	Hepatitis C virus
IL	Interleukin
L1, L2, L3, L4	Stage larvae
LB Agar	Luria broth Agar
LEV	levamisole
MHC	Major histocompatibility complex
miRISC	microRNA induced silencing complex
miRNA	microRNA
ML	Macrocyclic lactone
mRNA	messenger RNA
nt	Nucleotides
PACT	protein kinase RNA activator
PE	phycoerythrin
PMA	Phorbol 12-myristate 13-acetate
pri-miRNAs	primary miRNAs
RISC	RNA-induced silencing complex
RNA	Riboucleic acid
RNA-Seq	RNA sequencing
siRNA	small interfering RNA
SOC	Super optimal broth
SSC-A	side scatter
Th1	Type 1 T helper
Th2	Type 2 T helper
TRBP	trans-activation response RNA-binding protein
Treg	Regulatory T cells

tRNA	transfer RNA
TSBP	Thiol Sepharose binding protein
UTR	Untranslated region

1 General Introduction

Gastro-intestinal (GI) nematodes are a major constraint to the global sheep industry both in terms of welfare issues as well as production losses. Parasitic nematode infections in sheep result in significant economic loss to the industry from the direct costs of drugs and veterinary services to the indirect costs of reduced animal production (decreased growth rate and wool gain) (Holmes, 1985; Suarez *et al.*, 2009). In 2005, a study estimated that gastro-intestinal parasites cost the UK sheep industry £85 million per year (Nieuwhof and Bishop, 2005).

Aside from the economic issues, chronic infections with GI nematodes are also a major detriment to animal welfare. Infection with GI nematodes can result in weight loss, diarrhoea, anaemia and hypoproteinaemia. Heavy infections can also result in death (Holmes, 1985). The predominant species of nematodes that affect sheep belong to the superfamily Trichostrongyloidea and include nematodes such as *Haemonchus contortus* and *Teladorsagia circumcincta (*abomasum), *Trichostrongylus colubriformis, Nematodirus spp. and Cooperia spp.* (small intestine), and *Oesophagostumum spp.* (large intestine)(Roeber *et al.*, 2013).

1.1 Haemonchus contortus

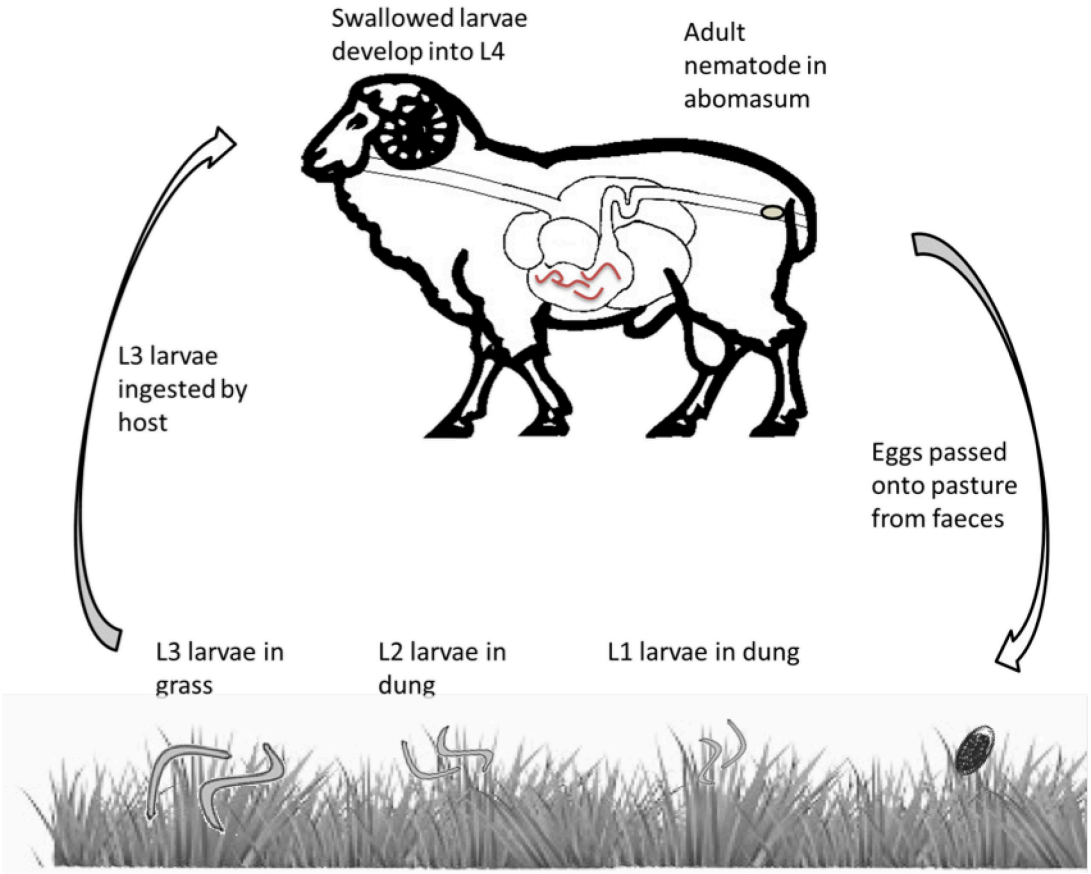
Haemonchus contortus, also known as the barber's pole worm, is one of the most common and most pathogenic parasites of sheep and goats. *H. contortus* is thought to have originated in Africa (Gilleard, 2013) and is most prevalent in warm and humid regions. However over the past 10 years, reports of haemonchosis have increased in the northern hemisphere in Swedish and Swiss sheep flocks (Höglund *et al.*, 2009; Murri *et al.*, 2014) and also within the UK sheep sector (Burgess *et al.*, 2012). The free living stages of *H. contortus* are susceptible to both cold and dry conditions (Roeber *et al.*, 2013) and the L4 larvae can undergo arrested development (hypobiosis) within the host to survive unfavourable conditions. The *H. contortus* female worm is approximately 18-30 mm long and has a distinctive 'barber pole' appearance, which is due to the white ovaries coiled around the blood-filled intestines. The males are much smaller at 10-20 mm long and have a well-developed copulatory bursa,

containing an asymmetrical dorsal lobe and a Y-shaped dorsal ray (Urquhart *et al.*, 1996).

These nematodes are the cause of significant economic losses for the livestock industry. Control of infection has been through the use of anthelmintics drugs, which work to expel the nematodes from the host. Resistance to the major anthelmintic classes is now common worldwide, although some breeds, such as the West African Dwarf goat are able to resist infection by *H. contortus* (haemonchotolerance) (Chiejina *et al.*, 2015).

1.1.1 Life Cycle

Figure 1.1 Life cycle of *H. contortus*



H. contortus has a direct life cycle typical of Trichostrongyloidea (Figure 1.1), consisting of both free-living stages and parasitic stages. Under favourable conditions, the whole cycle lasts approximately 21 days. The adult females are extremely prolific egg layers, releasing between 5,000 to 10,000 eggs per day in the host faeces. The first larval stage (known as L1) hatches on the pastures where it develops and moults to the second stage larvae (L2), then moults again to become the infective third stage larvae (L3). In ideal conditions, it takes 5

days for the larvae to develop from egg to L3, however depending on the conditions, L3 larvae can survive on pasture for up to 6 months, although the free living stages of *H. contortus* are susceptible to both cold and dry conditions (Roeber *et al.*, 2013). The L3 is a non-feeding stage as it remains enclosed in the cast cuticle of the L2 stage. Sheep become infected when they ingest grass contaminated with the infective L3. The L3 then sheds the protective L2 cuticle in the rumen and develops into the L4 stage. The larvae will then moult and develop into sexually mature male or female worms. The process from ingestion of the L3 to the mature adult stage takes 18-21 days (Urquhart *et al.*, 1996).

The L4 and juvenile adults pierce the mucosal lining of the abomasum and feed on host blood. Each adult female worm is capable of removing approximately 0.05 mL of blood each day (Urquhart *et al.*, 1996). Adult female worms begin to lay eggs approximately 18 days following infection (Urquhart *et al.*, 1996).

1.1.2 Hypobiosis

H. contortus infective stages are highly susceptible to low temperatures, and to enable the nematode to survive unfavourable conditions, the development of the L4 stage into the adult stage can be paused. This process, known as hypobiosis is the primary means of over-wintering for *H. contortus* (Blitz and Gibbs, 1972). The L4 larvae stop development and can remain in the mucosa for between three to four months with no apparent host response (Miller and Horohov, 2006). The nematode is then able to survive the winter in the hypobiotic form before resuming development when conditions are more favourable.

A large number of animals will succumb to haemonchosis in spring due to the re-emergence of the parasite from hypobiosis and the increased susceptibility of ewes to *H. contortus* in the periparturient period, due to the immunosuppression and metabolic demands of pregnancy (Baker *et al.*, 1998). While the precise mechanisms triggering hypobiosis are poorly understood, it is proposed to be influenced by both immune and environmental cues (Miller and Horohov, 2006). It is clear that hypobiosis is an important adaptation that contributes to the success and survival of *H. contortus*.

Hypobiosis has been compared to the dauer state of *C. elegans*, as both are stages where development is arrested to enable to the nematodes to exist for a period of time, before resuming development when conditions are more favourable (Klass and Hirsh, 1976). However, the most analogous stage of *H. contortus* to the dauer may be the infective L3. The infective L3 stage of *H. contortus* is a motile, free-living stage. Similar to the dauer stage, development in the infective L3 stage is arrested and resumes when the infective L3 stage enters the host animal where it becomes parasitic (Davies and McKerrow, 2003; Laing *et al.*, 2013). However, in *H. contortus*, the infective L3 stage is an obligatory stage whereas the dauer stage is an alternative developmental stage.

The morphology of dauer larvae is distinct from other stages: they have a specialized cuticle with alae (Cassada and Russell, 1975), the oral orifice is closed by an internal plug and the pharynx is constricted (Vowels and Thomas, 1992). The metabolism of the dauer stage also undergoes changes. From the L1 to L2 stages, the larvae shifts from the glyoxylate cycle to generate carbohydrates from lipid stores to aerobic respiration. Dauer larvae do not undergo this shift (Wadsworth and Riddle, 1989). Instead they have low levels of transcription, although they have elevated levels of certain mRNAs such as heat shock protein 90 (Dalley and Golomb, 1992), superoxide dismutase (Larsen, 1993) and catalase (Vanfleteren and De Vreese, 1995).

1.2 Pathogenesis

Infection with *H. contortus* results in haemonchosis with the severity of the disease depending upon the worm burden. Economic losses are incurred through morbidity and mortality as well as the decreased productivity and weight gain of the animals. Haemonchosis can be separated into three forms; the hyperacute form arises from a huge worm burden, up to 30,000 worms, with symptoms including severe anaemia and sudden death in apparently healthy sheep. The acute form (2,000 to 20,000 worms) is characterized by severe anaemia, hypoproteinaemia, oedema and death. Chronic haemonchosis develops during a prolonged dry season, where reinfection is low. The continuous loss of blood from a worm burden of several hundred worms is sufficient to produce clinical signs that are similar to those for poor nutrition, such as loss of weight, weakness and inappetence rather than anaemia (Urquhart *et al.*, 1996).

In general, lambs are the most at risk of haemonchosis, however, mature sheep that are undernourished or have a compromised immune system may also have fatal anaemia (Bethony *et al.*, 2006). Sheep in late pregnancy and lactation also experience a phenomenon where the faecal egg count rises, called the periparturient rise (Connan, 1968; O'Sullivan and Donald, 1970, 1973). This phenomenon is associated with a decrease in host immunity and ewes are less able to resist infection and expel worms (Beasley *et al.*, 2010; Falzon *et al.*, 2013). This increase in egg output ensures that eggs and consequently infective L3 will be available on pasture for the new season's lambs.

1.3 Anthelmintic control of *Haemonchus*

The most commonly used method for controlling *Haemonchus* and other worm infections in livestock is with broad-spectrum anthelmintics. They are used as both a curative and as a preventative treatment, however, the widespread usage of these drugs has resulted in the development of resistance (Kaminsky, 2003). There are 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). *H. contortus* was one of the first parasites reported to show anthelmintic resistance 50 years ago (Drudge *et al.*, 1957). Isolates of *H. contortus*, *T. circumcincta* and *Trichostrongylus colubriformis* that are resistant to all three classes of anthelmintic have now been reported throughout the world (Kaplan, 2004). In some cases, including in the UK (Sargison *et al.*, 2007) multi-drug resistance has become such a problem that farms have been abandoned and livestock culled. The decreasing viability of current anthelmintics has driven the development of alternative strategies for parasite control.

Benzimidazoles (BZs) are the oldest of the anthelmintic classes and were approved for use in 1961 with the first case of resistance being reported in 1964 (Kaplan, 2004). Upon exposure to BZs, microtubules disappear suggesting that BZs act by binding to nematode tubulin, causing the depolymerisation of microtubules and resulting in starvation, paralysis and expulsion of the nematode (Prichard, 1990). In *H. contortus,* resistance to BZs seems to be most commonly conferred by a single nucleotide polymorphism encoding a tyrosine

instead of phenylalanine at codon 200 in ß-tubulin isotype-I (Kwa *et al.*, 1994) or isotype II (Kwa *et al.*, 1993); expression of the *H. contortus* Phe200Tyr mutant in wild type *C. elegans* conferred BZ resistance (Kwa *et al.*, 1995). Phe to Tyr or Phe to His mutations at codon 167 are also sufficient to confer resistance (Li *et al.*, 1996; Prichard, 2001), while a SNP at position 198 (Glu to Ala) is the rarest of the different mutations globally but is associated with resistance in some *H. contortus* isolates (Chaudhry *et al.*, 2015).

Levamisole and other imidazothiazoles (LEV) were first approved in 1970 with resistance emerging nine years later. LEV is a nicotinic cholinergic agonist that opens sodium and potassium cation channels resulting in the depolarisation of muscle cell membranes (Prichard, 2001). In *C. elegans*, eight genes make up the functional levamisole-sensitive acetylcholine receptor; five genes, *unc-38, unc-63, lev-8, unc-29* and *lev-1* form the acetylcholine cation channel (Boulin *et al.*, 2008) with three others acting as accessory proteins. Mutations in any of the eight genes that make up the channel confers resistance to LEV. Alongside the genes of the channel, *lev-11* and *unc-22* encode tropomyosin and twitchin, respectively which are both involved in muscle contraction, *unc-68* a ryanodine receptor involved in intracellular calcium regulation and the transmembrane protein *lev-10,* which is involved in postsynaptic aggregation of nicotinic acetylcholine receptors, have all been shown to confer levamisole resistance (Jones *et al.*, 2005). In parasitic nematodes, the mechanism of resistance is still unclear. Attempts to identify the five genes that encode the acetylcholine channel in *H. contortus*, *T. circumcincta* and *T. colubriformis* found orthologues of all the genes except for *lev-8*, however these orthologues were not altered in LEV-resistant or LEV-susceptible isolates of the three nematodes and truncated transcripts of *unc-63* were found in the resistant strain only (Neveu *et al.*, 2010). It was hypothesised that the truncated *unc-63* acts as a dominant negative on AChR expression (Boulin *et al.*, 2011). Whilst *lev-8* was not identified, *acr-8* was found to be highly similar in the three parasitic nematodes compared to *C. elegans* and could replace the function of *lev-8* (Hernando *et al.*, 2012; Neveu *et al.*, 2010).

Macrocyclic lactones (MLs), such as ivermectin, were first approved in 1981 with the first report of resistance in 1988. In nematodes, MLs target the glutamate-

gated chloride channels (GluCl) in muscle and nerve cells of *C. elegans* causing paralysis and death (Dent *et al.*, 1997, 2000; Ghosh *et al.*, 2012) and studies on *Ascaris suum* muscle showed that ivermectin antagonised GABA receptors (Martin and Pennington, 1989). *C. elegans* has six GluCl genes, two of which are alternatively spliced suggesting that the channels may exist in multiple forms and that resistance may be multigenic (Yates *et al.*, 2003). Resistance in gastro-intestinal nematodes may also be multigenic. In *H. contortus*, orthologues of some of the *C. elegans* GluCl subunits have been identified however two genes, *Cel-glc-1* and *Cel-avr-15*, were not identified in *H. contortus*, whilst the parasite has two genes, *Hco-glc-5* and *Hco-glc-6*, which are not present in *C. elegans* (Beech *et al.*, 2010). Ivermectin has been shown to bind the *H. contortus* GluCl subunit (HcGluClα) in COS-7 cells with high affinity (Cheeseman *et al.*, 2001; Forrester *et al.*, 1999, 2002). However, there is currently no consensus between different strains of *H. contortus* as to the exact mechanism of ivermectin resistance.

Recently, a new class of anthelmintic, an amino-acetonitrile derivative was discovered that is effective against nematodes with multi-drug resistance (Kaminsky *et al.*, 2008b). This new drug, monepantel, has a broad-spectrum efficacy, low toxicity to mammals and a nematode specific target (Kaminsky *et al.*, 2008a), but resistance has already been reported to this compound (Mederos *et al.*, 2014). Another therapeutic is Startect® from Zoetis which contains derquantel (2-desoxoparaherquamide) from the anthelmintic class, the spiroindoles. Startect® is a combination therapy containing both abamectin (an ML compound) and derquantel which was able to reduce faecal egg counts by over 99% in mixed strongyle infection (Little *et al.*, 2010). Derquantel is an acetylcholine antagonist while abamectin increases opening of GluCls and a combination of the two were synergistic on the nicotinic acetylcholine receptors of the muscle (Puttachary *et al.*, 2013).

1.4 Vaccination as a means of controlling *H. contortus*

Vaccination against nematodes may be the most effective and efficient strategy for control and a substantial effort has been put into developing a vaccine against helminth parasites. Vaccines have a major advantage over anthelmintics as their use does not leave any chemical residues in treated animals that would

require withdrawal periods (Dalton and Mulcahy, 2001; Vercruysse *et al.*, 2004). However, it has proved extremely difficult to induce protective immunity against nematode parasites.

Due to the complex nature of nematodes, the immune system is exposed to a range of diverse antigens. In the draft version of the *H. contortus* genome, there are over 20,000 predicted genes with differential expression across life-cycle stages. In addition, whilst nematode antigens can be produced in recombinant formats, these usually fail to achieve the same protection as native antigens, possibly due to incorrect folding and/or important post-translational modifications such as glycosylation (Dalton *et al.*, 2003). As a means of control, vaccines have a much narrower spectrum of application than drugs, being limited to a single species or even strain of pathogen.

The first and the best example of a successful vaccine against livestock nematodes is the irradiated L3 vaccine for *Dictyocaulus viviparous* in cattle. The vaccine called Bovilis® Huskvac (formerly known as Dictol) was released in February 1959 and is comprised of two doses of 1000 irradiated larvae given four weeks apart (Jarrett *et al.*, 1958, 1959, 1960). Irradiated *H. contortus* L3 also show high levels of protection in older sheep, however, young lambs are not protected (Smith and Angus, 1980).

Attempts to vaccinate against *H. contortus* have either used excretory/secretory (ES) products or have adopted the so-called 'hidden antigen' approach. During infection, *H. contortus* excrete or secrete a variety of molecules into their host, collectively known as the ES products. These products are generally immunogenic and, as they are continuously released, the host will be constantly exposed resulting in a natural boosting effect. Vaccination with *H. contortus* ES products with molecular weights of 15 and 24 kDa was able to significantly reduce mean faecal egg counts and abomasal worm burden by 70% (Schallig *et al.*, 1997). ES proteins eluted from a thiol sepharose column were used to vaccinate calves against the related nematode, *Ostertagia ostertagi*, resulting in a 60% decrease in faecal egg count as well as an 18% reduction in total worm numbers and significant reduction in worm length (Geldhof *et al.*, 2002, 2004).

The ‘hidden antigen’ approach relies upon purification of proteins that are associated with the gut membranes of the worm. The advantage of this system is that the antigens used for immunisation are not normally ‘seen’ by the host immune system and so the parasite is less likely to evade the vaccines by mutating these antigens. However, because of the ‘hidden’ nature of these antigens, the response is not boosted by challenge infection, meaning that repeated vaccination is required. Significant protection has been achieved using native *H. contortus* protein extracts (Knox *et al.*, 2003; Newton and Munn, 1999). There are three main vaccine targets: H11, H-Gal-GP and Thiol Sepharose-Binding Protein (TSBP). H110D (H11) is an integral membrane glycoprotein from the intestinal microvilli of parasitic stages of *H. contortus*. Lambs vaccinated with enriched H11 from adult *H. contortus* showed between 70 to 89% reduction in total worm burden, depending on extraction method (Munn *et al.*, 1993). However, attempts to use a recombinant form of H11 expression baculovirus in insect cells resulted in only 30% protection (Reszka *et al.*, 2007).

H-Gal-GP was identified from glycoprotein fractions isolated from intestinal membranes of *H. contortus* and binds selectively to lectins with a specificity for N-acetylgalactosamine (Smith *et al.*, 1994). Sheep immunised with H-Gal-GP showed reduced worm burdens of up to 72% (Smith *et al.*, 1994). However, the antigen used exists as a glycoprotein complex and is predominantly formed of proteases, including two pepsin-like aspartyl proteases, four metalloendopeptidases, cysteine proteases, a cysteine protease inhibitor and thrombospodin (Skuce *et al.*, 2001; Smith *et al.*, 1999). Immunisation with a single or combination of recombinant candidate proteins did not produce the same level of protection as H-Gal-GP (Knox, 2011; Smith *et al.*, 2003), suggesting that protection is conveyed by the complex as a whole, or that the recombinant proteins do not possess the correct post-translational modifications.

Thiol Sepharose-Binding Protein (TSBP) is another membrane bound fraction from *H. contortus* gut tissue (Skuce *et al.*, 1999). Immunisation with protein extracts enriched for cysteine protease by Thiol Sepharose chromatography achieved a 47% reduction in final worm burden (Knox *et al.*, 1999). ES products

from adult *H. contortus*, also fractionated on a Thiol Sepharose column, reduced worm burden by 50% in immunised animals (Bakker *et al.*, 2004). Out of the fraction, three cathepsin B-like cysteine proteinases (hmcp1, hmcp14 and hmcp16) were identified as the protective proteins and recombinant versions were created with glutathione S-transferase. These fusion proteins were not able to reduce faecal egg counts; however, worm burdens were reduced by 38% (Redmond and Knox, 2004). A repeat trial with non-fusion recombinant proteins reduced faecal egg counts by 27% and worm burdens by 29% (Redmond and Knox, 2006).

Attempts to immunise using recombinant proteins have not shown the same protection as obtained with native antigen (Geldhof *et al.*, 2008; Roberts *et al.*, 2013). This could be due to incorrect post-translational modifications and/or folding (Smith and Zarlenga, 2006). There could also be a requirement for the antigen to be presented as a complex. As discussed below, natural immunity to gastro-intestinal parasites is slow to develop in large animals and the mechanism of immunity and the type of immune response required is complex and not fully understood (Roberts *et al.*, 2013). However, sheep immunised with eight recombinant *T. circumcincta* proteins had up to 92% fewer eggs in the peak egg shedding period and a 75% reduction in adult worm burdens at post-mortem. These proteins were selected based on different approaches, including L3 or L4 targets (cathepsin F-1, astacin-like metalloproteinase-1, protein with unknown function (Tci-ES20) and activation-associated secretory protein-1), homologue of a protective antigen from *Ancylostroma caninum* (Tci-SAA-1) and immunosuppressive molecules released by *T. circumcincta* larvae (macrophage migration inhibitory factor-1 (Tci-MIF-1), calcium-dependent apyrase-1 and a TGF- ß homologue) (Nisbet *et al.*, 2013). The success of this vaccine supports the hypothesis that more than one antigen may be required to elicit immunity.

In October 2014, Barbervax® (www.barbervax.com.au), the first commercial vaccine against *H. contortus,* was released. The vaccine contains native proteins purified from the lining of the *H. contortus* intestine and immunisation stimulates the production of antibodies against the proteins that circulate in the host blood. During infection, as *H. contortus* blood feed, they ingest antibodies,

which attach to the lining of the worm's intestine, blocking digestion and causing the worm to starve (Knox *et al.*, 2003; Smith, 1999).

1.5 Immune response to gastro-intestinal nematodes

1.5.1 The innate response

For an immune response to be initiated, the host must be able to recognise the pathogen. In the gastro-intestinal tract, intestinal epithelial cells (IECs) provide a physical barrier against pathogens. Specialised structures, such as the Peyer's patches, which are located in the small intestine, allow the host to sample antigen and present it to antigen presenting cells (APC), such as dendritic cells. In the abomasum, dendritic cells (DC) are located primarily below the epithelium. The DC take up antigen at this location then migrate to the draining abomasal lymph nodes to present the antigen to naïve T cells via major histocompatibility complex class II molecules (MHC), resulting in the initiation of an adaptive T cell response (Belkaid and Oldenhove, 2008; Diebold, 2008; McNeilly *et al.*, 2009).

IECs appear to be essential for the conditioning of DC and promoting Th2 responses. When mice with an IEC-specific deletion of intrinsic IκB kinase (IKK-ß) were infected with *T. muris*, expression of the thymic stromal lymphopoietin derived from IECs was lowered and a pathogen-specific Th2 response failed to develop. Furthermore, these animals produced increased levels of IL-12 from DC, and tumour necrosis factor-α, T-cell-derived interferon-γ and interleukin-17, and presented with intestinal inflammation (Zaph *et al.*, 2007).

IECs also release cytokines such as IL-25 and IL-33 that activate neighbouring immune cells including IEC, DC and macrophages, resulting in a Th2 type response. IL-33 is a chemoattractant (Komai-Koma *et al.*, 2007) and potent stimulator of Th2 cells (Schmitz *et al.*, 2005). IECs also produce chemokines that recruit mucosal mast cells. Mast cells produce Th2 cytokines such as IL-13, IL-4 and IL-5, helping to reinforce the Th2 response. These cytokines act on IECs inducing changes in permeability, and decreased absorption of glucose to increase the intraluminal fluid which facilitates worm expulsion (Madden *et al.*, 2004).

A common feature of helminth infection is eosinophilia. IL-5 is predominantly involved in the generation and activation of eosinophils. However, the exact role of eosinophils in host protection is unclear, although it is likely that these cells have a role in resistance against reinfection by larval stages (Maizels and Balic, 2004).

Other cells such as circulating basophils and macrophages also play an important role during early infection. Basophils are thought to be a source of early IL-4 that promotes Th2 cell differentiation during helminth infections (Min *et al.*, 2004). Similar to the Th1/Th2 response, macrophages can be split into classically activated (CAM) or alternatively activated macrophages (AAM). Alternatively activated macrophages are associated with the Th2 response and play a role in tissue repair, whereas the IFN-γ-mediated classically activated macrophages are associated with Th1 responses. AAM are characterised by low production of IL-1ß, IL-12, and IL-23 and produce urea and regulatory cytokines such as TGF- ß and increased arginase-1 activity, which favours l-arginine metabolism towards proline, polamines and urea production compared to the nitric oxide production in CAM (Reyes and Terrazas, 2007).

AAM are stimulated by high levels of IL-4 and IL-13 (Allen and Maizels, 2011) and are an important effector cell of the protective memory response. AAMs can be induced by all phyla of helminths although they have different functions depending on the species (Reyes and Terrazas, 2007). In *H. polygyrus*, AAMs have been shown to be essential in protection against secondary infections. This resistance is due to an immune response against the L3 larval stage of the parasite where memory CD4+ T cells accumulate at the site of infection and form a Th2 type granuloma and secrete IL-4 which induces AAMs. These macrophages are then able to impair the larvae through an arginase-dependent pathway. Depletion of macrophages using chlodronate or arginase inhibitor resulted in increased larval mobility and reduction in worm expulsion (Anthony *et al.*, 2006). There is also increasing evidence that AAMs are involved in limiting the Th1 inflammatory response. In *S. mansoni* infection, IL-4 receptor signalling in macrophages were required to sustain the Th2 cytokine response and occured through a IL-10 independent pathway (Herbert *et al.*, 2004).

1.5.2 The adaptive response

During helminth infection, both T and non-T cells, secrete Th2 cytokines and other co-stimulatory factors. IL-4 is considered to be vital in the Th2 response (Reynolds *et al.*, 2012). IL-4 phosphorylates the signal transducer STAT-6 and transcription factor GATA-3 and is necessary for the maturation of naïve CD4+ T cells into the Th2 phenotype. In *H. polygyrus*, immunity to secondary infection is diminished when anti-IL-4 antibody is administered; however, blockade of the IL-4 receptor results in complete abolition of immunity. These data suggest that another interleukin, IL-13, also signals through the IL-4 receptor and is able to compensate for the absence of IL-4 (Urban *et al.*, 1991a).

GATA-3 is the major transcription factor for Th2 cells, whilst T-bet serves the same role for Th1 cells. In experiments where *Gata3* was deleted at the onset of Th2 development, IL-4 was not expressed. However, the deletion of *Gata3* in mature Th2 cells does not impair IL-4 expression, suggesting that T-bet and GATA-3 might re-model genes to their permissive state (Reiner, 2005).

B cells and the antibodies they produce are also an important part of the immune response against helminth infections. *T. muris* infection of mice that lack B cells resulted in a Th1 type response and increased susceptibility. The protective Th2 response and resistance could be restored with naïve B cells or anti-IL-12 and treatment of B cell deficient mice with parasite specific IgG antibodies could prevent worm establishment (Blackwell and Else, 2001). On the other hand, infection with *N. brasiliensis* showed that B cells were not required for the initial activation of T cells but are important for the proliferation and differentiation of Th2 cells (Liu *et al.*, 2007). IL-4 stimulates antibody class switching to the IgE and IgG isotypes. IgE has been shown to be correlated with fecal egg counts and results in mast cell proliferation and degranulation (Murphy *et al.*, 2010). IgG antibodies were shown in *H. polygyrus* infection to limit egg production although only IgG and IgA antibodies that arise from multiple infections were able to prevent adult worm development (McCoy *et al.*, 2008).

1.6 The role of Th2 type responses in helminth infections

For the most part, the Th2 response is presumed to be protective against all helminth infections, however, whilst this is the response typically associated with helminths, there are some exceptions. In schistosomiasis, administration of adjuvants that are inducers of Th1 lymphocyte development and antigen from the larval stage of *Schistosoma mansoni* induced resistance to infection (Mountford *et al.*, 1996; El Ridi and Tallima, 2012). Likewise, in mouse models of *S. mansoni*, the initial immune response to cercarial infection is Th1 dominated, switching to a Th2 response coincident with egg laying by adult females (Pearce and MacDonald, 2002). Similarly, infection with *Trichinella spiralis* or *Nippostrongylus* initially produces a Th1 response characterised by an increase in mRNA for IFN-y in mitogen-stimulated mesenteric lymph node cells, however this response does not last and shifts to a Th2 response.

As referred to above, the simple Th1/Th2 dichotomy to explain susceptibility/resistance is not applicable to all parasitic helminths. Interestingly, in mice that are resistant to *T. muris*, a strong Th2 response was observed whereas mice with Th1 type responses were much more susceptible to infection (Cliffe and Grencis, 2004). Depletion of IL-4 could convert the resistant mice to susceptible and administration of IL-4 could convert susceptible mice to resistant. However, other factors are important as the level of parasite challenge can affect the development of a Th1 or Th2 response in *T. muris* infection (Grencis *et al.*, 2014). In infection with the gastro-intestinal nematode *Heligmosomides polygyrus*, antibodies against the Th2 cytokine IL-4 or IL-4 receptor limits protection, whilst IL-4 treatment of infected mice decreased egg production and terminated infection in *N. brasiliensis* infected mice (Urban *et al.*, 1991b, 1995) by promoting Th2 responses (Mearns *et al.*, 2008).

1.7 Treg and Th17 responses

The Th2 response is the predominant immune response responsible for the control of gastro-intestinal helminth infections; however, for a successful immune response, the strength and duration of the response needs to be controlled. In mouse models of *T. muris* infection, persistent infection resulted in the induction of regulatory T cells (Treg). Inhibiting Treg cells using antibodies

resulted in exacerbated pathology demonstrating that Treg cells are important for controlling the immune response (D'Elia *et al.*, 2009). In *H. polygyrus* infected mice, the initial Th2 response changed to a Treg dominated response by day 28, with an increase in Foxp3+ cells, and elevated levels of IL-10 and TGFß+ T cells (Finney *et al.*, 2007). Treg cells are able to suppress immune responses in three main ways: Tregs directly affect other T (effector) cells; they can interact with antigen presenting cells (APC) and affect ability of other T cells to engage and become activated by the same APC; or through the release of suppressor molecules such as CTLA-4 and TGF-ß (Tang and Bluestone, 2008).

Another Th subset, Th17 are also an important component in the immune response against helminth infections. This subset of cells is distinct from the Th1 or Th2 responses and their primary function is the clearance of pathogens that are ineffectively cleared by Th1 or Th2 cells (Korn *et al.*, 2009). Th17 cells achieve this through acute inflammation by recruiting neutrophils and macrophages and are important in controlling bacterial infections (Bi *et al.*, 2007; Harrington *et al.*, 2005, 2006). Th17 cells produces inflammatory cytokines such as IL-7, IL-17, TNFα (Park *et al.*, 2005) and IL-22 (Veldhoen *et al.*, 2008). However,Th17 cells have also been associated with autoimmune inflammatory diseases such as arthritis (Leipe *et al.*, 2010). Differentiation of naïve T cells to Th17 requires IL-6 (Mangan *et al.*, 2006; Veldhoen *et al.*, 2006). In mice that are deficient for IL-6, Th17 differentiation can also be induced by IL-21 and TGF-ß (Korn *et al.*, 2007). The involvement of TGF- ß in the differentiation of Th17 cells suggest that these cells show some similarity to T regulatory cells where TGF- ß induces differentiation of naïve T cells into T regulatory cells although analysis of the expression profile of cell surface molecules shows that Th17 cells have more similarity to Th1 cells than Th2 cells (Nakae *et al.*, 2007).

In *Schistosoma mansoni* infected mice, IL-17 levels were most directly associated with severity of hepatic granulomatous inflammation suggesting that pathology is controlled by Th17 cells (Rutitzky *et al.*, 2005). Antibody neutralization of TGF- ß lead to decreased IL-17 and a reduced worm burden, although this may have had additional effects, such as on Treg cells. The association between Th17 cells and other helminth infections is less clear. In *Trichinella spiralis,* Th17 has been linked to muscosal damage and hyper-contractility of the jejunum in

infected mice suggesting a role in worm expulsion (Fu *et al.*, 2009) whilst in *T. circumcincta*, Th17 cytokines were shown to correlate with susceptibility to infection (Gossner *et al.*, 2012).

1.8 Immune responses to nematode infection in ruminants

As can be seen from the brief review above, most studies on the immune response to nematode infections have been carried out in inbred mice, where the availability of multiple genetically defined animals, (knock-outs and knock in, plus GFP expressing animals), facilitates the identification of specific subsets of cells with important immune functions. However for ruminants such as sheep and cattle, the situation is much less advanced, although recent studies have identified various cytokines and cell surface markers of immune cells (Entrican *et al.*, 2015).

Protective immune responses do occur in sheep naturally infected with GI nematodes, although immunity has been difficult to induce by vaccination. When Blackface sheep were trickle infected with *T. circumcincta* L3, susceptible lambs showed extensive inflammatory infiltration of the mucosa and sub-mucosa and high levels of IL-6, IL-21 and IL-23A, suggesting that the failure to control larval colonization, adult infection and egg production is due to the activation of the Th17 subset (Gossner *et al.*, 2012).

In *Haemonchus*, a clear Th2-oriented immune response was observed in lambs challenged with *Haemonchus* L3 (Lacroux *et al.*, 2006). In certain cases, infection with *H. contortus* can result in self-cure and development of protective immunity against reinfection. This immunity is dependent on Th2 cytokines such as IL-4, IL-10 and mast cells and IgE (Miller and Horohov, 2006). In both resistant and non-resistant breeds, high IL-5 production was observed although levels were highest in the resistant sheep. Minimal Th1 type IFN-γ activity was observed. In lambs 28 days after infection, parasite specific IgG1 and IgE was higher in the resistant lambs, and higher densities of mast cells and eosinophils were observed in the mucosa (Gill *et al.*, 2000). In adult sheep, Th2 cytokines increase during *Haemonchus* infection, along with parasite specific IgE levels (Shakya *et al.*, 2009), however parasite specific IgE is not present in infected

lambs (Kooyman *et al.*, 2000). During a primary *H. contortus* infection, eosinophils, CD4+ T cells and B cells levels were increased in the abomasal mucosa (Balic *et al.*, 2000). Sheep that were challenged nine weeks after immunisation exhibited *H. contortus* L3 surrounded by eosinophils, and showed significant damage, suggesting that the L3 larvae are not expelled but are susceptible to eosinophil attack (Balic *et al.*, 2006).

Many studies on immunity to *H. contortus* have compared susceptible and resistant breeds of sheep. For example, in the *H. contortus* resistant St. Croix lamb, infection with *H. contortus* L3 resulted in lower numbers of established larvae. These sheep had increased numbers of neutrophils at the site of infection as well as increased levels of circulating white blood cells and IgA (Bowdridge *et al.*, 2015). Comparison of the resistant (Santa Ines) and the susceptible (Suffolk and Ile de France) breeds of sheep showed that inflammatory cells and parasite-specific IgA in the abomasum were inversely correlated with *H. contortus* worm burden, indicating that they are important in impairing *H. contortus* development and fecundity (Amarante *et al.*, 2005). In sheep vaccinated with *H. contortus* ES, elevated levels of parasite-specific IgE antibodies correlated with protection (Kooyman *et al.*, 2000). Elevated IgE has also been shown in infections with *T. circumcincta* (Huntley *et al.*, 2001; Pettit *et al.*, 2005).

Overall, the immune response to *H. contortus* is predominantly Th2 type in the abomasal mucosa, which protects the host from infection and is characterised by increases in Th2 cytokines, recruitment of eosinophils and production of parasite-specific IgA and IgE (reviewed by McNeilly and Nisbet, 2014). In *H. contortus* susceptible sheep, this immune response is not correctly established. However, gastro-intestinal helminths also produce a variety of immunomodulatory molecules that may also impact the ability of the host immune system to respond to infection.

1.8.1 Immunomodulation by gastro-intestinal nematode products

Gastro-intestinal helminths, such as *H. contortus* typically induce a Th2 response, which is the primary protective immune response against helminth infections. However, helminths are able to enhance their survival within the host

by suppressing host immunity. Immunomodulation by helminths is targeted at many components of the immune response including dendritic cells, T regulatory cells and cytokine release (IL-10 and TGF-ß) and primarily involves skewing the immune response away from a damaging pro-inflammatory response. However, Th2 responses can also become 'modified' and this may thwart the host's ability to expel the parasite. As mentioned previously, the Th2 immune response is not always protective against helminths and since the alternative Th1/Th17 response results in inflammation, a skew towards Th2 may be beneficial to the host as it could improve clinical outcome.

One of the most important sources of immunomodulatory products released by nematode parasites is the excretory-secretory product (ES) and substantial work has been dedicated to determining the content of the ES and the function of the immunomodulatory molecules. The majority of data so far has been attained from studies using parasites of mice and human.

The chronicity of many helminth infections suggests that they are able to modulate the immune response most likely through suppression of different Th types in favour of regulatory T cells (Kamal and El Sayed Khalifa, 2006; Rook, 2009). In response to helminth antigens, host TGF-ß is stimulated which causes the generation and expansion of T regulatory cells. Interestingly, TGF- ß homologues have been identified in the secretions of adult *Brugia malayi* and are able to bind mammalian TGF- ß receptors (Gomez-Escobar *et al.*, 1998, 2000). Homologues of TGF- ß have also been identified in *H. contortus*, *T. circumcincta* and *N. brasiliensis*. Expression in *H. contortus* and *N. brasiliensis* was highest in the larval stages and low in adults whereas expression in *H. polygyrus* and *T. circumcincta* was highest in the adult stages compared to infective larvae (McSorley *et al.*, 2010). Mice infected with *Brugia pahangi* L3 showed increased $Foxp3^+$ Treg cells and were able to modulate immune responses (Gillan and Devaney, 2005) and injection of mice with *H. polygyrus* excretory-secretory antigen is able to induce $Foxp3^+$ Treg cells (Grainger *et al.*, 2010).

In addition to a TGF-ß homologue, some parasitic nematodes also produce macrophage migration inhibitory factor (MIF) homologues (Vermeire *et al.*, 2008). MIF is a cytokine that is produced from T lymphocytes and macrophages

and is a key pro-inflammatory cytokine within the innate immune response. In *Brugia*, administration of MIF induced expression of AAM activation markers (Prieto-Lafuente *et al.*, 2009). MIF has also been identified in *T. circumcincta* and in adult stages is located exclusively in the gut tissue and has roles in inhibiting monocyte chemotaxis (Nisbet *et al.*, 2010).

Helminths are also able to suppress immune response via dendritic cells. Dendritic cells (DC) stimulated with *H. polygyrus* were less responsive to TLR ligation and suppressed both Th1 and Th2 antibody responses in mice while also inducing regulatory T cells (Segura *et al.*, 2007). Exposure of DC to the surface glycoplipid (CarLA) of *Trichostrongylus colubriformis* L3 did not result in classical maturation and only minor up-regulation of certain genes (Pernthaner *et al.*, 2012). Similarily, *H. contortus* ES products are able to only partially mature human DC and were shown to dampen the typically strong response evoked by stimulation of the DC with bacterial lipopolysaccharides (Rehman *et al.*, 2015).

As well as secreting cytokine mimics, parasitic helminths also secrete a variety of proteases in the ES, some of which aid in host invasion. In schistosomes, the ES secretions from infective larval stages contain serine proteases which have roles in tissue invasion (McKerrow *et al.*, 2006). These proteases may also function in evading the immune system. Incubation of schistosomes with human serum activated the complement system and a parasite protease, m28, cleaved C3, C3b, iC3b and C9 and protected the parasite from attack by leucocytes (Fishelson, 1995; Ghendler *et al.*, 1996). Treatment of schistosomes with a protease inhibitor, resulted in increased sensitivity to leucocyte attack (Fishelson, 1995). *S. mansoni* also produce an elastase-like serine protease which is able to cleave IgE but not IgA1, IgA2 or IgG1 (Pleass *et al.*, 2000). Within the nematodes, the blood feeding parasites, *H. contortus* and *N. americanus,* contain a number of proteases which are thought to be primarily involved in facilitating blood feeding. In *H. contortus*, cysteine- aspartic-, and metalloproteases were identified in the ES, whilst aspartic, cysteine and serine proteases were identified in *N. americanus*. These proteases are thought to degrade host tissue and/or prevent coagulation of the blood (Brown *et al.*, 1995; Karanu *et al.*, 1993). In *H. contortus*, cysteine proteases inhibitors were able to

block degradation of radiolabeled haemoglobin implicating a role in extra corporeal digestion of the blood (Fetterer and Rhoads, 1997).

In addition to proteases, parasitic helminths also possess protease inhibitors that have a range of functions, from protection against digestion, anticoagulation, to down-regulating T cell responses (Knox, 2007). Extracts of *A. suum* were shown to inhibit a range of host proteases that are able to target the parasite in the gastro-intestinal tract (Knox, 2007). However, proteomic analysis of ES products from larval *A. suum* identified serine protease inhibitors involved in blood coagulation, complement activation and inflammation (Wang *et al.*, 2013). Serine protease inhibitors have also been identified in a number of different parasites although these are not always secreted (Molehin *et al.*, 2012). In *B. malayi*, the serine protease inhibitor, Bm-SPN-2 is secreted by the microfilariae stage and targets neutrophil serine protease, elastase and cathepsin G (Stanley and Stein, 2003; Zang *et al.*, 1999). Cathepsin G is a chemoattractant for T cells and monocytes (Chertov *et al.*, 1997) whilst elastase breaks down parasite membrane proteins (Freudenstein-Dan *et al.*, 2003). In *H. contortus*, hc-serpin was also identified exclusively on the epithelial cells of the *H. contortus* gastro-intestinal tract. hc-serpin inhibits trypsin activity and prolongs the coagulation time of rabbit blood and may be important in the ability of *H. contortus* to blood feed (Yi *et al.*, 2010). Cysteine protease inhibitors (cystatins) have also been shown to modulate host immune responses. In *B. malayi* the cystatin homologue Bm-CPI-2 inhibits cysteine proteases in the endosomes/lysosomes of human B cells and inhibits the presentation of antigens via the MHC II pathway (Manoury *et al.*, 2001). Interestingly, while the *C. elegans* ortholgue Ce-CPI-2 has been shown to be essential in oocyte maturation and fertilization and mutating the *cpi-2* gene results in sterile worms, *C. elegans* cystatins do not have any effect on human or murine T cells (Hartmann and Lucius, 2003; Schierack *et al.*, 2003).

Metabolic enzymes are found in ES and the enzyme triose phosphate isomerase is a major component of *B. malayi* adult ES along with other glycolytic enzymes such as aldolase and GAPDH (Hewitson *et al.*, 2008). In the flukes *Fasciola hepatica* (Bernal *et al.*, 2004), *Schistosoma bovis* (de la Torre-Escudero *et al.*, 2010) and *Clonorchis sinesis* (Wang *et al.*, 2011b), enolase either in the ES or as

a surface protein binds human plasminogen and may contribute to parasite invasion of host tissues.

Many helminths proteins are heavily glycosylated which may also contribute to modulation of immune response. In *T. suis,* glycans present on ES proteins interact with DC via c-type lectin receptors. The glycans were also shown to be essential in the suppression of pro-inflammatory cytokines (such as TNF-α and IL-12) from human dendritic cells, as well as lipopolysaccharide induced pro-inflammatory chemokines such as C-X-C Motif Ligands 10 and C-C Motif Ligand 2 (Klaver *et al.*, 2013). The ability of glycans to polarize T cells has also been shown with schistosome egg antigens. Mice injected with antigen treated to alter the glycan structure failed to mount strong Th2 biased responses characteristic of immunisation with schistosome eggs (Okano *et al.*, 1999).

A further mechanism by which *H. contortus* may evade host immune responses is by inhibiting the host complement pathway. Adult worms are reported to produce a complement C3 binding glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Sahoo *et al.*, 2013) and a Calreticulin protein that binds to C1q (Suchitra and Joshi, 2005) and to C-reactive protein (Suchitra *et al.*, 2008). *H. contortus* also produces a 66kDa antigen in the ES and decreases production of hydrogen peroxide and nitric oxide in monocytes. Antibodies against the antigen caused loss of motility in adult worms *in vitro* and resulted in death of the parasite (Rathore *et al.*, 2006). The galectin Hco-gal-m is present on *H. contortus* worm gut surfaces but has also been identified as part of the excretory/secretory products of *H. contortus*, *Ostertagia ostertagi* and *T. circumcincta* (Hewitson *et al.*, 2009).

1.9 Genome of *H. contortus*

In 2013, Laing *et al.* and Schwartz *et al.* both published draft versions of the *H. contortus* genome. The two papers utilised different strains of *H. contortus* and employed different techniques to generate the genome. Laing *et al.* used an inbred drug susceptible strain called MHco3(ISE) (Laing *et al.*, 2013; Redman *et al.*, 2008). The genome published by Schwartz *et al.,* used a different strain: an Australian linguiform-macrocyclic lactone susceptible *H. contortus* strain called McMaster (Schwarz *et al.*, 2013). Overall, the genomes from both strains are

predicted to be a similar size with a similar number of predicted protein coding genes (Table 1.1).

Table 1.1 Summary statistics from both genome projects.

Assembly:	Haemonchus_contortus_MHco3-2.0, GCA_000469685.1 (Laing *et al.*, 2013)	Hco_v4_coding_submitted (Schwarz *et al.*, 2013)
BioProject ID	PRJEB506	PRJNA205202
Base Pairs:	369,846,877	319,640,208
Genome Size:	369,846,877	319,640,208
Data source:	Wellcome Trust Sanger Institute	University of Melbourne
Scaffolds:	23,860	14,419
Gene counts		
Coding genes:	21,869	23,610
Gene transcripts:	26,367	24,942

With the availability of the genome, a range of genetic and genomic resources can be developed for this parasite. However, both *H. contortus* genomes are currently very poorly annotated. Comparison of the genome with well-annotated genomes such as that of *C. elegans* will also allow prediction of *H. contortus* gene functions. From the gene models, the untranslated regions of the mRNAs can also be identified as well as the prediction of different species of RNAs such as ribosomal, transfer and the small non-coding RNAs.

In *C. elegans*, a widely used method of identifying gene function has been through the use of RNA interference (RNAi). RNAi is a naturally occurring mechanism whereby foreign double stranded RNA (dsRNA), from a virus or transgene, is recognised by the cell and processed by Dicer to form double-stranded fragments of 20–25 base pairs with a 2-nucleotide overhang at the 3' end. These are called small interfering RNA (siRNA). The sense strand of the siRNA is degraded, leaving the antisense strand that has the complementary sequence to the mRNA. In fruitflies and mammals, the antisense strand is incorporated into the RISC complex, targets complementary mRNA sequences and degrades them, silencing the gene. This pathway shares many similarities with the miRNA pathway (see Chapter 1.10). However, one interesting

difference in animals is that miRNA only require a partial mRNA match, while siRNA require an exact complementary sequence (Mack, 2007). In worms and plants, the antisense siRNA is amplified by pairing up with a complementary mRNA and synthesis of a new dsRNA. Dicer then processes the new dsRNA to produce different siRNAs that bind to different sites on the mRNA. The siRNA then binds and degrades target genes (Novina and Sharp, 2004; Vermeulen *et al.*, 2005; Zamore *et al.*, 2000).

In *C. elegans*, RNAi has been achieved by injection of adult worms with dsRNA or with dsRNA expression plasmids, soaking in a solution containing dsRNA or by feeding worms bacteria expressing dsRNA (Fire *et al.*, 1998; Tabara, 1998; Timmons and Fire, 1998). Attempts to replicate RNAi in parasitic nematodes have been reported with varying success in species including *Nippostrongylus brasiliensis*, *Brugia malayi*, *Ascaris suum* and *Trichostrongylus colubriformis* (Aboobaker and Blaxter, 2003; Hussein *et al.*, 2002; Islam *et al.*, 2005; Issa *et al.*, 2005).

In *H. contortus*, RNAi has been attempted in L1 and L3 stages. Initial findings suggest that RNAi is viable but depends upon the gene targeted and its location. RNAi by soaking of exsheathed L3 resulted in decreased transcript levels of beta-tubulin and *sec-23,* but not of the other nine genes tested. Electroporation of *H. contortus* L1 worked for two of four genes (decreased transcript levels for *Hc-ben-1* and *Hc-sod-1*) but feeding of L1-L2 stages did not decrease expression at all (Geldhof *et al.*, 2006). Soaking of exsheathed L3, L4 and adult worms in medium containing ß-tubulin dsRNA resulted in a significant decrease in the targeted gene, however changes were also detected in non-target genes and the RNAi treated L3 showed decreased viability and reduced development to the L4 stage (Kotze and Bagnall, 2006). In a further study, genes localised to the *H. contortus* L3 intestine, excretory or amphid cells were more susceptible to RNAi, suggesting that the site of gene expression was important with accessiblity to the environment optimizing the chances of success (Samarasinghe *et al.*, 2011).

With only certain genes being susceptible to RNAi, its use as a tool to identify gene function is limited in *H. contortus*. With the lack of such tools, ascribing gene function in *H. contortus* is difficult. This is compounded by the fact that RNAi in other animal parasitic nematodes, such as *H. polygyrus* also shows

limited success (Lendner *et al.*, 2008). While *C. elegans* is a good model for conserved genes, it has limited use in determining the function of parasite-specific genes.

1.10 microRNAs

microRNAs (miRNAs) are small, 19 to 25 nucleotide non-coding RNAs that have an important role in negatively regulating post-transcriptional gene expression. Vertebrate genomes are predicted to encode as many as 1000 unique miRNAs (Berezikov *et al.*, 2005) which potentially regulate expression of at least 30% of genes (Lewis *et al.*, 2005). miRNAs are able to suppress gene expression by base pairing to the 3' UTR, 5' UTR or the open reading frame of target mRNA and lead to transcriptional repression or cleavage of the mRNA (Bartel, 2009; Selbach *et al.*, 2008), although the 3' UTR is the best characterised binding site for most miRNAs. In animals, miRNAs exhibit partial complementarity to their mRNA targets. Target specificity is determined by the seed sequence, a 6 to 8 nucleotide region at the 5' end of the miRNA (Lewis *et al.*, 2003, 2005). Each miRNA is able to regulate multiple mRNAs (Lewis *et al.*, 2003; Lim *et al.*, 2005) and targeted mRNAs can be regulated by many different miRNAs (Pillai, 2005) demonstrating the complexity of this regulatory system. The first two miRNAs identified, *lin-4* and *let-7*, were discovered through genetic screens in *Caenorhabditis elegans* (Lee *et al.*, 1993a; Reinhart *et al.*, 2000). Both miRNAs were shown to be essential for the correct developmental timing of *C. elegans* and had extremely small, ~22nt, active transcripts which were derived from much longer hairpin structured RNA precursors. These short transcripts target the RNA-induced silencing complex (RISC) to partially complementary sequences in the 3' UTR of the target mRNA, resulting in gene silencing by either degradation of the mRNA or inhibition of translation.

Gastro-intestinal nematodes such as *Haemonchus contortus* are closely related to *C. elegans* both in body plan and developmental programme. Whether or not microRNAs are as important in parasitic nematodes is currently unknown. Winter *et al.,* (2012) sequenced small RNA libraries with the aim of identifying microRNAs in two species of parasitic nematode, *H. contortus* and *Brugia pahangi*. A number of miRNAs were found that were conserved between *C.*

elegans and *H. contortus* as well as many that were unique to *H. contortus* (Winter *et al.*, 2012).

1.10.1 miRNA Biogenesis

The miRNA genes are usually transcribed by RNA polymerase II as long, stem loop containing primary transcripts (pri-miRNAs), which have a 5' 7-methyl guanosine cap and are poly-adenylated at the 3' end (Lee *et al.*, 2004)(Figure 1.2). Each of the stem-loop structures is composed of about 70 nucleotides, which are flanked by sequences responsible for efficient processing.

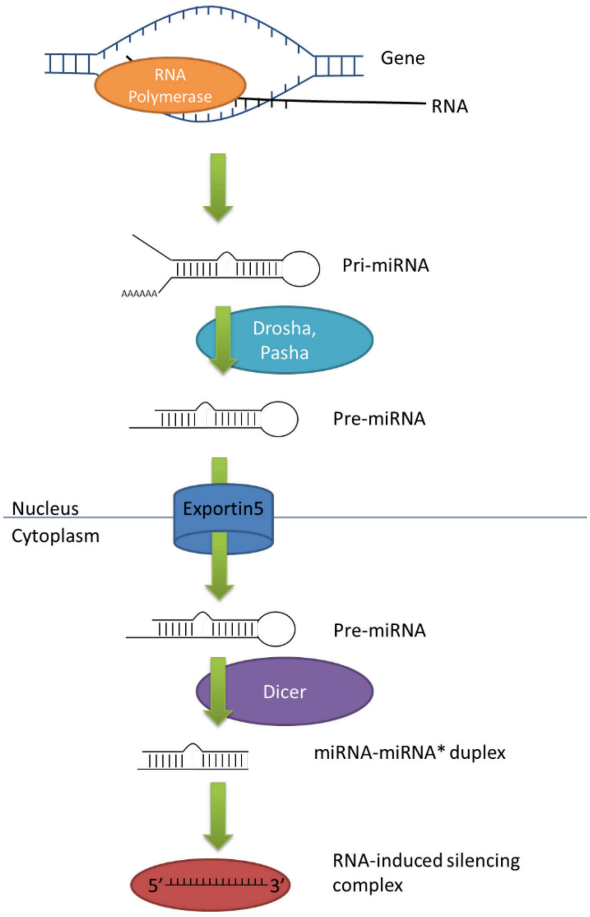


Figure 1.2 Biogenesis of microRNA. In the nucleus, RNA is transcribed by RNA polymerase II into pri-miRNAs. This is then processed by Drosha or Pasha into pre-miRNAs which are exported out of the nucleus by Exportin5. Pre-miRNas are cleaved by Dicer into miRNA duplexes and one strand is used in the microRNA-induced silencing complex. Figure adapted from Kim *et al.*, (Kim, 2005).

Drosha, a 160kDa RNaseIII endonuclease contains an RNAse III domain, an RNA binding domain and requires interaction with a cofactor, DiGeorge syndrome critical region gene 8 (DGCR8) in humans or Pasha in *D. melanogaster* and *C.*

elegans, to form a complex known as the microprocessor complex (Denli *et al.*, 2004; Han *et al.*, 2004; Landthaler *et al.*, 2004). The microprocessor complex cleaves the pri-miRNAs into a 60-70 nucleotide structure called the pre-miRNA. DGCR8 is believed to assist Drosha in recognising the substrate (Lee *et al.*, 2003). The Drosha microprocessor complex can measure the length of the stem of the hairpin pri-miRNA and cleave approximately two helical turns (~22 nucleotides) from the terminal loop (Zeng and Cullen, 2005).

The pre-miRNA has a 2 nucleotide overhang at the 3' end which is recognized by Exportin5, which exports the pre-miRNA from the nucleus into the cytoplasm via the nuclear pore complexes in a Ran guanosine triphosphate-dependent manner (Lund *et al.*, 2004). In the cytoplasm, the pre-miRNA is cleaved by the RNaseIII enzyme Dicer together with TRBP/PACT proteins. The loop that joins the 3' and 5' arms leaves a short imperfect double stranded miRNA:miRNA* duplex about 22 nucleotides long (Hutvágner *et al.*, 2001).

Either strand can act as a functional miRNA although only one is usually incorporated into the microRNA-induced silencing complex (miRISC) where the miRNA and its target interact. The strand used is selected on the basis of its thermodynamic stability. The strand with a relatively unstable base pair at the 5' end remains and the unselected strand is usually degraded (Khvorova *et al.*, 2003; Mourelatos *et al.*, 2002). miRNAs are denoted with either a -5p or -3p suffix, indicating if the miRNA originates from either the 5' or 3' arm of the pre-miRNA respectively. Sequences that have similar mature sequences have a lettered suffix, such as Hco-miR-40a and Hco-miR-40b.

A major component of the RISC complex is argonaute, an RNA binding protein that is essential for miRNA-induced silencing. Argonaute contains two RNA binding domains: a PAZ domain that can bind the single stranded 3' end of the mature miRNA and a PIWI domain that structurally resembles ribonuclease-H and functions to interact with the 5' end of the guide strand. The PAZ domain forms a specific binding module for the 2 nucleotide 3' overhang generated by Dicer (Parker and Barford, 2006; Song and Joshua-Tor, 2006). A third domain has been discovered between the PAZ and PIWI domains, called the MID domain (Parker *et al.*, 2005; Yuan *et al.*, 2005). This domain contains a highly basic pocket, which

is able to anchor small RNAs onto the argonaute proteins as it is able to bind the 5' phosphate.

When bound to the mature miRNA, the RISC is able to recognise complementary messenger RNA molecules and can degrade them or repress translation, resulting in decreased levels of protein translation. The mechanism that is employed by the RISC complex depends on the degree of complementarity between the miRNA and the mRNA. A high degree of complementarity generally results in silencing due to endonucleolytic cleavage.

1.10.2 microRNAs and disease

microRNAs are now known to be involved in a wide range of different processes such as metabolism, development, muscle differentiation and stem cell division (Hatfield *et al.*, 2005; Naguibneva *et al.*, 2006; Rottiers and Näär, 2012; Sayed and Abdellatif, 2011). With such important roles in a wide variety of biological processes, dysregulation of miRNAs can play a roles in a number of clinical diseases, such as cancer and autoimmune diseases (Hayes *et al.*, 2014; Pauley *et al.*).

The miRNA, miR-155, has been of considerable interest as it is involved in key biological processes such as haematopoiesis and immunity. miR-155 has been found to be overexpressed in a variety of tumours including thyroid carcinomas, breast, colon, pancreatic, cervical and lung cancer (Faraoni *et al.*, 2009). In pancreatic ductal adenocarcinoma (PDAC), the expression of the proapoptotic p53 gene, *tumor protein 53-induced nuclear protein 1* (TP53INP1) was reduced when miR-155 is overexpressed in PDAC cells (Gironella *et al.*, 2007). Furthermore, interaction between the TP53INP1 3' UTR and miR-155 was shown by luciferase assay (Gironella *et al.*, 2007) suggesting that miR-155 suppresses TP53INP1 which then fails to induce cell cycle arrest or apoptosis.

miR-21 is highly expressed in cultured human hepatocellular cancer cells and luciferase assay showed that miR-21 can target the 3' UTR of the tumour suppressor phosphate and tensin homolog (PTEN). Inhibition of miR-21 in human hepatocellular cancer cells caused increased expression of PTEN and decreased tumour cell proliferation, whilst transfection of cells with miR-21 increased

tumour cell proliferation (Meng *et al.*, 2007). The tumour suppressor protein, programmed cell death 4 (PDCD4), was also confirmed as a target of miR-21 by luciferase assay and in miR-21 inhibited MCF-7 cells, cellular proliferation was reduced. This could be rescued by inhibiting PDCD4 by siRNA (Frankel *et al.*, 2008).

Studies have also shown that the hepatitis C virus requires miR-122 expression. In humans, miR-122 is highly expressed in the human liver and is a key factor in liver development, differentiation, homeostasis and metabolic function (Bandiera *et al.*, 2015). Due to its essential role, absence of miR-122 in mice results in steatohepatitis and fibrosis (Hsu *et al.*, 2012; Tsai *et al.*, 2012), suggesting that this miRNA has anti-inflammatory and anti-fibrotic properties. miR-122 also has anti-tumour properties as decreased levels of miR-122 are associated with poor prognosis and metastasis (Bai *et al.*, 2009; Xu *et al.*, 2012; Zeng *et al.*, 2010) and miR-122 knockout mice had abnormal expression of genes involved in cell growth and death and developed liver tumours (Hsu *et al.*, 2012; Tsai *et al.*, 2012). This miRNA also has a role in control of liver-specific pathogens, the best described example of which is the hepatitis C virus (HCV). On the hepatitis C virus RNA genome, there are four binding sites for miR-122 (Nasheri *et al.*, 2011) although miR-122 appears to have a positive effect on viral translation and replication (Jopling *et al.*, 2005; Villanueva *et al.*, 2010). This observation is counterintuitive as binding of miRNAs to viral RNA is usually deleterious. The binding sites for miR-122 on the HCV genome also differs from most miRNAs as it targets the 5′ UTR instead of the 3′ UTR (Jangra *et al.*, 2010). The miR-122 binding sites are upstream of the HCV internal ribosome entry sites and the binding of miR-122 to these sites has been shown to enhance the association of ribosomes to the viral RNA (Jangra *et al.*, 2010). This association also protects the viral RNA from degradation by exonucleases (Conrad *et al.*, 2013; Li *et al.*, 2013). miR-199a also binds to the HCV genome and inhibits HCV replication (Murakami *et al.*, 2009). In addition, miR-122 and miR-199a have been shown to target the hepatitis B virus and repress HBV replication and mRNA expression (Chen *et al.*, 2011; Zhang *et al.*, 2010).

Miravirsen (previously known as SPC3649) is a 15 base oligonucleotide that is complementary to miR-122 and can inhibit miR-122. Miravirsen has been

demonstrated to have antiviral activity against all HCV genotypes *in vitro* (Li *et al.*, 2011). In an HCV infected chimpanzee model, miravirsen was able to suppress HCV RNA levels and no evidence of viral resistance or side effects to the therapy were detected (Lanford *et al.*, 2010). Phase 2 trials have also been conducted on patients with chronic HCV infection resulting in prolonged dose-dependent reductions in HCV RNA levels with some patients having undetectable levels of HCV RNA (Janssen *et al.*, 2013). Miravirsen is the first therapy to exploit a microRNA target and demonstrates the potential of anti-microRNA therapy.

1.10.3 Nematode microRNAs

miRNAs have been identified in at least 16 nematodes including *H. contortus* (Cai *et al.*, 2015a). An analysis of 11 free-living and animal or plant parasites from the phylum Nematoda found that there is some conservation of miRNAs across the different species with at least 20% of *C. elegans* miRNAs being conserved (Sarkies *et al.*, 2015). The deep sequencing work carried out by Winter *et al.*, demonstrated that in *H. contortus* and *B. pahangi*, only 28% and 40% respectively of the miRNAs identified were conserved in any other organisms (Winter *et al.*, 2012). However, in more closely related species such as between *C. elegans*, *C. remanei* and *C. briggsae*, many of the miRNAs are highly conserved (Shi *et al.*, 2013; de Wit *et al.*, 2009). Interestingly, comparison of *H. contortus* miRNAs with *Ascaris suum* miRNAs identified Hco-miR-5352 as the only miRNA conserved between the two, whilst seven miRNAs were identified to be conserved between *Ascaris* and *Brugia* (Winter *et al.*, 2012). It is possible that the lack of conservation in parasitic species is due to the absence of sufficient genome data on more closely related species.

In addition to small RNA sequencing projects, whole genome sequence data is becoming more available with WormBase ParaSite (www.parasite.wormbase.org/species.html) which contains 80 nematode genome assemblies. The presence of genome data will enable *de novo* identification of miRNAs by small RNA sequencing projects and mapping RNA reads to the genome using programmes such as miRDeep2 and mireap (Winter *et al.*, 2012).

In addition to identifying miRNAs, the RNA silencing pathway requires the presence of RISC and the argonaute (ago) protein. The argonaute protein is essential in the RISC, as it is involved in recognising and binding miRNAs and mediating gene silencing pathways (Peters and Meister, 2007). Homologues (or orthlogues) of the Ago-like gene 1 and 2 can be identified in all the fully sequenced nematode genomes (Buck and Blaxter, 2013).

Currently, *C. elegans* miRNAs are the most well characterised. Most of the work to date on parasitic nematode miRNAs is descriptive and does not probe the function of specific genes. The let-7 miRNA plays a key role in larval development in *C. elegans* (Ren and Ambros, 2015). In *Trichinella spiralis,* let-7 homologues can be identified and their expression is highest in the eggs and larvae respectively. Furthermore, the *Brugia pahangi* bpa-miR-5364 is also a member of the let-7 family and is significantly upregulated post-infection and may have a role in the transition of *B. pahangi* L3 stage from mosquito to mammalian host (Winter *et al.*, 2015). In *Schistosoma japonicum*, miR-31 and bantam are highly expressed, predominantly in the ovaries of female *S. japonicum* as are the mRNA targets of these miRNAs. Suppression of miR-31 and bantam in female worms resulted in 70% of the females with defects in their ovaries with numerous vacuoles and damaged oocytes, demonstrating the importance of these two miRNAs in schistosome reproduction (Zhu *et al.*, 2016a).

Luciferase reporter assays, Western blot and quantitative real-time PCR (qRT-PCR) are the most common methods used to determine individual miRNA:mRNA interactions. In a luciferase reporter assay, the 3' UTR of the target gene of interest is cloned downstream of luciferase and the binding of a given miRNA to the mRNA target site should cause a reduction in luciferase mRNA and protein. Modulation of the mRNA by the miRNA should also affect the amount of target protein being produced and Western analysis using an antibody against the target protein can also be used to validate a miRNA/mRNA interaction. Finally, qRT-PCR can be performed using total RNA isolated from cells and probes or primers for a specific miRNA and mRNA target (Kuhn *et al.*, 2008).

Functional characterization of a specific miRNA has usually involved differential expression of a single miRNA. A miRNA can be over-expressed by transfection of

a synthetic miRNA mature or precursor or by introduction of a miRNA expression construct. Transient transfection of a synthetic miRNA precursor is a commonly used for over-expressing the miRNA of interest (Martinez-Sanchez and Murphy, 2013; Thomson *et al.*, 2011), however, the precursor must be incorporated into the RISC complex to become active. The high level of synthetic precursor may saturate the RISC complex and displace endogenous miRNAs leading to potential false positive results (Khan *et al.*, 2009). Instead, miRNA mimics are now a more common approach as they bypass the normal miRNA biosynthesis. miRNA mimics are chemically synthesized RNA duplexes that are designed to have only a single active miRNA strand.

On the other hand, it is also possible to inhibit miRNAs through the use of chemically modified antisense oligonucleotides (anti-miRs). Anti-miRs are designed to specifically target miRNAs and bind with high affinity and thus relieve inhibition of genes by the miRNA of interest (Li and Rana, 2014).

1.10.3.1 Secreted helminth miRNAs – roles and functions

Modulation of a host system by miRNAs has been shown in viruses where miRNAs produced by the virus are able to prolong longevity of infected cells and evade the immune system (Kincaid and Sullivan, 2012). The large number of novel parasitic helminth miRNAs, and the importance of certain miRNAs in the host immune response, suggests that parasite specific miRNAs may play important roles in the host-parasite relationship.

As microRNAs can be secreted, several studies have investigated whether parasite-derived miRNAs can be detected in the serum of infected animals or human patients. In mice and humans infected with *S. mansoni*, parasite-derived miRNAs were detectable in the serum (Hoy *et al.*, 2014). 62 miRNAs were also identified in nodule fluid from cattle infected with *Onchocerca ochengi*, 59 of which were identical to miRNAs from *Onchocerca volvulus* (Quintana *et al.*, 2015). In gastro-intestinal nematodes, *H. polygyrus* miRNAs can be identified in the ES, however these miRNAs could not be detected in serum of infected mice (Buck *et al.*, 2014) suggesting that the miRNAs released by the nematode may act locally at the mucosal surface. Furthermore, secreted miRNAs can be split into two groups: those that are enriched in the supernatant, and those that are

enriched in extracellular vesicles (exosomes). Extracellular vesicles (EV) are structures formed from the plasma membrane that are approximately 30 to 100 nm diameter and are able to transport DNA, RNA and proteins (Raposo and Stoorvogel, 2013a; Yáñez-Mó *et al.*, 2015a). In *H. polygyrus*, exosomes are thought to originate from the worm intestine and are able to protect the miRNAs from degradation and can be internalised by mouse intestinal epithelial cells. This phenomenon does not seem to be limited to gastro-intestinal nematodes, as exosomes containing miRNAs have also been identified in other species (Hu *et al.*, 2012) and have been shown to be able to modulate host gene expression (reviewed by Zhang *et al.*, 2015).

In *S. mansoni,* analysis of the EV secreted by schistosomula identified 109 proteins and 205 miRNAs (Nowacki *et al.*, 2015). While the targets and function of these miRNAs are unknown, three of the miRNAs have been previously identified in serum from infected mice, humans or rabbits. In addition, a number of *S. mansoni* miRNAs can be found in similar abundance between the EV depleted or EV enriched fractions suggesting that those miRNAs that are not encapsulated by EV are stabilised by proteins, such as argonaute. EV containing RNA can also be identified from the L1 of the porcine whipworm *Trichuris suis*. The exact content of RNA inside the EV was not investigated although the short length of the RNA suggests that they may be miRNAs (Hansen *et al.*, 2015). The adult fluke, *Dicrocoelium dendriticum* has also been shown to produce EV. These EV contain 84 proteins with a large proportion of proteins involved in vesicle transport and known exosome components such as heat shock protein 70, as well as a number of miRNA sequences that show high homology to known *S. mansoni, S. japonicum and E. multilocularis* miRNAs (Bernal *et al.*, 2014)

1.10.4 Biomarkers of disease

Secreted miRNAs may also have potential as biomarkers of infection. Current methods to diagnose gastro-intestinal nematodes rely on faecal egg counts. However, whilst this is relatively simple and inexpensive, faecal egg counts are not an accurate indication of adult worm numbers. As miRNAs are stable in most clinical samples such as serum and as some miRNAs are well conserved between different species (Etheridge *et al.*, 2011), they may be useful as diagnostic tools. In humans, a number of microRNAs have already been proposed as possible

diagnostic biomarkers for a range of diseases, such as expression of miR-141 in the serum of prostate cancer patients (Mitchell *et al.*, 2008); mir-29a and mir-92a are elevated in plasma from patients with colorectal cancer (Huang *et al.*, 2010) while miR-1 is elevated in patients with acute myocardial infarctions (Ai *et al.*, 2010).

In helminth infections, parasite-specific miRNAs could serve as possible biomarkers as they would be absent from the uninfected host. In dogs infected with the filarial nematode *Dirofilaria immitis*, miR-71 and miR-34 in the serum can be used to differentiate infected from uninfected animals (Tritten *et al.*, 2014). Deep sequencing of small RNAs from *Schistosoma japonicum* infected rabbits identified five parasite specific miRNAs (sja-miR-8185, Bantam, miR-3479, miR-10 and miR-3096) with all except miR-3096 being detectable in the plasma by RT-PCR (Cheng *et al.*, 2013). In humans infected with *S. mansoni*, miR-277, miR-3479-3p and bantam could be detected in the serum and can be used to determine if a patient was egg positive or negative with a high specificity and sensitivity (Hoy *et al.*, 2014). Further experiments confirmed that miR-277 and miR-3479-3p, but not bantam could be detected in serum of *S. japonicum* infected mice, independent of mouse strain (Cai *et al.*, 2015b).The expression of these two miRNAs correlated with egg burden and liver fibrosis (Cai *et al.*, 2015b). In the filarial nematode, *Litomosoides sigmodontis* parasite-derived miRNAs could also be detected in the serum of mice, with miR-100, bantam and miR-86 being the most abundant (Buck *et al.*, 2014). However, the usefulness of parasite-derived miRNAs may be limited to those parasites living in the tissues as in mice infected with *H. polygyrus* or sheep infected with *H. contortus*, the serum from infected animals did not contain detectable levels of miRNAs even those that are highly expressed in the nematode (Britton *et al.*, 2015; Buck *et al.*, 2014).

Another potential source of miRNAs could be in the stools. miRNAs have been characterised in human and mouse faeces. These miRNAs are also present in extracellular vesicles and are able to target and regulate the gut microbiota and may be a host defence mechanism(Liu *et al.*, 2016). In humans, miRNA-135b and miR-20a were found to be significantly higher in patients with colorectal cancer which was reflected in the stool samples (Wu *et al.*, 2014; Yau *et al.*, 2015).

In addition to the presence of parasite derived miRNAs, the presence of the parasite may affect the expression of host miRNAs, which could also be used as a biomarker of disease. In *S. japonicum* infections, host miR-223 is significantly elevated but returns to normal after the mice were treated with praziquantel (He *et al.*, 2013). Decreased serum miR-223 is also a potential biomarker for sepsis (Wang *et al.*, 2010b) as well as having roles in inflammation, haematopoietic differentiation and in cancer (Haneklaus *et al.*, 2013).

The large amount of data on miRNAs from parasitic helminths have shown that they are novel regulators of important biological functions such as drug resistance and development and as such could serve as potential targets of therapeutics (Britton *et al.*, 2014). Parasite miRNAs have also been shown to be present in the ES and can be found enclosed in extracellular vesicles. These secreted miRNAs have been found to be a potential regulator of host genes and may aid in the parasites' ability to evade or modulate the host immune system. This may have the potential to lead to interesting therapeutics both in terms of reducing helminth survival within a host as well as potentially developing a therapy that utilises extracellular vesicles to deliver miRNAs.

1.11 Aims of the project

The overall objective of this project was to determine if the *H. contortus* mir-5352 cluster plays an important role in the host-parasite interaction. The specific aims were as follows:

1. Investigate the Hco-miR-5352 locus, the only miRNA identified to which is conserved between *H. contortus* and the parasitic nematode *Ascaris suum*.

2. Characterise the miR-5352 cluster and its potential function in the host-parasite interaction.

3. Identify mammalian targets of the hco-miR-5352 cluster and validate using experimental methods.

4. Characterise the miRNA content of *H. contortus* ES using small RNA libraries and the presence of extracellular vesicles in the ES.

5. Investigate effects of hco-miR-5352 on the global gene expression in a mouse intestinal epithelial cell.

2 Methods and Materials

2.1 Collection of *H. contortus* worms

Adult or L4 *Haemonchus contortus* (MHco3(ISE) strain) were collected from the Moredun Research Institute from sheep infected with 5000 L3, on day 7 (for L4) or day 28 (for adults). Worms were washed 3 times in RPMI supplemented to a final concentration with 100 units per ml and 100 µg per ml of penicillin/streptomycin and 1.25 µg/mL amphotericin B. Worms were then transported to the University of Glasgow in the above solution, maintaining them as close to 37ºC as possible.

2.2 *In vitro* culture of adult and L4 stages of *H. contortus*

Approximately 100 adult worms (mix of male and female worms) were cultured in 25 ml of RPMI supplemented with 1% glucose and 100 µg per ml penicillin/streptomycin. After 24 hours, spent medium was pipetted out and replaced with fresh medium. The spent medium was then centrifuged at 1500 g for 5 minutes to remove any eggs and L1 and then filtered through a 0.22 µm filter and stored at -80°C. The total volume was concentrated from approximately 50 ml to 1ml using Vivaspin 20 ml 10,000 MWCO sample concentrators. Spin tubes were first rinsed with deionized water before filling with sample material. Tubes were spun at 3500 rpm at 4°C until sample material was concentrated to approximately 1 ml. The concentrated medium, henceforth referred to as ES, were then stored at -80°C.

2.3 RNA extraction and precipitation

2.3.1 Qiagen Qiazol method

To 200 µl of concentrated ES (Chapter 2.2), 1 ml of QIAzol reagent was added and left at room temperature for 5 minutes, then RNA extraction was carried out according to the manufacturer's instruction (Qiagen). To precipitate the RNA, 1 µl of glycogen, 20 µl of 3M NaCl and 500 µl of 100% EtOH was added to the sample and left at -80°C overnight. The next morning, samples were spun at 14,000 g for 25 minutes at 4°C. Supernatant was replaced with 0.5 ml of 75%

EtOH and spun at 7,500 g for 5 minutes. The resulting RNA pellet was resuspended in DEPC H_20 until the RNA pellet had dissolved.

2.3.2 Trizol method - Homogenizing samples

2.3.2.1 Tissues and Whole Worms

Tissues and whole worms were ground to a fine dust in a liquid nitrogen cooled pestle and mortar. 1ml of Trizol was then added and left to liquefy.

2.3.2.2 Cells in suspension

1 x 10^6 cells in suspension were spun down at 1000 g for 5 minutes and resuspended in 1 ml of Trizol. Cells were subjected to repetitive pipetting before vortexing for 3 minutes.

2.3.2.3 Adherent Cells

1 ml of Trizol was added to each well directly and lysed by repetitive pipetting and then vortexed for 1 minute.

2.3.3 RNA extraction from sheep abomasal and lymph node tissue.

30 mg of abomasal tissue or 10 mg of lymph node tissue were added to a hard tissue homogenizing CK28 tube (Precellys). Tissues were obtained from sheep that had been infected 28 days previously with L3 of *H. contortus* or from uninfected, pathogen-free sheep. At this time point, worms are adult and the females are producing eggs. 1 ml of Trizol reagent was added to each tube. Tissue was homogenised using a Precellys 24 high-throughput tissue homogenizer. Lymph node tissue samples were homogenised for 2 x 50 second cycles at 6000 rpm. Abomasal tissues were homogenised for 5 x 23 second cycles at 6200 rpm. Between each cycle, samples were left on ice for 2 minutes. Homogenised tissues were then centrifuged at 14,000 rpm for 10 minutes at 4°C to pellet tissue debris. 0.2 mL of chloroform was added per 1 ml of Trizol Reagent and RNA extracted as detailed in section 2.3.4

2.3.4 Trizol method - RNA isolation

RNA was extracted from the samples using Trizol (Life Technologies) according to the manufacturer's recommendations. Briefly, 0.2 ml of chloroform was added per 1 mL of Trizol reagent and shaken thoroughly. Tubes were then incubated for 3 minutes at room temperature before centrifuging at 12,000 g for 15 minutes at 4°C. The upper aqueous phase was separated into a new tube and 0.5 mL of room temperature 100% isopropanol was added. Tubes were then incubated at room temperature for 10 minutes before centrifuging at 12,000 g for 10 minutes at 4°C. The supernatant was removed and 1 mL of 75% ethanol was added. The sample was vortexed briefly, then centrifuged at 7,500 g for 5 minutes at 4°C. The wash was discarded and tubes were air-dried for 5 to 10 minutes. The RNA pellet was then resuspended in RNase free water and the solution was passed several times through a pipette tip. Yield and purity were assessed using a Nanodrop spectrophotometer.

2.3.5 miRNA qRT-PCR

The RNA extracted was first DNase I treated (Life Technologies) according to the manufacturer's instructions. Briefly, 2 µg of RNA was added to 1 µl 10X DNase I buffer, 5 µl DNase and made up to 50 µl with RNase free water and incubated for 30 minutes at 37°C. 1 µl of 50 mM EDTA was added and incubated at 75°C for 10 minutes to inactivate the DNase I.

In an RNAse-free microcentrifuge tube, 4 µl of total RNA was added, 4 µl of 5 x poly A polymerase buffer, 1 µl of 10 mM rATP and made up to 20 µl with RNAse-free water. 1 µl of *E. coli* poly A polymerase was added and the reactions incubated at 37°C for 30 minutes followed by 5 minutes at 95°C to terminate the adenylation

For each RNA sample, cDNA synthesis reaction was prepared by adding the following components to an RNase free microcentrifuge tube: 2 µl of 10X AffinityScript RT buffer, 4 µl of the polyadenlyation reaction, 0.8 µl of 100 mM dNTP mix, 1 µl of 10 µM RT adaptor primer,1 µl of AffinityScript RT/RNase Block enzyme mixture and made up to 20 µl final volume with RNase-free water. Reactions were gently mixed and incubated at 55°C for 5 minutes, then 25°C for

15 minutes and 42°C for 30 minutes. To terminate the reverse transcription, reactions were incubated at 95°C for 5 minutes. Samples were stored at -20°C.

qRT-PCR was performed following the miRNA qRT-PCR Master Mix protocol from Agilent Technologies. Samples were read in the Agilent Mx3005P qRT-PCR System and analyzed using MxPro QRT-PCR software (Agilent Technologies). All qRT-PCR reactions were carried out in triplicate and mean values plotted using Microsoft Excel.

Table 2.1 Table of qRT-PCR primers used. Oar-miR-122a, Oar-miR-26a and Oar-miR-103 were adapted from their bovine homologues (Coutinho *et al.*, 2007).

miRNA	miRBase ID	Sequence
Hco-miR-5895	MIMAT0023337	CGT AGC ATC TGT ATG TCT
Hco-miR-5960	MIMAT0023464	GGA GTC GGA GGG TTA T
Hco-miR-45	MIMAT0023339	CTA GAG ACA CAT TCA GCT
Hco-miR-61	MIMAT0023335	TAG ACT GTT ACT CGG CGT
Hco-miR-43	MIMAT0023336	TAT CAC AGT GTC ATT GGG T
Hco-miR-50	MIMAT0023329	TGATATGTCTGATATTCTTGG
Oar-miR-122a	Adapted from Bta-miR-122a MIMAT0003849	TGG AGT GTG ACA ATG GTG
Oar-miR-26a	Adapted from Bta-miR-26a MIMAT0003516	TTC AAG TAA TCC AGG ATA GG
Oar-miR-103	Adapted from Bta-miR-103 MIMAT0003521	AGC AGC ATT GTA CAG GG

2.4 Extracellular vesicle (EV) isolation by ultracentrifugation

ES material was collected as described in Chapter 2.2. The diameter of extracellular vesicles are reported to be between 30 and 100 nm and are able to pass through the 0.22 µm filter. 12.5 ml of ES was added to a 14 ml ultracentrifuge tube and centrifuged at 100,000 g for 2 hours in a Sorvall Discovery 90SE using a Surespin 630 swing bucket rotor with 17 ml buckets. The supernatant was aspirated and transferred into a fresh 50 ml Falcon tube and stored at -80°C (EV-depleted fraction). Centrifuge tubes were refilled with additional ES sample and centrifuged again. After aspirating the supernatant

again, the EV pellets were washed twice with 12.5 ml of PBS. Before the final wash, the pellets were resuspended by pipetting and pooled into a single tube. After the final wash, most of the supernatant was removed except for 100-200 µl of PBS which was used to resuspend the pellet, which was then stored at -80°C.

2.4.1 Estimation of protein concentration

Protein concentrations were estimated using the Qubit® Protein Assay Kit (Thermo Fisher Scientific). 200 µl of Qubit® Working Solution was prepared by diluting the Qubit® Protein Reagent 1:200 in Qubit® Protein Buffer. Three assay tubes (0.5 mL thin-walled PCR tubes) were prepared for the three standards with additional tubes for each sample being assayed. For the standards, 10 µl of each Qubit® standard was added to 190 µl of the Qubit® Working Solution. Between 1 µl and 20 µl of sample was added to each tube and made up to 200 µl with the Qubit® Working Solution. All tubes were vortexed and incubated at room temperature for 15 minutes before protein concentrations were read using a Qubit® 2.0 Fluorometer.

2.4.2 Electron microscopy

Electron microscopy was carried out at the University of Edinburgh by Thomas Tzelos. Briefly, 8 µl of resuspended EV pellet was fixed with 8 µl of 4% paraformaldehyde. 5 µl of resuspended pellet were deposited on Formvar-carbon coated EM grids. Membranes were covered and left to absorb for 20 minutes in a dry environment. 100 µl of PBS was dropped onto a sheet of Parafilm and the grids were transferred onto the Parafilm (membrane side down) to wash and then onto a 50 µl drop of 1% glutaraldehyde for 5 minutes. Grids were then transferred onto a 100 µl drop of distilled water and left to stand for 2 minutes, a procedure repeated a further seven times. To contrast and embed samples, the grids were transferred to a 50 µl drop of methyl cellulose-UA for 10 minutes on ice. The grids were then blotted to remove excess fluid on Whatman No 1 filter paper so that a thin film was left behind over the EV side of the grid and left to air dry for 5 to 10 minutes. EV were observed under the electron microscope at 80 kV. (http://users.ugent.be/~kraemdon/Exosomes-microvesicles/Thery%20et%20al.,%20Exosome%20Protocols.pdf)

2.4.3 Small RNA library preparation and sequencing

For all samples, total RNA was extracted using the Qiagen miRNeasy kit. For the total ES library from adult worms, ES from 24 and 48 hours was combined. To optimise the amount of RNA available for the EV libraries from both L4 and adult worms, ES was collected from 24, 48 and 72 hour cultures and combined for RNA extraction. 200 µl of concentrated ES material (prepared as described above) was mixed with 1 ml of Qiagen QIAzol lysis buffer and vortexed for 10 seconds. The mix was then left at room temperature for 5 minutes before adding 240 µl of chloroform. Samples were inverted several times and left at room temperature for 3 minutes before centrifuging at 12,000 g for 15 minutes at 4°C. The upper aqueous phase was transferred to a new collection tube and 1.2 ml of 100% ethanol was added. 700 µl of sample was then transferred into the Qiagen RNA easy Mini spin column. Columns were spun at 8,000 g for 15 seconds at room temperature. On column DNase digestion was carried out with the Qiagen RNase-Free DNase kit. 350 µl of the RWT buffer was added to the spin column and centrifuged at room temperature for 15 seconds at 8,000 g. 10 µl of DNase I stock was mixed with 70 µl of Qiagen RDD buffer and placed carefully onto the column and incubated at room temperature for 15 minutes. 350 µl of RWT buffer was then added and centrifuged at 8,000 g for 15 seconds. Columns were then washed with 500 µl of Qiagen RPE buffer and centrifuged for 15 seconds at 8,000 g. A further 500 µl of RPE buffer was added and columns centrifuged for 2 minutes at 8,000 g. 30 µl of RNase free water was added to the column membrane and centrifuged for 1 minute at 8,000 g to elute. Elution was repeated a second time with 30 µl of water to elute as much RNA as possible. Buffers RWT, RDD and RPE are proprietary Qiagen buffers and are included in the Qiagen miRNeasy kit

RNA integrity was determined using an Agilent 2100 Bioanalyzer by the University of Glasgow Polyomics Facility. Small RNA libraries were generated from RNA samples using the Illumina Trueseq™ Small RNA Preparation kit according to the manufacturer's instructions. Small RNA library preparation for the adult total ES sample was carried out by the University of Glasgow Polyomics Facility and small RNA libraries of the EV and EV-depleted supernatant from L4 and adult ES were prepared by LC Sciences (LC Sciences, Houston, TX).

2.4.3.1 EV from T. circumcincta

Concentrated EV was obtained from the L4 stage of *T. circumcincta* exactly as detailed for *H. contortus* (Chapter 2.4) by Thomas Tzelos (Moredun Research Institute, Edinburgh). Total RNA was extracted from *T. circumcincta* EV as described above and a small RNA library created by the Glasgow Polyomics Facility.

2.4.4 Data Analysis

Illumina sequence reads from the University of Glasgow Polyomics Facility were processed to clip adaptor sequences and identify duplicate reads. miRDeep2 was used to map the reads to a list of miRNA precursor, mRNA, tRNA and rRNA sequences obtained from Dr Winter (Winter *et al.*, 2012). Sequences were mapped to the RNA sequences with perfect alignments, allowing a maximum of one mismatch in the 3' end of the alignment.

Data from the libraries generated by LC Sciences and the data from the library generated by Glasgow Polyomics Facility were processed together using a proprietary pipeline script by LC Sciences. Briefly, mappable sequences from raw reads were extracted from the raw data. Reads were mapped to *H. contortus* miRNAs from miRBase (version 21, 2014) and the *H. contortus* genome. Reads that did not map to *H. contortus* were mapped to other nematode miRNAs in miRBase. Mapped reads were then normalised against each other. The normalisation method calculated the geometric mean for each miRNA across the five samples. The raw reads were then adjusted based on the geometric means. The median value of the adjusted read counts for each library were calculated (library size parameter) and the raw reads were then normalised based on the library size parameter.

2.5 HEK293 cell culture

To set up the dual luciferase assay to investigate the potential interaction of miR-5352 with target gene CD69, HEK293 (human embryonic kidney) cells were used and maintained in Dulbecco's Modified Eagle Medium (Invitrogen) supplemented with 5 mM L-Glutamine (Invitrogen), 10% heat inactivated foetal bovine serum (HI FBS, Invitrogen), 100 µg per ml penicillin/streptomycin (Life

Technologies). Cells were grown in 75 cm^2 vented flasks (Corning) at 37°C with 5% CO2. Cells were passaged after reaching 80-90% confluence. The medium in the flask was aspirated and 2 ml of pre-warmed 0.25% trypsin-EDTA (Invitrogen) was added and left for 3 minutes to detach the cells. 8 ml of fresh warmed medium was added to stop trypsinization and the cells were transferred to a 15 ml Falcon tube. Cells were pelleted by centrifuging at 1200 rpm for 5 minutes, the supernatant decanted and the cells resuspended in 10 ml of fresh medium. Cells were counted and 1 x 10^6 total cells transferred to a new flask for routine passage. Cells were also plated into a 96-well plate for the dual luciferase assay, as described in 2.8

2.6 Plasmid purification

Competent *E. coli* TOP10 cells (Life Technologies) were removed from -80°C and thawed on ice. 10 ng of plasmid was mixed gently with 50µl of competent cells and incubated on ice for 30 minutes. Cells were heat shocked for 45 seconds at 42°C and placed back on ice for 2 minutes. 800 µl of SOC medium was added to each tube and incubated at 37°C for 1 hour. 10 µl and 100 µl of the suspension were plated out on LB agar plates containing the appropriate antibiotic and incubated in a 37°C incubator overnight. Single colonies were picked from the plate and sub-cultured in 250 ml of LB Broth and grown overnight in a 37°C shaking incubator. Plasmid purification was carried out using the Qiagen EndoFree® Plasmid Maxi Kit according to the manufacturer's instruction and the DNA concentrations were checked on a Genequant pro.

2.7 Dual Luciferase Reporter Assay in HEK cells

The dual luciferase assay (Promega) requires two plasmids; a plasmid containing the Firefly luciferase gene and the 3′ UTR of interest, and a Renilla luciferase plasmid to act as a transfection control. The Firefly luciferase plasmid containing the 3′ UTR of CD69 (pMirTarget) was purchased from Origene, while the Renilla expressing plasmid (phRG-TK Renilla) was a kind gift from Dr Jane Kinnaird (University of Glasgow). The Hco-miR-5352 binding sites (AUGUGCAA) on the CD69 3′ UTR was scrambled to (GTCAAGT). The mutated CD69 3′ UTR was then synthesised and inserted into the pMirTarget plasmid by OriGene. Two

plasmids, pCMV and pEGFP-N1 containing the Hco-miR-5352 cluster were gifts from Dr Alan Winter (University of Glasgow).

A description of each plasmid is presented in Table 2.2.

Table 2.2 Description of plasmids used.

Plasmid Name	Fluorescent tag	Antibiotic Resistance	Concentrations used (ng/96 well)
pMirTarget (containing the 3′ CD69 UTR)	Red Fluorescent Protein	Neomycin/Kanamycin	25
pMirTarget (containing the 3′ CD69 mutated UTR)	Red Fluorescent Protein	Neomycin/Kanamycin	25
pHRG-TK Renilla (PGL3 basic vector backbone)	No fluorescent protein as it can affect the renilla reading	Ampicillin	1
microRNA expression plasmids			
pCMV	EGFP	Ampicillin	50
pEGFP-N1	EGFP	Neomycin/Kanamycin	50

24 hours prior to transfection, 1 x 10^4 cells were plated out in a 96-well plate in 100 µl of complete growth medium. For each well, a maximum of 100 ng of DNA was mixed with 20 µl of Opti-MEM® I Reduced Serum Media (ThermoFisher Scientific) without serum and 0.35 µl of Lipofectamine® LTX® Reagent (ThermoFisher Scientific). Mixtures were incubated for 30 minutes and 20 µl of the DNA-Lipofectamine® LTX® Reagent complexes were added to each well. The plate was gently rocked before incubating at 37ºC in a 5% CO_2 incubator for up to 48 hours. Cells were transfected with different ratios of plasmids (see results for details) in order to optimise the read-out while maintaining viability.

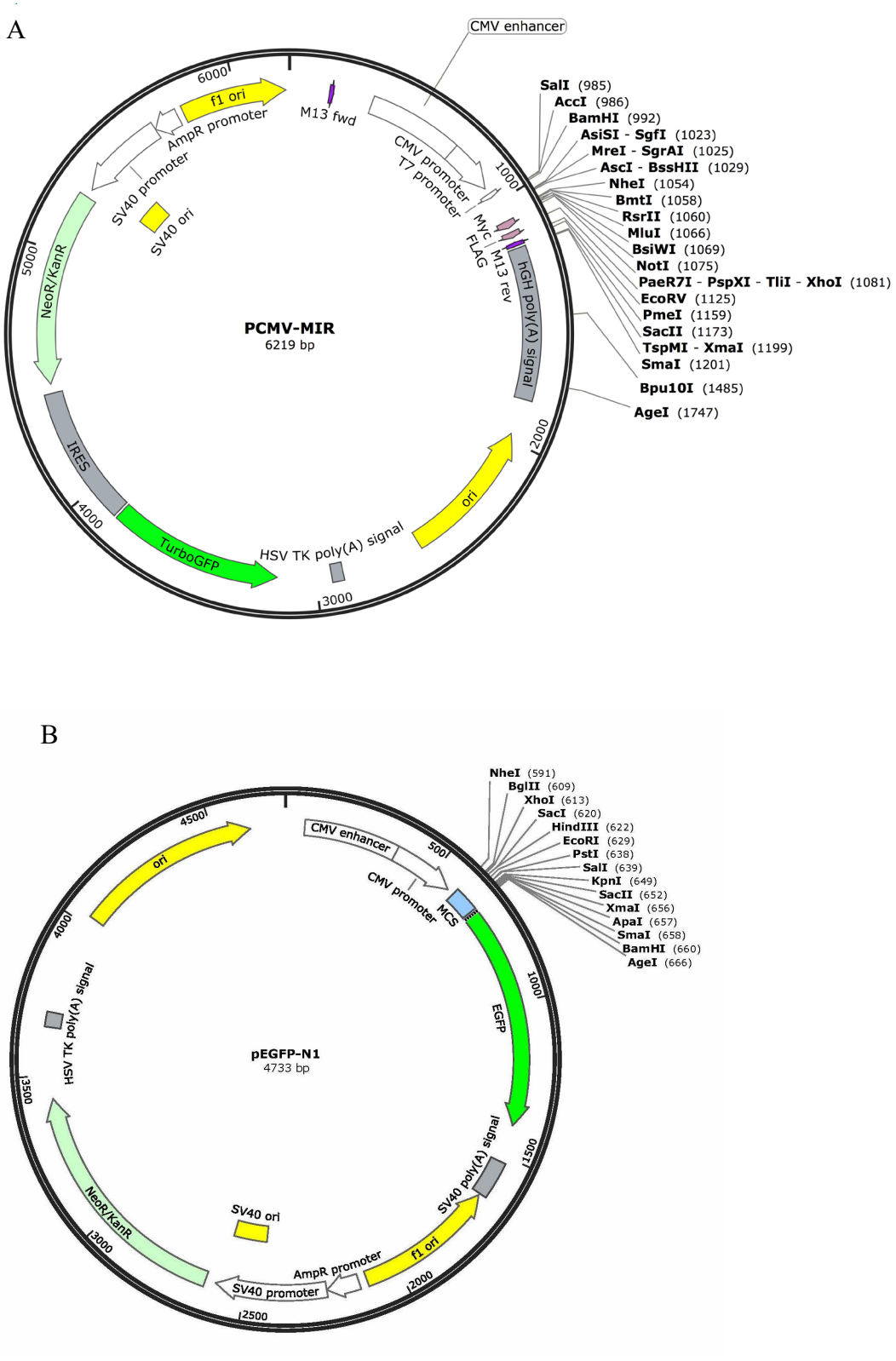


Figure 2.1 Schematic of the two miRNA expression plasmids, A) pCMV, Hco-miR-5352 cluster insert cloned into Sgf I/Not I and B) pEGFP-N1, Hco-miR-5352 insert cloned BamHI/BamHI. Plasmid schematics were adapted from www.addgene.org and www.origene.com.

The dual luciferase assay was carried out as specified in the manufacturer's protocol and adjusted for the plate size. Briefly, medium was removed from the cultured cells, which were washed once with PBS and then the supernatant removed. 20 µl of 1X passive lysis buffer was added to each well and shaken for 15 minutes at room temperature. Cell lysates were transferred into a white 96 well plate and 100 µl of Luciferase Assay Reagent II added to measure the Firefly luminescence first, before adding the Stop & Glo® reagent to quench the Firefly luminescence and initiate the Renilla signal. Luminescence was measured using a GloMax® 96 luminometer and data exported into Microsoft Excel. In each experiment, untransfected cells were used to determine background fluorescence. The averaged value for Firefly and Renilla background signal was subtracted from all other readings. As the Renilla signal acts as a transfection control, the Firefly signal was normalized to the Renilla signal to give a Firefly/Renilla ratio.

A double stranded RNA molecule was designed to mimic the sequence of native Hco-miR-5352 (Dharmacon). The miRIDIAN microRNA mimic (sequence UUGCACAUGAUGUACGACCUGCA) was chemically enhanced with the Dharmacon proprietary ON-TARGET modification protocol to preferentially programme RISC. This mimic was used in some experiments instead of the Hco-miR-5352 expression plasmids. For these experiments, 150 nM of the Hco-miR-5352 mimic was transfected into HEK293 cells with the Firefly and Renilla plasmids and either 150 nM siGLO Red transfection indicator (Dharmacon) or 150 nM of the miRIDIAN mimic negative control, Cel-miR-67 (Dharmacon). Cel-miR-67 was chosen as a negative control as it is commonly used in mammalian cell assays.

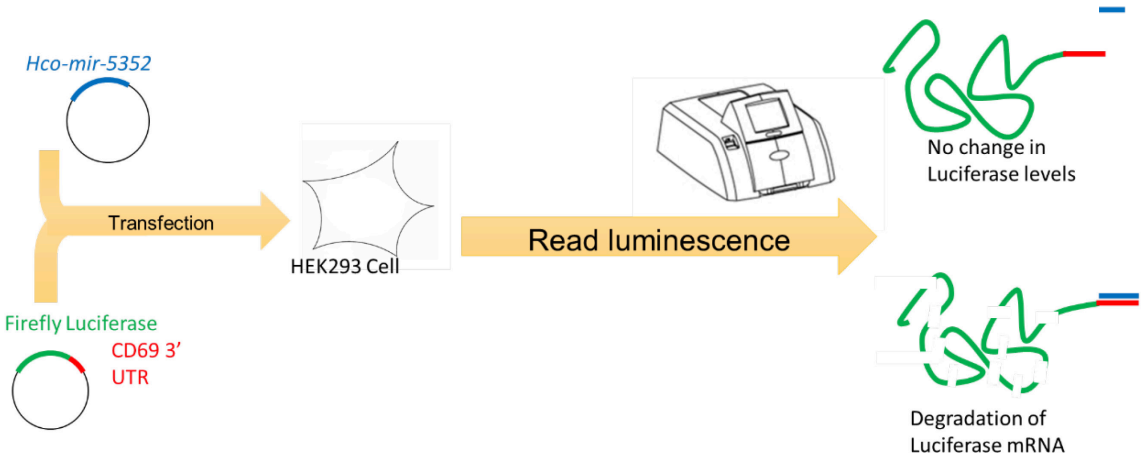


Figure 2.2 Format of the Dual Luciferase Assay. HEK293 cells were plated into tissue culture plates. After 24 hours, cells were transfected with the appropriate plasmids and left for up to 48 hours before lysis and addition of luciferase substrate. If the miRNA binds to the 3' UTR of CD69, cloned downstream of the luciferase reporter gene, the mRNA will be degraded causing a reduction in the level of luminescence.

2.8 Jurkat Cells and Transfection assays

To further investigate a possible interaction between miR-5352 and the CD69 3' UTR, Jurkat T leukemia cells (clone E6-1) were used, as it is possible to induce expression of CD69 on these cells. Jurkat cells were grown in RPMI 1640 medium (ATCC modification) supplemented with 10% HI-FBS (Invitrogen). Cells were grown in 75 cm^2 vented flasks (Corning) at 37°C with 5% CO_2. Cells were passaged 2-3 times a week and grown to a maximum density of 1.5 x 10^6 cells/ml. Transfection assays were performed using electroporation with the Amaxa™ transfection kit, according to the manufacturer's recommendations. Cells were sub-cultured for two days before nucleofection. On the day of transfection, 1 x 10^6 cells were aliquoted for each nucleofection reaction, pelleted at 1000g for 5 minutes and the supernatant removed. The cell pellet was then carefully resuspended in 100 µl of Nucleofector® Solution. 150 nM of miRNA mimic or 2 µg of miRNA expression plasmids was added to the cell suspension and transferred to an Amaxa™ certified cuvette. The cuvette was inserted into the Nucleofector® Cuvette holder and a Nucleofector programme was applied. The manual suggested the use of two programmes, programme 1 for high viability, and programme 5 for high expression and both were tested. The sample was then incubated for 10 minutes at room temperature. 500 µl of

prewarmed medium was added to the cuvette and then the cells transferred into 12-well plates containing 1 ml of prewarmed medium. In some experiments, Jurkat cells were transfected with miRNA mimic using lipofectamine essentially as described in 2.8.

2.9 Phorbol 12-myristate 13-acetate (PMA) stimulation of Jurkat cells for CD69 induction

To induce CD69 expression, Jurkat cells were exposed to differing concentrations of PMA (Sigma) from 2.5 ng/ml to 20 ng/ml for one hour. Cells were then spun at 1000 g for 5 minutes, the supernatant removed and cells resuspended in 5 ml of fresh RPMI ATCC medium with 10% HI FBS and left at 37 ^{o}C for 24 hours. Following this incubation, cells were processed for FACS as described below0.

2.10 Fluorescence-activated cell sorting (FACS)

Jurkat cells were washed twice with FACS Buffer (PBS containing 2% HI FBS, 0.2% sodium azide), incubated with phycoerythrin (PE) labeled mouse anti-human CD69 (BD Biosciences) or PE Mouse IgG1, κ Isotype control (BD Biosciences) for 30 minutes at 4^{o}C then washed twice with FACS Buffer by centrifugation at 1000 g for 5 minutes. Immunostaining was analyzed by flow cytometry (BD Accuri™ C6 flow cytometer with BD Accuri™ C6 software). Live cells were gated by forward scatter (FSC-A) versus side scatter (SSC-A). A fluorescence histogram was created based on the live cell gate and background fluorescence determined from isotype-labelled unstimulated (no PMA) cells. Cells with fluorescence above the threshold were counted.

2.11 Hierarchical clustering

GENE-E version 3.0.240 (http://www.broadinstitute.org/cancer/software/GENE-E/) was used to create a hierarchical clustering of miRNA expression across the *H. contortus* life stages. Data was normalised using the formula $z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$ where $x_i = (x_1, \ldots, x_n)$ and $z_i$is the normalised data. Normalised data was imported into Gene-E. Row distance was calculated using the one-minus Pearson

correlator metric, complete linkage and clustered by row. All other settings were default.

2.12 Visualisation of RNA transcript data

Artemis Genome Browser and Annotation Tool from the Wellcome Trust Sanger Institute (WTSI) version 16 (http://www.sanger.ac.uk/science/tools/artemis)(Rutherford *et al.*, 2000) was used to visualise RNA transcript data from Laing *et al.*, (2013). *H. contortus* scaffold 1000.1 (Haemonchus_v1.fa with annotation file Hc_rztk_1+2+8+9.augustus.gff). BAM files containing the RNA transcript data from eggs (eggs.bam), sheathed and exsheathed L3 (L3_she.bam, L3_ex.bam), L4 (L4.bam) and adult males and females (adult_male.bam, adult_female.bam) were obtained from the Sanger FTP website (ftp://ftp.sanger.ac.uk/pub/project/pathogens/Haemonchus/contortus/) and imported into the Artemis programme. Genes were annotated based on the location of RNA reads and the reading frame. Annotated genes were then searched against the NCBI database to identify closely related homologues or domains.

2.13 Conservation in parasitic worms

WormBase ParaSite (http://parasite.wormbase.org/) contains both published and unpublished genomes from 90 different species. To identify homologues of the Hco-miR-5352 cluster in the other nematodes, the full length *H. contortus* cluster sequence was used to query the ParaSite database using BLASTN with the default search parameters. Sequences obtained from BLAST search analyses were then aligned against the *H. contortus* sequence to determine pairwise similarity.

2.14 Sequence alignments

Geneious version 6 (http://www.geneious.com (Kearse *et al.*, 2012)) was used to perform sequence alignments. Sequences were selected and Geneious Alignment settings were as follows:

Cost Matrix: 65% Similarity; Gap open penalty: 12; Gap extension penalty: 3; Alignment type: Local alignment (Smith Waterman)

2.15 RNA folding

CID-miRNA (http://mirna.jnu.ac.in/cidmirna/) was installed on a Microsoft® Windows 8™ desktop computer (Intel® Core™ i3-3220, 8GB RAM, AMD Radeon HD 7470) running Ubuntu 14.04 LTS (www.ubuntu.com) in Virtualbox VM manager 4.3.20 (www.virtualbox.org/). The potential miRNA sequences of interest were run using CID-miRNA to identify novel miRNAs. Mfold 3.4 (http://mfold.rna.albany.edu) was used to produce a prediction of the secondary structure of the miRNA.

2.16 Tree Builder

To describe the evolutionary relationship between a set of sequences, the programme Geneious version 6 was used to build phylogenetic trees. Sequences were selected and aligned using the Geneious Tree Builder with the following options:

Cost Matrix: 65% Similarity; Gap open penalty: 12; Gap extension penalty: 3; Alignment type: Global alignment with free end gaps; Genetic Distance Model: Tamura-Nei; Tree build Method: Neighbour-joining; Outgroup: no outgroup.

2.17 Creation of the 3’ UTR database

The 3' UTR database was created by combining two databases created by Axel Martinelli (Wellcome Trust Sanger Institute). The 3' UTR consists of sequences downstream from the stop codon of the annotated *H. contortus* genes whose length was determined based on transcriptomic data. The 3' DS (downstream) database was generated by taking 1,000 nucleotides downstream of the stop codon. This 3' UTR sequences were created based on an earlier version of the genome annotations (Hc_rztk_1+2+8+9.gff) and contains 21,779 annotated genes.

2.18 Target Prediction

Target predictions were performed on a Windows 8 PC (Intel® Core™ i3-3220, 8GB RAM, AMD Radeon HD 7470) running Ubuntu 14.04 LTS (www.ubuntu.com) in Virtualbox VM manager 4.3.20 (www.virtualbox.org/). PITA (http://genie.weizmann.ac.il/pubs/mir07/index.html), Miranda (http://www.microrna.org/microrna/getDownloads.do) and RNAhybrid (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid) were installed in Ubuntu as per their instructions. All miRNAs in miRBase labelled as *H. contortus* miRNAs (http://www.mirbase.org/cgi-bin/mirna_summary.pl?org=hco) were downloaded in fasta format, a total of 195 miRNAs and run against each 3' UTR sequence in the database using all three programmes with the following settings:

1. PITA was run with default settings. For the analysis, PITA results were filtered for a seed sequence of 8 bases, ΔΔG less than -10
2. miRanda was filtered for scores greater than 145 and energy less than -10
3. RNAhybrid was filtered for p-values less than 0.1 and energy less than -22

2.19 Identifying putative homologues of *H. contortus* miRNA target genes

C. elegans homologues of *H. contortus* genes were identified by obtaining the protein sequence of the genes and using BLASTP to compare with the NCBI nr (non-redundant) protein sequence database, selecting specifically for *C. elegans* sequences. Proteins were considered homologues if the query coverage and identity were above 40%.

2.20 Gene ontology analysis

C. elegans homologues of *H. contortus* genes were submitted to AmiGO 2 Term enrichment service (http://amigo2.geneontology.org/amigo powered by http://pantherdb.org/). Data was analysed for biological processes, molecular function and cellular component. AmiGO 2 also calculated a P-value which is the GO terms identified in the query list compared to the GO terms in the whole genome. For the most part, none of the gene lists submitted had a p-value lower than 0.05 in which case, the GO terms with the most number of associated genes were obtained.

2.21 Qiagen Ingenuity Pathway Analysis

The raw data from the RNA-Seq were submitted to the Qiagen Ingenuity Pathway Analysis® program (Ingenuity Systems, Redwood City, CA, USA;www.ingenuity.com). An excel spreadsheet consisting of the gene IDs in the Gene Symbol – mouse (Entrez Gene) format, the log2(fold change) values, the p-values and the q-values, was uploaded into the IPA software. The data was analysed using the "Core Analysis" function and various cutoffs were applied as detailed in the results section. Query genes were mapped onto canonical pathways or used to create a network of known associations between genes and proteins, independent of pathways.

2.22 Transcriptomic data for *H. contortus* genes

Transcriptomic data was obtained from Laing *et al.*, (2013) using the MHco3(ISE) inbred strain of *H. contortus*. The data consists of RNA-Seq data from eggs, L1, sheathed and exsheathed L3, L4, adult male and female and adult female gut tissue. RNA-Seq data was generated based on the Augustus gene prediction software (Stanke *et al.,* 2006) and gene expression data was performed using DESeq by Axel Martinelli (Wellcome Trust Sanger Institute)(Laing *et al.,* 2013). All *H. contortus* material used in this chapter for ES production were harvested as described in Chapter 2.1 using the MHco3(ISE) strain.

2.23 MODE-K cell culture and transfection

The murine intestinal epithelial MODE-K cells was maintained at 37°C in an atmosphere of 5% CO2 in air in Dulbecco's minimal essential medium (DMEM) supplemented with 10% heat-inactivated FBS, 2mM L-glutamine (Sigma), 100 U/ml penicillin and 100 µg/ml streptomycin. For experiments, MODE-K cells were cultured until 80% confluent, the supernatant was carefully aspirated and the monolayer washed twice with PBS before adding 5 ml of 0.25% EDTA-trypsin (Sigma) for 2 minutes. 5 ml of medium was then added and the cells transferred to a 15 ml Falcon tube, which was centrifuged at 400 g for 5 minutes. The supernatant was aspirated before resuspending the pellet in 4 ml of medium. Cells were counted using a haemocytometer and 1 x 10^5 cells were transferred to each well of a 12 well plate in 1 ml of medium. Cells were allowed to grow

for 24 hours before adding 200 µl of Opti-MEM® reduced serum media without serum, 1.5 µl of LTX™ reagent and 300 nM of either the Hco-miR-5352 miRNA mimic or miRIDIAN microRNA Mimic Negative Control (Cel-miR-67). This control is not found in vertebrates, has minimal sequence identity to other mammalian miRNAs and is commonly used as the negative control in miRNA mimic experiments. After a further 24 hours, the supernatant was aspirated and total RNA extracted using the Qiagen miRNeasy mini kit as detailed in Chapter 2.3.1.

2.24 Gene Expression Analysis and qRT-PCR.

Qiagen RT^2 Profiler PCR arrays were used to examine the expression patterns of 84 mouse immune cell surface markers (Qiagen Catalogue number PAMM-055Z). or 84 key innate and adaptive immune response genes (Qiagen Catalogue number PAMM-052Z). (See Appendix Section 9.3 and 9.4 for gene lists). Two plates of each array were used, one was probed with RNA from MODE K cells that had been transfected with miR-5352 mimic, the other with RNA from cells transfected with the control mimic.

For each 96 well plate, 0.5 µg of RNA was used and processed with the RT^{2}First Strand Kit (Qiagen). RNA was incubated with 2 µl of Buffer GE and made up to 10 µl with RNase free water (Qiagen). The genomic DNA elimination mixture was added for 5 minutes at 42 ^{o}C then transferred onto ice for at least 1 minute. 4 µl of 5x Buffer BC3, 1 µl of Control P2, 2 µl of RE3 Reverse Transcriptase mix and 3 µl of RNase free water were then added to the genomic DNA elimination mix and mixed gently by pipetting. The reverse-transcriptase mix was incubated for 15 minutes at 42 ^{o}C before incubating at 95 ^{o}C for 5 minutes to halt the reaction. 91 µl of RNase free water was then added and mixed.

Probing of the array was carried out according to the manufacturer's instructions using the specific RT^2 SYBR Green Rox qPCR Master Mix (Qiagen). The 1350 µl of the RT^2 SYBR Green MasterMix was briefly centrifuged before mixing with 102 µl of the cDNA synthesis reaction and 1248 µl of RNase free water. 25 µl of the mixture was then pipetted into each well in the PCR array and sealed with optical thin wall 8-cap strips. Plates were centrifuged for 1 minute at 1000 g at room temperature to remove bubbles.

Real-time PCR amplification was carried out on a Stratagene Mx3005p machine. In the MxPro program, the REAL TIME > SYBR Green with Dissociation curve was selected and thermal profile adjusted to:

Segment 1: 1 cycle of 95 ^{o}C for 10 minutes

Segment 2: 40 cycles of 95 ^{o}C for 15 seconds then 60 ^{o}C for 1 minute

Segment 3 (dissociation melt): 1 cycle of 95 ^{o}C for 1 minute then 55 ^{o}C for 30 seconds and finally 95 ^{o}C for 30 seconds.

Data analysis was performed using the Qiagen GeneGlobe Data analysis centre website (http://www.qiagen.com/gb/shop/genes-and-pathways/technology-portals/browse-qpcr/qrt-pcr-for-mrna-expression/data-analysis/~/link.aspx?_id=93C80B99536C4FCDBD1E5FC0759F1324&_z=z). C_T values were obtained from the MxPro software and uploaded into the Qiagen website and quantification using the $\Delta\Delta C_T$ method. This method calculates the fold change as the difference of the C_T value of each gene of interest compared to the reference genes (β-Actin, β-2 microglobulin, glyceraldehyde-3-phosphate dehydrogenase, Glucuronidase-β and heat shock protein 90 α (cytosolic) class B member 1).

$$\Delta Ct\ (treated) = Ct\ (Gene\ of\ interest\ treated) - Ct\ (Reference\ genes\ treated)$$
$$\Delta Ct\ (control) = Ct\ (Gene\ of\ interest\ control) - Ct\ (Reference\ control)$$
$$\Delta\Delta\ Ct = \Delta\ Ct\ (treated) - Ct\ (control)$$

The normalized target gene expression level is calculated as $2^{(-\Delta\Delta CT)}$ and fold change is the Hco-miR-5352 values divided by the control values.

The Qiagen GeneGlobe can also provide data quality checks. PCR array reproducibility was determined by comparing the average C_T values for the PCR positive control wells (PPC). If the two arrays have a difference of less than two, the array is given a pass. Reverse transcriptase efficiency is calculated by subtracting the average C_T value for the reverse transcriptase control wells from the average PPC wells. If the difference is less than 5, the check passes. The final check is for genomic DNA contamination and is determined by calculating if the C_T value for the genomic DNA contamination well is less than 35.

2.25 RNA-Sequencing and analysis

RNA was extracted from MODE-K cells transfected with Hco-miR-5352 or Cel-miR-67 exactly as described above for the RT^2 Profiler Array (Chapter 2.24). Three replicates were obtained for each miRNA transfection. RNA integrity was determined using an Agilent 2100 Bioanalyzer by University of Glasgow Polyomics Facility. RNA libraries were generated from RNA samples using the Illumina Truseq™ Small RNA Preparation kit according to Illumina's Directional mRNA-Seq Sample Preparation (LC Sciences, Houston, TX).

Data processing was performed by LC Sciences (LC Sciences, Houston, TX) as follows: sequence reads were mapped to the mouse genome (version mm9) using the Tophat program (Trapnell *et al.*, 2009) and abundance estimated at the gene level using Cufflinks (Trapnell *et al.*, 2013). Abundance was normalized and differential expression analysis was performed using Cuffdiff (Trapnell *et al.*, 2013). The fold change value was calculated as the ratio between the average normalized read count for each gene. All fold changes values shown are listed as the log2(fold change). The q-value, false discovery rate adjusted p-value, was also calculated.

Genes showing significant difference (q-value less than 0.05) were selected for GO Term enrichment analysis using GAGE (Luo *et al.*, 2009). Genes that were highly expressed in Hco-miR-5352 transfected cells compared to the Cel-miR-67 transfected cells (up-regulated) along with lower expressed genes (down-regulated) were combined and grouped according to GO Terms. For each GO Term, the number of down-regulated genes relative to the number of up-regulated genes was analyzed to determine whether the GO Term was enriched for up- or down-regulated genes.

3 Characterisation of the miR-5352 cluster

3.1 Introduction.

The first miRNAs were discovered in the free-living nematode *Caenorhabditis elegans* and were found to have important roles in the regulation of development in the worm (Lee *et al.*, 1993b; Wightman *et al.*, 1993). Since then, 35,828 mature miRNAs in 223 species have been registered with the miRBase miRNA database (http://www.mirbase.org/ Release 21). Many miRNAs are key post-transcriptional regulators of gene expression and have been shown to have important roles in a range of diverse processes including immune responses (Gantier *et al.*, 2007) and cancer (Hayes *et al.*, 2014).

Parasitic nematodes, such as *H. contortus* are closely related to *C. elegans* (Blaxter *et al.*, 1998). Both are members of nematode clade V and have similar developmental processes and body plans. *C. elegans* is used as a model organism to understand conserved biological aspects of *H. contortus* and other nematodes. *C. elegans* has been used to investigate the mode of action of the major anthelmintic drugs and identify mechanisms of drug resistance (Holden-Dye and Walker, 2007). However, *C. elegans* does not have a parasitic lifestyle and its value in modelling parasite-specific processes is therefore less powerful. The similarities between *C. elegans* and *H. contortus* are also evident from comparative genomic studies. The *C. elegans* genome was published in 1998 (https://www.ncbi.nlm.nih.gov/pubmed/9851916) and contains 20,470 protein coding genes, approximately 35% of which have human homologues. Draft versions of the *H. contortus* genome were published in 2013 (Laing *et al.*, 2013; Schwarz *et al.*, 2013). Laing *et al.*, (2013) showed that the *H. contortus* genome is 370 Mb long, consisting of 26,044 scaffolds containing 67,687 contigs and is predicted to have 21,799 protein coding genes. Nearly six thousand orthologous clusters were identified between the *H. contortus* and *C. elegans* genomes, five thousand of which were one-to-one orthologues. The two genomes also share similar genome structure and content. In addition, the *H. contortus* genome is also very poorly annotated which therefore limits the analysis of *H. contortus* genes as well as limiting the effectiveness of many bioinformatic tools. To work

around this limitation, homologues of *H. contortus* genes can be identified in other, more well character organisms such as *C. elegans*. However, the *H. contortus* genome also shows that there are differences between the *H. contortus* and *C. elegans* genomes with a number of genes appearing to be unique to *H. contortus*.

With the availability of genome sequence data for *H. contortus*, Winter *et al.*, (2012) used deep-sequencing and bioinformatics approaches to identify *H. contortus* miRNAs. In total, 192 miRNAs were identified and appeared to be predominantly species-specific, with only 54 of the 192 miRNAs being conserved in other organisms and only 44 miRNAs conserved between *H. contortus* and *C. elegans*. It should be noted that these miRNAs may not account for all the miRNAs in *H. contortus*, as they were identified from L3 and from adult stage worms, and from homology to miRNAs in miRBase. Other life cycle stages may express different miRNAs. The miRNAs were also mapped to an early version of the *H. contortus* genome, consisting of 73,256 scaffolds of much smaller length than the current version. The scaffold N50 is the length of the scaffold in which 50% of the genome is contained in scaffolds of equal or greater length; the larger this value, the better the genome assembly. The N50 values for the old and new genomes is 9.9 kbp and 83 kbp, respectively. To put this in context, the N50 for *C.* elegans is 17,493 kbp (Laing *et al.*, 2013). The large number of short contigs in the early version of the *H. contortus* genome introduces multiple breaks in the assembly where genes or miRNAs could occur. The improvement in the *H. contortus* genome assembly and the availability of other parasite genomes (http://parasite.wormbase.org/species.html) will greatly aid our identification and characterisation of parasitic nematode miRNAs.

One particularly interesting observation from Winter *et al.* (2012) was the discovery of miRNAs that are present in the *H. contortus* genome as clusters. It is known that several miRNAs can be transcribed as a single pri-miRNA and are therefore co-regulated (Schmeier *et al.*, 2009). Winter *et al.*, (2012) were able to identify eight clusters of miRNAs consisting of 23 miRNA loci in *H. contortus*. Three of these clusters consist of four miRNA loci that are all orientated in the same direction and are grouped closely together, with the largest of these clusters spanning a region of 680 nucleotides. One cluster, referred to as the

Hco-miR-5352 cluster, is of particular interest. miR-5352 sequence was initially identified from *Ascaris suum* (Wang *et al.*, 2011a) and is present in miRBase. However, it is not present in *Caenorhabditis* species, which are phylogenetically more similar to *H. contortus* than is *A. suum*, which is a clade III species. This suggested that miR-5352 and possibly other members of the cluster may have a specific role related to parasitism. The cluster encodes four miRNAs: Hco-miR-61, Hco-miR-5352, Hco-miR-43 and Hco-miR-5895. miR-5895 was identified as unique to *H. contortus*, while miR-61 and miR-43 show some conservation between *H. contortus* and *C. elegans*.

The aims of this chapter were to characterise the miR-5352 cluster, the expression of the Hco-miR-5352 cluster across different life cycle stages, the genomic location of the cluster and to determine whether these miRNAs are conserved in other parasitic nematode genomes. Important to defining a potential role of the miR-5352 cluster in parasitic infection and host-parasite interaction, the presence of miR-5352 in *H. contortus* excretory-secretory (ES) products was also examined. Finally, endogenous *H. contortus* targets of the cluster were also identified.

3.2 Results

3.2.1 Hco-miR-5352 cluster expression across *H. contortus* life cycle stages

The Hco-miR-5352 cluster comprises four miRNAs; Hco-miR-61, Hco-miR-5352, Hco-miR-43 and Hco-miR-5895. The four miRNAs are located extremely close together on the genome, spanning a region of 422 bases.

Previous experiments had involved the construction of a microarray containing 609 predicted *H. contortus* miRNAs plus 238 *C. elegans* miRNAs identified from miRBase release 15 (Winter *et al.*, 2012). The microarray had been probed with RNA isolated from four life cycle stages (L3, L3 exsheathed and cultured at 37^{o}C for 48 hours, L4 taken 7 days post-infection and adults 21 days post infection) and the dissected gut of *H. contortus* female worms. These data were re-analysed to determine the expression of each of the miRNAs in the cluster (Figure 3.1). Minimal expression was observed in stages other than adult worms,

and female worms appear to express elevated levels of all the miRNAs in the cluster compared to male worms.

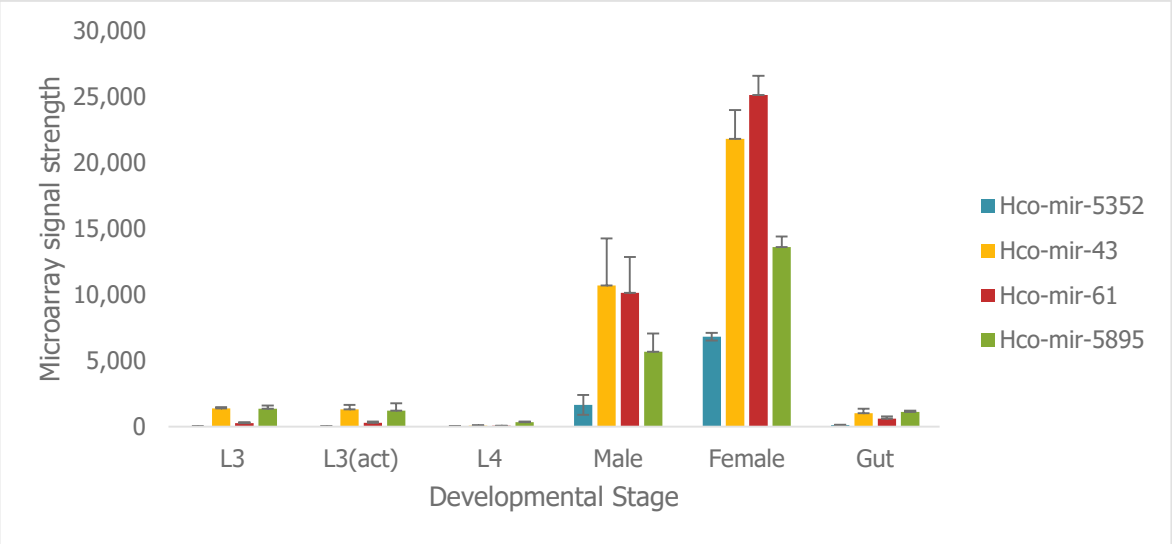


Figure 3.1 Level of expression of the Hco-miR-5352 cluster across the *H. contortus* life cycle as assessed by microarray. L3 are sheathed L3, L3(act) (exsheathed and cultured at 37ºC for 48 hours), L4 stage, 7 days post-infection, adult males and females, 21 days post-infection. Microarray signals shown as the mean +/- SD of three biological replicates, except for L4 samples, where n=2.

For all four miRNAs, both a 5p and 3p version were identified from the original sequence data of Winter *et al.*, (2012). From the microarray data, the 3p versions of Hco-miR-61, Hcso-miR-5352, Hco-miR-43 and the 5p version of Hco-miR-5895 showed high levels of expression and were accepted by miRBase as miRNAs (Figure 3.2). Only one version, the dominant miRNA from each *H. contortus* hairpin was submitted to miRBase. Hco-miR-43-5p the alternative sequence, whilst not in miRBase shows a similar profile to the other miRNAs in the cluster and a similar level of expression to Hco-miR-5352-3p. This suggests that Hco-miR-43-5p may also be functional.

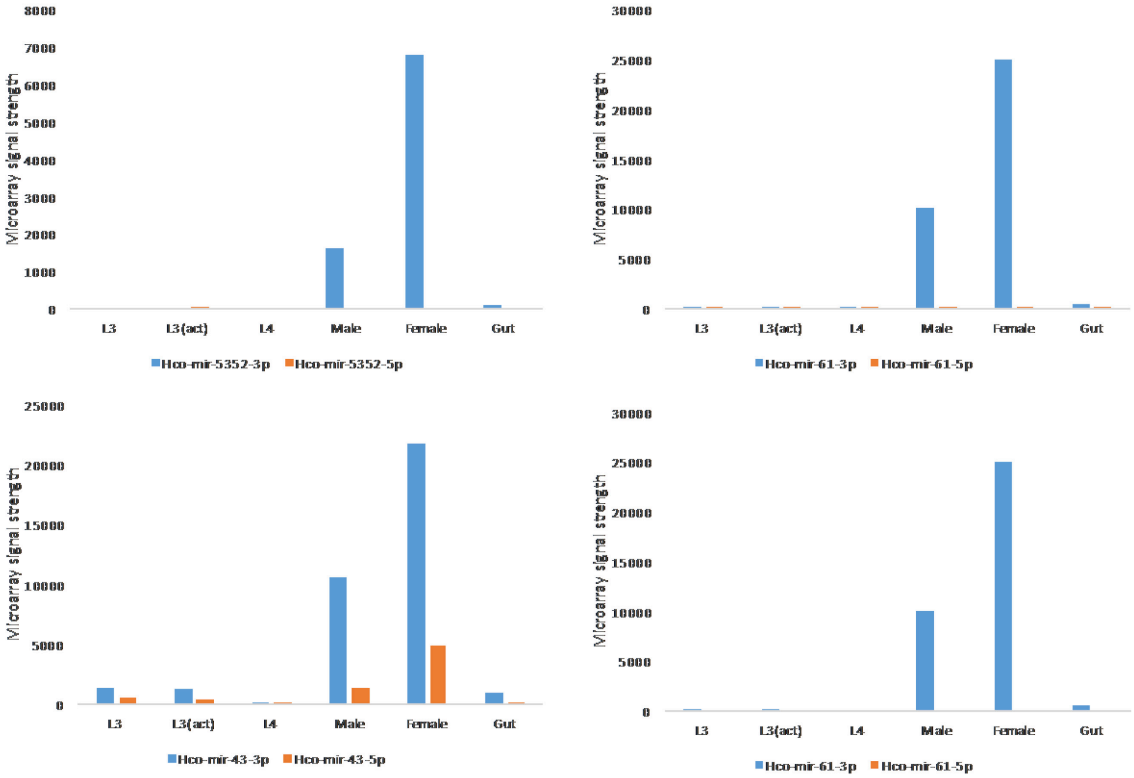
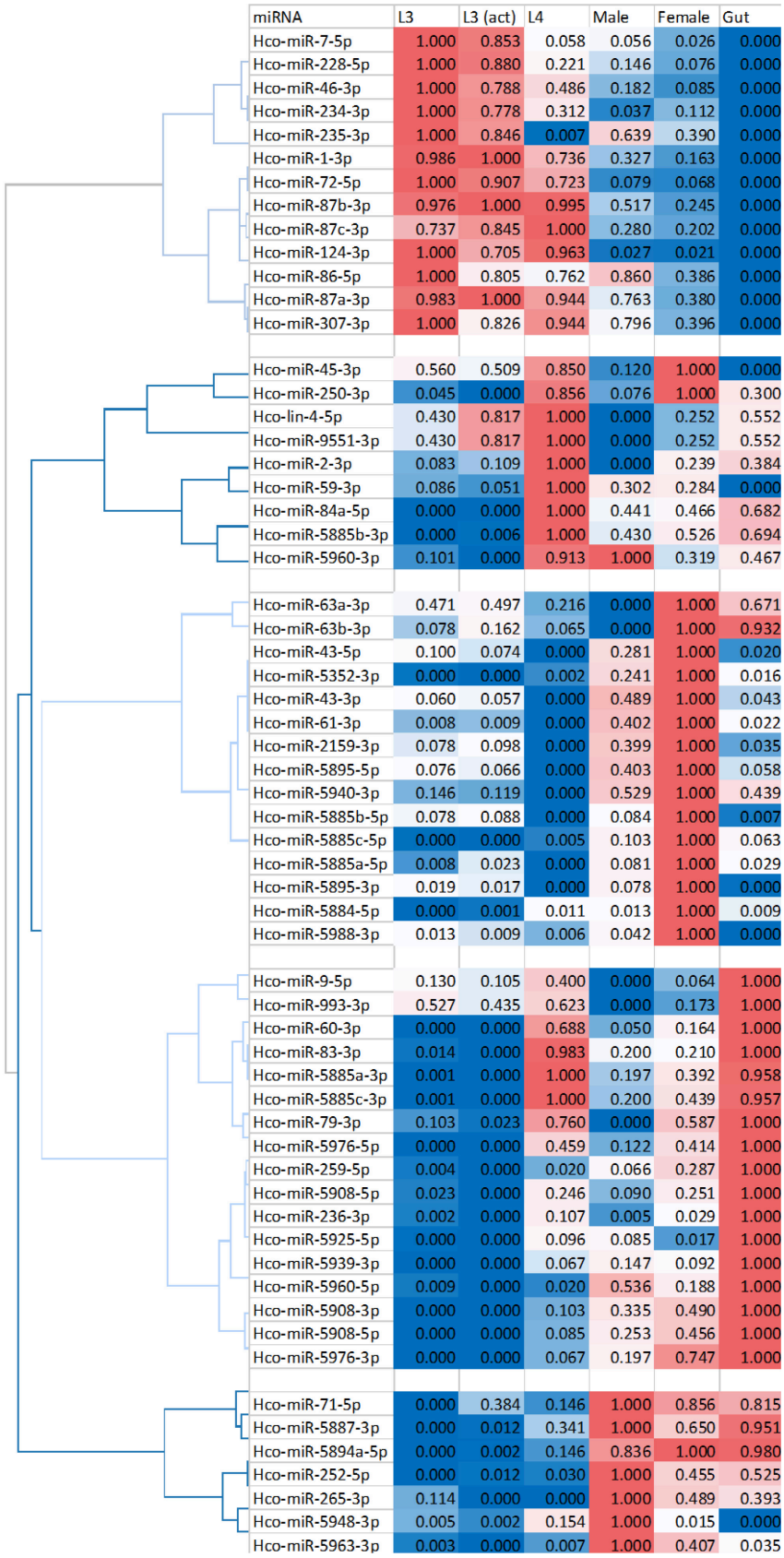


Figure 3.2 Microarray expression data for the 5p and 3p versions of the four miRNAs in the Hco-miR-5352 cluster across the *H. contortus* life cycle and from the gut. The stages analysed are: L3, L3(act) (activated L3, exsheathed and cultured at 37°C for 48 hours), L4 removed seven days post-infection with L3, adult males and females at 21 days post-infection. In addition, gut tissue was dissected from adult female worms. The y-axis shows the signal from the array +/- SD of the mean of triplicate biological samples except for L4 where n=2.

To investigate how the expression profile of miRNAs in the cluster compares to other miRNAs, a hierarchical clustering program was used to cluster the miRNAs by their microarray expression profile (Figure 3.3). The programme groups miRNAs based on the pair-wise distance between life stages. The four miRNAs in the Hco-miR-5352 cluster are located on the same branch along with an additional miRNA, Hco-miR-2159-3p. All of these miRNAs have low expression in both L3 stages, the L4, and the gut tissue. Expression is highest in the female worms with an intermediate level in the male worms. Hco-miR-2159-3p is not located close to the Hco-miR-5352 cluster on the current version of the genome.



miRNA	L3	L3 (act)	L4	Male	Female	Gut
Hco-miR-7-5p	1.000	0.853	0.058	0.056	0.026	0.000
Hco-miR-228-5p	1.000	0.880	0.221	0.146	0.076	0.000
Hco-miR-46-3p	1.000	0.788	0.486	0.182	0.085	0.000
Hco-miR-234-3p	1.000	0.778	0.312	0.037	0.112	0.000
Hco-miR-235-3p	1.000	0.846	0.007	0.639	0.390	0.000
Hco-miR-1-3p	0.986	1.000	0.736	0.327	0.163	0.000
Hco-miR-72-5p	1.000	0.907	0.723	0.079	0.068	0.000
Hco-miR-87b-3p	0.976	1.000	0.995	0.517	0.245	0.000
Hco-miR-87c-3p	0.737	0.845	1.000	0.280	0.202	0.000
Hco-miR-124-3p	1.000	0.705	0.963	0.027	0.021	0.000
Hco-miR-86-5p	1.000	0.805	0.762	0.860	0.386	0.000
Hco-miR-87a-3p	0.983	1.000	0.944	0.763	0.380	0.000
Hco-miR-307-3p	1.000	0.826	0.944	0.796	0.396	0.000
Hco-miR-45-3p	0.560	0.509	0.850	0.120	1.000	0.000
Hco-miR-250-3p	0.045	0.000	0.856	0.076	1.000	0.300
Hco-lin-4-5p	0.430	0.817	1.000	0.000	0.252	0.552
Hco-miR-9551-3p	0.430	0.817	1.000	0.000	0.252	0.552
Hco-miR-2-3p	0.083	0.109	1.000	0.000	0.239	0.384
Hco-miR-59-3p	0.086	0.051	1.000	0.302	0.284	0.000
Hco-miR-84a-5p	0.000	0.000	1.000	0.441	0.466	0.682
Hco-miR-5885b-3p	0.000	0.006	1.000	0.430	0.526	0.694
Hco-miR-5960-3p	0.101	0.000	0.913	1.000	0.319	0.467
Hco-miR-63a-3p	0.471	0.497	0.216	0.000	1.000	0.671
Hco-miR-63b-3p	0.078	0.162	0.065	0.000	1.000	0.932
Hco-miR-43-5p	0.100	0.074	0.000	0.281	1.000	0.020
Hco-miR-5352-3p	0.000	0.000	0.002	0.241	1.000	0.016
Hco-miR-43-3p	0.060	0.057	0.000	0.489	1.000	0.043
Hco-miR-61-3p	0.008	0.009	0.000	0.402	1.000	0.022
Hco-miR-2159-3p	0.078	0.098	0.000	0.399	1.000	0.035
Hco-miR-5895-5p	0.076	0.066	0.000	0.403	1.000	0.058
Hco-miR-5940-3p	0.146	0.119	0.000	0.529	1.000	0.439
Hco-miR-5885b-5p	0.078	0.088	0.000	0.084	1.000	0.007
Hco-miR-5885c-5p	0.000	0.000	0.005	0.103	1.000	0.063
Hco-miR-5885a-5p	0.008	0.023	0.000	0.081	1.000	0.029
Hco-miR-5895-3p	0.019	0.017	0.000	0.078	1.000	0.000
Hco-miR-5884-5p	0.000	0.001	0.011	0.013	1.000	0.009
Hco-miR-5988-3p	0.013	0.009	0.006	0.042	1.000	0.000
Hco-miR-9-5p	0.130	0.105	0.400	0.000	0.064	1.000
Hco-miR-993-3p	0.527	0.435	0.623	0.000	0.173	1.000
Hco-miR-60-3p	0.000	0.000	0.688	0.050	0.164	1.000
Hco-miR-83-3p	0.014	0.000	0.983	0.200	0.210	1.000
Hco-miR-5885a-3p	0.001	0.000	1.000	0.197	0.392	0.958
Hco-miR-5885c-3p	0.001	0.000	1.000	0.200	0.439	0.957
Hco-miR-79-3p	0.103	0.023	0.760	0.000	0.587	1.000
Hco-miR-5976-5p	0.000	0.000	0.459	0.122	0.414	1.000
Hco-miR-259-5p	0.004	0.000	0.020	0.066	0.287	1.000
Hco-miR-5908-5p	0.023	0.000	0.246	0.090	0.251	1.000
Hco-miR-236-3p	0.002	0.000	0.107	0.005	0.029	1.000
Hco-miR-5925-5p	0.000	0.000	0.096	0.085	0.017	1.000
Hco-miR-5939-3p	0.000	0.000	0.067	0.147	0.092	1.000
Hco-miR-5960-5p	0.009	0.000	0.020	0.536	0.188	1.000
Hco-miR-5908-3p	0.000	0.000	0.103	0.335	0.490	1.000
Hco-miR-5908-5p	0.000	0.000	0.085	0.253	0.456	1.000
Hco-miR-5976-3p	0.000	0.000	0.067	0.197	0.747	1.000
Hco-miR-71-5p	0.000	0.384	0.146	1.000	0.856	0.815
Hco-miR-5887-3p	0.000	0.012	0.341	1.000	0.650	0.951
Hco-miR-5894a-5p	0.000	0.002	0.146	0.836	1.000	0.980
Hco-miR-252-5p	0.000	0.012	0.030	1.000	0.455	0.525
Hco-miR-265-3p	0.114	0.000	0.000	1.000	0.489	0.393
Hco-miR-5948-3p	0.005	0.002	0.154	1.000	0.015	0.000
Hco-miR-5963-3p	0.003	0.000	0.007	1.000	0.407	0.035

Figure 3.3 Hierarchical clustering of microarray data for the miRBase accepted microRNAs. Only the 60 sequences with variation across life stages are shown Data normalised on a scale of 0 (no expression) to 1 (high expression) across the life cycle stages. Blue boxes indicate low signals, white indicates medium expression and red boxes indicate high signals.

3.2.2 Differences between genome assemblies

In the original publication of Winter *et al.*, (2012), miRNAs were identified based on an extremely early version of the *H. contortus* genome project from 2008. Supercontig_0041242 was identified as the contig containing the Hco-miR-5352 cluster and it was only 1,031 bp in length. The *H. contortus* genome assembly (haemonchus_V1.gtr100.fa_nicenames) and functional annotation (Hc_finalgenemodels_renamed.gff3) were downloaded from the Sanger ftp website (ftp://ftp.sanger.ac.uk/pub/project/pathogens/Haemonchus/contortus/genome/).

From the most recent release of the *H. contortus* genome, the scaffold identified with the miR-5352 cluster is now 94,939 bp long (Laing *et al.*, 2013). The cluster region is presented graphically in Figure 3.4 (older version of the genome) and Figure 3.5 (more recent version).

Figure 3.4 Precursor and mature sequences for the miRNAs in the Hco-miR-5352 cluster obtained from miRBase aligned to the *H. contortus* assembled contigs from 21/08/08 (all reads).

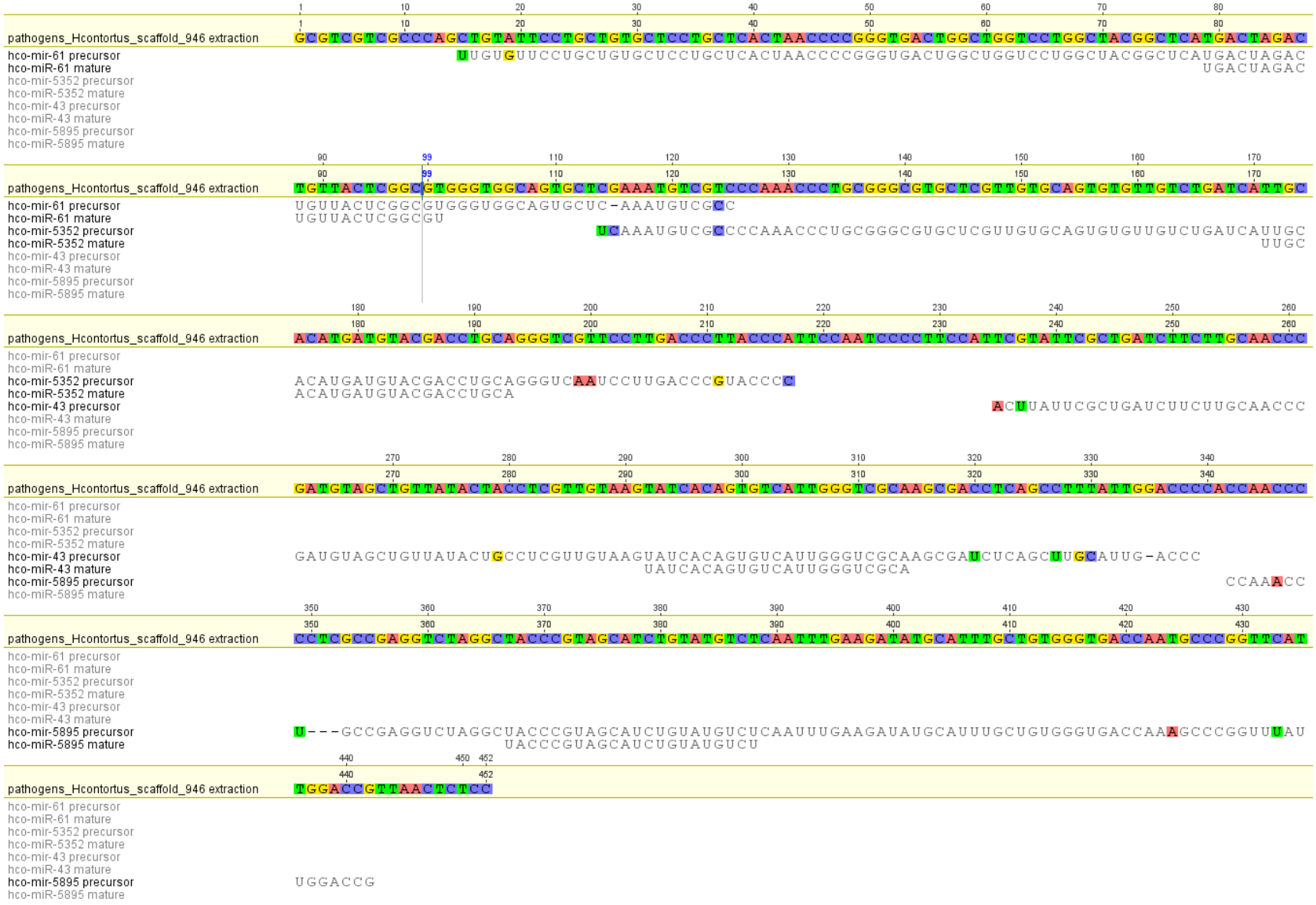


Figure 3.5 Precursor and mature sequences for the miRNAs in the Hco-miR-5352 cluster obtained from miRBase aligned to the most recent version of the *H. contortus* genome scaffold, pathogens_Hcontortus_scaffold_946 (ftp://ftp.sanger.ac.uk/pub/project/pathogens/Haemonchus/contortus/genome/). Bases that are highlighted in colour show mismatches between the miRBase sequence and the genome assembly. Dashes (-) represent gaps in the alignment.

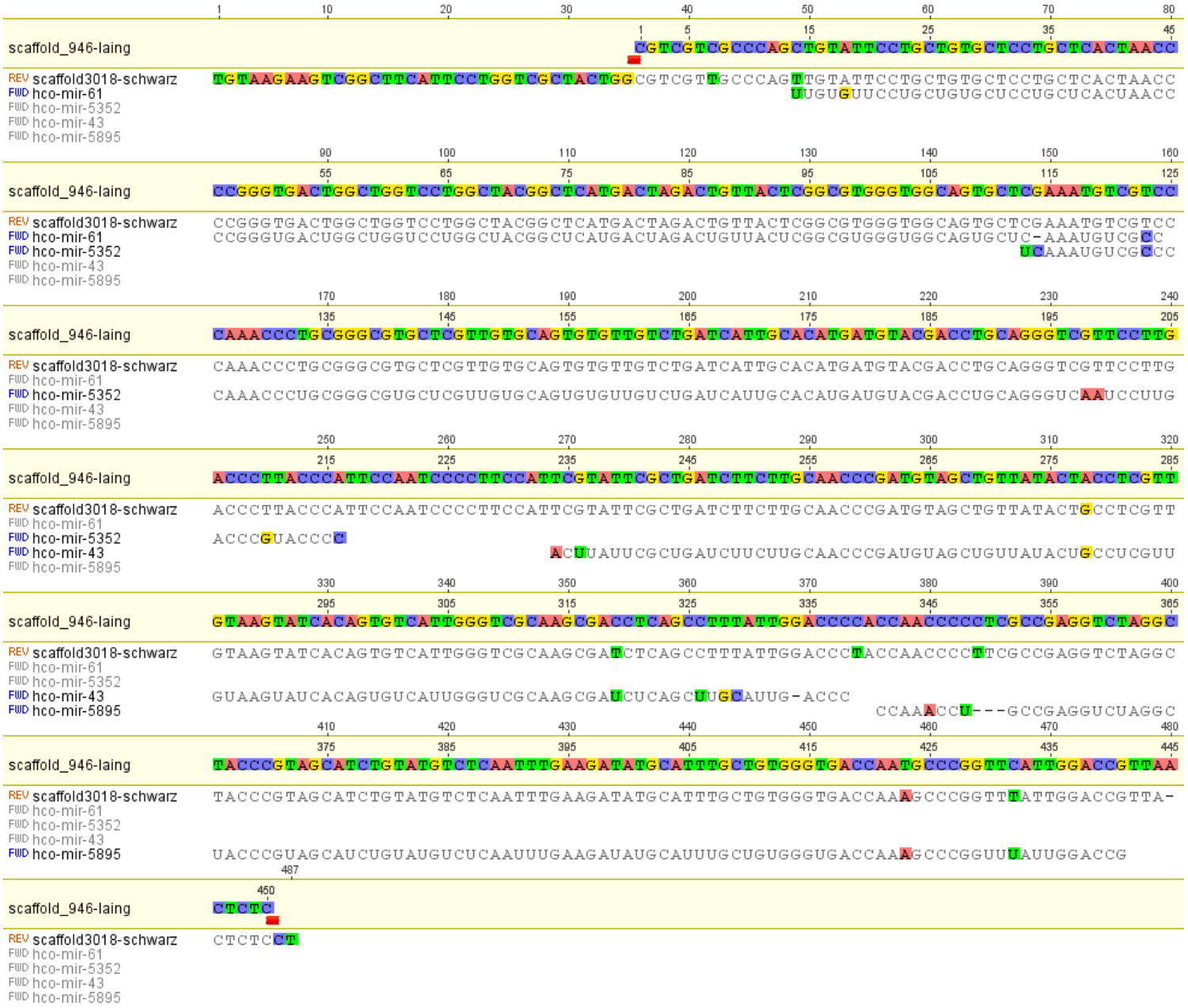


Figure 3.6 Precursor sequences for the miRNAs in the Hco-miR-5352 cluster obtained from miRBase, scaffold3018 from the Schwarz *et al.*, genome (2013) and scaffold946 from the Laing *et al.*, genome (2013).

In the alignment between the Hco-miR-5352 precursor and the most recent genome version, there are seven base mismatches all located at the extreme ends of the precursor. However, the mature sequence is still exactly the same. Figure 3.7shows a comparison between the two miR-5352 precursor sequences. Figure 3.7A shows the folding of the miRBase accepted precursor and B shows the precursor sequence from the new genome assembly. Both have exactly the same structure and the seven mismatches are all located within the loops of the precursor. Both structures also have very similar initial free energies (dG), -44.20 and -44.90 respectively. This indicates that the sequence variations at the ends of the precursor sequence may be slightly more stable compared to the miRBase accepted sequence. The Hco-miR-5352 can also be identified in the Schwarz genome project on scaffold3018 (Schwarz *et al.,* 2013). Figure 3.6 shows the alignment between the two miR-5352 containing scaffolds from the Schwarz and Laing genome and the miRBase accepted miRNAs. The two scaffolds are very similar with only eight mismatches. The presence the Hco-miR-5352 cluster precursor sequences can also be identified and whilst there are some mismatches within the precursor sequences, the mature sequence is identical.

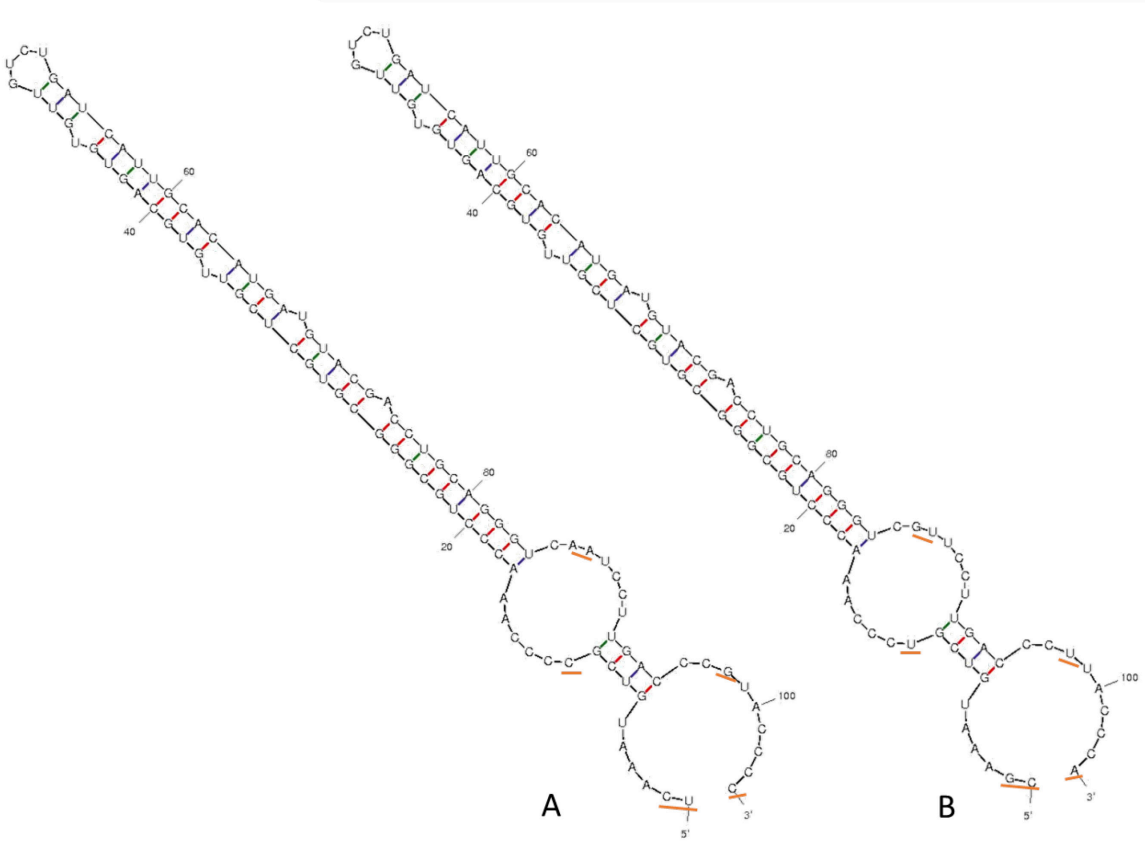


Figure 3.7 mfold diagrams of the A) miRBase accepted Hco-miR-5352 precursor and B) recent genome version of the Hco-miR-5352 precursor. Differences have been underlined in orange. A has an initial free energy (dG) of -44.20 and B has a dG of -44.90.

3.2.3 Location of Hco-miR-5352 cluster on the *H. contortus* genome

The scaffold containing the Hco-miR-5352 cluster was identified and imported into Artemis along with RNA transcript data for eggs, L1, L3 sheathed and L3 exsheathed, L4, adult males and females, and gut tissue (Laing *et al.*, 2013).

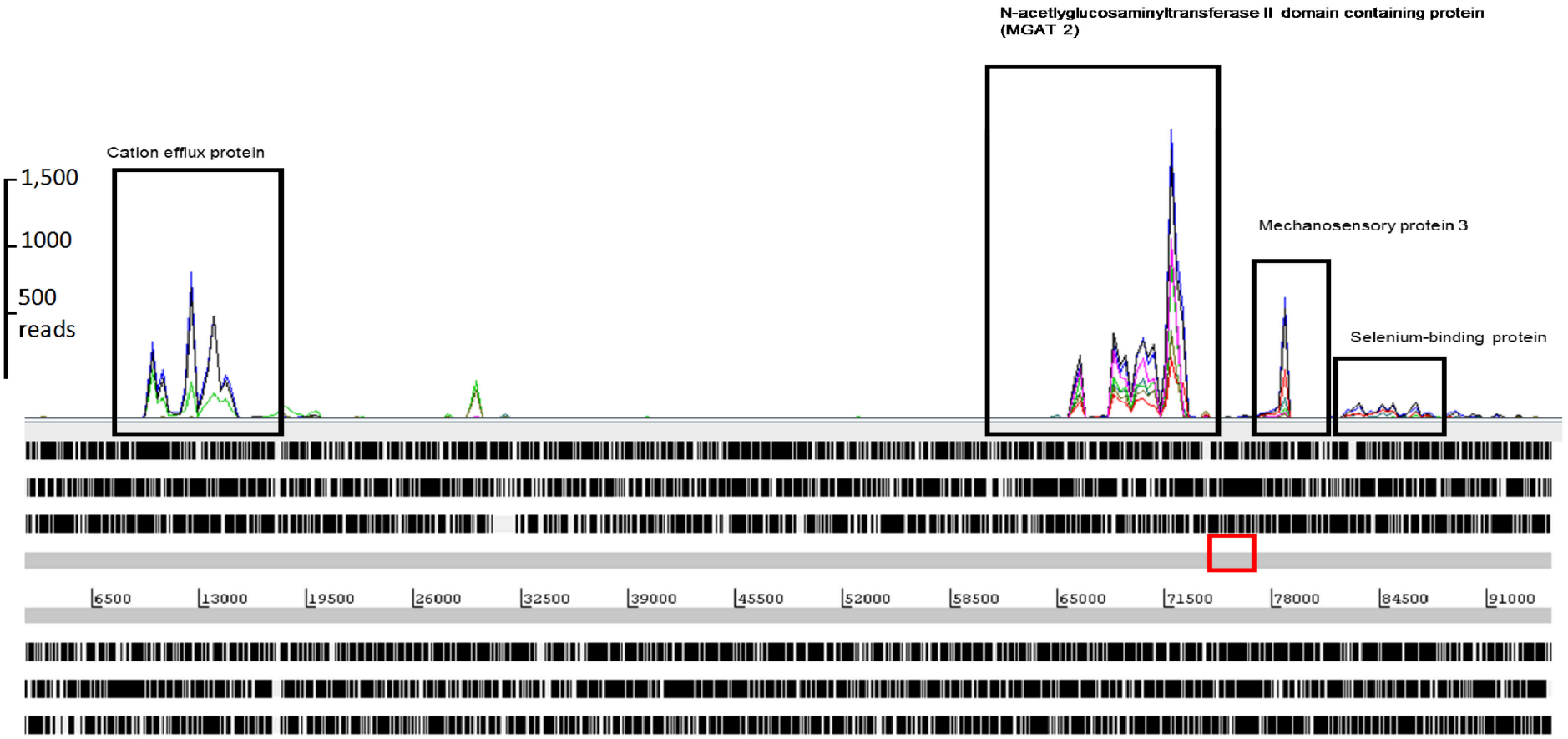


Figure 3.8 Screenshot of Artemis showing the miR-5352 containing *H. contortus* scaffold 1000.1 (Laing *et al.*, 2013) **with RNA transcript data from eggs, L3 sheathed and exsheathed L3, L4 and adult males and females. The six rows of black lines show the stop codons in each of the six reading frames, the top three rows show the three forward frame and the bottom three rows, the reverse frame. The scaffold is 95,000 bases long and contains four regions that show high levels of RNA transcripts (black boxes). The Hco-miR-5352 cluster is located between two of these regions (red box).**

This analysis was carried out to identify protein coding genes near to the Hco-miR-5352 cluster which may potentially be regulated by these miRNAs. The scaffold contains four regions that have RNA transcript data and contain annotated genes. The cation efflux protein and MGAT2 show the highest level of transcript data (Figure 3.8). The region between the cation efflux protein and MGAT2 has one small peak at 29-30,000 bases, however, the sequence shows very poor similarity to a Metridin ShK toxin domain containing protein (63% query coverage and 53% identity). Next to MGAT2 is a zinc finger and homeobox domain containing protein that has some similarity to a *C. elegans* mechanosensory protein 3 (98% query coverage, 56% identity) and the last annotated gene is a selenium binding protein on the reverse frame.

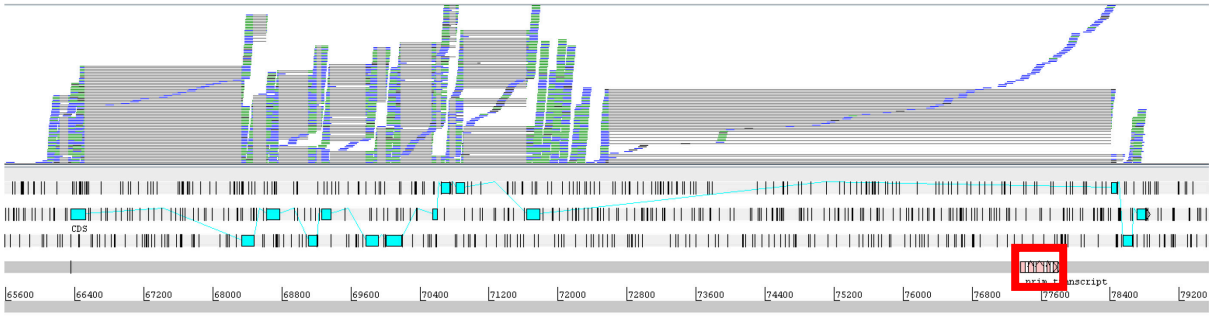


Figure 3.9 Close up view of the MGAT2 protein. The miR-5352 cluster is shown in red. Blue boxes along the three forward reading frames indicate coding regions (Laing *et al.,* 2013). **Blue and Green lines represent the RNA reads; grey lines indicate paired reads.**

Figure 3.9 shows a close up view of the MGAT2 domain containing protein with the location of the coding regions according to Laing *et al.,* (2013). MGAT2 is predicted to encode a Golgi enzyme that is essential in converting oligomannose to complex N-glycans. The Hco-miR-5352 cluster is located towards the 3' end of the gene and is situated within a 6.5kb intron, although this intron has been removed in the most recent version of the *H. contortus* genome. BLAST searches of the encoded protein identified it in several other organisms and it shows 58% identity to a *C. elegans* protein called GLY-20.

To determine if the miR-5352 cluster may target any of the genes on the same scaffold, the target prediction programme Probability of interaction by target accessibility (PITA) was run against all the predicted genes on the scaffold. No

targets were found on the 3' UTR of any of the genes, however, the MGAT2 gene has 8 target sites for all four of the miRNAs in the miR-5352 cluster, all of which are located towards the end of the gene (Figure 3.10). From the Artemis screenshot, the gene is highly transcribed in the L3 stage and less so in the adult males, which correlates well with the increased levels of cluster expression in adult worms. These data suggest that the MGAT2 gene may be controlled by the miR-5352 cluster.

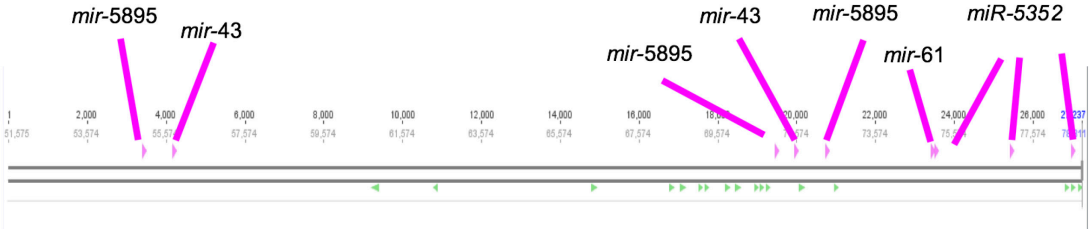


Figure 3.10 Diagram showing the predicted miRNA binding sites in the *H. contortus* MGAT2 gene (pink arrows) and the coding regions (green arrows). None of the predicted miRNA binding sites are located in the coding regions.

PITA was also run against the other genes on the scaffold to determine if these genes could be potential targets of the miRNA cluster. While none of the other mRNAs had any binding sites for any of the miRNAs in the cluster, they all had predicted binding sites for other miRNAs (Figure 3.11)

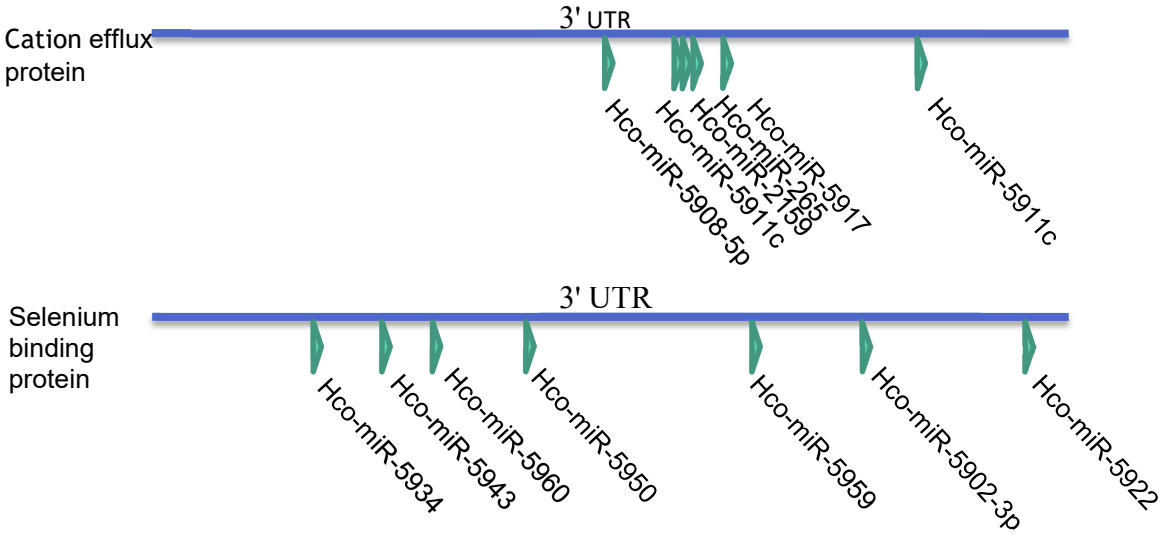


Figure 3.11 Diagram showing the location of PITA predicted miRNA binding sites on the 3' UTR of the Cation efflux protein and the Selenium binding protein (green triangles).

The binding sites for the cation efflux protein seem to congregate towards the middle of the UTR, whereas binding sites in the selenium binding protein are more evenly spread across the UTR. The majority of miRNAs identified for these two genes are also unique to *H. contortus*. Hco-miR-265 and 2159 can be identified in a few other organisms, however their function is unknown. The MEC3 3' UTR did not have any binding sites that had a binding energy over the minimum. The unnamed protein with poor similarity to Metridin ShK had 14 binding sites for Hco-miR-7 and 71 (Figure 3.12). The binding sites are all predicted to be grouped relatively close together towards the 3' end of the UTR. Since the seed sequence of the two miRNAs is identical apart from one mismatch, all the sites identified are predicted to bind to both miRNAs.

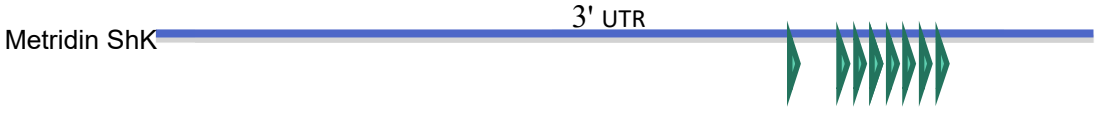


Figure 3.12 Diagram showing the binding sites for Hco-miR-7 and 71 on the 3' UTR of Metridin ShK. All sites shown (green triangles) apply to both miRNAs due to the similarity of the seed sequence between them.

PITA was also run to look for binding sites across the whole gene for each of the miRNAs in the miR-5352 cluster. Over 70 miRNA binding sites were found for each of the four genes. However, MEC3 and Selenium binding protein had only one additional binding site for Hco-miR-5895 outside of the 3' UTR. The absence of miRNA binding sites in the other genes of the scaffold seems to suggest that MGAT2 is the most likely candidate to be targeted by miRNAs of the miR-5352 cluster.

Figure 3.13A shows the read counts for miR-5352 and the MGAT2 mRNA transcript in the larval and adult stages of *H. contortus* (correlation coefficient of -0.046), suggesting there is no correlation between MGAT2 and Hco-miR-5352 (Figure 3.13B).

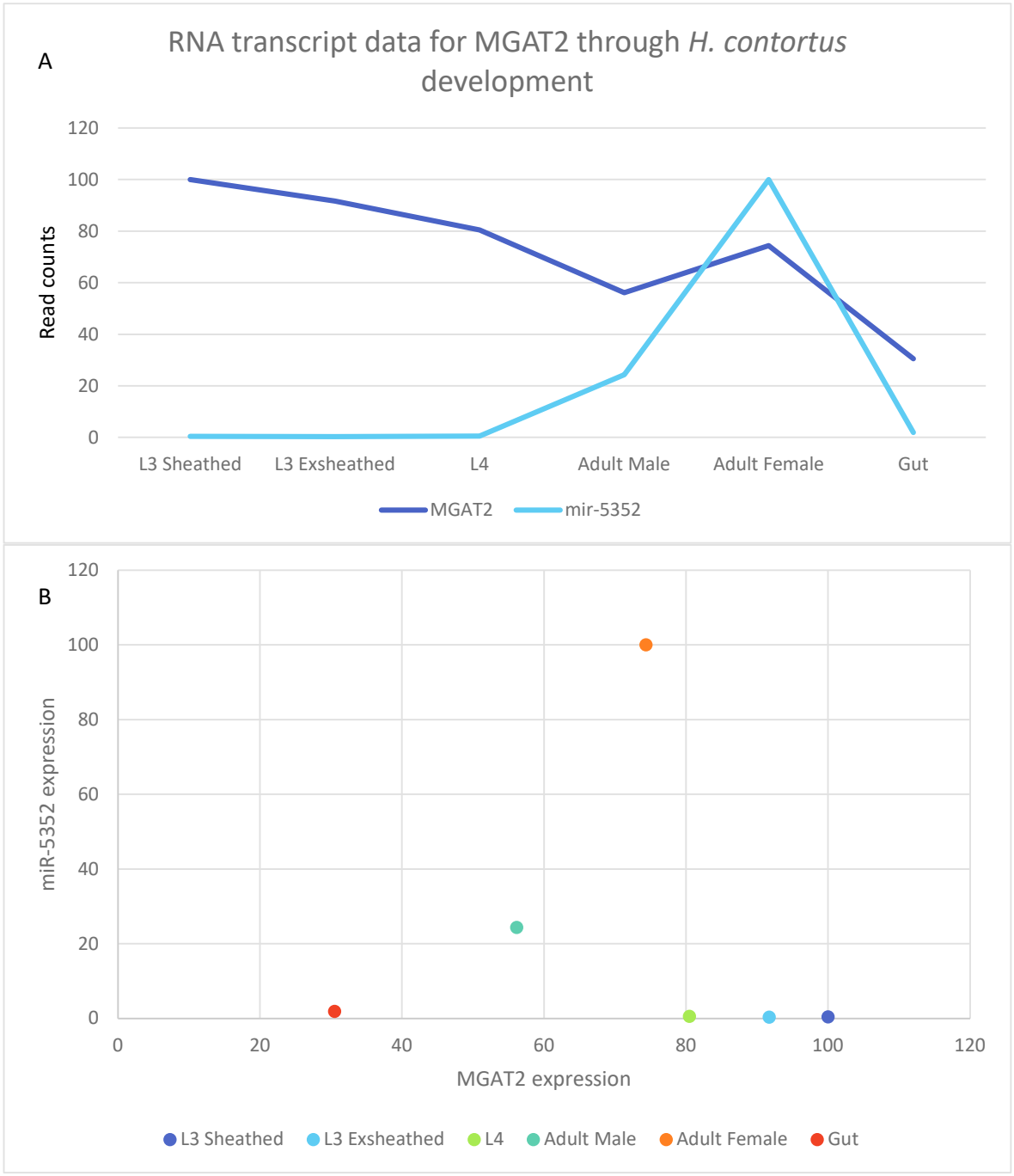


Figure 3.13 Normalised reads from MGAT2 RNA transcript data and miR-5352 microarray data. A) Graph showing the expression of MGAT2 RNA transcripts and miR-5352 microarray signal across different life cycle stages. MGAT2 shows a decreasing level of expression whereas miR-5352 expression increases. B) Scatter plot of the MGAT2 expression compared to the miR-5352 expression with a correlation coefficient of -0.046.

3.2.4 Conservation of the miR-5352 cluster in other parasitic nematodes

Using BLAST analysis, Winter *et al.,* investigated the presence of the miR-5352 cluster in other nematodes (Winter *et al.,* 2012). Only in *Ascaris suum* could miR-5352 and miR-43 be identified in miRBase. This analysis was further expanded using WormBase ParaSite to investigate the presence of the cluster in the genomes of other parasitic worms that have since been sequenced. Out of the 90 nematode species for which information is present in WormBase ParaSite, the miR-5352 cluster was identified in nine species, namely; *Oesophagostomum dentatum, Strongylus vulgaris, Nippostrongylus brasiliensis, Necator americanus, Teladorsagia circumcincta*, *Heligmosomoides bakeri, Cylicostephanus goldi, Ancylostoma duodenale* and *Ancylostoma ceylanicum*. The specific databases from which members of the miR-5352 cluster could be identified and the location of these parasites within the host is shown in Appendix Table 8.1. Figure 3.14 shows the multiple alignment of the cluster sequence from all nine organisms along with the sequence of the Hco-miR-5352 cluster. For all four miRNAs, there is significant similarity between the precursor sequences across all nine species, ranging between 61.9 to 78.8% pairwise identity. This number increases to over 95% when only the mature sequences are considered. None of these sequences have been reported as miRNAs and are not present in miRBase. All of these sequences were annotated manually and RNA folding was used to check the folding structure of these precursor sequences (data not shown).

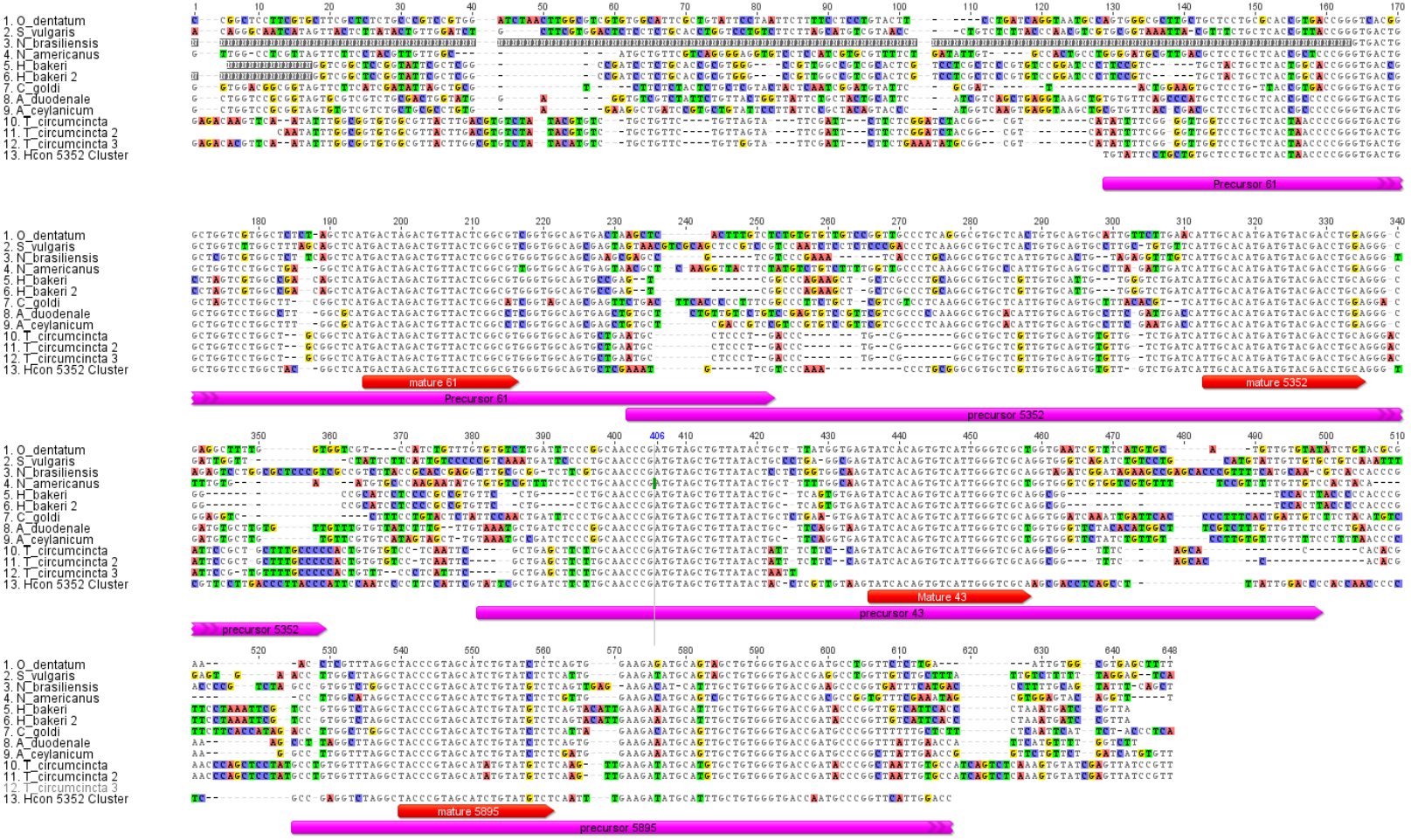


Figure 3.14 BLAST searches were used to identify the scaffolds in WormBase ParaSite that were most similar to the *H. contortus* miR-5352 cluster. The regions were then extracted and a multiple sequence alignment was carried out with all the sequences. Bases that are highlighted in colour show mismatches between the miRBase sequence and the genome assembly. Dashes (-) represent gaps in the alignment. The purple bars indicate the miRBase precursor sequence and the red bars indicate the miRBase mature sequence. Multiple sequences were obtained from *T. circumcincta* and *H. bakeri* as indicated by the number after the name.

From the BLAST searches, multiple sequences were identified from *H. bakeri* and *T. circumcincta* that contain the Hco-miR-5352 cluster. From the contig location coordinates, *T. circumcincta* sequences 1 and 2 are located at 144782-144965 and 144994-145200. With *H. bakeri*, the locations are also extremely close together at 48067-48260 and 48281-48489. For both sets of sequences, they show near 100% similarity to each other. The high similarity between the sequences and their location on the genome suggests that these sequences may have been duplicated. *T. circumcincta* sequence 3 is located on a different contig and is both shorter and shows a number of mismatches to sequence 1 and 2, suggesting it may have arisen through duplication and divergence (Figure 3.15).

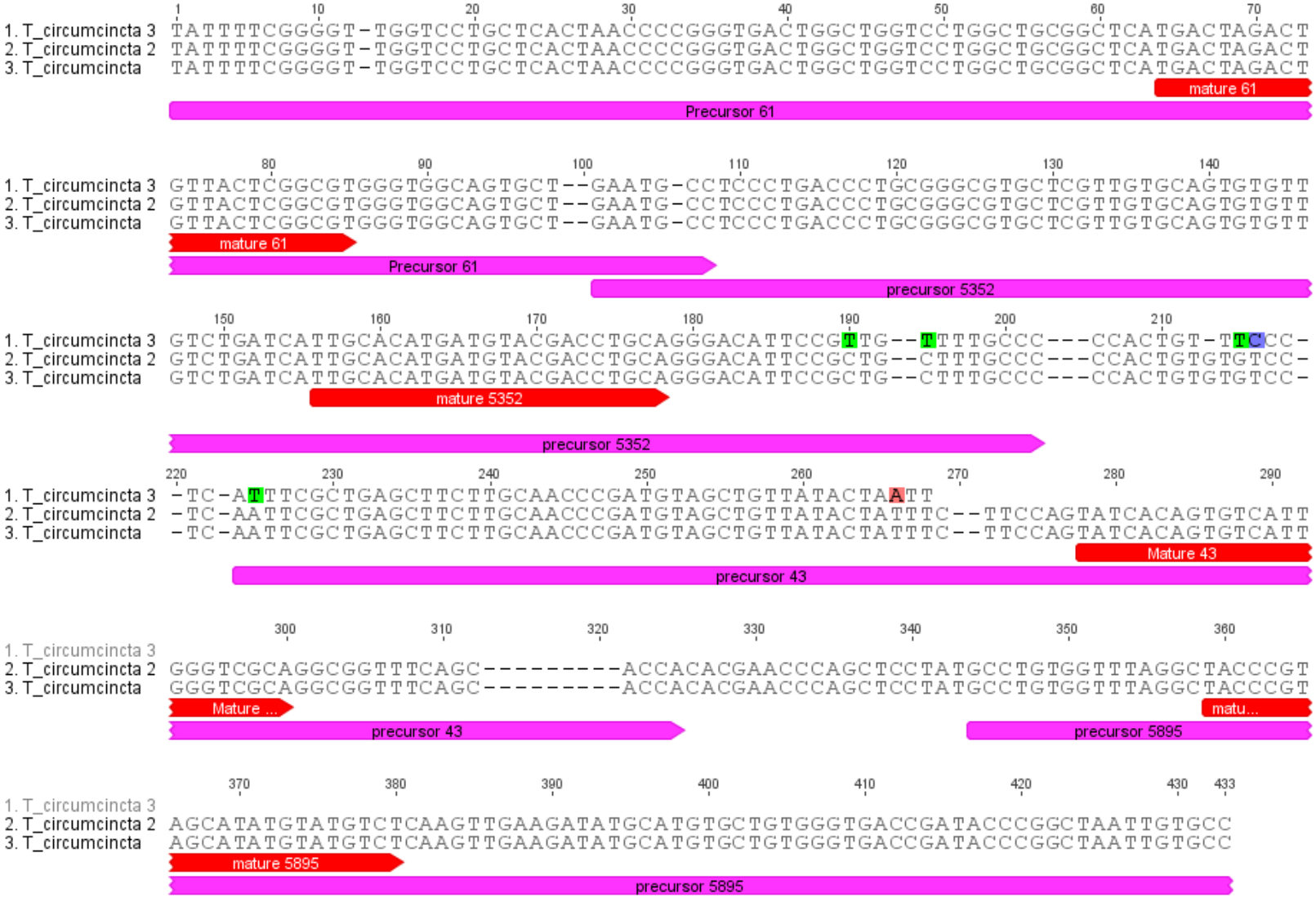


Figure 3.15 Alignment of the three *T. circumcincta* miR-5352 cluster sequences obtained from WormBase ParaSite. Highlighted nucleotides are divergent and dashes (-) indicate gaps in the alignment.

A consensus tree was generated using the sequences from WormBase ParaSite (Figure 3.16). Out of all the parasitic nematodes with genome data, only the sub-order Strongylida contained the complete Hco-miR-5352 cluster. The exception is the super-family Metastrongyloidea: neither *Angiostrongylus costaricensis* nor *A. cantonensis* contain the Hco-miR-5352 cluster as a whole, nor do they contain any of the individual miRNAs. Part of the cluster is present in *Ascaris,* belonging to the Ascaridida family (see below).

H. contortus is a clade V nematode and is phylogenetically, closely related to the free-living nematode *Caenorhabditis elegans*. This phylogenetic relationship facilitates the use of comparative analysis for genome annotation and functional studies (Gilleard, 2004; Laing *et al.,* 2011). Although the two organisms are closely related, the Hco-miR-5352 cluster was not identified in *C. elegans*, however, *C. elegans* does possess a cel-miR-43 and cel-miR-61. The seed sequences of the *C. elegans* miRNAs are identical to the *H. contortus* sequences, although the miRNAs are on different scaffolds and together do not form a cluster (data not shown). *Ascaris*, a clade III nematode of the gastro-intestinal tract has a miR-5352 sequence but *Brugia malayi,* a clade III nematode of the lymphatic system does not. Thus the conservation of this miRNA cluster does not appear to be clade-specific. Each of the organisms that have the miR-5352 cluster are parasitic nematodes that infect a range of hosts. From this initial analysis, those nematodes with a conserved miR-5352 cluster seem to be localised to the gastro-intestinal tract (Figure 3.16).

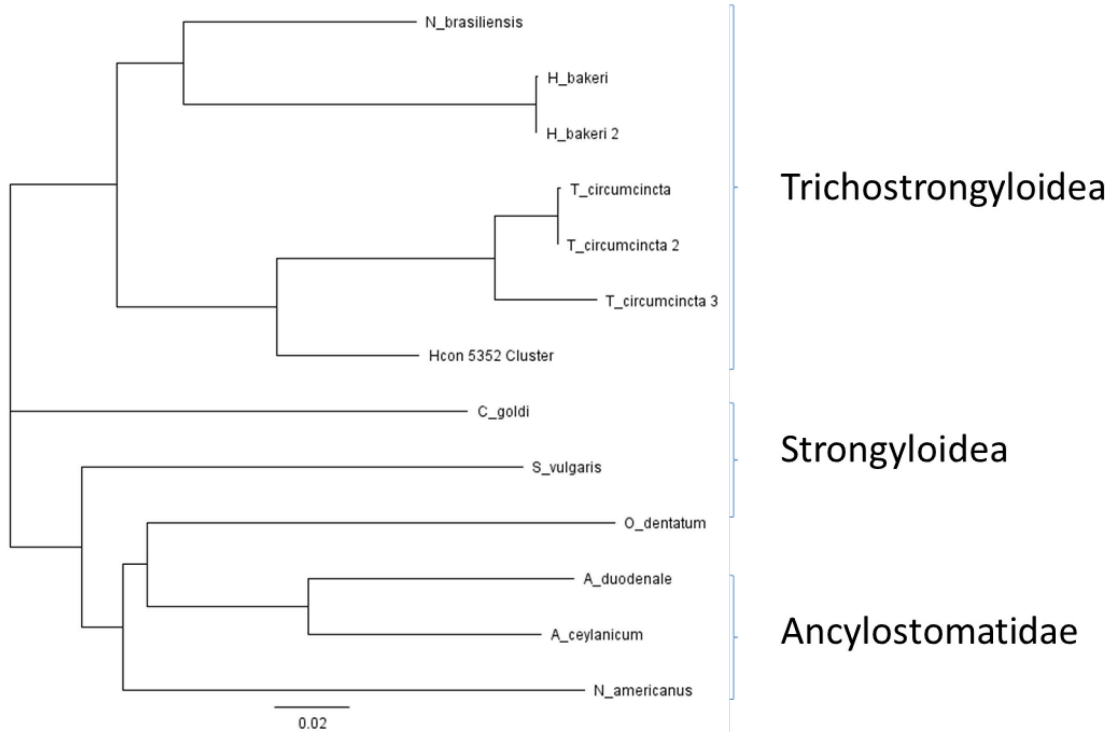


Figure 3.16 Phylogenetic tree of the Hco-miR-5352 cluster sequences obtained from BLAST analysis. Sequences were only identified in the sub-order Strongylida, but not in the super-family Metastrongyloidea. Scale shows nucleotide substitutions per site.

BLASTN searches of WormBase ParaSite did not identify the cluster in *Ascaris suum*. This was surprising as Asu-miR-5352 is an accepted miRNA by miRBase, with a 70% pairwise identity to the *H. contortus* sequence (Figure 3.17). It is also the only other miRNA in miRBase numbered as miR-5352.

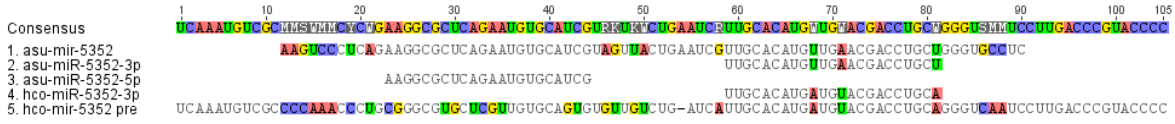


Figure 3.17 Multiple alignment of the Hco-miR-5352 precursor and mature sequence with the *Ascaris suum* miR-5352 sequences. Highlighted nucleotides show divergence between the two species. Pairwise identity between the two precursor sequences is 62% and 73.7% between the two 3p sequences and 5p respectively.

However, a search of the Hco-miR-5352 precursor sequence in WormBase ParaSite identified a hit in the *Ascaris suum* and *A. lumbricoides* genome assembly. Out of the 97 *A. suum* miRNA sequences in miRBase, 85 had 100% similarity to an *A. lumbricoides* scaffold. The other 12 *A. suum* sequences have

similarity scores between 90 to 98.8% to *A. lumbricoides*. For the Hco-miR-5352 cluster, six versions of miR-43 were identified in both *Ascaris* genome scaffolds, however, miR-61 and miR-5895 were not found in miRBase or in the *Ascaris* genome.

Hco-miR-5352 cluster precursor and mature sequences along with *A. suum* miR-43 and 5352 sequences were aligned to both *Ascaris* genomes that contained miR-5352. From Figure 3.16, the mature *Ascaris* sequences are very similar, differing by only one base (99.8% pairwise similarity). Hco-miR-5352 has a 70% pairwise identity to the Asu-miR-5352 precursor. Hco-miR-43 has a 69% identity *to Asu-miR-43d*.The *Ascaris* miRNAs are also in the same order as in *H. contortus*, albeit with a much larger gap between the two precursors (Figure 3.18). In conclusion, it appears that while the *Ascaris* genome contains miR-5352 and miR-43, other members of the cluster cannot be identified at present.

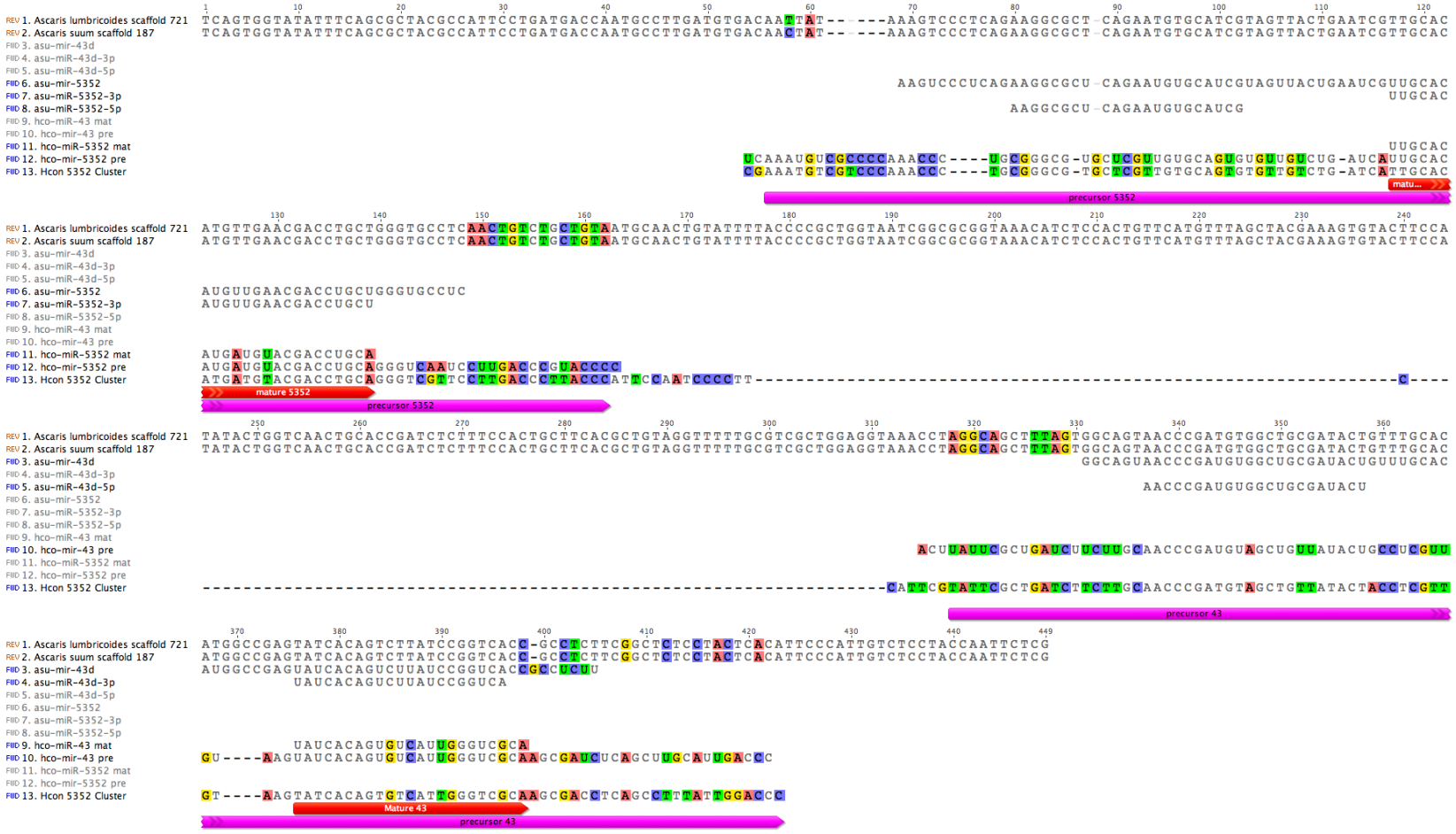


Figure 3.18 Multiple alignment of the Hco-miR-5352 and Hco-miR-43 precursor and mature sequence and *Ascaris suum* miR-5352 and miR-43 sequences against the *A. suum* genome scaffold (scaffold:AscSuum_1.0_submitted:Scaffold187:163878:164556:-1) and *A. lumbricoides* genome scaffold (A_lumbricoides_Ecuador_v1_5_4, scaffold ALUE_scaffold0000721). Highlighted nucleotides show divergence in the alignment and dashes (-) indicate gaps.

In addition to *Ascaris*, searches of the most recent release of WormBase ParaSite databases for the mature Hco-miR-5352 sequence identified *Dictyocaulus viviparus* as a hit. The region of alignment is 23 bp long and has 95.7% similarity to the *H. contortus* sequence (Figure 3.19). Further searches identified a 100% match to Hco-miR-43 and a 90.9% match to Hco-miR-61. From the location, the three matches described above were extremely close together and extracting the surrounding regions and doing more in depth analysis identified sequence aligning to Hco-miR-5895 but with only 87.9% identity. This indicates that the full cluster is present in *D. viviparus* but is less conserved than in the other Trichostrongylid species. Figure 3.20 shows the phylogenetic tree in Figure 3.16 with the *A. suum* and *D. viviparus* sequences included. Interestingly, even though *D. viviparus* contains the whole miR-5352 cluster, the phylogenetic tree suggests that the *D. viviparus* sequence is very distantly related to the other sequences. *Ascaris* is also very distantly related. Mfold analysis of the predicted precursor sequences seems to indicate that the miRNAs are able to fold into hairpin loops (Figure 8.1). In contrast to all the other nematode scaffolds encoding the miR-5352 cluster, the direction of all these miRNAs in *D. viviparus* is reversed.

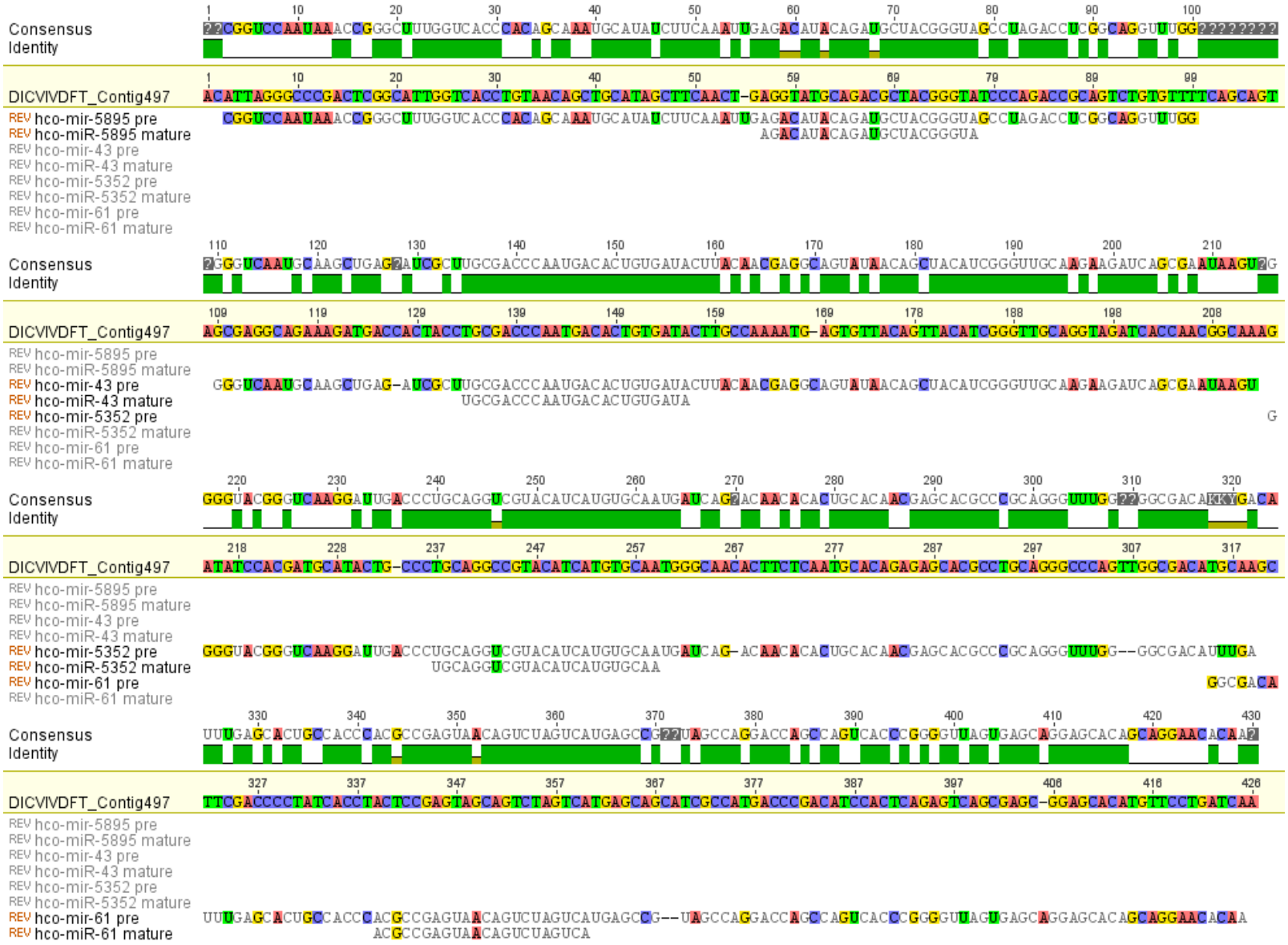


Figure 3.19 Multiple alignment of the *Dictyocaulus viviparus* contig 497 with the precursor and mature sequences of the *H. contortus* Hco-miR-5352 cluster. Highlighted nucleotides show divergence between the *D. viviparus* sequence and the *H. contortus* sequences. Dashes (-) indicate gaps between the alignment.

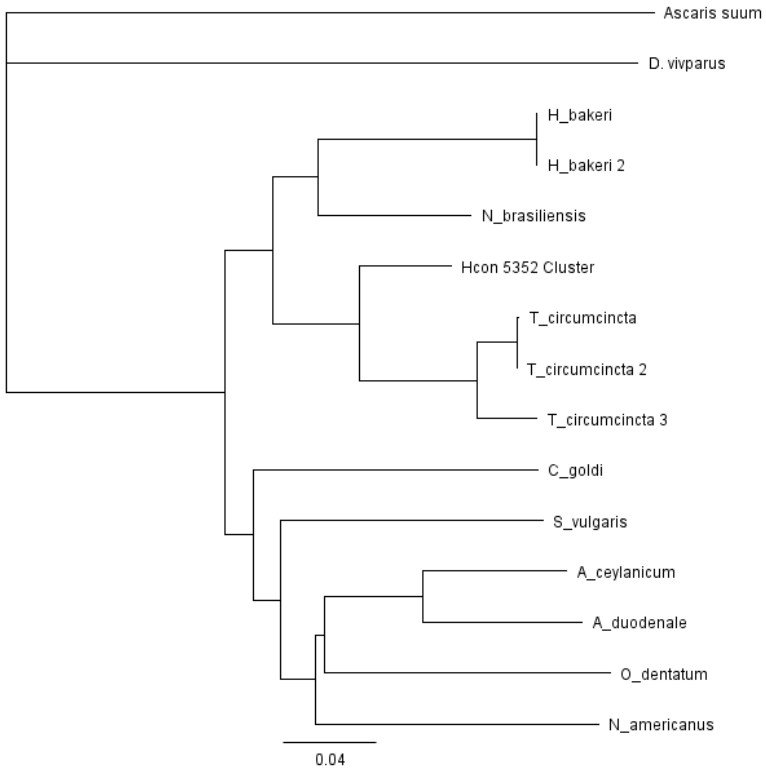


Figure 3.20 Phylogenetic tree of the Hco-miR-5352 cluster sequences obtained from WormBase ParaSite including *Ascaris suum* and *Dictyocaulus viviparus.* Scale shows nucleotide substitutions per site.

3.2.5 Specific miRNAs are detected in ES by qPCR

The presence of the miR-5352 cluster only in nematodes that parasitise the gastro-intestinal tract suggests that these miRNAs could function in the host-parasite interaction. The recent studies of Buck *et al.*, showed compelling evidence that parasitic nematode miRNAs can interact with host cells to alter host gene expression (Buck *et al.*, 2014). If *H. contortus* miRNAs are able to function in a similar manner to *H. polygyrus* ES products, the miRNAs must be able to leave the worm to interact with the host. One possibility is that the miRNAs are released in the nematode ES products. Since *H. contortus* is a blood feeder, there is also the possibility that miRNAs may find their way into the host from the *H. contortus* gut through released gut material and/or regurgitation.

To investigate whether miR-5352 could be detected in parasite ES products, *H. contortus* was cultured *in vitro*. ES was collected from mixed sex adult worms

over a 24-hour period, miRNA extracted and cDNA synthesised (as described in Chapter 2.2). RT-PCR was carried out for miR-5352 in adult ES with whole adult female *H. contortus* cDNA being used as a positive control. In addition, several worms were removed prior to *in vitro* culture and were dissected by Neil Marks to remove the worm intestine, from which RNA was prepared. The data shown in Figure 3.21 confirmed the presence of miR-5352 in the ES but, perhaps surprisingly, not in the *H. contortus* gut tissue. In this experiment, the data has not been normalised as it was unknown which miRNA to use as the normaliser. Hco-miR-50 was included in the experiment as microarray data showed that it was expressed at a relatively constant level across all the life stages and could be used to demonstrate that the RNA samples are PCR competent.

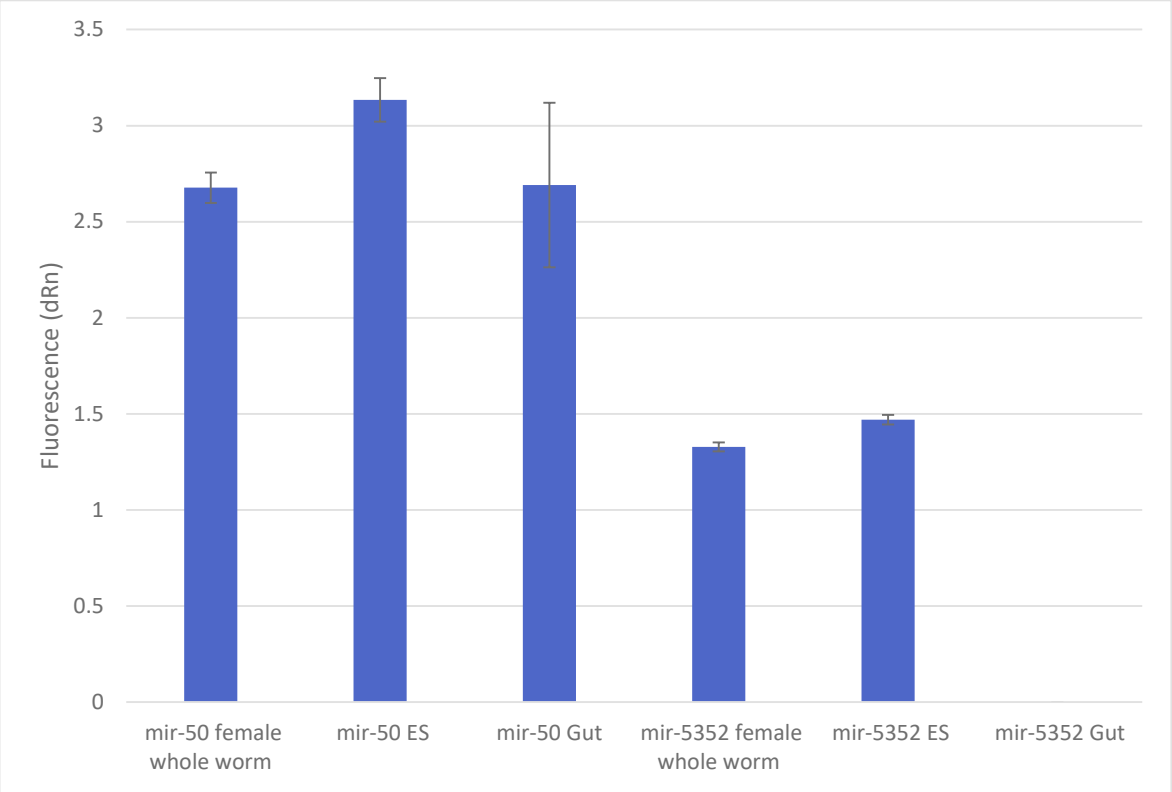


Figure 3.21 RT-PCR amplification analysis of *Hco-miR-50* and Hco-miR-5352 expression in adult female whole worms, adult ES and adult gut samples. Results shown represent the mean ± SD of three technical replicates.

3.2.6 miR-5352 is also present in eggs/L1

When collecting the adult worm ES, there was a large number of eggs and hatched L1 in the medium. The ES was centrifuged and filtered to remove these stages and eggs and L1 collected to determine if these life cycle stages express Hco-miR-5352. Eggs and L1 collected by centrifugation were ground in liquid nitrogen. cDNA was then synthesised from the RNA and used alongside adult female whole worm cDNA in RT-PCR. miR-50 was used for quality control of the samples. It is unknown if miR-50 is present in eggs and L1, however, the microarray data shows that miR-50 levels are similar from L3 to adult. As such, it should be amplified in the whole worm sample, if nothing else. Figure 3.22shows that the expression of Hco-miR-50 between adult female whole worms and eggs/L1 is very similar, as is the expression of Hco-miR-5352.

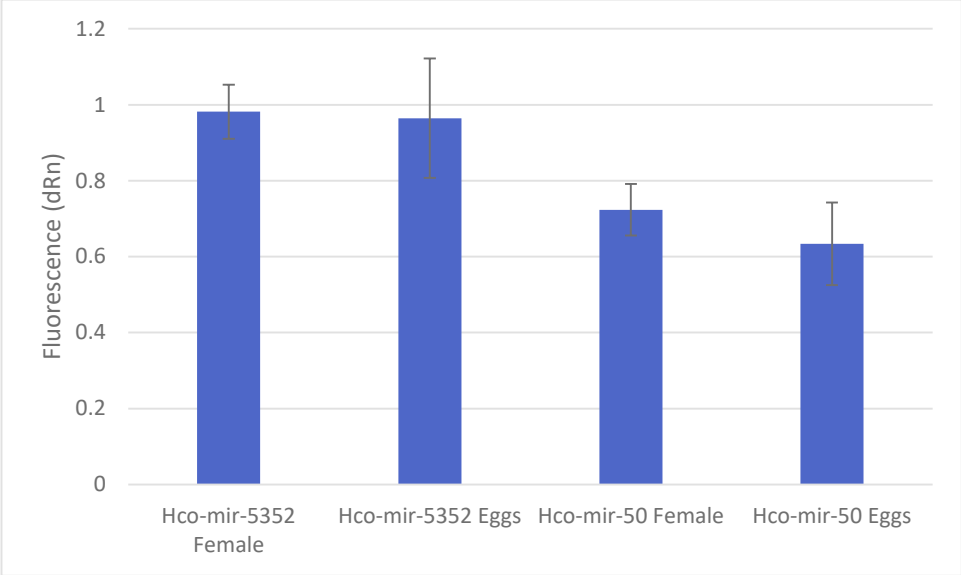


Figure 3.22 RT-PCR amplification analysis of *Hco-miR-50* and Hco-miR-5352 expression in adult females and egg/L1 sample. Results shown represent the mean ± SD of three technical replicates.

3.2.7 Creation of the *H. contortus* 3′ UTR database

With the availability of high quality genome data, target prediction for the miR-5352 cluster can be performed using an *H. contortus* gene target database. To this end, a 3' UTR database was created based on transcriptomic data consisting of 14,507 sequences, which accounts for approximately half of the annotated

genes. The length of these sequences was quite varied ranging from 1 nucleotide to 14,000 nucleotides (Figure 3.23).

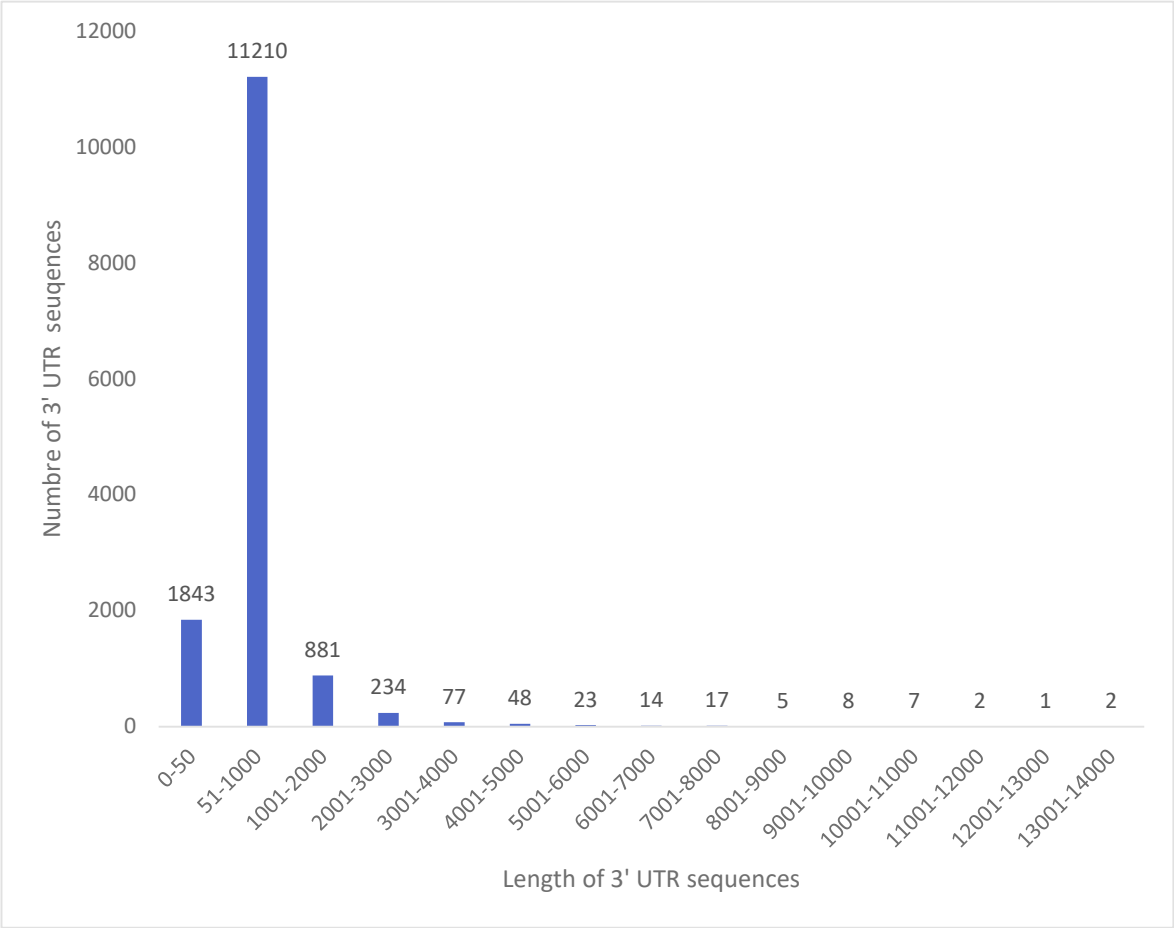


Figure 3.23 Summary of the 3' UTR lengths in nucleotides as predicted based on transcriptomic data.

Since the 3' UTR sequence database lacks a large number of *H. contortus* genes and contains sequences that are unlikely to be UTRs, a second database, the 3' downstream (DS) was created by taking 1,000 nucleotides downstream of each annotated stop codon. This database contains 93% of the annotated genes.

Table 3.1 shows the length distribution of the 3' DS database and while the majority of sequences are between 950-1049 nt long, there are a number which are shorter. These were due to the gene being annotated towards the end of a scaffold and there being no more sequence downstream of the stop codon.

The final database was created by combining the 3' UTR and 3' DS databases. The 3' UTR sequences was taken if the sequence was greater than 50 nucleotides or less than 1,000 nucleotides. Those sequences that did not fulfil these criteria or did not have a corresponding 3' UTR sequence were replaced with the 3' DS sequence.

Table 3.1 Summary of the 3' UTR sequences in the 3' DS database. The database was created by extracting 1,000 nucleotides downstream of the annotated stop codon or until the end of the scaffold, whichever came first.

Length (nt)	Number of UTR sequences
0-50	83
51-100	48
101-200	97
201-300	76
301-400	84
401-500	81
501-600	68
601-700	83
701-800	73
801-900	58
901-1000	19582
Grand Total	**20333**

3.2.8 PITA target identification

PITA was used to identify miRNA targets as it focuses on the target site accessibility based upon the mRNA secondary structure and calculating the thermodynamics of the interaction. This is important as a thermodynamically unfavourable interaction, such as a strong 3' UTR structure, will interfere with the binding of the miRNA.

The four miRNAs in the Hco-miR-5352 cluster and all the 3' UTRs in the UTR database were run through PITA to identify possible targets. Table 3.2 shows the results obtained for each of the miRNAs. The targets were then filtered for ΔΔG (< -10) and seed length (8 base match with one G:U wobble or one mismatch). ΔΔG is calculated as ΔG(duplex), the hybridization energy of the miRNA to the binding site, minus, ΔG(open), the energy required to open the secondary structure around the binding site.

The large number of targets predicted by PITA makes analysis difficult (Table 3.2). To narrow down the lists of targets, several different methodologies were applied. Initially, targets in common between the four miRNAs of the Hco-miR-5352 cluster were analysed.

Table 3.2 The four miRNAs in the Hco-miR-5352 cluster were run through PITA to identify 3' UTR targets. Targets were filtered for ΔΔG less than -10 and only those potential interactions with a seed sequence of 8 bases with one G:U wobble or one mismatch were considered

microRNA	Targets filtered by ΔΔG and seed length
Hco-miR-43	1980
Hco-miR-5352	930
Hco-miR-5895	321
Hco-miR-61	884

3.2.8.1 Common 3' UTR targets of the four miRNAs in the Hco-miR-5352 cluster predicted by PITA

It is known that miRNAs are able to form clusters (Altuvia *et al.*, 2005) and that these clusters are usually co-ordinately controlled by transcription factors, implying a coordinated regulation (Schmeier *et al.*, 2009; Wang *et al.*, 2010c; Woods *et al.*, 2007). However, analysis of the *H. contortus* genes is complicated by incomplete annotation of the genome.

The target results for all four miRNAs of the Hco-miR-5352 cluster were pooled to identify any targets that were present in all four target lists. 1230 unique 3' UTR targets were pooled, of which 129 had binding sites for two of the miRNAs. Only 3 sequences, g1903 and g2005 (common element in Hco-miR-43, 5895 and 5352) and g8098 (common element in Hco-miR-43, 61 and 5352), were targeted by three miRNAs (see Figure 3.24). No target had binding sites for all four miRNAs of the cluster.

As these *H. contortus* genes are poorly annotated, BLAST was used to identify homologues in *C. elegans*. No *C. elegans* homologue for g2005 was found, whilst g1903 is similar to the *C. elegans* gene C18B2.3, however no function is known.

G8098 shows similarity to cnt-2, an Arf GTPase activating protein (GAP) of the AGAP family which is required for asymmetric cell division.

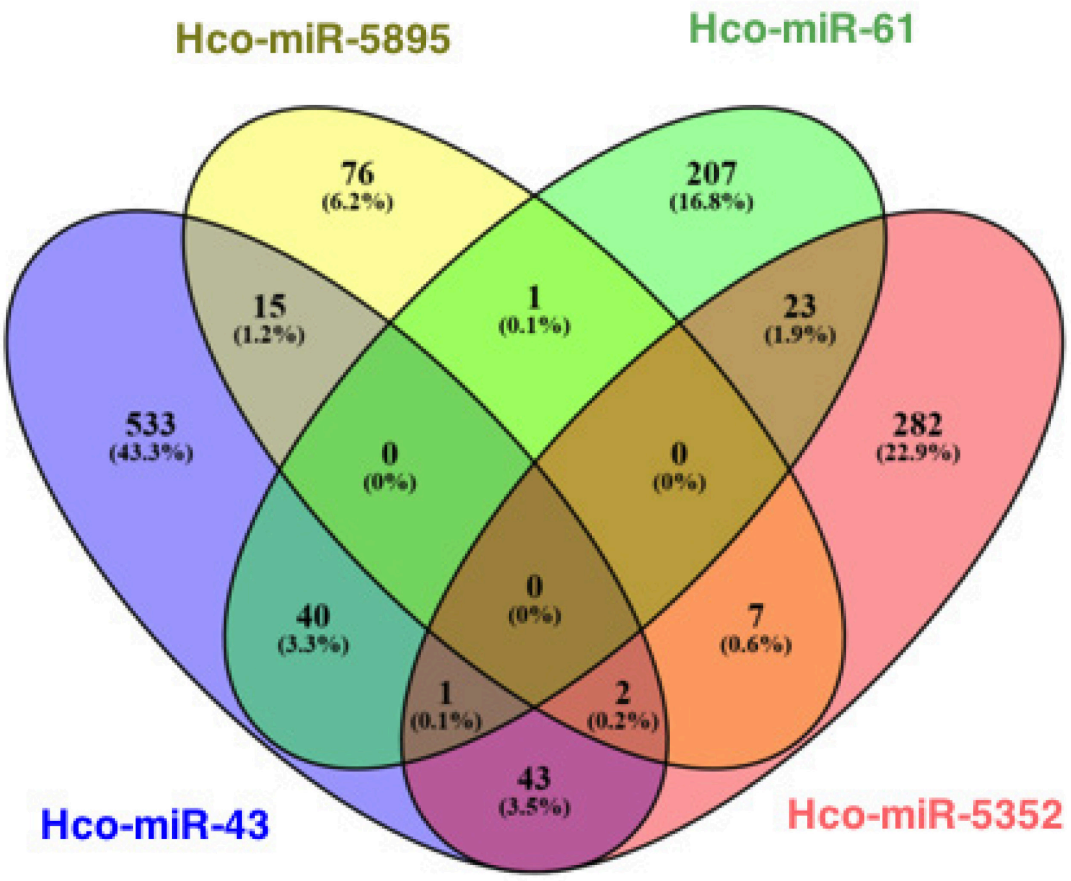


Figure 3.24 Venn diagram showing the distribution of genes in the target lists of the four miRNAs of the Hco-miR-5352 cluster. No target appears in the target list for all four miRNAs, three targets appear in the list of three genes.

G2005 can be excluded from further analysis as the read count for the miRNA in the different life cycle stages was less than ten. G1903 and g8098 are both predicted targets of three miRNAs (Hco-miR-5352, 43 and one other). The correlation coefficient between g1903 or g8098 and miR-5352 was calculated. G1903 or g8098 show a correlation coefficient of 0.98 and 0.14 respectively, showing that they are not inversely related to miR-5352, although g1903 expression correlates with miR-5352 (Table 3.3).

Table 3.3 Hierachical clustering of the transcriptome of *H. contortus* genes showing common predicted targets of three of the miRNAs in the Hco-miR-5352 cluster. Gene expression across the life cycles stages are clustered together. Blue squares indicate low expression, white squares indicate medium expression and red squares indicate high expression.

Gene ID	L3she	L3ex	L4	Male	Female	Gut
g2005	0	0	2	2	1	3
g1903	590	696	875	1822	16006	1582
g8098	324	317	440	1354	615	956

3.2.8.2 3' UTR targets with multiple Hco-miR-5352 binding sites predicted by PITA

As very few genes were common between multiple miRNA target lists, a further analysis focused on just the targets of Hco-miR-5352. However, the filtered lists still predict 930 unique *H. contortus* 3' UTR as potential targets. It has been demonstrated that increasing the number of miRNA target sites increases the degree of gene repression (Doench *et al.*, 2003) and thus the number of binding sites was used to narrow down the list of targets.

From the 930 targets, only 20 possessed two Hco-miR-5352 binding sites with a further two mRNAs possessing three binding sites (Table 3.4). For the two mRNAs with three binding sites, one was identified as a theoretical protein and had no similarities with any sequence in NCBI. The other sequence (g18228) was identified as a Neurotransmitter-gated ion-channel ligand-binding and "Neurotransmitter-gated ion-channel transmembrane region domain containing protein [*Haemonchus contortus*]" (Figure 3.25).

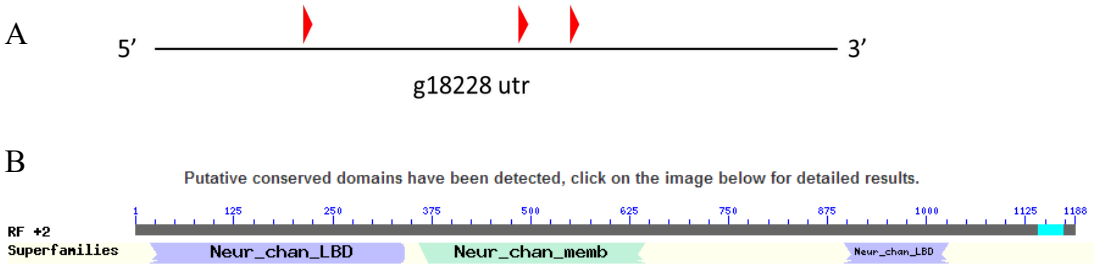


Figure 3.25 A) Diagram showing the binding sites of miR-5352 on the 3' UTR of target protein g18228. B) Putative conserved domains according to NCBI blast of the most significant protein target (g18228) of miR-5352 identified by PITA

The Neurotransmitter protein (g18228) has three domains that show high similarity to the Neur_chan_LDB and member superfamilies with E-values ranging between 6.25e-05 and 5.62e-08. The protein contains one possible transmembrane region and two ligand binding domains. RNA transcript data in *H. contortus* across development shows the highest level of expression is in the L3 stages. Expression drops off in the L4 with no expression in the adult female. While this profile is inversely correlated to miR-5352 expression (correlation coefficient of -0.56). (Figure 3.26), it is possible that g18228 is a larval specific gene controlled by other factors. Neither miR-5352 nor g18228 can be detected in adult gut samples.

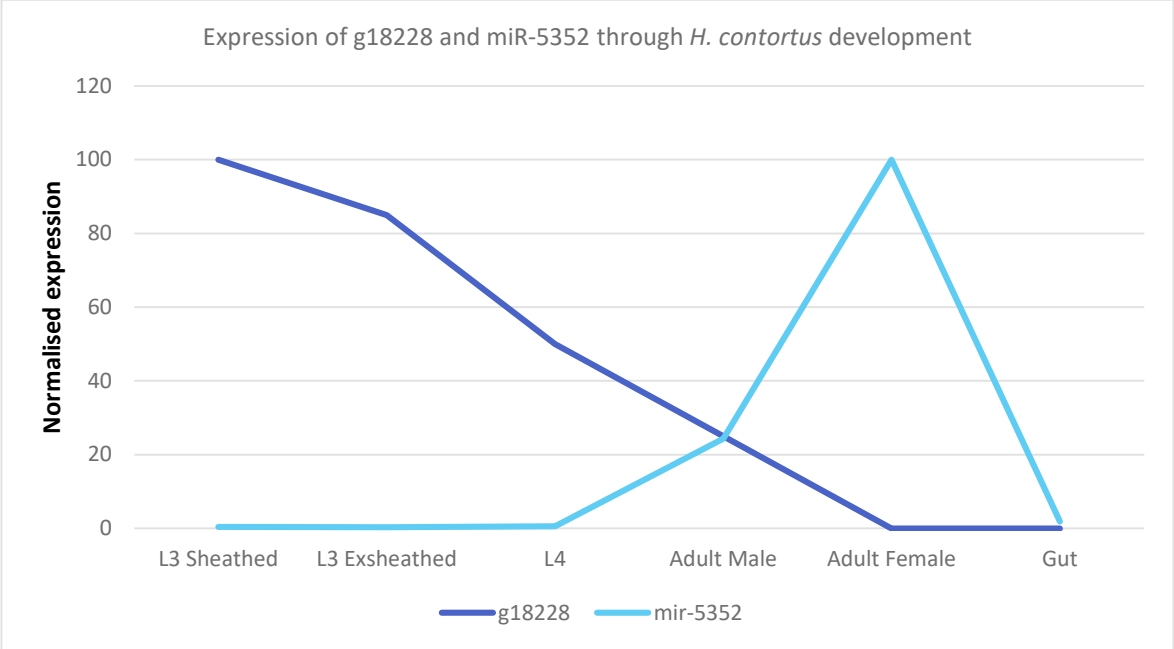


Figure 3.26 Number of normalised reads from RNA transcript data for g18228 and miR-5352 through *H. contortus* development.

Table 3.4 List of *H. contortus* genes with multiple predicted binding sites for Hco-miR-5352 using the PITA miRNA target prediction programme.

H. contortus genes	Number of binding sites for Hco-miR-5352
g18228	3
g20583	3
g20464	2
g14383	2
g7181	2
g9959	2
g14440	2
g2931	2
g16627	2
g4610	2
g13604	2
g10601	2
g15020	2
g745	2
g18278	2
g566	2
g20120	2
g19468	2
g6664	2
g3824	2
g2608	2
g10239	2

Transcriptomic data was then obtained for the remaining 22 unique predicted targets of Hco-miR-5352 and compared to the expression of Hco-miR-5352. Genes which show an inverse correlation with Hco-miR-5352 were clustered using Gene-E (Table 3.5). The clustering shows that along with g18228, g15020 (Tropomodulin domain containing protein) and g745 (unknown *Haemonchus* protein) all have similar expression patterns across life cycle stages.

Unlike g18228, g6664 and g3824 both have high expression in the gut where Hco-miR-5352 is low and high expression in the adult stages where Hco-miR-5352 expression is high. Both g6664 and g3824 contain the Sec1 domain. Mammalian Sec1 proteins share 44-63% sequence similarity with nematode UNC-18 and have been reported to be involved in general secretion and synaptic transmission

(Halachmi and Lev, 2002). The target prediction combined with the transcriptome data suggests that g18228, g15020, g745, g6664 and g3824 could all be possible targets of Hco-miR-5352.

Table 3.5 Hierachical clustering of the transcriptome of 13 *H. contortus* genes that have multiple sites on their 3' UTR that were identified by PITA as potential binding sites.The correlation coefficient was calculated for each gene and compared to Hco-miR-5352. Only genes with a correlation coefficient less than -0.5 were selected. Gene expression across the life cycles stages are clustered together. Blue squares indicate low expression, white squares indicate medium expression and red squares indicate high expression. Data was not normalised for each miRNA across the life cycle stages to show the low read counts for certain miRNAs

Gene ID	L3she	L3ex	L4	Male	Female	Gut
g10601	25	25	27	1	1	2
g14440	47	51	36	1	1	0
g18228	20	17	10	5	0	0
g15020	3080	2956	2051	2010	1626	1326
g745	1756	1605	1203	1171	902	882
g2608	222	225	124	184	124	125
g14383	19	12	52	26	2	9
g13604	244	268	218	199	161	271
g3824	1396	1183	1210	694	825	1243
g6664	113	100	84	58	47	102
g20464	247	231	106	62	9	387
g19468	16	14	9	10	7	25
g9959	267	233	136	152	140	333

As miRNAs often target different genes within a canonical pathway, it was of interest to determine whether the 13 genes listed in Table 3.5 might belong to a particular pathway. To this end, it was necessary to identify *C. elegans* homologues of the *H. contortus* genes. Homologues were identified for 10 of the 13 *H. contortus* genes of interest. To identify any pathways in common, the *C. elegans* homologues were submitted to IPA for pathway analysis. Out of the 15 homologues, only three, (g13604, g9959 and g19468) were present in the IPA database and no canonical pathway was identified that linked those three genes. Gene ontology analysis of the 15 homologues also did not identify any GO terms that were significantly enriched.

3.2.9 *H. contortus* targets predicted by PITA, miRanda and RNAhybrid.

There are a large number of resources available that can perform miRNA target prediction and use different features such as seed match, phylogenetic conservation, free energy and site accessibility. These programmes are tested using experimentally validated data, however, no single programme can explain all the variance in experimental data (Min and Yoon, 2010). This could be due to a lack of understanding of the complete process or a failure to incorporate important components of the miRNA-mRNA interaction into the prediction algorithm.

PITA, miRanda and RNAhybrid were used in combination to predict binding sites of all *H. contortus* miRNAs on the *H. contortus* 3' UTR database. RNAhybrid is a tool that calculates the minimum free energy of hybridizing a short RNA sequence to a long RNA sequence, whereas miRanda predicts targets based on sequence complementarity, position on the 3' UTR and species conservation (species conservation was not utilised in the current analysis). Therefore, a combination of different programmes could potentially improve the accuracy of target prediction. The programmes were run using the default settings. In total, 4 million hits were predicted by the three programmes covering 20,975 of the *H. contortus* 3' UTRs.

From Figure 3.24, RNAhybrid predicted the highest number of targets for a single miRNA (Hco-miR-5885 having over 30,000 predicted targets) and in general, appeared to predict slightly higher numbers than the other programmes (Figure 8.2). The median and mean targets for PITA is the lowest of all three programmes although both PITA and RNAhybrid have greater variation compared to miRanda (Table 3.6).

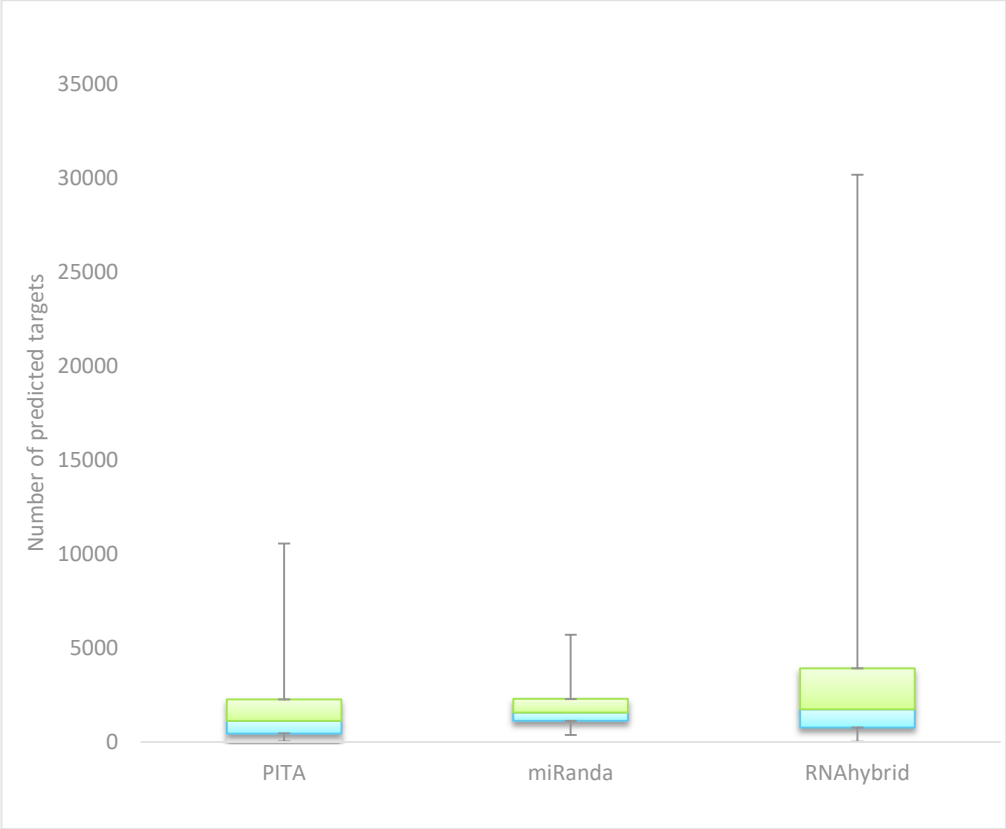


Figure 3.27 PITA, miRanda and RNAhybrid were used to predict potential miRNA 3' UTR targets for all 195 *H. contortus* miRNAs in miRBase. Boxplot shows the distribution of predicted targets for the three programmes .

Target lists from all three programmes were then combined and 3' UTRs that appeared in all three lists were selected and analysed to determine if the sites identified by the programmes overlap each other (Table 3.6). The miRNA with the lowest number of targets in common between the three programmes is Hco-miR-5893 (2 targets), whereas Hco-miR-265 has the most targets at 2,102. The average number of targets was 348.

Table 3.6 Descriptive statistics of the targets that were predicted by PITA, miRanda, RNAhybrid and the number of targets that are in common between the three programmes .

List of Targets	PITA	miRanda	RNAhybrid	Hits in three programmes
Minimum	33	371	11	2
1st Quartile	448	1119	762	62
Median	1125	1573	1732	180
Mean	1689	1824	3133	348
3rd Quartile	2260	2289	3922	455
Max	10559	5698	30177	2102

3.2.9.1 Common 3' UTR targets of the Hco-miR-5352 cluster predicted by PITA, miRanda and RNAhybrid

Target lists for the four miRNAs of the Hco-miR-5352 cluster were obtained and targets that had sites predicted by all three programmes that also overlapped were filtered. Of the 1281 unique targets predicted, 37 of the genes appear in multiple lists. g6507 (Q65ZG7) is the only gene that appears in three lists, specifically Hco-miR-5895, 61 and 43.

From Figure 3.28, 15 targets are common between Hco-miR-5352 and Hco-miR-43, five are common between Hco-miR-5352 and Hco-miR-61, with none being common between Hco-miR-5352 and Hco-miR-5895.

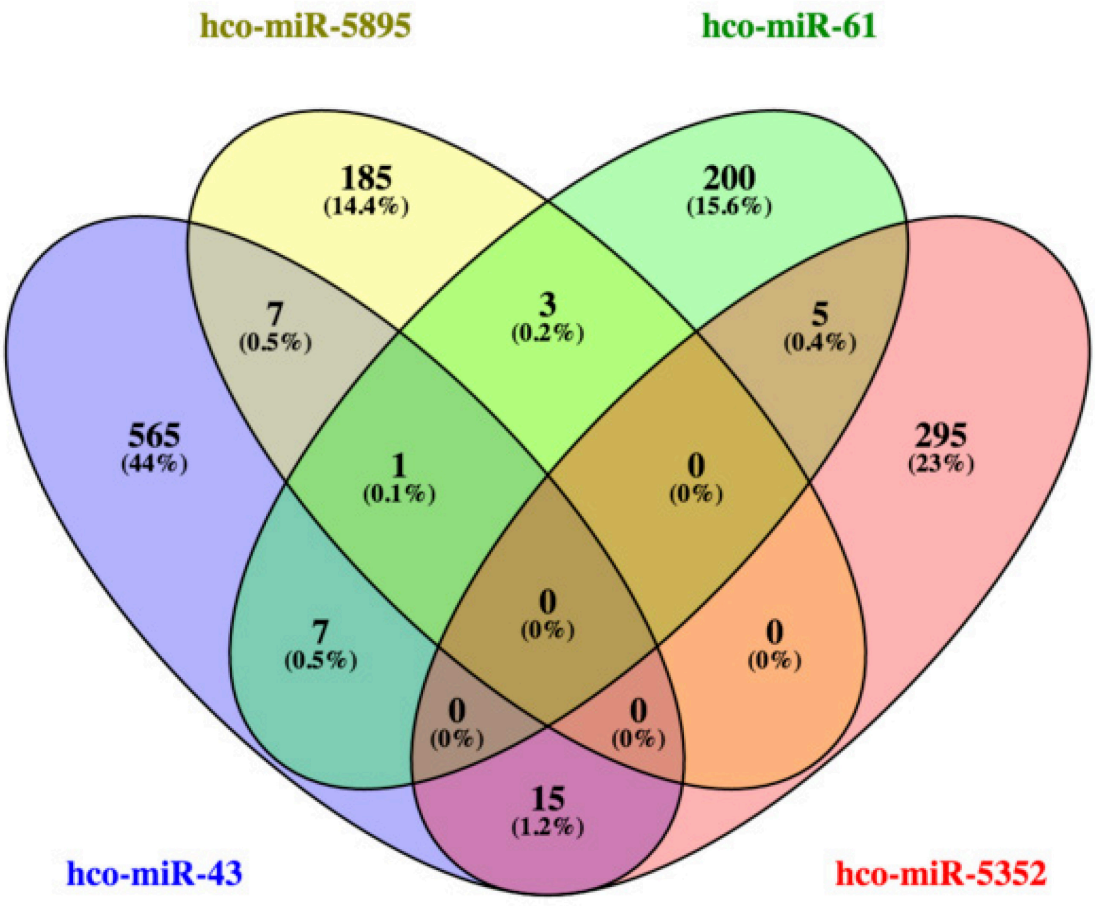


Figure 3.28 Venn diagram showing the number of common targets between the four miRNAs.

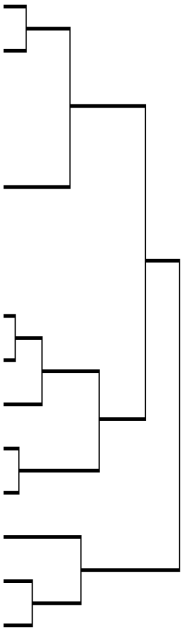
The expression profile of these common predicted targets in the *H. contortus* life cycle is shown in Figure 3.29. Some of the miRNAs show a very low level of

expression in the life cycle stages. Two genes, g18699 and g13463 have a correlation coefficient of -0.41 when compared to Hco-miR-5352 suggesting there may be a slight inverse correlation. From Figure 3.29 only 10 of the genes have *C. elegans* homologues, of which, only g9447 and g13463 were present in IPA but did not belong to any canonical pathways.

Figure 3.29 Hierarchical clustering of transcriptomic data across five life cycle stages and the gut of the genes in common between A) Hco-miR-43 and Hco-miR-53532 and B) Hco-miR-61 and Hco-miR-5352. Transcriptomic reads are shown in a heat map (blue indicates low abundance, white indicates median abundance and red indicates high abundance).

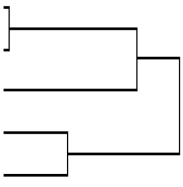
A) Genes in common between Hco-miR-43 and Hco-miR-5352

Gene ID	*C. elegans* homologue	L3she	L3ex	L4	Male	Female	Gut
g9024	Q8MNT0	853	109	145	702	702	499
g6292	Q94287	494	243	245	342	623	516
g7840	Q09278	0	3	0	0	1897	1
g14324		0	0	0	0	333	0
g21268		0	0	0	0	744	0
g10923		1	16	12	17	161	22
g10924		65	15	18	41	6547	750
g10742		484	75	72	206	182	197
g18699		1226	293	296	726	253	900
g9633	Q9XTD6	3	0	1	1	1	0
g13463	O44995	93	84	90	25	23	3
g9447	Q18214	52	41	46	40	35	12
g19672		54	27	21	297	40	5
g4802	Q21274	88	467	433	383	393	200
g12725		462	832	761	926	439	327



B) Genes in common between Hco-miR-61 and Hco-miR-5352

Gene ID	*C. elegans* homologue	L3she	L3ex	L4	Male	Female	Gut
g12342		598	478	428	435	553	967
g2972	Q9XW73	358	289	260	280	404	500
g10736		149	5	8	179	8	3
g7175	Q21146	293	786	770	587	903	597
g3652	O44399	20	183	138	319	140	36



3.2.9.2 3' UTR targets of Hco-miR-5352 predicted by PITA, miRanda and RNAhybrid

For Hco-miR-5352, a total of 4,460 targets were predicted. Of these, 112 UTRs (105 genes) were predicted by all three programmes with predicted overlapping sites, with slightly higher number of targets if considering two programmes (Figure 3.30). None of the targets predicted by two or more programmes had more than one binding site.

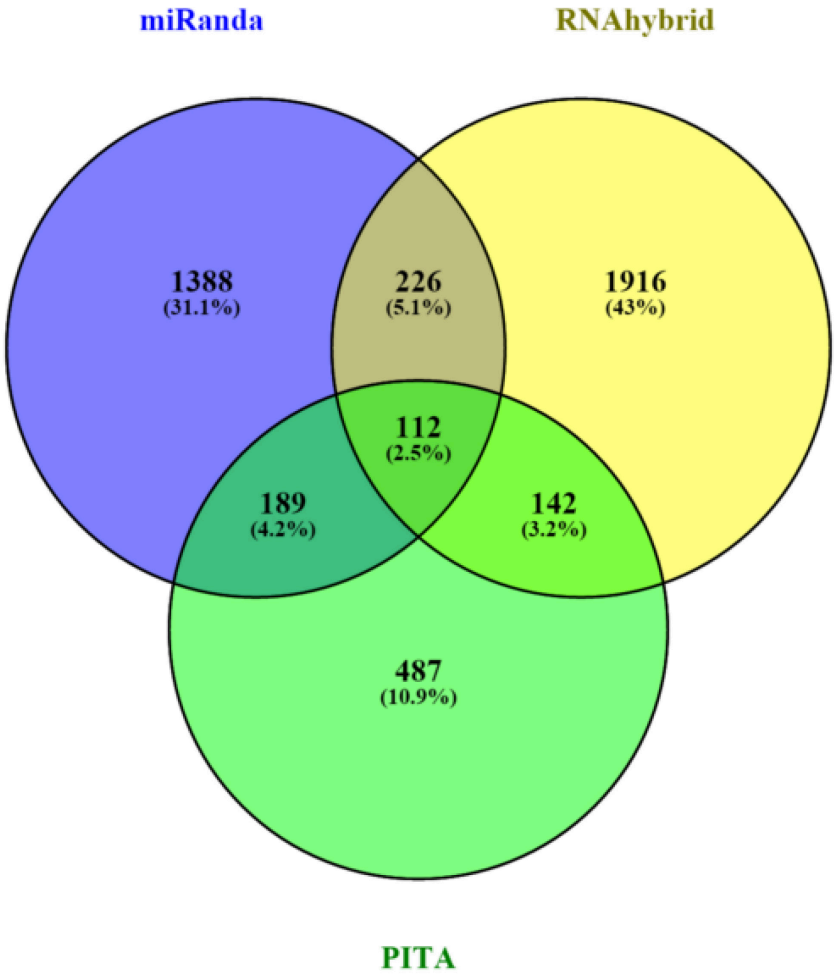


Figure 3.30 PITA, miRanda and RNAhybrid were used to identify Hco-miR-5352 binding sites in the *H. contortus* 3' UTR database. UTRs that are predicted targets from all three programmes were checked to determine if the predicted binding sites between the three programmes overlap. 112 UTRs (105 unique genes) had overlapping binding sites.

Transcriptomic data for the 105 *H. contortus* genes that are predicted targets of Hco-miR-5352 using all three programmes were obtained and clustered based on expression across *H. contortus* life cycle stages. Table 3.7 shows the 20 genes

from the hierarchical clustering that have an inverse expression profile to Hco-miR-5352 (correlation coefficient of less than -0.4). 13 *C. elegans* homologues were identified by BLAST, of which eight homologues are uncharacterised proteins. Gene ontology analysis did not identify any terms that were significantly enriched.

Table 3.7 Hierarchical clustering of transcriptome data of the 105 *H. contortus* genes with binding sites predicted by all three programmes focusing on the 20 genes that show an inverse expression profile to that of Hco-miR-5352, of which only 13 have *C. elegans* homologues.

Gene ID	*C. elegans* homologues	L3she	L3ex	L4	Male	Female	Gut
g11158		177	148	50	100	35	84
g17463		5	4	2	4	2	3
g17048	A0T4E6	126	105	27	88	1	6
g2608	O45200	222	225	124	184	124	125
g12386		2528	2541	17	205	279	22
g17058	G5EGB4	727	750	21	5	0	1
g1421	Q21603	180	137	17	0	0	10
g10222	Q22900	498	459	122	114	11	33
g4597	Q20101	54	53	14	5	5	2
g15439		180	186	48	43	25	53
g9447	Q18214	84	90	25	23	3	37
g20615	O62169	156	199	14	2	19	5
g2330		55	85	10	2	1	4
g19077	Q4W5P0	10223	10908	5667	1808	2867	2121
g19497	G5EDQ6	2	2	1	1	0	0
g9030		48	45	18	13	0	0
g10910	G5ECC4	69	87	70	15	2	24
g21379		15	26	18	2	2	1
g13390	A8W765	136	110	107	54	14	5
g12725	Q21274	832	761	926	439	327	281

The expression level for some of the genes listed in Table 3.7 is relatively low (for example g17463, g19497 and g4015). g19077 (Q4W5P0 in *C. elegans*, homologous to Drosophila Squid protein) had the highest expression of all genes in every stage. The function of Q4W5P0 in *C. elegans* is unknown, however, in Drosophila, Squid is required for anteroposterior polarity in the Drosophila oocyte (Steinhauer and Kalderon, 2005).

3.3 Discussion

This chapter presents a detailed analysis of the miR-5352 cluster in *H. contortus*. miRNAs have been identified in a range of parasitic nematodes. Analysis of the sequence shows that while a number of the miRNAs are conserved between different species, the majority of the miRNAs are unique to the nematodes (Wang *et al.*, 2011; Winter *et al.*, 2012). In *H. contortus*, Winter *et al.*, identified 54 of the 192 *H. contortus* miRNAs as unique when compared to miRBase miRNAs (Winter *et al.*, 2012), suggesting that the miRNAs may have a conserved function. It is likely that the number of conserved miRNAs will rise when these miRNAs are compared to more closely related nematodes such as *T. circumcincta*. It is interesting to note, that in *Ascaris suum,* 54 out of the 97 identified miRNAs are conserved (Wang *et al.*, 2011). Whilst the percentage of conserved miRNAs is much higher, *H. contortus* has nearly double the number of identified miRNAs. In both studies, miRNAs were identified that were highly expressed in particular stages of the parasite life cycle suggesting that they may have important functions in development or interaction with the environment. In *C. elegans*, lin-4 and let-7 are both temporarily regulated during development and loss of function mutations resulted in retarded development (Lee *et al.*, 1993; Reinhart *et al.*, 2000). Whilst not all miRNAs are essential, deleting all members of the miR-35 or miR-51 families in *C. elegans* resulted in embryonic death (Alvarez-Saavedra and Horvitz, 2010).

The initial analysis performed by Winter *et al.,* (2012) identified Hco-miR-5352 as an interesting miRNA. While it could be identified in *H. contortus* and had been sequenced from *A. suum* (Wang *et al.*, 2011a), it was not conserved in the closely related free-living nematode *C. elegans*. An important aspect of the current work was to determine how widely conserved this miRNA cluster is within parasitic nematodes. This was aided by the increased availability of genome data, with 99 genomes representing 90 parasitic helminths searchable in WormBase ParaSite, thus providing a clearer picture of gene conservation. Such comparative genomic studies are an important step in helping determine gene function. In the four years since the initial *H. contortus* miRNA analysis, there has been a huge increase in the miRNA data available. miRBase itself had 14,197 miRNA sequences across 133 species in 2010 (release 15) and has nearly doubled

to 28,645 in 223 species in 2014 (release 21). However, there is still little annotated data on miRNAs in parasitic nematodes and it was therefore necessary to identify and annotate these manually.

The Hco-miR-5352 cluster was identified in ten different parasitic nematodes. For all, the order of the miRNAs and their orientation are completely conserved. Nine of the species identified are clade V parasitic nematodes but all have different hosts and not all are blood feeders. Interestingly, the miR-5352 cluster is not present in all clade V nematodes. Based on current data, it is absent in the free-living nematode *C. elegans*, the necromenic nematode *Pristionchus pacificus* and the lungworm *Angiostrongylus*. Interestingly, *Angiostrongylus cantonensis* is a lungworm of rats which are infected by ingesting L3 larvae, which pass through the intestines and penetrate the intestinal wall and migrate to the brain then to the lungs (Cowie, 2013). In contrast to the other nematodes in which the miR-5352 cluster can be identified, adult worms of *Angiostrongylus* do not reside in a lumenal site, but within the pulmonary arteries, suggesting that the miR-5352 cluster may have a function related to the location of the parasite in the host.

The miRNA cluster is also not restricted to clade V species as the clade III intestinal nematode *Ascaris* also contains a Hco-miR-5352 homologues. Therefore, with the exception of *D. viviparus*, a common feature is that the adult parasites all reside within the gastro-intestinal tract (Table 8.1). In this study, a cluster containing all four miRNAs was also identified in *D. viviparus*, the adult worms of which reside within the lungs. The pairwise similarity of the mature miRNA sequences to the *D. viviparus* contig are all over 87% and the sequence is able to fold into miRNA hairpins, suggesting that *D. viviparus* contains the full miR-5352 cluster. The impact of being in the reverse orientation is unknown. The mature miRNA sequence and the minor product (3p and 5p) are also extremely well conserved. Whilst the precursor sequence is less well conserved, RNA folding algorithms are able to produce a hairpin structure suggesting that these miRNA sequences may be functional (Figure 8.1).

In *Ascaris*, the cluster is shorter, encoding only miR-5352 and miR-43, and missing miR-61 and miR-5895 at the 5’ and 3' ends of the cluster, respectively. It

is possible that *Ascaris* has some redundancy and does not require miR-61 and miR-5895 as part of the cluster, as other related miRNAs may perform their function. To examine this, the stringency settings in ParaSite (sensitivity to Distant Homologies) were reduced but did not yield any sequences related to miR-61 or miR-5895 in *Ascaris*. BLAST searches identified only a number of 15bp or shorter alignments. These sequences do not form the typical hairpin loop structure associated with miRNAs and are unlikely to be miRNAs. Winter *et al.*, (2012) identified Hco-miR-5895 as belonging to the miR-55 family. It is possible that another member of the miR-55 family may be substituting for miR-5895. To investigate the presence of another member of the miR-55 family, a BLAST search was carried out in the *Ascaris suum* and *A. lumbricoides* genome. No significant hits were obtained. The reason for the absence of miR-61 and miR-5895 from the cluster is unclear.

A role for the miRNAs of the cluster in parasitism is suggested by their presence in parasite ES products and by their expression profile, as shown by microarray analysis. All four miRNAs of the cluster are most abundant in the *H. contortus* adult parasitic stages, with higher expression in adult females compared to male worms. Without expression profile data for the other parasites, it is unknown if this temporal pattern is conserved as well. The higher expression of Hco-miR-5352 in adult females could be due to its presence in eggs as determined by RT-PCR. As *H. contortus* females can produce up to 10,000 eggs daily, this could account for a significant proportion of the Hco-miR-5352 signal.

If the miRNAs of the cluster are involved in parasitism, they could either regulate genes expressed by the parasite itself or target host gene expression. The genomic location of the microRNA may give an indication to its function. It is predicted that as many as 40% of miRNAs lie within the introns of genes and the expression of these microRNAs often correlates with the transcription of the gene (Baskerville and Bartel, 2005; Rodriguez *et al.*, 2004). For example, in human alveolar basal epithelial cells, miR-301 is located in the first intron of *spindle and kinetochore-associated protein 2 (ska2)* and was shown to target MEOX2 to affect the ERK/CREB pathway, which in turn regulates the expression of *ska2* (Cao *et al.*, 2010).

H. contortus RNA transcript data shows that the scaffold on which the Hco-miR-5352 cluster is located on contains four genes but the cluster does not appear to be intronic. Examination of the expression profile of nearby genes using RNA-Seq data shows that all four predicted mRNAs have the highest expression in the L3 and L3 activated stages with decreasing levels in the adult stages. This expression profile is the inverse of the miRNA cluster, suggesting that these mRNAs may be targets. To determine if any of the miRNAs have binding sites for the four mRNAs, the miRNA target prediction program PITA was used to search for potential binding sites in the 3′ UTR or the rest of the mRNA sequence. Binding sites for all four members of the Hco-miR-5352 cluster were identified on the MGAT2 sequence, and although these sites were not present in the 3′ UTR, the miR-5352 cluster could target MGAT2.

Global analysis of potential regulation of *H. contortus* genes by all *H. contortus* miRNAs was investigated bioinformatically using a single program (PITA) and by using a combination of three programmes (PITA, miRanda and RNAhybrid). Initial results from these three programmes identified a large number of potential targets. Due to the incomplete knowledge regarding the interaction between mRNAs and miRNAs, it is difficult to conclusively state that a miRNA targets a given mRNA. Based on current knowledge, predicted targets were narrowed down by looking for UTRs with multiple miRNA binding sites (Doench *et al.,* 2003), targets in common between the three programmes, or between the four miRNAs of the Hco-miR-5352 cluster, looking at pathways and IPA networks that contain the targets or at gene ontology to determine if the miRNA targets genes with similar functions. Combining the data with *H. contortus* transcriptomic data gives a good indication as to whether the expression profile of the genes follows the inverse expression of the miRNAs (Laing *et al.,* 2013).

When identifying target with multiple miR-5352 binding sites, the gene g18228 was found. The expression profile of this gene was inversely related to that of miR-5352, but was extremely low throughout the life cycle. Two additional predicted gene targets showed a similar expression pattern to g18228. These were g6664 and g3824 which show homology to the *C. elegans* UNC-18. The *C. elegans* UNC-18 enables the docking of vesicles to synaptic regions and promotes vesicle exocytosis (McEwen and Kaplan, 2008). The two UNC-18 targets are

particularly interesting due to their roles in the docking of vesicles as miRNAs in the excretory-secretory products of *H. polygyrus* are located in extracellular vesicles/exosomes (Buck *et al.*, 2014).

However, this approach failed to identify any convincing *H. contortus* targets of the miR-5352 cluster. In part, this may be due to the lack of annotation of the *H. contortus* genome which hampers attempts to identify miRNA target genes. Attempting to address this using *C. elegans* homologues was only partially successful. Not every *H. contortus* gene has a *C. elegans* homologue meaning that the homologue identification removed many genes from the final list. While *C. elegans* is the best annotated and most closely related organism to *H. contortus*, there are still *C. elegans* genes that have an unknown function. Since *C. elegans* is much better annotated and a lot is known regarding its genes, bioinformatics prediction could work for genes or pathways that are highly conserved between *H. contortus* and *C. elegans* and possibly the conserved *H. contortus* miRNAs as well.

It is possible that the Hco-miR-5352 cluster targets host genes, perhaps those involved in the immune response against helminths. It is well documented that nematode parasites are able to modify the host immune response. Research to date on parasite immunomodulatory molecules has focused largely on proteins. One of the main sources for these proteins is the parasite ES products, which are thought to represent molecules also released *in vivo* and potentially important in host-parasite interactions. Importantly, Buck *et al.,* (2014) identified the presence of miRNAs in the secretions of *H. polgyrus* and *Litomosoides sigmodontis*. Here it was shown that miR-5352 can be detected in adult *H. contortus* ES. Detailed characterisation of miRNAs in exosomes and in ES supernatant of *H. contortus* will be presented in Chapter 4.

4 Characterisation of microRNAs in *H. contortus* ES material

4.1 Introduction.

As demonstrated in Chapter 3, RNA could be successfully extracted from *H. contortus* adult excretory secretory (ES) products. Further, it was possible to amplify Hco-miR-5352 sequences from the RNA using RT-PCR. The presence of *H. contortus*-specific miRNAs in the ES material and their expression in the adult worms, suggests that miRNAs derived from adult *H. contortus* may function outside of the worm. The recent studies of Buck *et al.* (2014) showed convincing evidence that parasitic nematode miRNAs are released into the ES and are able to alter host gene expression.

There is increasing evidence that miRNAs from a wide range of sources are able to enter and interact with host cells. For example, *Onchocerca* miRNAs have been identified in plasma or nodule fluid of animals infected with filarial nematodes (Quintana *et al.*, 2015). Even miRNAs from the consumption of plants are reported in serum and tissues of mice (Liang *et al.*, 2014; Zhang *et al.*, 2012), although, subsequent studies have shown that transfer of miRNAs orally is not a prevalent event nor a robust method of miRNA delivery (Snow *et al.*, 2013).

In mammalian systems, miRNAs in body fluids have either been found associated with proteins (such as Argonaute) or present in extracellular vesicles (EV) (possibly exosomes) (Turchinovich *et al.*, 2011, 2013). Buck *et al.,* identified 263 unique miRNAs from the secretions of *H. polygyrus*. Transmission electron microscopy (TEM) identified vesicle-like structures between 50 and 100 nm in diameter and proteomic analysis identified an enrichment of EV-associated proteins such as Alix, which is associated with EV biogenesis (Baietti *et al.*, 2012; Buck *et al.*, 2014; Théry *et al.*, 2001). These nematode EV were able to suppress T–helper 2-type responses in mice and Dusp1, a regulator of MAPK signalling, was found to be repressed by nematode miRNAs in a luciferase assay (Buck *et al.*, 2014).

While the results presented in Chapter 3 demonstrated that Hco-miR-5352 sequences could be amplified by RT-PCR from adult worm ES, without an appropriate normaliser it was difficult to tell if this miRNA is relatively highly abundant. There are an additional 195 known *H. contortus* miRNAs and investigating their presence or absence in the ES might give a greater insight into the role of these miRNAs and their mechanism of secretion. It is also important to determine whether the miRNA profile of the ES is similar to that in the whole adult worm and whether the presence of miRNAs might simply be due to the death of *H. contortus* worms during *in vitro* culture. Unlike *H. polygyrus*, *H. contortus* does not survive well in serum-free medium and as a result, ES samples from *H. contortus* were all taken from the first 72 hours of culture.

In this chapter, RNA-Seq was used to investigate the total profile of small RNAs present in the ES products of *H. contortus*. This method is able to detect all RNAs from the sample including non-coding RNAs. With RNA-Seq, any novel transcripts not identified by Winter *et al.,* (2012) can also be detected. RNA-Seq can quantify discrete read counts, which offers a broader dynamic range and easier detection of low abundance miRNAs than microarray analysis. Most importantly, the RNA-Seq can accommodate single nucleotide variants, indels and other changes between the accepted sequence and the sample. From Chapter 3, there are slight variations between some of the original miRNA sequences reported in miRBase and the sequence from the most up to date version of the *H. contortus* genome. Whilst the changes are not present in the mature sequence, the ability of RNA-Seq to detect these differences will give an indication into the possible sequence variations, if any, that occur.

4.2 Results

4.2.1 Analysis of a small RNA library prepared from *H. contortus* ES

Data presented in Chapter 3 demonstrated using qPCR that Hco-miR-5352 could be detected in adult worm ES material. In order to investigate which other miRNAs are also present in the *H. contortus* ES, a small RNA library was created (as described in Chapter 2.4.3) and sequenced at the University of Glasgow Polyomics Facility. Over 15 million reads were sequenced and were mapped to previously identified *H. contortus* RNA sequences (Winter *et al.*, 2012)(Figure 4.1). Alignments were accepted if the read was longer than 8 base pairs and had no more than one mismatch.

The initial study by Winter *et al.*, (2012) identified 462 potential miRNA sequences from adult whole worm and L3 stages of *H. contortus*, of which 211 were submitted to miRBase and 192 accepted (Winter *et al.*, 2012). This number has been revised in the current version of miRBase (release 21, http://www.miRBase.org/ , accessed 10/01/2015) and now consists of 195 miRNAs. Of the 195 accepted *H. contortus* microRNAs, 85 were found in the ES small RNA library. In addition, a further 83 sequences were identified that mapped to miRNAs that had been rejected from the miRBase list for a variety of reasons including: not folding according to miRBase programs or the Computational Identification of microRNA program (CID-miRNA) (Tyagi *et al.*, 2008), similarity to existing rRNA or tRNA sequences or low counts (<10 reads) in the adult library.

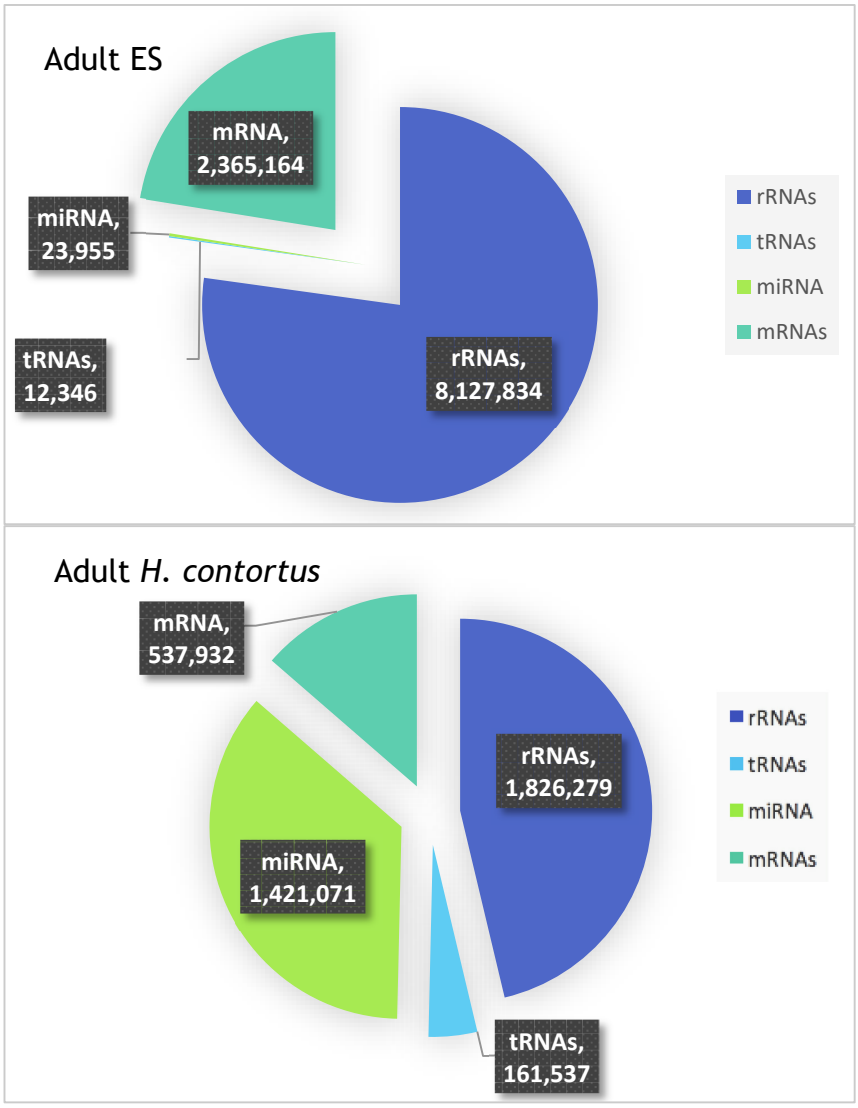


Figure 4.1 Reads from the *H. contortus* adult ES small RNA library and *H. contortus* adult whole worm small RNA library (Winter *et al*, 2012), were mapped to *H. contortus* RNA sequences. Numbers indicate number of unique reads for each RNA category

All the reads from both the adult whole worm and the ES small RNA libraries were mapped to all 462 potential miRNA sequences as well as tRNA, rRNA and mRNA sequences obtained from Winter *et al.*, (2012). In the small RNA library derived from ES material, 77% of the total reads map to ribosomal RNA sequences, with mRNAs being the second most abundant class with 22% of the reads (Figure 4.1). Transfer RNAs (tRNAs) comprise the smallest group (0.1%) and miRNA reads (0.2%) represent approximately double that of the tRNAs. In comparison, in the adult whole worm small RNA library, rRNAs make up 32% of

the reads, followed by miRNAs (14%). Reads from tRNAs are the smallest group and mRNAs (14%) are roughly three times greater than tRNAs (4%).

Due to the difference in total read numbers (Figure 4.1), the miRNAs were ranked by the number of reads. Table 4.1 shows the ten most abundant miRNA sequences for each library, including miRNA sequences that were not accepted by miRBase. In the adult ES library, five of the ten abundant sequences are not accepted miRNAs; in the adult whole worm library, only Hco0325 is not an accepted miRNA. Hco0129, 0135 and 0134 were rejected as miRNAs as they had less than 10 reads each in the adult whole worm library. However, in the ES library, their abundance is fourth, fifth and sixth respectively and all show read numbers >500.

Three of the entries in the adult ES library, Hco-miR-61, Hco-miR-5895 and Hco0325 appear in the top 10 list for the adult whole worm library. Hco-miR-87b, on the other hand, is the tenth most abundant miRNA in the adult whole worm library but does not appear at all in the ES library. The large variation in the relative abundance of these miRNAs suggests that specific miRNAs may be secreted in the ES. Table 4.2 shows a comparison of the rankings showing only the miRBase-accepted miRNAs. The miRNAs of the Hco-miR-5352 cluster are very highly abundant in both libraries and the rankings are somewhat conserved with expression of miR-5352 being the lowest of the four, then miR-43. In the whole worm adult library, miR-61 expression is higher than that of miR-5895, but in the ES library, the reverse is true.

Table 4.1 A) Top 10 miRNAs, ranked by total reads, found in the adult ES small RNA library. B) Top 10 miRNAs by total reads in the adult whole worm library (Winter *et al.*, 2012)**. Both tables include sequences that were discarded from the original paper (i.e. have no miRBase name).**

A) Adult ES library

	Ranking	Name	Reads	Rank in Adult Library
Hco0322	1	hco-mir-5960	16743	55
Hco0325	2		2401	9
Hco0023	3	hco-mir-5895	1093	4
Hco0129	4		738	347
Hco0135	5		700	354
Hco0134	6		503	328
Hco0324	7		433	27
Hco0021	8	hco-mir-61	406	1
Hco0025	9	hco-mir-45	403	22
Hco0089	10	hco-mir-228	393	13

B) Mixed sex Adults

	Ranking	miRBase name	Adult Whole Worm Reads	Rank in ES library
Hco0021	1	hco-mir-61	543826	8
Hco0043	2	hco-mir-71	54876	24
Hco0371	3	hco-mir-5983	47494	53
Hco0023	4	hco-mir-5895	25614	3
Hco0059	5	hco-mir-84	19745	102
Hco0001	6	hco-mir-5884	18952	17
Hco0022	7	hco-mir-43	18561	11
Hco0079	8	hco-mir-5352	15943	12
Hco0325	9		12689	2
Hco0035	10	hco-mir-87b	12230	#N/A

Table 4.2 Comparison of the top 10 miRBase miRNAs in the adult ES library and their corresponding ranking in the adult whole worm library. The miRNAs are ranked by total normalised reads.

MicroRNA	Rank in Adult ES RNA library	Rank in Adult whole worm RNA library
hco-mir-5960	1	48
hco-mir-5895	2	4
hco-mir-61	3	1
hco-mir-45	4	21
hco-mir-228	5	12
hco-mir-43	6	7
hco-mir-5352	7	8
hco-mir-5938	8	139
hco-mir-40b	9	31
hco-mir-5885c	10	15

4.2.2 Mapping reads to the miRNA precursor sequences

In the initial analysis above, the reads obtained from the small RNA ES library were mapped to the original *H. contortus* miRNA precursor sequences. This was done to allow for slight variations in the sequence length and to allow for comparisons with the Adult whole worm data obtained from Winter *et al.*, (2012). It is therefore possible that the sequences that were mapped do not correlate with the mature miRNA sequence. Analysis of the library shows that the data obtained from the sequencing were of high quality and the high phred score (Figure 8.3 and Figure 8.4) indicates that there was a low probability of errors. From Table 4.1, Hco0129, Hco0135, and Hco0134 were all significantly enriched in the ES library, however, these sequences were originally rejected as miRNAs due to their low read number (<10 reads) in the adult whole worm library. This suggests that there are differences between the adult whole worm and adult ES libraries and that there are sequences that could still be miRNAs even though their expression in adults is low.

Mapping of the reads from the library to the miRNA precursor sequences could result in reads being assigned to a miRNA that is dissimilar to the mature miRNA sequence. To further investigate where the reads mapped the precursor or mature miRNA sequence, ES library reads were aligned to the precursor sequence. Figure 8.5 shows the alignment of the miRBase-accepted miRNA, Hco-miR-5895. Figure 8.6 to Figure 8.10 show the alignment of the 5 'unofficial' miRNAs in the top 10 of the ES small RNA library (Hco0129, Hco0134, Hco0135, Hco0324 and Hco0325 respectively). Identical read sequences have been grouped together and the number of times they appear in the library is denoted by the number after "Reads_".

The majority of library reads that mapped to Hco0129 and Hco0135 mapped very close to their predicted mature sequence. A small number of the Hco0134 reads mapped to the mature sequence, the remainder map 30 bases downstream of the mature sequence. This second group of miRNAs comprised the bulk of the reads and were very well conserved; however, searching miRBase for this sequence did not yield significant hits in any species.

Hco0324 was rejected originally because of similarity to a rRNA. BLAST searches using the precursor sequence showed a 99% sequence identity to an *Oesophagostomum venulosum* 28S rRNA and a 98% sequence identity to a *H. contortus* 28s rRNA. Roughly a third of the total reads for Hco0324 map to the predicted mature sequence. The rest of the sequences have less than 10 reads each and are highly variable in their length. Unlike Hco-miR-5895, the reads for Hco0324 map across the whole precursor. Whilst 37% of the reads do map close to the predicted mature sequence, the lower specificity suggests that this sequence may not be a miRNA. If the reads that are similar to the predicted mature Hco0324 sequence are considered, the miRNA drops from position 7 in the ES library to 15th. This alignment profile is very similar to the alignment seen in ribosomal RNA matches, suggesting that whilst Hco0324 does fold into a stem-loop structure, it may be a rRNA and not a miRNA.

Hco0325 was rejected due to similarity to a tRNA and Figure 8.10 shows that only three reads in the whole worm library mapped to the predicted mature region. The rest of the sequences mapped to the 3′ end and whilst they were highly conserved, were not similar to any miRNAs in miRBase. The high degree of conservation in the 3′ sequences suggests that the predominant strand in Hco0134 and Hco0325 is in the 3′ end and not the 5′.

From the alignment data, there is a possibility that a number of the miRNAs that were rejected from the original list could be true miRNAs. Hco0129 and Hco0135 seem to be extremely promising miRNA candidates, as does the 3′ section of Hco0134 and Hco0325 whereas Hco0324 is unlikely to be a miRNA.

4.2.3 *H. contortus* ES material contains extracellular vesicles (EV)

Analysis of secreted miRNAs from a range mammalian species has demonstrated that they are present either associated with specific proteins or are contained within extracellular vesicles. Analysis of ES material from *H. polygyrus* found that the nematode produces vesicles containing miRNAs and homologues of mammalian proteins associated with EV biogenesis (Buck *et al.*, 2014). *H. contortus* ES, ultracentrifuged and processed for transmission electron microscopy, was shown to contain a large number of round circular structures

very similar to those identified in *H. polygyrus* ES (Figure 4.2), suggesting that *H. contortus* ES also contain extracellular vesicles.

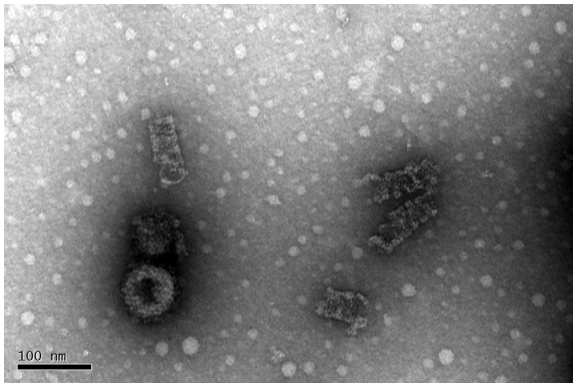


Figure 4.2 Transmission electron microscopy of ultracentrifugation pellet (0.7 µg/µl total protein) from adult *H. contortus* ES products. Scale indicates 100 nm. Blue arrows show EV-like structures.

4.2.4 Comparative analysis of small RNA libraries prepared from EV from *H. contortus* ES and from the EV-free supernatant

To investigate the RNA content of the EV and the EV-free supernatant, separate small RNA libraries were created from the ES products of adult and L4 stages of *H. contortus* (as described in Chapter 2.4.3) and sequenced by LC Sciences. The number of total reads obtained from these samples were lower than that from the total ES RNA library, although both the L4 and adult EV libraries contained around 1.5 times more raw reads than the corresponding EV-free supernatant libraries (Table 4.3).

Table 4.3 Number of raw reads obtained from the RNA libraries.

Adult EV- free supernatant	L4 EV-free supernatant	Adult EV	L4 EV
8,771,077	7,909,495	13,688,137	10,909,509

The reads were then mapped to *Haemonchus contortus*, *Caenorhabditis elegans*, *Caenorhabditis brenneri*, *Caenorhabditis briggsae*, *Ascaris suum*, *Pristionchus pacificus*, *Brugia malayi*, *Strongyloides ratti* and *Panagrellus redivivus* miRNAs from miRBase database version 21 (released June 2014). The mapping was done by LC Sciences using a proprietary script. Raw reads were aligned to a list of sequences obtained from miRBase as well as the *H. contortus* genome (h

contortus.PRJEB506.WS248.genomic.fa.gz released on 04/09/2015). The script aligns reads to the mature miRNA sequences and allows for up to three mismatches in the 3' end of the alignment.

In total, the five libraries (including the previously described adult total ES small RNA library) contained over 56 million reads, with only 35 million reads (62%) being mappable (Figure 4.3). These 35 million reads are further analysed in Figure 4.4; the majority of reads do not map to miRNA sequences from miRBase; only 254,000 reads map to mature miRNA sequences from miRBase. A further 7 million reads map to positions on the *H. contortus* genome. Interestingly, ~195,000 reads (224 unique sequences) map to locations on the genome that can fold into a hairpin suggesting that there may be novel miRNA sequences that are present in the most recent version of the genome that were not detected in the older versions. The rest of the reads map to the genome but do not form hairpins. These sequences are most likely other RNA species that have not been annotated.

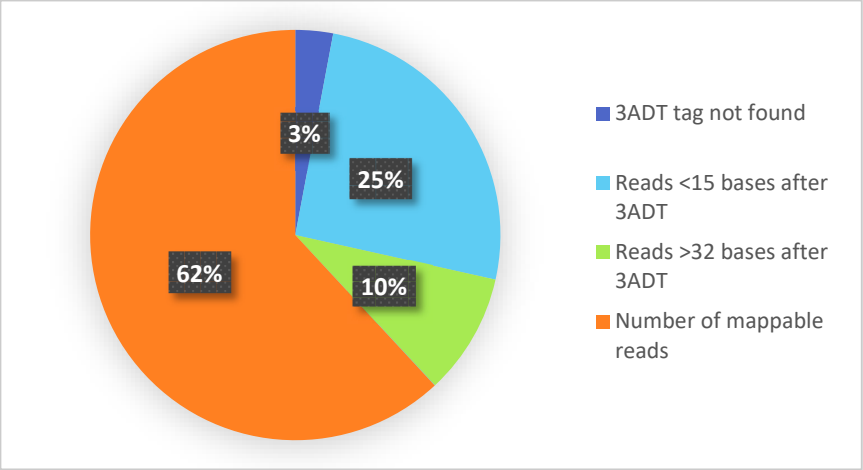


Figure 4.3 Pie chart showing the percentage of reads filtered from the five separate small RNA ES libraries based on the presence of the 3' adapter (3ADT).

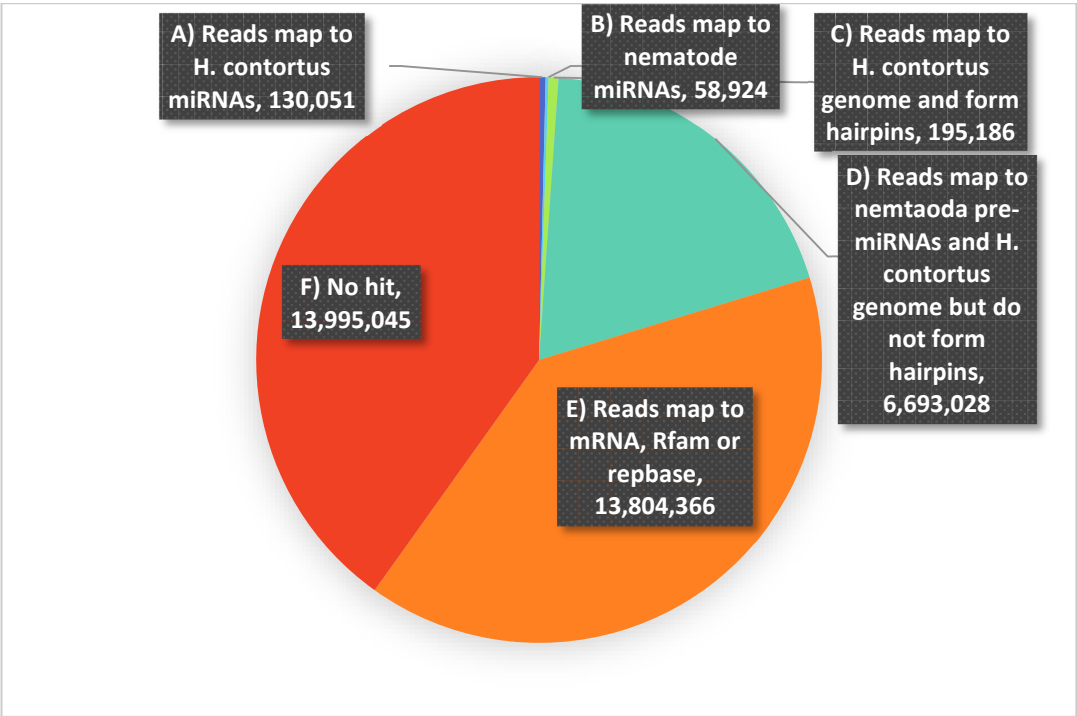


Figure 4.4 Pie chart showing the statistics of mappable reads (Figure 4.3) from all five libraries. A=Reads map to *H. contortus* miRNAs, B=Reads map to nematode miRNAs, C=Reads map to *H. contortus* genome and form hairpins, D=Reads map to nematode pre-miRNAs and *H. contortus* genome but do not form hairpins, E=Reads map to mRNA, Rfam (repetitive DNA elements) or repbase (non-coding RNA database), F=Does not map to any known sequences.

Table 4.4 shows the normalised reads for the miRBase miRNAs in the EV and EV-free supernatant libraries and the adult total ES library. This list also includes alternate strand miRNAs such as Hco-miR-5960-3p, which, while not specifically stated in the miRBase database, can be inferred from the precursor sequence.

Table 4.4 microRNA reads in the normalised small RNA libraries. List of miRBase microRNAs identified in the small RNA libraries. microRNAs that had less than 10 reads in two or more of the libraries were discarded from the final list. Highlighted reads indicate those present in the Hco-miR-5352 cluster.

MicroRNA	Adult EV-depleted	Adult EV	L4 EV-depleted	L4 EV	Adult Total ES
Hco-miR-5960-5p	4062	637	17160	1336	25958
Hco-miR-5960-p3	405	115	63	186	4734
Hco-miR-5895-5p	2267	2289	78	82	3047
Hco-miR-5939-3p	166	36	7	126	2548
Hco-miR-45-3p	1160	1681	1567	179	1125
Hco-miR-61-3p	924	422	3	0	1114
Hco-miR-228-5p	163	48	66	258	1067
Hco-miR-43-3p	659	157	2	3	829
asu-miR-100a-5p	5212	3181	2463	55492	683
Hco-miR-5352-3p	720	462	2428	31	492
Hco-miR-40b-3p	541	531	3	0	322
Hco-miR-5884-5p	25	122	17	113	243
Hco-miR-5885a-3p	16	664	260	4347	205
Hco-miR-9551-3p	308	178	18	0	160
Hco-miR-43-p5	795	457	12	50	152
Hco-miR-5885b-3p	28	672	113	1058	150
Hco-miR-63a-3p	393	187	81	104	144
Hco-miR-71-5p	87	282	355	317	126
cel-let-7-5p	34	596	1041	13532	107
prd-miR-7911c-5p	17	38	13	195	92
Hco-miR-5885c-3p	12	378	114	1356	82
Hco-miR-5899-3p	1678	1991	5606	4565	50
Hco-miR-5960-p5	8	1	47	25	49
Hco-miR-63b-3p	461	307	113	129	42
Hco-miR-87a-3p	57	20	160	25	34
Hco-miR-5908-3p	20	936	111	1009	31
cel-miR-50-5p	37	96	135	198	21
Hco-miR-993-3p	6	9	20	25	21
cel-mir-5592-1-p3	7	26	4	120	18
Hco-miR-2-3p	30	13	66	123	16
Hco-miR-236-3p	3	269	31	503	16
Hco-miR-5976-5p	11	144	40	28	16
Hco-miR-250-3p	30	14	6	3	14
Hco-lin-4-5p	62	614	4092	11107	8
Hco-miR-60-3p	2	35	16	53	8
Hco-miR-83-3p	19	346	341	2341	5
Hco-miR-84a-5p	3	250	319	66	5
Hco-miR-259-5p	1	129	53	327	3
bma-mir-36b-p5	0	12	12	31	0
cbn-mir-64f-p5	0	20	13	72	0
prd-miR-7911a-5p	0	24	28	187	0

To determine if there are major differences in abundance between the five ES libraries, scatterplots were made plotting normalised reads of one library against the other. A total of 890 miRNAs were successfully mapped across the five libraries, only 40 miRNAs were both recognised by miRBase and had more than 10 reads in three or more of the libraries (Figure 4.5 and Figure 4.6) The red line in the graphs show where the read numbers are equal between the two libraries. The further the miRNA dot is from the line, the greater the difference in the read counts. Members of the miR-5352 cluster were identified and scatterplot analysis showed that Hco-miR-5895-5p was present at similar levels in both the adult EV and adult ES supernatant libraries (Figure 4.5). Other miRNAs showing similar abundance in the adult ES supernatant and EV include Hco-miR-5899-3p, Hco-miR-45-3p and Hco-miR-40b-3p. There are also a number of miRNAs that are abundant in the EV reads only with a negligible read count in the supernatant library. These are Hco-miR-5885a-3p, Hco-miR-5885b-3p, cel-let-7-5p, Hco-miR-5885c-3p, Hco-miR-2-3p, Hco-miR-250-3p, Hco-lin-4-5p, Hco-miR-60-3p, Hco-miR-83-3p and Hco-miR-84a-5p and Hco-miR-87a-3p. There are fewer miRNAs that are predominantly found in the supernatant. These miRNAs such as Hco-miR-43-3p and Hco-miR-5960-5p, all have some expression in the EV although not as high as in the supernatant.

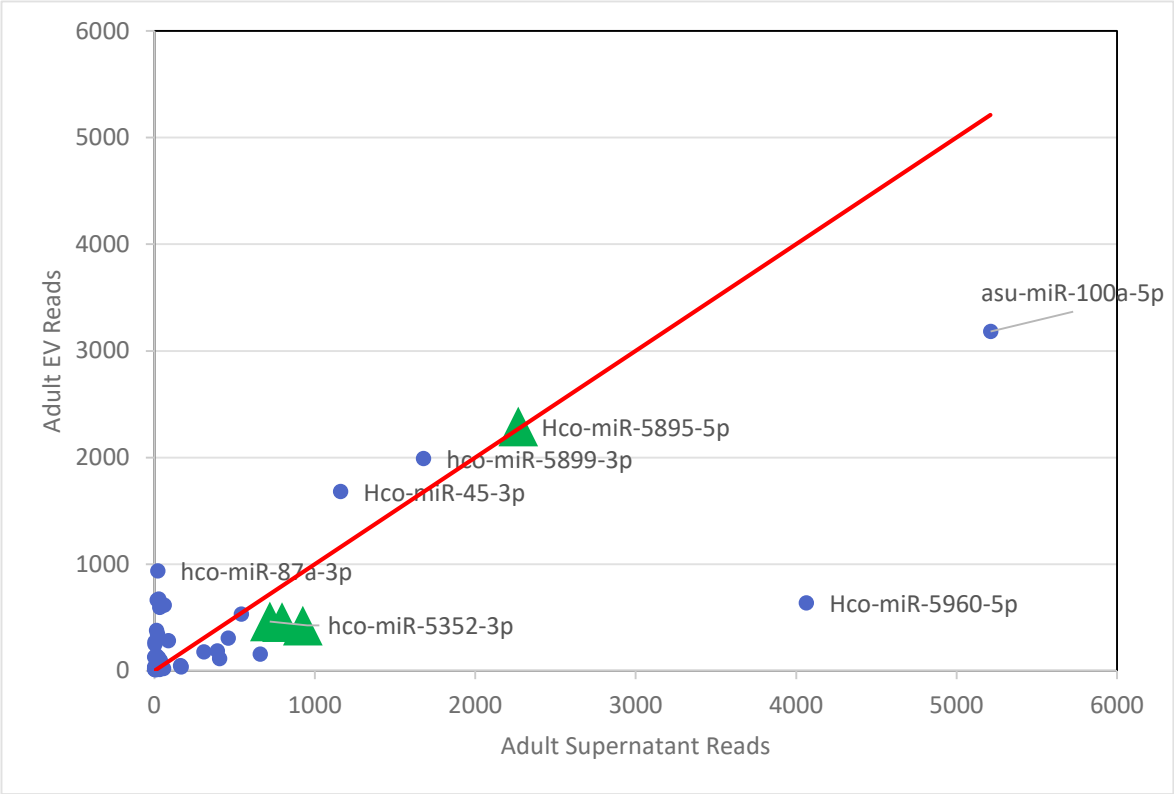


Figure 4.5 Graph comparing the normalised adult supernatant reads against the adult EV reads for individual miRNAs. The red line indicates that the number of reads are equal in each library. Green triangles indicate the members of the Hco-miR-5352 cluster.

Figure 4.6 Graph comparing the normalised L4 supernatant reads against the L4 EV reads for individual miRNAs. The red line indicates that the number of reads are equal in each library. Green triangles indicate the position of the Hco-miR-5352 cluster.

From Figure 4.6, the abundance of the Hco-miR-5352 cluster is extremely low, with only Hco-miR-5352-3p above the threshold and only in the L4 supernatant. Hco-miR-45-3p was also mostly identified in the supernatant. Unlike the adult libraries, there are far fewer miRNAs that have uniform abundance between the EV and EV-free supernatant with Hco-miR-5899-3p being the closest. There are a number of miRNAs that are predominantly found in the EV, such as Hco-miR-5885a-3p, Hco-miR-5885b-3p and Hco-miR-83-3p, all of which are also EV enriched in the adult stages. Hco-miR-5908-3p in the L4 EV is 9 times higher than the L4 supernatant however in the adult stages, this difference drops to 2.6 times.

From Figure 4.5 and Figure 4.6, some of the miRNAs show differences in their localisation and are more abundant in either the EV or the EV-depleted supernatant. There are, however, a number of miRNAs, including the Hco-miR-5352 cluster where the reads are split between the two libraries. Interestingly, Hco-miR-5895 has similar read numbers in both the EV libraries versus the corresponding EV-depleted libraries. The read numbers are also higher in the adult ES libraries compared to the L4 ES libraries which correlates with the miRNA microarray data from the whole worm libraries, except for Hco-miR-5352-3p which is much higher in the L4 EV-depleted library than any of the other libraries. The other members of the cluster all have slightly higher values in the EV-depleted libraries compared to the EV libraries, except for Hco-miR-43-p5 L4 libraries where the reverse is true. Hco-miR-61-3p and Hco-miR-43-3p both had read numbers in the L4 libraries below the threshold

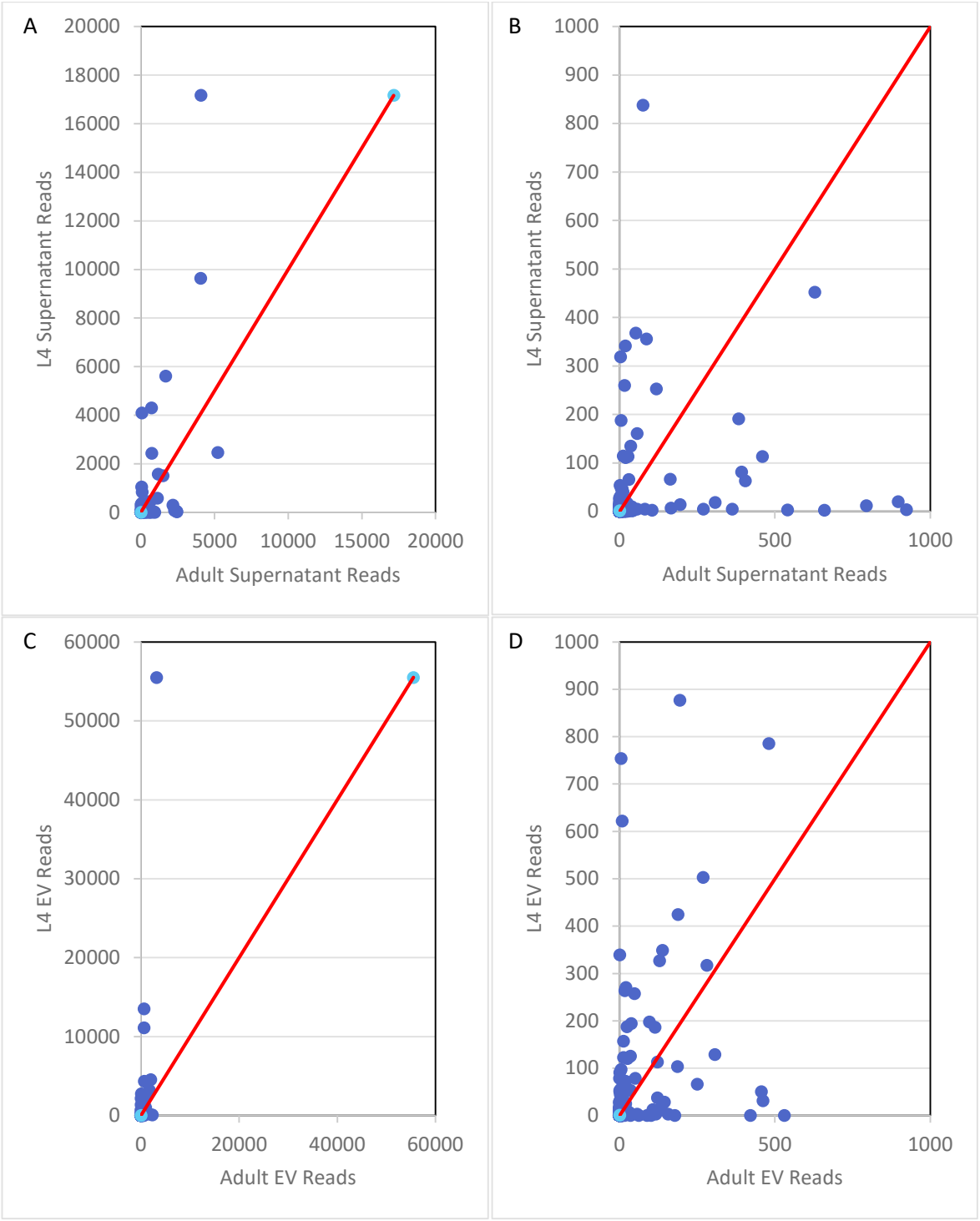


Figure 4.7 Comparison of miRNA expression in EV or the supernatant between life cycle stages. B and D are zoomed in sections of A and C respectively. Scale for both X and Y axis varies between the four graphs.

Figure 4.7 shows a comparison of the miRNA reads between the adult and L4 stages for the supernatant or EV small RNA libraries. The graphs show that the L4 stage had a much higher maximum read value compared to the adult stage with over 55,000 reads being recorded for the most abundant miRNA, Asu-miR-

100 in the L4 EV (note variation in scale in both axes). In contrast, the most abundant read for the adult EV library was 3,000, also from Asu-miR-100. The figure also shows that the miRNA profile for the EV and supernatant varies quite substantially between the two life stages with some miRNAs being predominantly expressed in one stage or the other.

4.2.5 Correlating small RNA library data with microarray data

Table 4.2 highlighted differences in the expression profile of miRNAs between the adult total ES and the adult whole worm libraries with regard to the read numbers. However, this table does not show the changes in the miRNA abundance across life stages. It was therefore of interest to determine if their expression in the ES correlates with their developmental expression in different life-cycle stages.

The data from the small RNA libraries were cross-referenced with the microarray data generated by Alan Winter and Neil Marks (unpublished data). Each library was sorted by maximum reads and the ten most abundant miRNAs were selected and compared to the microarray signal. A number of miRNAs did not have corresponding microarray data and were excluded from the analysis.

Figure 4.8 The expression profile across life cycles for the ten most abundant miRNAs in the adult total ES library. Data for different life-cycle stages and the gut tissue indicates the microarray signal.

Adult Total ES	L3	L3(act)	L4	Male	Female	Gut
Hco-miR-5960-5p	2725	2565	2924	12404	6021	20908
Hco-miR-5960-3p	9767	8033	23777	25282	13540	16081
Hco-miR-5895-5p	1350	1220	339	5683	13604	1111
Hco-miR-45-3p	3961	3736	5239	2024	5900	1494
Hco-miR-61-3p	279	287	67	10133	25106	619
Hco-miR-228-5p	3071	2705	696	466	253	21
Hco-miR-43-3p	1395	1319	82	10688	21793	1015
Hco-miR-5352-3p	27	24	37	1656	6805	130
Hco-miR-40b-3p	43	40	8	13	223	54
Hco-miR-5884-5p	26	34	151	170	11413	132

Figure 4.8 shows the ten miRNAs that were most highly expressed in the adult total ES library. Hco-miR-5960-5p and 3p are the most abundant miRNAs in the adult total ES library and are both highly expressed in the gut, suggesting that

the gut may be the source of these miRNAs. However, none of the other miRNAs that are abundant in the adult total ES are highly expressed in the gut. Of the remaining eight miRNAs, only Hco-miR-228-5p was highly expressed in the L3 stage and decreases in the L4 and adult stages. The remaining seven miRNAs have the highest expression in the adult female stages.

Figure 4.9 The expression profile across life cycle stages for the ten most abundant miRNAs in the A) Adult EV-depleted supernatant and B) Adult EV libraries. Numbers indicated the normalised read counts from miRNA microarray analysis of *H. contortus* worm stages.
A

Adult Supernatant	L3	L3(act)	L4	Male	Female	Gut
Hco-miR-5960-5p	2725	2565	2924	12404	6021	20908
Hco-miR-5960-3p	9767	8033	23777	25282	13540	16081
Hco-miR-5895-5p	1350	1220	339	5683	13604	1111
Hco-miR-45-3p	3961	3736	5239	2024	5900	1494
Hco-miR-61-3p	279	287	67	10133	25106	619
Hco-miR-43-3p	1395	1319	82	10688	21793	1015
Hco-miR-5352-3p	27	24	37	1656	6805	130
Hco-miR-40b-3p	43	40	8	13	223	54
Hco-miR-5899-3p	11697	14170	14231	13024	11559	11007
Hco-miR-63b-3p	485	693	452	291	2775	2606

B

Adult EV	L3	L3(act)	L4	Male	Female	Gut
Hco-miR-5960-5p	2725	2565	2924	12404	6021	20908
Hco-miR-5895-5p	1350	1220	339	5683	13604	1111
Hco-miR-45-3p	3961	3736	5239	2024	5900	1494
Hco-miR-5352-3p	27	24	37	1656	6805	130
Hco-miR-40b-3p	43	40	8	13	223	54
Hco-miR-5885a-3p	341	305	37413	7610	14857	35869
Hco-miR-5885b-3p	1622	1910	53743	24053	29032	37793
Hco-miR-5899-3p	11697	14170	14231	13024	11559	11007
Hco-miR-5908-3p	62	55	4226	13619	19883	40502
Hco-lin-4-5p	7043	8657	9419	5249	6298	7551

The ten most abundant miRNAs in the adult EV-depleted library are mostly the same as the adult total ES library with the exception of Hco-miR-5899-3p, which shows a very consistent and high level of expression across all the life-cycle stages examined. Between the adult EV library and adult EV-depleted library, Hco-miR-40b-3p, Hco-miR-45-3p, Hco-miR-5352-3p, Hco-miR-5895-5p, Hco-miR-

5899-3 and Hco-miR-5960-5p are present in both libraries. The adult EV-depleted library contains Hco-miR-63b-3p as well as two members of the Hco-miR-5352 cluster; Hco-miR-43-3p and Hco-miR-61-3p, and Hco-miR-5960-3p which is the most abundant miRNA in the adult total ES library (Table 4.1). On the other hand, the adult EV library contains Hco-lin-4-5p and Hco-miR-5908-3p, which both had extremely low reads in the adult total ES library, and Hco-miR-5885a-3p, Hco-miR-5885b-3p.

Analysis of the expression profile of these miRNAs shows that of the ten miRNAs that are highly expressed in the adult EV-depleted library, seven are most highly expressed in the adult female stage (Figure 4.9). Only Hco-miR-5960-3p and Hco-miR-5899-3p show a different pattern, with both being abundant in the L4 stage. In the adult EV library, only four of the miRNAs are adult female enriched (Hco-miR-40b-3p and three members of the Hco-miR-5352 cluster; Hco-miR-5895-5p, Hco-miR-45-3p and Hco-miR-5352-3p). Hco-miR-5960-5p, Hco-miR-5885a-3p, Hco-miR-5885b-3p and Hco-miR-5908-3p are all highly expressed in the gut, which is a potential source of ES EV, with an intermediate level of expression in the adult female. Two of these gut enriched miRNAs, Hco-miR-5885b-3p and Hco-miR-5885a-3p are also extremely highly expressed in the L4 stage, Hco-miR-5885b-3p reaching almost 54,000 reads, which is about two times greater than the highest signal of any miRNA in the adult EV-depleted miRNAs.

Comparison of the ten most abundant miRNAs in L4 EV-depleted supernatant and the L4 EV libraries identifies only five miRNAs in common between the two libraries; Hco-lin-4-5p, Hco-miR-5885a-3p, Hco-miR-5899-3p, Hco-miR-5960-5p and Hco-miR-83-3p. Out of the five common miRNAs, all appear in the adult EV library except for Hco-miR-83-3p. For both L4 libraries, there are a number of miRNAs that peak in the L4 stage. However, certain miRNAs such as Hco-miR-5352-3p, Hco-miR-71-5p, Hco-miR-5960-5p and Hco-miR-236-3p have low expression in both L3 stages and the L4 stage, even though there are extremely abundant in the L4 EV and L4 EV-depleted libraries. The origin of these miRNAs is unclear although it cannot be ruled out that the L4 maintained *in vitro* were undergoing a limited amount of development. Compared to the adult stages, a larger proportion of the miRNAs in the L4 ES libraries have high expression in the *H. contortus* gut

Figure 4.10 The expression profile across life cycle stages for the ten most abundant miRNAs in the A) L4 Supernatant and B) L4 EV .

A

L4 Supernatant	L3	L3(act)	L4	Male	Female	Gut
Hco-miR-5960-5p	2725	2565	2924	12404	6021	20908
Hco-miR-45-3p	3961	3736	5239	2024	5900	1494
Hco-miR-5352-3p	27	24	37	1656	6805	130
Hco-miR-5885a-3p	341	305	37413	7610	14857	35869
Hco-miR-71-5p	13003	18613	15138	27606	25500	24902
Hco-miR-5899-3p	11697	14170	14231	13024	11559	11007
Hco-lin-4-5p	7043	8657	9419	5249	6298	7551
Hco-miR-235-3p	2762	2347	80	1787	1113	60
Hco-miR-83-3p	1364	1231	10639	3146	3240	10799
Hco-miR-84a-5p	6	16	22187	9785	10352	15138

B

L4 EV	L3	L3(act)	L4	Male	Female	Gut
Hco-miR-5960-5p	2725	2565	2924	12404	6021	20908
Hco-miR-5885a-3p	341	305	37413	7610	14857	35869
Hco-miR-5885b-3p	1622	1910	53743	24053	29032	37793
Hco-miR-5885c-5p	15	13	59	911	8750	563
Hco-miR-5899-3p	11697	14170	14231	13024	11559	11007
Hco-miR-5908-3p	62	55	4226	13619	19883	40502
Hco-miR-236-3p	41	30	576	55	178	5114
Hco-lin-4-5p	7043	8657	9419	5249	6298	7551
Hco-miR-83-3p	1364	1231	10639	3146	3240	10799
Hco-miR-1-3p	3225	3271	2411	1077	541	9

4.2.6 Potential miRNAs discovered from the small RNA libraries

From Table 4.4, 41 of the miRBase recognised miRNAs were detected in the small RNA libraries prepared from ES products. There is a possibility that a number of miRNAs sequenced here were not identified by Winter *et al.*,(2012) in the original study due to the use of an early version of the *H. contortus* genome. As the ES libraries were mapped to miRBase, a number of sequences were identified that are present in other organisms but were not identified previously in *H. contortus*. For example, a large number of reads from the adult EV and supernatant libraries mapped perfectly to the Asu-miR-100a-5p mature sequence

from *Ascaris suum* (Wang *et al.*, 2011a). BLAST searches of the supercontig assembly used by Winter *et al.*, (2012) did not yield any results for this sequence. To determine if this fragment is a possible miRNA, the Asu-miR-100a-5p sequence was searched against the current *H. contortus* genome and a single match was identified. Regions flanking the BLAST hit were extracted and mfold used to fold the structure. From Figure 4.11, the *H. contortus* sequence folds into a stem loop structure, suggesting that this particular miRNA does have a *H. contortus* homologue. Due to the significant improvement in genome quality since 2012, it is possible that there are additional miRNAs that were not identified from the initial genome assembly.

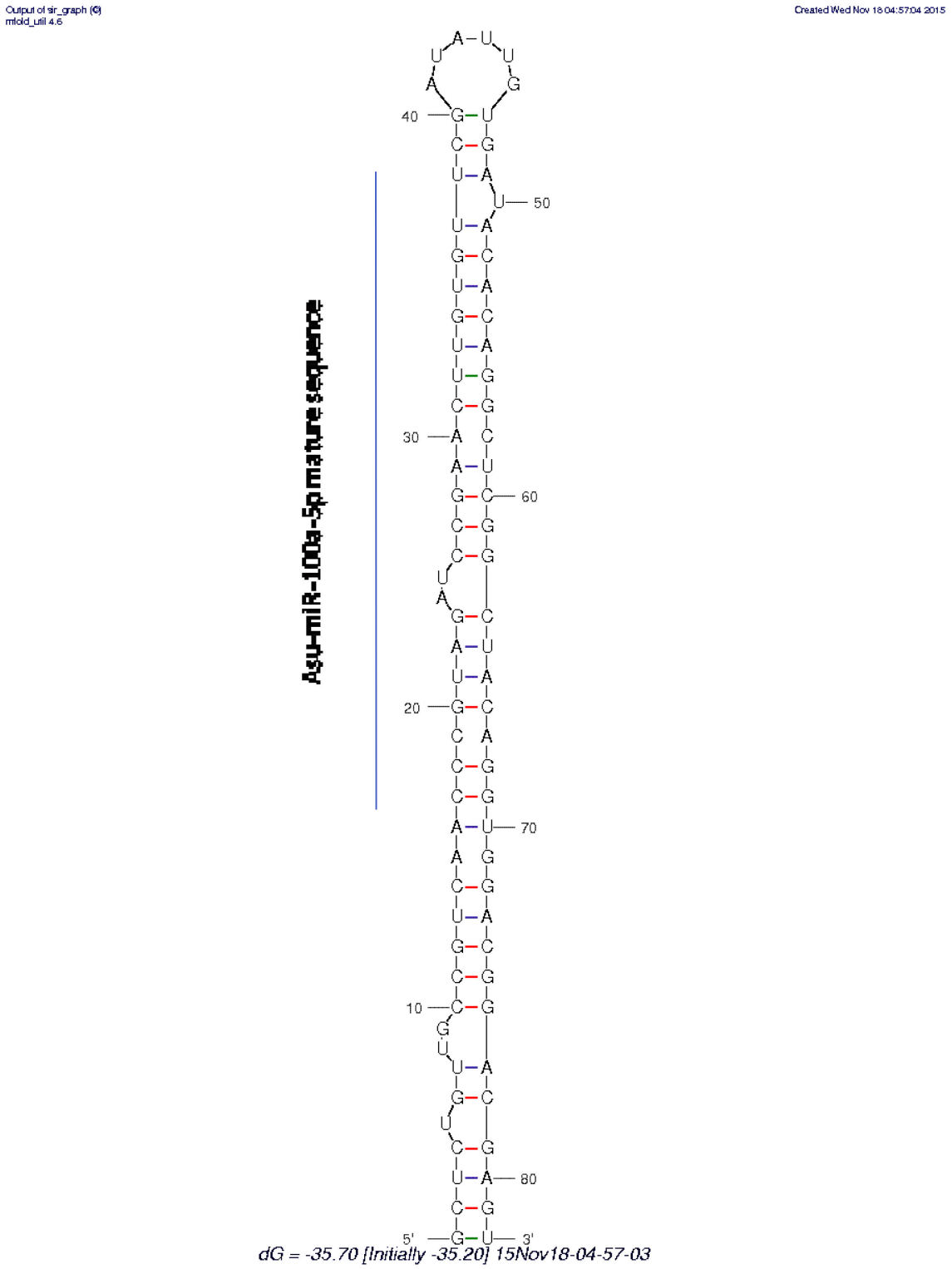


Figure 4.11 miRNA identified in *H. contortus* with high similarity to the asu-miR-100a-5p sequence was folded using the programme mfold and shows the predicted stem loop structure. The asu-miR-100a-5p sequence is highlighted.

4.2.7 EV content of *Teladorsagia circumcincta*

The presence of miRNAs in EV is not limited to *H. contortus* as these have also been identified in a number of different helminth parasites (Coakley *et al.*, 2015). Whilst both are Clade V nematodes, *H. polygyrus* is a parasite of rodents. To investigate whether EV containing miRNAs are also present in a closely related nematode, EV from the L4 stage of *Teladorsagia circumcincta*, another sheep parasite (a kind gift from Thomas Tzelos, Moredun Research Institute, Edinburgh) were obtained and a small RNA library created exactly as described for *H. contortus*. From Table 4.5, the number of reads in the library is very similar to the small RNA library created from the *H. contortus* L4 EV data.

Table 4.5 Summary of the reads that map to *T. circumcincta* RNA sequences using BLAST and the mapper program from miRDeep.

Teladorsagia circumcinta small RNA library	Total Reads / Unique Reads
Small RNA library	11,177,763/4,251,705
Mapping to *T. circumcincta* ribosomes (mapper)	20,480 / 5,223
Mapping to the *T. circumcincta* genome (mapper)	89,647 / 26,354
Mapping to *T. circumcincta* ribosomes (BLAST)	294,723 / 139,369
Mapping to the *T. circumcincta* genome (BLAST)	364,316 / 160,808

With the absence of any *T. circumcincta* miRNA data, reads from the small RNA library were mapped to the entire miRBase hairpin database including *H. contortus* sequences. Out of the 28,645 sequences in the miRBase database, the mapping program could only identify a total of 138 sequences. From those, only one *H. contortus* sequence, Hco-miR-5947, and no *C. elegans* sequences were identified in the database. This is likely due to the high stringency settings of the mapping program (only one mismatch allowed in the 3' end of the read).

Table 4.6 Number of miRNAs identified in the *T. circumcincta* library and the number of corresponding reads identified.

miRNAs identified	138
Total number of reads mapped	465

Instead of using the mapping program, BLAST was used to query the miRBase mature sequence database with the reads from the *T. circumcincta* small RNA library. This method identified 13,637 miRNAs across 190 different species.

Table 4.7 Sumamry of miRNAs idetnified in the *T. circumcincta* exosomes

Nematode	Number of miRNAs identified
Ascaris suum	70
Brugia malayi	52
Caenorhabditis elegans	107
Haemonchus contortus	95
Homo sapiens	1035
Mus musculus	866
Schistosoma japonicum	23
Schistosoma mansoni	9

In the *T. circumcincta* small RNA library, 95 *H. contortus* homologues were identified in the database, accounting for a total of 5,099 reads. Table 4.8 shows the comparison between the *T. circumcincta* L4 EV and the *H. contortus* L4 EV. Of the 95 *H. contortus* miRNAs, only 41 had reads over the threshold in the *T. circumcincta* library (10 reads). Of these, only 19 miRNAs were present in the *H. contortus* L4 EV, while 11 miRNAs that are present in the *H. contortus* L4 EV were not present in the *T. circumcincta* L4 EV. The relative abundance of miRNAs between the two L4 EV was also very different. The most abundant miRNA in the *T. circumcincta* L4 EV, Hco-miR-235-3p was not present in the comparable *H. contortus* library, whilst the most abundant *H. contortus* miRNA, asu-miR-100a-5p, only had 144 reads in *T. circumcincta* (9^{th} in the library).

Whilst the small RNA library sizes are very similar, the highest reads for any miRNA is 3,812 for *T. circumcincta* and 55,492 in *H. contortus*, suggesting that the two L4 EV libraries are quite different in composition.

Table 4.8 Comparison between the T. circumcincta and *H. contortus* L4 EV. Numbers indicated the normalised read numbers from the respective libraries. Data shown if the reads in one library exceeded the cut-off of 10.

	T. circumcincta L4 EV	*H. contortus* L4 EV
Hco-miR-235-3p	3812	Not Found
Hco-miR-45-3p	2072	179
Hco-miR-236-3p	1331	503
Hco-miR-228-5p	1309	258
Hco-miR-72-5p	1156	Not Found
Hco-miR-5885c-3p	531	1356
Hco-miR-5885a-3p	528	4347
Hco-miR-59-3p	521	Not Found
asu-miR-100a-5p	452	55492
Hco-miR-5885b-3p	377	1058
Hco-miR-60-3p	374	53
Hco-miR-71-5p	342	317
cel-let-7-5p	305	13532
Hco-lin-4-5p	298	11107
Hco-miR-259-5p	220	327
Hco-miR-5960-5p	220	1336
Hco-miR-993-3p	201	25
Hco-miR-5978-5p	192	Not Found
Hco-miR-5922-5p	185	Not Found
Hco-miR-790-5p	182	Not Found
Hco-miR-5899-3p	166	4565
Hco-miR-50-5p	157	Not Found
Hco-miR-5991-3p	144	Not Found
Hco-miR-87a-3p	104	25
Hco-miR-87b-3p	100	Not Found
Hco-miR-9-5p	100	Not Found
Hco-miR-84a-5p	97	66
Hco-miR-87c-3p	91	Not Found
Hco-miR-252-5p	88	Not Found
Hco-miR-307-3p	88	Not Found
Hco-miR-63a-3p	88	104
Hco-miR-5908-3p-3p	82	Not Found
Hco-miR-86-5p	79	Not Found
Hco-miR-43-3p	66	3
Hco-miR-5948-3p	66	Not Found
cel-miR-50-5p	60	198
Hco-miR-83-3p	57	2341

Hco-miR-5884-5p	41	113
Hco-miR-79-3p	41	Not Found
Hco-miR-40b-3p	31	Not Found
Hco-miR-46-3p	31	Not Found
Hco-miR-5900-3p	28	Not Found
Hco-miR-5921-5p-5p	28	Not Found
bma-mir-36b-p5	22	31
cbn-mir-64f-p5	6	72
cel-mir-5592-1-p3	3	120
prd-miR-7911a-5p	3	187
Hco-miR-2-3p	Not Found	123
Hco-miR-43-p5	Not Found	50
Hco-miR-5352-3p	Not Found	31
Hco-miR-5895-5p	Not Found	82
Hco-miR-5908-3p	Not Found	1009
Hco-miR-5939-3p	Not Found	126
Hco-miR-5960-p3	Not Found	186
Hco-miR-5960-p5	Not Found	25
Hco-miR-5976-5p	Not Found	28
Hco-miR-61-3p	Not Found	Not Found
Hco-miR-63b-3p	Not Found	129
Hco-miR-9551-3p	Not Found	Not Found
prd-miR-7911c-5p	Not Found	195

4.2.8 Protein concentration of the libraries

The protein concentration of all the samples were estimated using the Qubit® protein assay kit prior to the RNA extraction and precipitation. Table 4.9 shows the amount of protein identified in each of the samples. The sum of the *H. contortus* adult EV and adult EV-free supernatant protein concentrations is very similar to the adult Total ES. The protein concentrations from the *H. contortus* L4 samples were all slightly higher than their corresponding adult values. *T. circumcincta* L4 EV protein concentrations were the lowest, under half the value of the *H. contortus* L4 EV concentration. The level of protein differs between the different libraries, however, the difference in protein concentration does not match up with the number of reads identified from the various libraries. This suggests that the protein concentration of the ES is not indicative of the amount of small RNA present.

Table 4.9 Protein concentrations of the various small RNA library samples prior to RNA extraction

Small RNA library samples	Protein Concentration prior to RNA extraction (µg/ml)	Number of total reads (millions of reads)
H. contortus Adult Total ES	136	15
H. contortus Adult EV	82	13.3
H. contortus Adult EV-free supernatant	61	8.7
H. contortus L4 EV	122	11
H. contortus L4 EV-free supernatant	93	7.9
T. circumcincta L4 EV	50	11.1

4.3 Discussion

In this chapter, the small RNA content of *H. contortus* ES material was analysed. During infection, *H. contortus* excrete or secrete a variety of molecules, collectively known as the excretory/secretory (ES) products. The protein content of ES from parasitic nematodes has been studied and found to have immunomodulatory properties. It is speculated that release of these molecules could be an immune evasion strategy or a homeostatic mechanism to minimize pathology (Grainger *et al.*, 2010). Current research into this area has mostly concentrated on proteins, however there is now increasing evidence that the ES of nematodes also contains miRNAs (Buck *et al.*, 2014; Nowacki *et al.*, 2015; Tritten *et al.*, 2014; Zamanian *et al.*, 2015).

In this work, several different libraries were prepared from *H.contortus* adult and L4 stage ES and the small RNA content compared. In the adult total ES small RNA library, the number of reads that could be mapped to miRNAs is relatively small compared to the reads mapping to other RNA species, with approximately 0.2% of the total library being miRNAs. In comparison, the miRNA content of the adult whole worm small RNA library accounts for a significant percentage of the total reads (14%) and is much higher than the tRNAs and nearly three times more than the mRNAs. The reason for the low level of miRNAs in the ES is unknown. Comparison of the small RNA libraries from whole worms and from ES suggests that there is selective release of specific miRNAs into the ES. The significant decrease in total miRNAs reads could be due to the differing starting amount of RNA in the two small RNA libraries or may reflect the absence of certain miRNAs in the ES library.

The RNA library obtained from *H. contortus* total ES contained a significant number of reads that mapped to ribosomal RNA sequences. In studies of total RNA obtained from cells such as cultured mammalian cells (e.g. HeLa), as well as a range of different organisms from bacteria to human tissue, small RNA libraries all show that a high proportion (upwards of 80%) of the total RNA is ribosomal species (Chen and Duan, 2011; Giannoukos *et al.*, 2012; Lodish *et al.*, 2000; Morlan *et al.*, 2012). The second most abundant RNA species are tRNAs, accounting for around 15% of the total RNA content (Lodish *et al.*, 2000; Warner,

1999), with the final fraction being composed of mRNA and other RNA species. A similar distribution of RNA species was also seen in hepatic schistosomula and the adult male and female worms of *S. mansoni*, although the distribution in the egg, cercariae and lung-stage show different RNA species distribution (Cai *et al.*, 2013). Total RNA from the adult *H. contortus* shows a slightly different distribution with high levels of rRNA with miRNAs being the second most abundant, followed by mRNAs and tRNAs (Winter *et al.*, 2012). One issue with these analysis is that the number of reads that map depends on the number of known RNA sequences from the organism. However, the high level of rRNAs appears to be a trend across various different organisms.

Along with the high level of ribosomal RNA sequences, there are also a large number of mRNA sequences and a smaller number of tRNAs. The function of these sequences in the ES is unclear. In *Schistosoma* ES products, Nowacki *et al.* (2015) were able to identify a class of tRNA-derived small RNAs (tsRNAs), which are a possible source of small non-coding RNAs (Anderson and Ivanov, 2014; Nowacki *et al.*, 2015). These tRNA fragments are derived from mature tRNAs and are able to interact with the RNAi machinery to down-regulate target genes, suggesting a possible function of these tRNA sequences in the ES (Anderson and Ivanov, 2014; Nowacki *et al.*, 2015).

Winter *et al.* (2012) discarded a number of miRNAs due to sequence similarity with ribosomal, messenger or transfer RNA sequences. However, some of these sequences were relatively highly represented in the total worm ES library. Analysis of the alignment of reads with these 'unofficial' miRNAs shows that some of these miRNAs could be true miRNAs due to the predicted mature sequence. If this is true, then the number of reads mapping to miRNA sequences may be higher than shown, although it is unlikely to account for all of the other RNA sequences. BLAST searches of the miRBase database identified only two of the 10 most abundant 'unofficial' miRNAs in other species. Hco0325 and Hco0297 were previously both discarded due to similarity with a tRNA and were not submitted to miRBase, however, they show similarity to two plant sequences aqc-miR166b and stu-miR8021 respectively. Figure 4.9 shows the alignment between Hco0297 to *Solanum tuberosum*-miR8021 and Hco0325 to the *Aquilegia caerulea* -miR166b. For both alignments, the *H. contortus* sequence is much

longer and results in an overhang at both 3' and 5'. The seed sequences of the two miRNAs are therefore different. If we assume that the miRNA prediction in *H. contortus* is slightly inaccurate and only consider the sequence that aligns perfectly, there are a still a number of mismatches across the whole miRNA with a number of them appearing in the seed region again. This suggests the miRNAs are not the same as any currently known miRNAs.

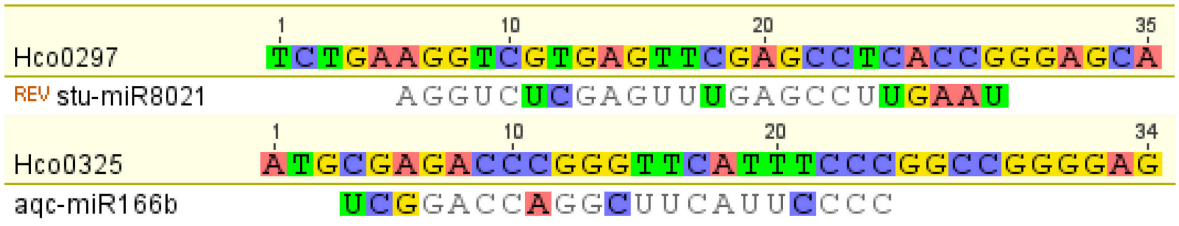


Figure 4.12 Alignment of the mature sequences of Hco0297 to the *Solanum tuberosum* miRNA stu-miR8021 and Hco0325 to the Aquilegia caerulea miRNA aqc-miR166b. Both alignments have multiple mismatches. The *H. contortus* mature miRNAs are also longer than their matches and their seed regions do not align.

In this study, there were also a large number of reads that did not map to any known *H. contortus* RNA sequences. The majority of these reads were rejected due to the high stringency settings for the mapping. However, detailed analysis of these rejected reads showed that the majority of them had high pairwise identity to *H. contortus* sequences (data not shown).

From the adult total ES library, Hco-miR-5960 was the most abundantly expressed miRNA in the library, accounting for 70% of the reads. Of all the miRNAs, Hco-miR-5960-5p is the only miRNA to appear in the top 10 of all four EV or EV-depleted libraries. This miRNA is also highly abundant in the gut sample of *H. contortus*. Its expression in adult female worms, whilst relatively high, is 3.5 times lower than in the female gut tissue, indicating enriched expression in gut tissue. This was also found for some other Hco-miRNAs including Hco-miR-83-3p, Hco-miR-5885a-3p, Hco-miR-236-3p and Hco-miR-5908-3p (see Figure 4.9) and suggests that they may have a special function within the parasite and that the gut may be the source of certain miRNAs.

In order to take into account small variations in the sequence length of RNA reads, the adult total ES small RNA library was mapped to the precursor sequence of miRNAs instead of the mature sequence. Since the precursor is

longer, it is possible that the reads from the library are mapping to a region of the precursor that is not an exact match to the mature miRNA sequence. To determine where the reads are mapping, multiple alignments were created to visualise the mapping data. For the large majority of miRNAs, the majority of the reads map exactly to the predicted mature sequence. Whilst the 5' seed region is extremely well conserved, there is a huge amount of variation in the length and the start and stop position of the reads. It has been well established that a miRNA gene can give rise to a group of heterogeneous products (isomiRs) that have variable lengths and end-sequences rather than to one singular miRNA (Morin *et al.*, 2008). The biological significance of isomiRs is unclear, however, isomiRs do not arise randomly (Lee *et al.*, 2010) and are due to imprecise cleavage by Drosha and Dicer (Starega-Roslan *et al.*, 2011; Warf *et al.*, 2011).

There is a growing amount of evidence that miRNAs are able to exist in bodily fluids and are protected from degradation by being packaged in EV or exosomes. In this thesis, these are referred to as EV as further characterisation would be required to classfiy these as exosomes. Circulating exosomes have also been shown to be an important method of intercellular communication (Turchinovich *et al.*, 2013). Work by Buck *et al.,* (2014) identified EV enriched in exosome markers in the ES products of *Heligmosomoides polygyrus.* These EV were internalized by mouse cells and could suppress murine Type 2 innate immune responses (Buck *et al.*, 2014). In *H. polygyrus*, the majority of miRNAs are present in the exosomes fractions (Buck *et al.*, 2014). In the present study, small RNA libraries were created from the EV and EV-free supernatant of *H. contortus* adults and L4 stage. The distribution of miRNAs between the *H. contortus* adult EV and the EV-free supernatant is not as pronounced as in *H. polygyrus*, as the number of miRNA reads is approximately equal between the two. Interestingly, the L4 libraries are much more similar to *H. polygyrus*, with 70% of the miRNA reads in the EV/exosomes. The majority of miRNAs show some preference to either the EV or the EV depleted fraction, suggesting that there may be two groups of miRNAs; the EV associated and the non-EV associated. In the adult libraries, there is clear enrichment of certain miRNAs in the EV, however, there are also miRNAs that are enriched in the EV depleted fraction. The miRNAs of the Hco-miR-5352 cluster are present in both fractions, albeit skewed more towards the EV depleted fraction. Studies on the two different sets of

extracellular miRNAs suggests that the non-EV miRNAs are associated with Argonaute-containing protein complexes (Arroyo *et al.*, 2011) or lipoprotein complexes (Vickers *et al.*, 2011) although these miRNAs may not be as stable as EV associated ones (Köberle *et al.*, 2013). Whether miRNAs contained within EV or exosomes have differing functions to those that are present in the supernatant is the subject of ongoing studies (Arroyo *et al.*, 2011; Turchinovich *et al.*, 2011). It is notable that of the ten most abundant miRNAs in either the adult or L4 *H. contortus* EV fraction, approximately half are also highly expressed in gut tissue. This suggests these miRNAs may be released from this site, possibly through regurgitation during feeding, although their function requires further analyses.

Material was also available from *T. circumcincta* L4 EV and a small RNA library was constructed and analysed. There was very little correlation in the miRNAs and their expression between *H. contortus* and *T. circumcincta*, although the quantity of material available for *T. circumcincta* may have limited the quality of the analysis. The difference between the *T. circumcincta* L4 EV and the corresponding *H. contortus* library could be partially due to lack of miRNA data for *T. circumcincta*, or that *H. contortus* and *T. circumcincta* miRNA sequences were sufficiently different that they were identified as different sequences.

In conclusion, the data presented in this chapter show the profile of miRNAs in the ES of *H. contortus*. The difference in miRNA content and abundance between the adult whole worm library and the ES libraries suggests that the miRNAs in the ES arise through some method of selection. The miRNAs in the ES can also be further divided into miRNAs present in the EV or the EV depleted fraction. In adult ES, the number of reads in the two fractions are approximately equal, however in the L4, 70% of the miRNA reads are in the EV fraction. The presence of stable miRNAs in the EV depleted fraction suggests that there is another method of miRNA stabilization possibly through association with protein or lipoprotein complexes (Arroyo *et al.*, 2011; Vickers *et al.*, 2011). This could be confirmed by characterization of the protein profile of the different fractions. The miRNAs of the Hco-miR-5352 cluster are highly expressed in both fractions, with a slightly higher abundance in the EV depleted fraction. The presence of

these four miRNAs in the ES suggests that they may be targeting a host gene. The genes that may be targeted are investigated in Chapter 5.

5 Identification of mammalian targets of the Hco-miR-5352 cluster

5.1 Introduction

In chapter 3, the Hco-miR-5352 cluster was shown to be highly expressed primarily in the adult stages of *H. contortus*. This particular cluster can also be identified in a number of other parasitic nematodes that infect a range of different mammalian hosts. This suggests that the miRNA cluster may be involved in parasitism and may possibly target a host mRNA. Analysis of the excretory-secretory products identified a number of *H. contortus* miRNAs, with members of the Hco-miR-5352 cluster being present in the ES. The difference in abundance between the ES miRNAs and the whole adult miRNAs suggests that the ES miRNAs are selectively released and that a number of these miRNAs may be contained in exosomes. Exosome or extracellular vesicles containing miRNAs have also been identified in a number of other parasitic nematodes including the secretions of the human parasite, *Brugia malayi* (Zamanian *et al.*, 2015), the porcine parasite, *Trichuris suis* (Hansen *et al.*, 2015), from mice infected with *H. polygyrus* (Buck *et al.*, 2014), from cattle and humans infected with *Onchocerca* (Quintana *et al.*, 2015), the protozoan parasite *Leishmania* (Lambertz *et al.*, 2015) as well as from *Schistosoma* (Zhu *et al.*, 2016b).

Identification of the mRNA targets of miRNAs is very complex. Unlike plant miRNAs, animal miRNA-mRNA interactions do not require perfect base pairing. Animal miRNAs are able to identify mRNAs based on an imperfect match to the seed sequence (6-8 nucleotide match). There are a number of features that influence miRNA-mRNA interactions and the effect that miRNA binding can have on target gene expression. Firstly, miRNAs often target the 3′ UTR of mRNA (Zisoulis *et al.*, 2010), as the number of AU residues increases the accessibility of the miRNA to the target site (Bartel, 2009). In addition, the number of target sites within a 3′ UTR sequence is related to the ability of the miRNA to downregulate the mRNA (Doench *et al.*, 2003). However, most miRNAs target multiple different mRNA and a given mRNA may be regulated by multiple different miRNAs (Brennecke *et al.*, 2005), which can make it difficult to identify the most likely or most important targets.

Since the rules governing miRNA-mRNA interactions were determined, computational methods have been devised to help identify miRNA targets. These algorithms were developed to identify potential mRNA targets based on current knowledge of miRNA-mRNA interactions (Peterson *et al.*, 2014). They generally focus on a short region (nucleotides 2-8) at the 5′ end of the miRNA and a complementary sequence on the 3′ UTR of the mRNA. Since the match length is so short, algorithms utilise additional criteria to identify targets such as: 1) the free energy of binding between the miRNA and the target, 2) cross-species conservation of the miRNA and the target sites and 3) the accessibility of the site to which the miRNA can localize and hybridize (Peterson *et al.*, 2014). In this study, TargetScan was chosen to examine potential targets of Hco-miR-5352 due to its ease of use and detailed database of annotated 3′ UTRs of genes in 23 vertebrate species. However, these algorithms are not completely accurate. A recent study used TargetScan to identify miRNAs that regulate the immunoregulatory protein B7-H3 (Nygren *et al.*, 2014). Of the 13 miRNAs identified as targeting B7-H3, eight were confirmed to be targets by a 3′ UTR luciferase assay. However, four miRNAs were also confirmed to target B7-H3 experimentally but were not predicted by Targetscan or miRanda, demonstrating the limitations of the prediction algorithms (Nygren *et al.*, 2014).

One of the challenges was to determine biologically relevant miRNA-target interactions. There are many additional methods that have been used for miRNA target identification, ranging from whole genome screens to gene-specific techniques. In general, the effect of a miRNA is primarily to reduce the expression of target mRNAs. However, in the broader techniques, such as genetic screens, it is difficult to differentiate between direct and indirect results of miRNA modulation. For example, a specific miRNA may target a transcription factor that then influences the expression of several downstream genes that are not the direct target of the miRNA. Experimental methods, such as the HIgh-Throughput Sequencing of RNA isolated by CrossLinking ImmunoPrecipitation (HITS-CLIP), rely on immunoprecipitating the miRISC containing the miRNA-mRNA duplexes (Zhao *et al.*, 2015; Zisoulis *et al.*, 2010). However, although the number of potential targets is reduced, the miRNA binding sequence cannot be directly identified and requires further computational approaches. On the other hand, gene-specific assays such as the

luciferase reporter assay have been extensively used to demonstrate a direct interaction between a specific miRNA and a 3′ UTR construct (Kong *et al.*, 2008; Kuhn *et al.*, 2008; Lytle *et al.*, 2007). The specificity of interaction can be demonstrated by altering the sequence of the 3′ UTR target site in the luciferase reporter construct.

In this chapter, the aim was to characterise the miR-5352 cluster in more detail with a view to determining possible host gene targets. To this end, experiments were conducted to investigate whether miRNAs can be detected in infected sheep tissues, indicating they are released *in vivo*. Bioinformatic tools were used to predict mammalian targets of Hco-miR-5352 and the other miRNAs in the cluster. Finally, the luciferase reporter system was used to experimentally validate one host mRNA-miRNA interaction, that between the CD69 3' UTR and Hco-miR-5352.

5.2 Results

5.2.1 Detecting *H. contortus* miRNAs in sheep tissue

The four miRNAs in the Hco-miR-5352 cluster are most abundant in the L4 and adult stages and, as both stages are found in the abomasum of the host sheep, one potential function may be to interact with host cells. If this is the case, it may be possible to detect miRNAs in sheep that have been infected with *H. contortus*. To this end, abomasal epithelial and abomasal lymph node tissues were obtained from two sheep infected with *H. contortus* and two uninfected sheep. RNA was extracted from these tissues and analysed by qRT-PCR to determine if any of the ES-enriched miRNAs could be detected in *ex-vivo* material.

qRT-PCR requires a normaliser gene to control for different input quantities of RNA. Coutinho *et al*. (2007) profiled bovine microRNAs from immune-related, small intestinal and embryonic tissues and found Bta-miR-26a, Bta-miR-103 and Bta-miR-122a to be highly expressed in all five tissues examined. BLAST sequence searches in miRBase and NCBI (*ovis aries* genome data oar_v3.1) identified three miRNAs in sheep with an exact sequence match (Clop *et al.*, 2006; Coutinho *et al.*, 2007). RNA was extracted from sheep tissue samples and qRT-PCR was carried out to determine the expression of the three ovine miRNAs. Primer sequences are given in Chapter 2.5. All three miRNAs were detected in both tissues (data not shown). However, Oar-miR-26a, the sheep version of Bta-miR-26a was chosen as a normaliser, as it showed the most consistent values across the four different tissue samples (Figure 5.1).

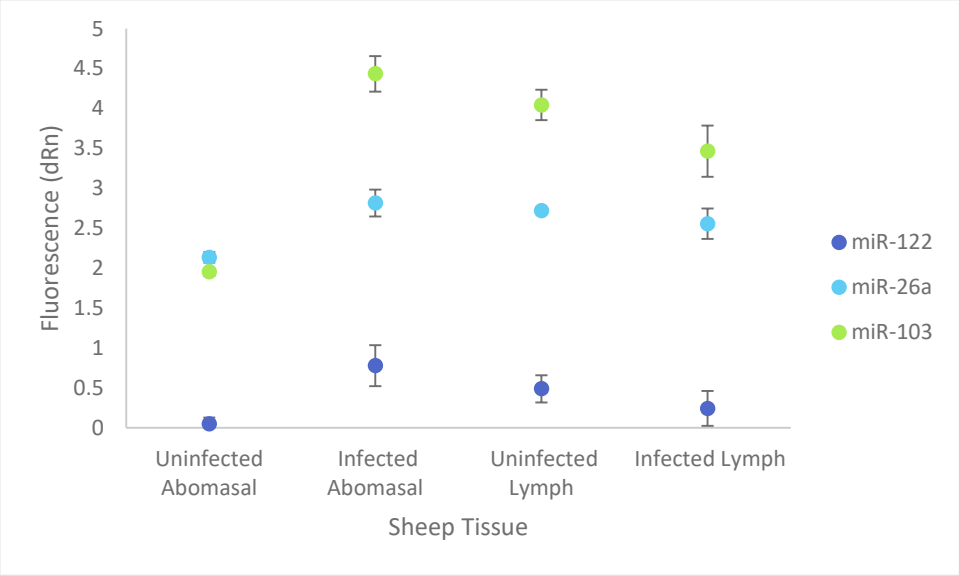


Figure 5.1 qRT-PCR to detect presence of Oar-miR-26a, Oar-miR-122 and Oar-miR-103 in infected and pathogen-free sheep abomasal and draining lymph node tissue. RT-PCR was carried out in triplicate and values shown as fluorescence (mean ± SD, n=3).

Three *H. contortus* miRNAs were selected to assay, based on their high abundance in the ES as well as their lack of similarity to sheep miRNAs. These were two members of the cluster, Hco-miR-5352 and Hco-miR-5895, and Hco-miR-5960, which is the most abundant miRNA found in the ES (see Table 4.4). The data shown in Figure 5.2 is the fluorescence normalised to the corresponding value for Oar-miR-26a in the respective tissue. For both the lymph node and abomasal tissue samples, Hco-miR-5352 showed an increase between infected and uninfected animals. In contrast, Hco-miR-5960 showed a minimal increase in ratios. The profile of Hco-miR-5895 also shows a high abundance in the infected tissues compared to the uninfected (Figure 5.2). Interestingly, the value for the infected abomasal tissue from sheep 4 is approximately three times higher than the abomasal tissue from sheep 3, although this same difference is not seen in the corresponding lymph node samples.

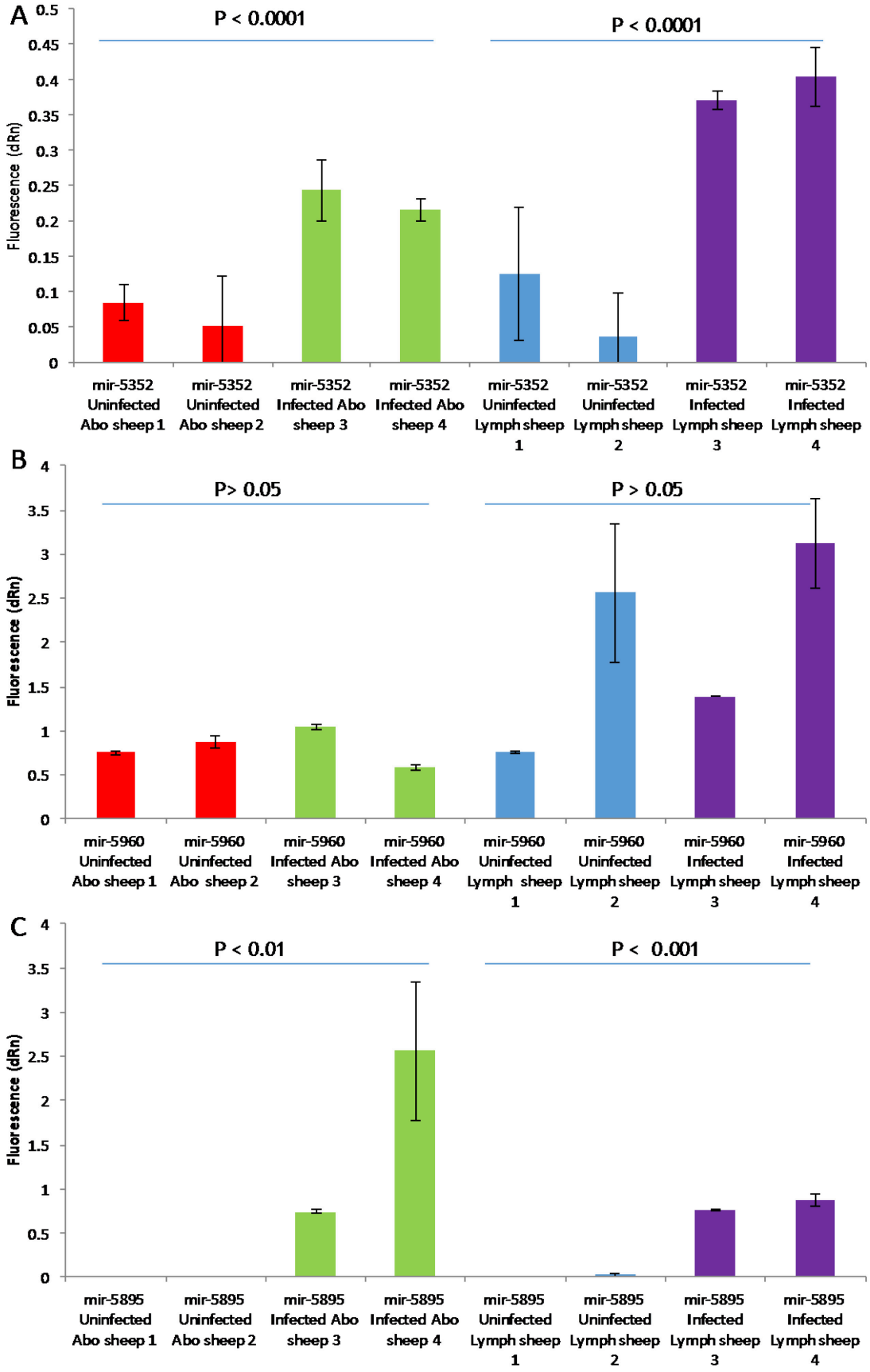


Figure 5.2 qRT-PCR to detect presence of A) Hco-miR-5352, B) Hco-miR-5960 and C) Hco-miR-5895 in infected and pathogen-free sheep abomasal (abo) and draining lymph node (lymph) tissue. Two samples were obtained from the abomasum and lymph nodes and qRT-PCR was carried out in triplicate and values averaged. Values were normalised to oar-miR-26a and shown the normalised fluorescence. P values shown were calculated using an analysis of variance test.

The data show that Hco-miR-5352 and Hco-miR-5895 levels are increased in infected sheep tissues compared to uninfected tissues, an observation that would be consistent with the release of these miRNAs *in vivo*. However, there is little difference in the expression of Hco-miR-5960 in infected and uninfected sheep tissues, despite the observation that this miRNA was the most abundant miRNA detected in the ES library (see Chapter 4). It is possible that the presence of Hco-miR-5960 in both infected and uninfected tissues may relate to non-specific amplification of a sequence related to miR-5960.

5.2.2 Bioinformatic search for mammalian microRNA targets

From the qRT-PCR data described above, specific *H. contortus* miRNAs could be detected in infected sheep tissue, indicating that these miRNAs may be involved in the host-parasite interaction perhaps by targeting a sheep mRNA. To investigate possible host targets, bioinformatic searches were conducted using the Targetscan release 5.2 (http://www.targetscan.org/vert_50/seedmatch.html) to identify potential mammalian targets of the four miRNAs in the Hco-miR-5352 cluster (miR-5352, miR-43, miR-5895, miR-61). Targetscan identifies potential targets by searching the 3′ UTR for sequences that are complementary to the seed sequence of the query miRNA. Table 5.1 shows the number of predicted targets obtained from Targetscan that had an exact match to positions 2-8 of the miRNA mature sequence (seed sequence) and showed good conservation amongst other mammals. Both Hco-miR-43 and Hco-miR-5352 had a large number of targets whereas Hco-miR-61 and Hco-miR-5895 showed many fewer targets.

Table 5.1 Number of targets predicted by Targetscan for the four miRNAs of the Hco-miR-5352 cluster

Hco-miR-43	Hco-miR-5352	Hco-miR-61	Hco-miR-5895
405	500	69	40

In total, 943 unique genes were predicted as targets of the Hco-miR-5352 cluster, a number of which are targeted by multiple miRNAs. Figure 5.3 shows the distribution of the genes between the different miRNA target lists. Since the

miRNAs of the cluster may be co-ordinately expressed, there is a possibility that they may target common genes, or distinct components of a single pathway. However, while no gene was a target of all four miRNAs, three genes ADAMTS5, CNOT6L and ZNF1 were predicted targets of Hco-miR-43, 61 and 5352 (Figure 5.3).

All three common target genes have metal ion-binding properties but have different functions. ADAMTS5 has been shown to be involved in inflammatory arthritis and osteoarthritis (Glasson *et al.*, 2005; Stanton *et al.*, 2005). Interestingly, the Human ADAMTS5 gene and the chimpanzee version both have a single Hco-miR-5895 binding site, however this site is not conserved across the other species in the Targetscan UTR database. The active site of a second predicted target, CNOT6L, is able to recognise RNA and to deadenylate RNAs (Mittal *et al.*, 2011; Wang *et al.*, 2010a), while the third potential target, ZNF148, has been associated with colorectal and renal cancer (Cai *et al.*, 2012; Gao *et al.*, 2013). Figure 5.4 shows the 3′ UTR of each of these genes with the predicted miRNA binding sites annotated.

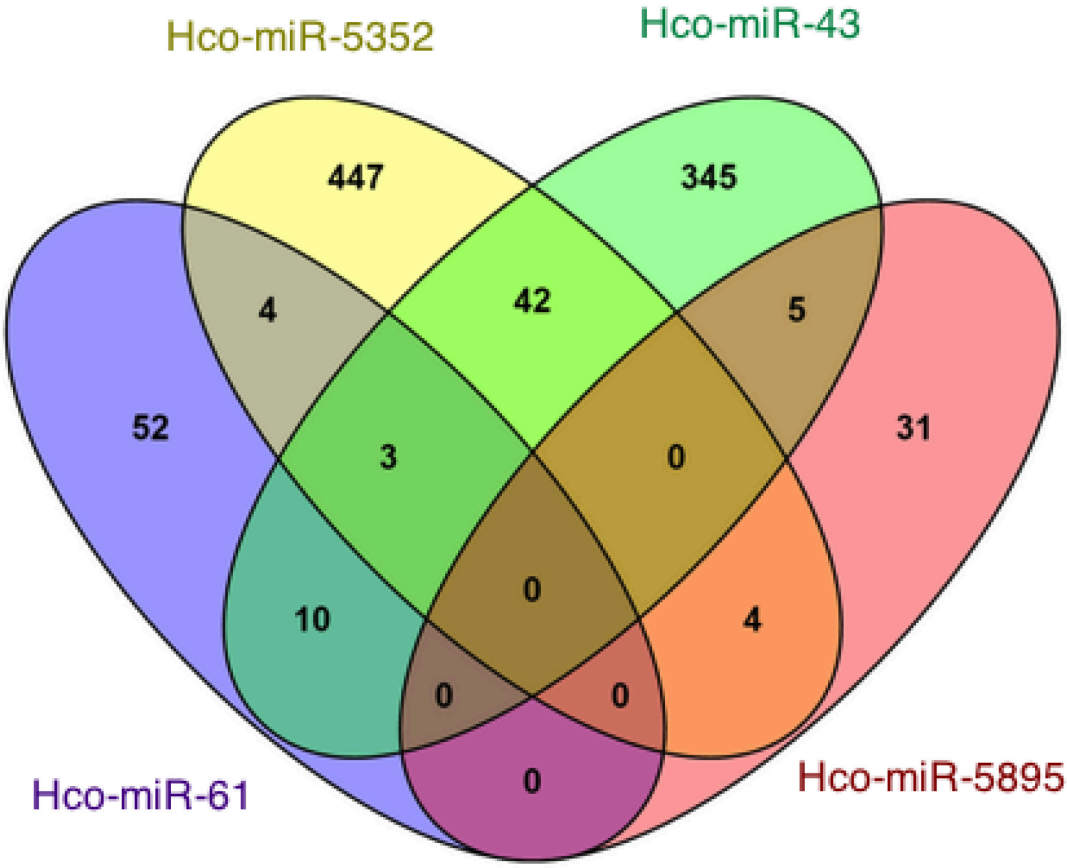


Figure 5.3 Venn diagram representing targets in common between the four miRNAs of the Hco-miR-5352 cluster.

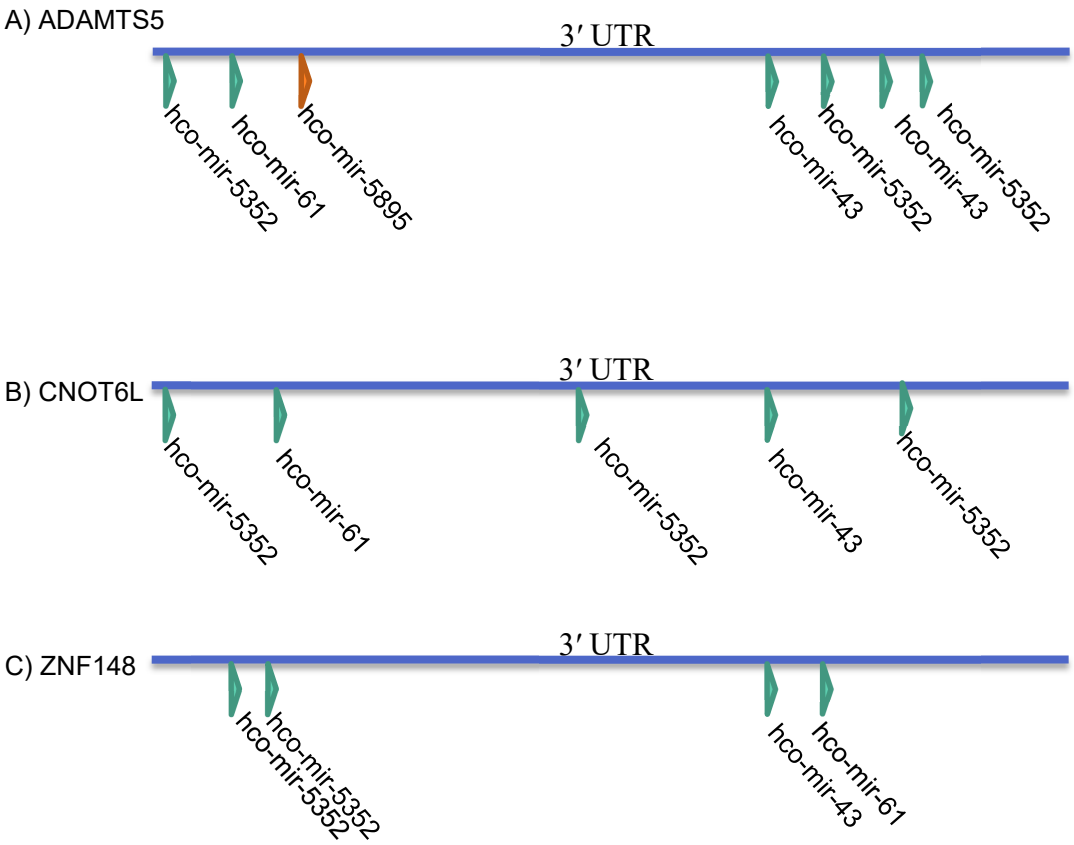


Figure 5.4 Diagram showing the location of Targetscan predicted binding sites (yellow triangles) in the 3′ UTR of the A) ADAMTS5 B) CNOT6L and C) ZNF148 gene. The non-conserved Hco-miR-5895 binding site in ADAMTS5 is shown as a blue triangle.

The 943 target genes from Targetscan were also analysed using Ingenuity Pathway Analysis (IPA Qiagen). The gene symbols were mapped to canonical pathways from the IPA library. The significance of the association between the input list and the canonical pathway was determined based on the number of genes from the list that mapped to the pathway, divided by the total number of genes in the pathway. A Fisher's Exact Test was used to determine the probability that the number of genes from the list that map to a pathway is due to chance. Table 5.2 shows a list of pathways targeted by all four miRNAs and the P-value. From the IPA analysis, Wnt/ß-catenin signalling has the highest possibility that it is targeted by the Hco-miR-5352 cluster, however, the P-values are all greater than 0.05, indicating that the association is most likely random. Whilst there are a few pathways with a P-value of less than 0.05, no pathway has a significant P-value for all four miRNAs.

Table 5.2 Ingenuity pathway analysis separated into the individual miRNAs. Numbers indicate the p-value. The P-value is calculated using Fisher's Exact Test and determines if

the proportions that map to the pathway are significantly different between the reference set and the sample set. A P-value of less than 0.05 indicates a statistically significant, non-random association.

Canonical Pathway	Hco-miR-5895	Hco-miR-5352	Hco-miR-61	Hco-miR-43
Wnt/β-catenin Signalling	0.08	0.11	0.07	0.06
Factors Promoting Cardiogenesis in Vertebrates	0.04	0.19	0.24	0.06
Protein Kinase A Signalling	0.10	0.13	0.41	0.02
HIPPO signalling	0.43	0.43	0.08	0.05
PPARα/RXRα Activation	0.25	0.29	0.18	0.12
ILK Signalling	0.58	0.15	0.19	0.13
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis	0.04	0.23	0.75	0.40
Dopamine Receptor Signalling	0.14	0.38	0.21	0.30
Cardiac β-adrenergic Signalling	0.20	0.35	0.31	0.15
B Cell Receptor Signalling	0.25	0.13	0.70	0.18
Epithelial Adherens Junction Signalling	0.53	0.28	0.15	0.19
Amyotrophic Lateral Sclerosis Signalling	0.46	0.12	0.57	0.14
Breast Cancer Regulation by Stathmin1	0.26	0.10	0.41	0.44
Insulin Receptor Signalling	0.20	0.15	0.64	0.25
TGF-β Signalling	0.43	0.17	0.23	0.34
Glucocorticoid Receptor Signalling	0.15	0.21	0.79	0.33
PI3K/AKT Signalling	0.50	0.20	0.30	0.34
Axonal Guidance Signalling	0.25	0.51	0.46	0.18
Tight Junction Signalling	0.55	0.36	0.17	0.36
Colorectal Cancer Metastasis Signalling	0.13	0.61	0.76	0.24
CDK5 Signalling	0.16	0.50	0.25	0.78
ERK/MAPK Signalling	0.58	0.22	0.40	0.30
Netrin Signalling	0.32	0.53	0.40	0.24
AMPK Signalling	0.26	0.41	0.34	0.44
Production of Nitric Oxide and Reactive Oxygen Species in	0.57	0.30	0.19	0.55

Macrophages				
Actin Cytoskeleton Signalling	0.61	0.22	0.75	0.19
NRF2-mediated Oxidative Stress Response	0.57	0.20	0.39	0.70

5.2.3 CD69 is a likely target of Hco-miR-5352

Table 5.3 shows the ten most likely mammalian mRNA targets predicted by Targetscan for Hco-miR-5352. The top hit is the CD69 3′ UTR which has exact matches to positions 2-8 of the mature miRNA as well as being a component of the host immune system. In humans, this UTR is 990 bases long and contains two Hco-miR-5352 binding sites that are exact matches to the full length of the Hco-miR-5352 mature seed sequence. Both sites are also fairly well conserved across different mammalian species (Figure 5.6A), while other targets sites are less well conserved across different species. Targetscan did not predict any binding sites for the other miRNAs of the cluster within the CD69 3' UTR, however, a sequence search did identify partial matches for Hco-miR-61 and Hco-miR-43. These sites match to part of the 8 bp seed sequence with 6 bases of Hco-miR-61 matching and 7 bases of Hco-miR-43 matching. Both these miRNA sites are conserved between humans and chimpanzee. In addition, Hco-miR-43 is also present in rhesus macaque and Hco-miR-61 in armadillos (Figure 5.6B and C). However, the conservation of these locations is very poor across the different species, suggesting that they are unlikely to be true miRNA binding sites.

One miRNA that does share some similarity to the Hco-miR-5352 mature sequence is miR-25 which has a seed sequence of GUGCAAUA, the first six nucleotides match the last six nucleotides of Hco-miR-5352 (Figure 5.5). The CD69 3' UTR also possess three predicted miR-25 sites, compared to the two Hco-miR-5352 sites. miR-25 has been identified as a switch activated in an overworked heart that drives the onset of heart failure (Bush and van Rooij, 2014; Wahlquist *et al.*, 2014) and is reported to promote reprogramming of somatic cells to induced pluripotent stem cells (Lu *et al.*, 2012). It does not appear to have any immunological function so the significance of the overlapping seed sequence is unclear.

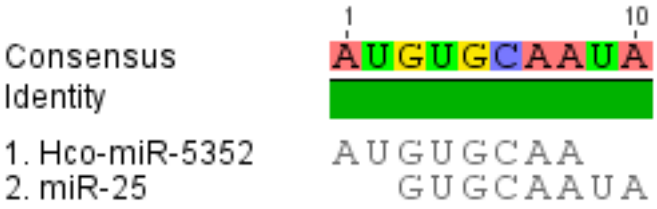


Figure 5.5 Sequence alignment of Hco-miR-5352 and miR-25 (also known as miR-32/92/92AB/363/367)

Table 5.3 The top 10 Targetscan predicted targets of Hco-miR-5352

Target gene	Gene name	Total	8mer	7mer-m8*	7mer-1A**
CD69	CD69 molecule	2	2	0	0
PCDH17	protocadherin 17	2	2	0	0
SYN2	synapsin II	2	2	0	0
USP47	ubiquitin specific peptidase 47	3	1	1	1
C18orf25	chromosome 18 open reading frame 25	2	1	1	0
MAP2K4	mitogen-activated protein kinase kinase 4	2	1	1	0
SERTAD2	SERTA domain containing 2	2	1	1	0
SLC12A5	solute carrier family 12, (potassium-chloride transporter) member 5	2	1	1	0
BACH2	BTB and CNC homology 1, basic leucine zipper transcription factor 2	2	1	0	1
GATAD2B	GATA zinc finger domain containing 2B	2	1	0	1

*** 7mer-m8: An exact match to positions 2-8 of the mature miRNA (seed sequence and position 8)**

****7mer-A1: An exact match to positions 2-7 of the mature miRNA (seed sequence) followed by an Adenine**

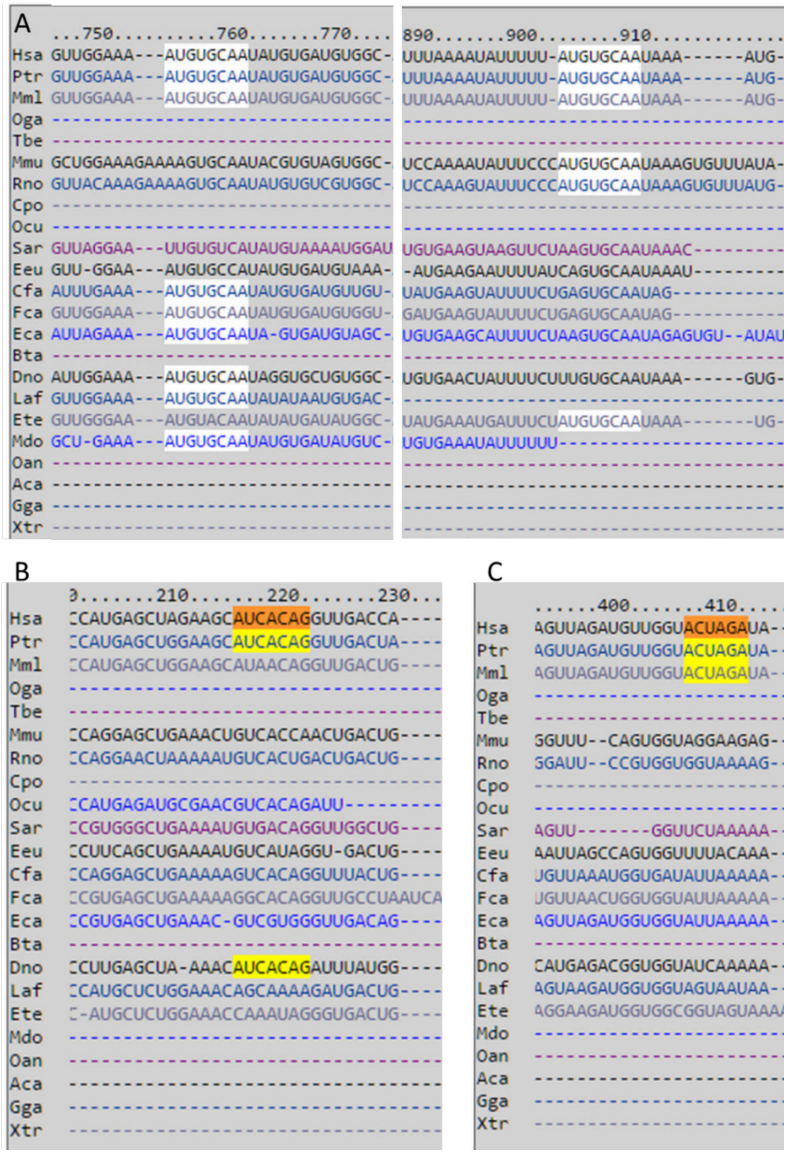


Figure 5.6 Targetscan alignment of the CD69 3′ UTR in different mammalian species. A) Targetscan predicted Hco-miR-5352 binding sites highlighted in white. B) Sequence search for Hco-miR-43 binding sites, showing a 7 base match and C) Hco-miR-61 site with a 6 base match. The Hco-miR-43 and Hco-miR-61 binding sites are not recognised by Targetscan as valid targets. Hsa = *Homo sapiens*, Ptr = *Pan troglodytes*, Mml = *Macaca mulatta*, Oga = *Otolemur garnetti*, Tbe = *Tupaia belangeri*, Mmu = *Mus musculus*, Rno = *Rattus norvegicus*, Cpo = *Cavia porcellus*, Ocu = *Oryctolagus cuniculus*, Sar = *Sorex araneus*, Eeu = *Erinaceus europaeus*, Cfa = *Canis familiaris*, Fca = *Felis catus*, Eca = *Equus caballus*, Bta = *Bos Taurus,* Dno = *Dasypus novemcinctus*, Laf = *Loxodonta africana*, Ete = *Echinops telfairi*, Mdo = *Monodelphis domestica*, Oan = *Ornithorhynchus anatinus*, Aca = *Anolis carolinensis*, Gga = *Gallus gallus*, Xtr = *Xenopus tropicalis*. Numbering along the top shows the nucleotide position within the 3' UTR from the stop codon.

5.2.4 Experimental validation of CD69 as a target of Hco-miR-5352

CD69 (Cluster of Differentiation 69) is a member of the calcium dependent lectin superfamily of type II transmembrane receptors (Sánchez-Mateos and Sánchez-Madrid, 1991). Activation of T lymphocytes induces the expression of CD69 and it appears to be one of the earliest inducible markers of T lymphocyte activation (López-Cabrera *et al.*, 1993). If Hco-miR-5352 is targeting a host mRNA, this particular target may be important in the host-parasite interaction and could influence the survival of *H. contortus* within the host. In an attempt to experimentally confirm this prediction, a dual luciferase assay was used to investigate whether an interaction could be demonstrated between Hco-miR-5352 and CD69 *in vitro*.

HEK293 cells were transfected with a combination of three plasmids; a miRNA expression plasmid containing the Hco-miR-5352 cluster, a Firefly plasmid containing the 3′ UTR of CD69 and a Renilla plasmid. The Renilla plasmid serves as an internal control and provides a measurement of the transfection efficiency and the baseline transcription levels, and normalises differences in cell viability, and the efficiency of cell lysis. If Hco-miR-5352 targets the 3′ UTR, it should cause the CD69-Firefly mRNA to degrade reducing the expression of Firefly luciferase signal.

5.2.4.1 Preliminary experiments to optimise the system

The phRG-TK Renilla plasmid was obtained from Dr. Jane Kinnaird and cloned into TOP10 *E. coli* cells. In preliminary experiments, 0, 12.5, 25, 50 and 100 ng of both the Firefly and Renilla plasmids were transfected into HEK293 cells with a view to optimising the signal strength while maintaining cell viability.

The lipofectamine LTX protocol suggests using a maximum of 100ng of DNA to transfect the cells. The lower the amount of DNA, the less toxic the conditions would be to the cells. From preliminary experiments, 25 ng of Firefly plasmid and 1 ng of Renilla plasmid gave a high luciferase reading and did not appear to affect cell viability (data not shown) and these concentrations were used in subsequent experiments. The luciferase plasmids were transfected into HEK293 cells along with 50 ng of either of the Hco-miR-5352 expressing plasmids or the

corresponding empty plasmid. Both miR expression plasmids contain the entire Hco-miR-5352 cluster on either a pCMV backbone or a pEGFP backbone (see Table 2.2 Chapter 2.7). As both miRNA plasmids contain a fluorescent GFP marker, transfection efficiency was determined using FACS analysis. In the representative experiment shown in Figure 5.7, over 70% of cells were fluorescent.

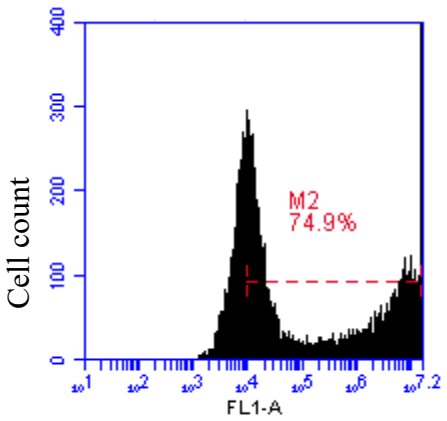


Figure 5.7 FACS analysis of HEK293 transfected cells. HEK293 cells were transfected with 50 ng of pEGFP empty vector. Cells were processed for FACS. The red line indicates the level of fluorescence above the background fluorescence determined from untransfected cells.

Several experiments were carried out using these conditions and Figure 5.8 shows the results of three separate experiments in which HEK293 cells were transfected with a combination of the Firefly (25 ng), Renilla (1 ng) and the Hco-miR-5352 expressing plasmids (50 ng). As expected, if miR-5352 targets the CD69 3' UTR, transfection with the pEGFP miR-5352 vector showed a 67% decrease in the ratio of Firefly to Renilla (P= 0.0217, Mann-Whitney test). However, the pCMV miR-5352 vector showed an increase in the ratio of Firefly/Renilla countcompared to the empty vector (P= 0.0047, Mann-Whitney test).

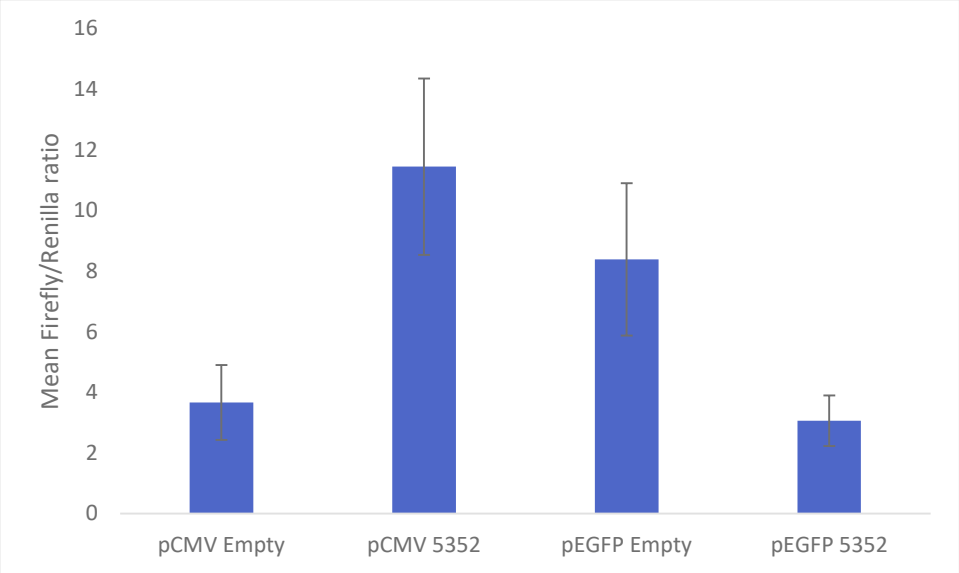


Figure 5.8 A dual luciferase assay to assess the interaction between miR-5352 and the 3′ UTR of CD69. Cells were co-transfected with 25 ng pMirTarget-CD69 Firefly, 1 ng phRG-TK Renilla and 50 ng of either pCMV or pEGFP plasmids containing the Hco-miR-5352 cluster or the empty vector. Each combination was carried out in triplicate and the experiment was repeated three times. 48 hours after transfection, cell extracts were prepared to measure CD69 expression, which is expressed as ratio of Firefly /Renilla luciferase. Graph shows the mean ratio +/- standard deviation of three separate experiments.

The presence of a Firefly and Renilla luciferase signal confirms that both plasmids have been successfully transfected into the cell, but the two miRNA-expressing plasmids gave opposing results. Using the pEGFP vector results in a significant decrease in luciferase levels, suggesting that the Hco-miR-5352 cluster is indeed able to downregulate CD69 expression. However, the pCMV vector showed the opposite effect. It is unclear what underlies this difference, as both plasmids use the same Cytomegalovirus (CMV) promoter. One possible explanation is that one plasmid may not produce any miRNA following transfection. Both the pCMV and pEGFP vectors contain a GFP tag that should be visible in transfected cells. To verify that both miRNA plasmids are being successfully transfected into the cells and transcribed, transfected cells were viewed under a fluorescence microscope 24 hours after transfection. Figure 5.9 shows UV microscopy images of HEK293 cells, with panels A and B showing the empty pCMV and pEGFP vectors, respectively. In both, a significant proportion of the cells are fluorescent demonstrating the success of transfection. However, panels C and D show the pCMV and pEGFP vectors containing the Hco-miR-5352 cluster where no detectable fluorescence was observed.

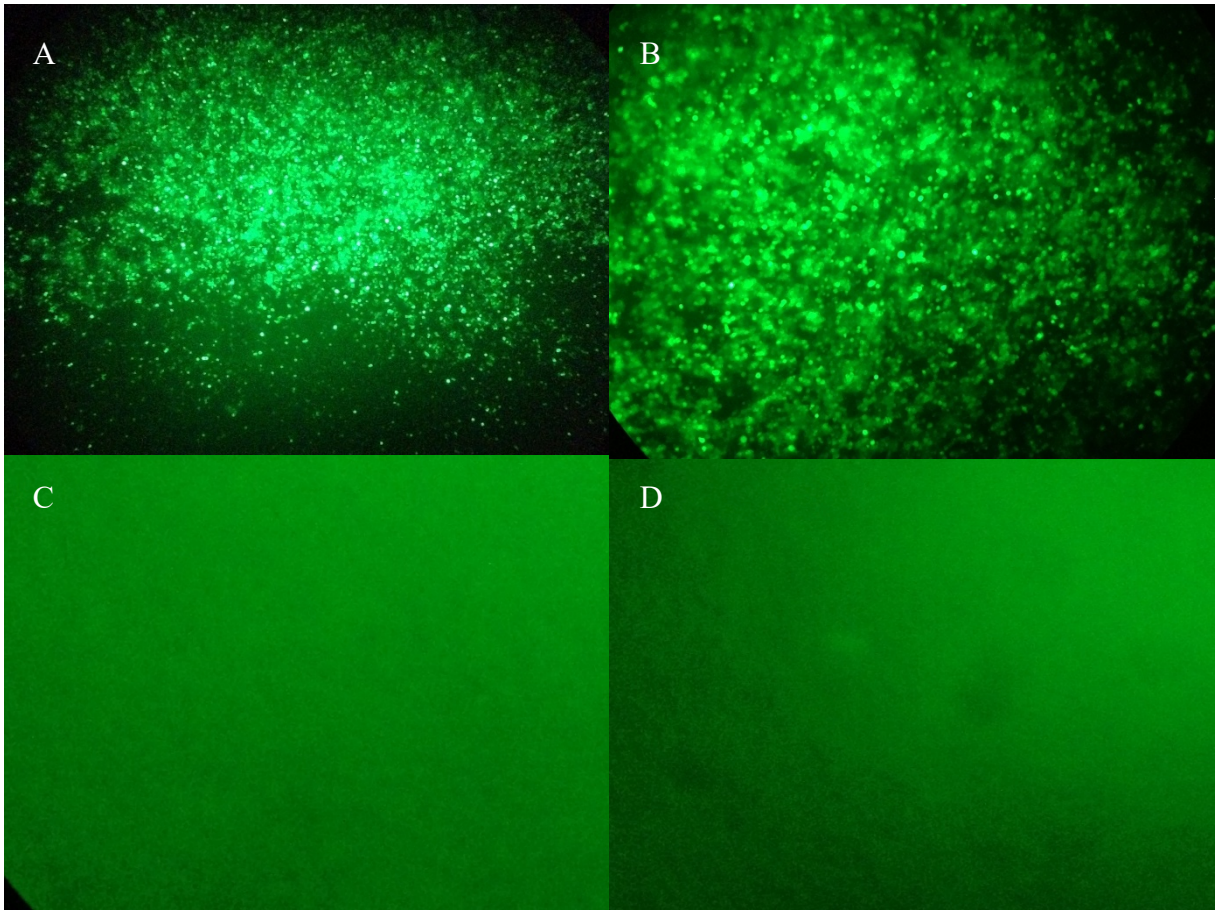


Figure 5.9 Fluorescence microscope images of HEK293 cells 24 hours after transfection with 25 ng pMirTarget-CD69, 1 ng phRG-TK Renilla and 50 ng of Hco-miR-5352 plasmids. A) Cells transfected with the empty pCMV vector, B) Cells transfected with empty pEGFP vector, C) pCMV-Hco-miR-5352 vector, D) pEGFP Hco-miR-5352 vector. Images were captured using a Leica DM IRB inverted microscope with a BP 545/30 nm excitation filter.

The absence of a GFP signal for both Hco-miR-5352-containing vectors suggests either that the Hco-miR-5352 plasmids are not being successfully transfected into the cells or that the GFP is not active. If the presence of the miRNA disrupts the expression of GFP, the plasmids may still be producing Hco-miR-5352. To examine this possibility, cells were transfected with the empty vector or the Hco-miR-5352 containing vectors and analysed by qPCR. In this experiment, an additional positive control was included, an Hco-miR-5352 microRNA mimic (Dharmacon). Furthermore, a siGLO red transfection indicator was also included as a visible measure of transfection efficiency and as a negative control for miR-5352 mimic expression as the mimic is not fluorescent.

Fluorescence microscopy was first used to confirm that the indicator siGLO can be transfected into HEK293 cells (Figure 5.10). Panel A shows a significant

amount of red fluorescence. Flow cytometry analysis identified that 74% of the HEK293 cells were successfully transfected (Figure 5.10B) indicating that the transfection conditions were working efficiently.

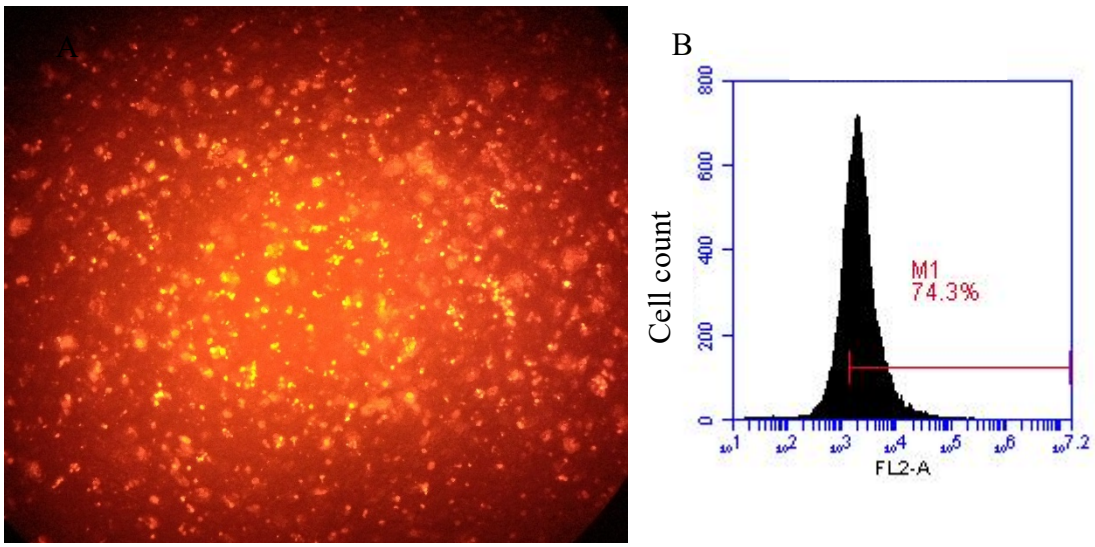


Figure 5.10 HEK293 cells can be efficiently transfected with siGLO red transfection indicator. Cells were transfected for 24 hours with 150nM siGLO, and then excess siGLO washed off prior to viewing. A) Microscopy photograph of siGLO transfected cells, red spots indicate fluorescent cells, photographed using a Leica DM IRB inverted microscope with a BP 620/60 excitation filter. Cells were then processed for FACS analysis. B) Red line indicates fluorescent cells, showing 74.3% of live cells expressed detectable fluorescence.

HEK293 cells were then transfected with 25 ng of pMirTarget-CD69, 1 ng phRG-TK Renilla and 50 ng of miRNA plasmid vectors or 150 nM of miRNA mimic. RNA was extracted for qRT-PCR analysis to determine if Hco-miR-5352 could be detected *(*see Figure 5.11). The negative controls showed no expression of Hco-miR-5352 and while the pCMV vector did produce some Hco-miR-5352, the level was extremely low compared to the Hco-miR-5352 mimic. However, the results do not explain the luciferase data. The pEGFP vectors show no Hco-miR-5352 expression, yet the pEGFP-Hco-miR-5352 vector was able to down-regulate pMirTarget-CD69. The experiment suggests that the miRNA plasmid vectors are not an effective means of driving miRNA expression. In contrast, cells transfected with the Hco-miR-5352 mimic showed an extremely high signal in the RT-PCR compared to the plasmid vectors and so the mimic was used in all subsequent experiments.

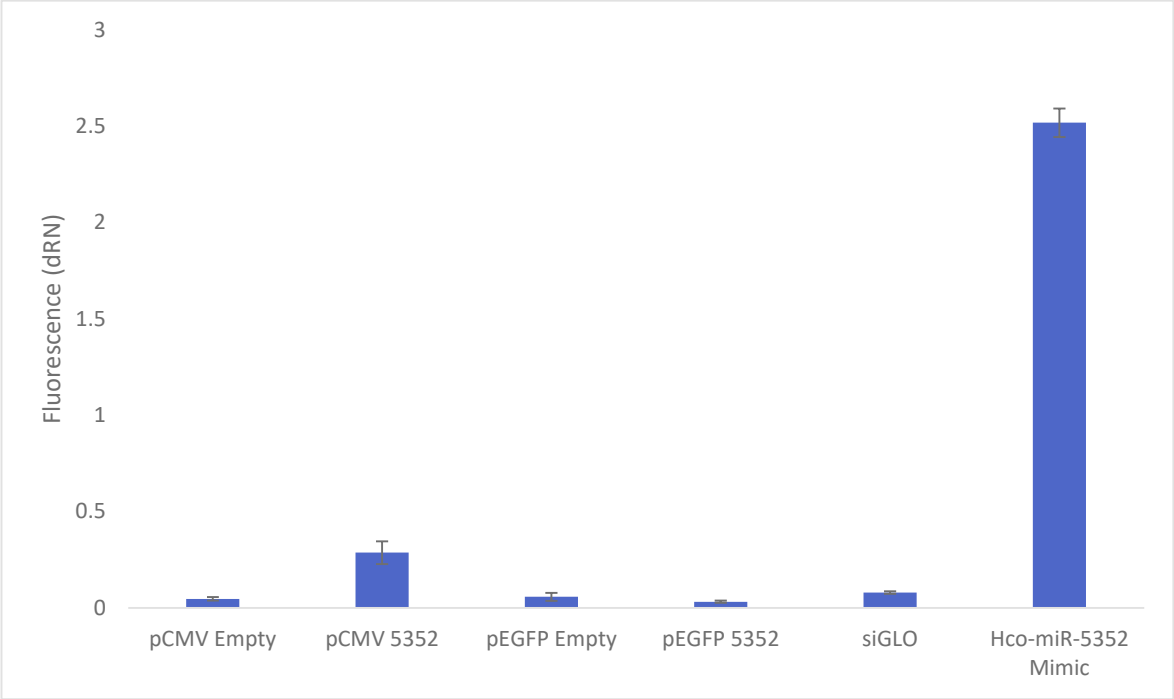


Figure 5.11 qRT-PCR to detect presence of Hco-miR-5352. Cells were co-transfected with 25 ng pMirTarget-CD69, 1 ng phRG-TK Renilla and 50 ng of either pCMV or pEGFP plasmids containing Hco-miR-5352 or the empty vector. Cells were also transfected with 150 nM of microRNA mimic Hco-miR-5352 or siGLO red transfection indicator. 48 hours after transfection, cells were lysed and RNA extracted. RT-PCR was carried out with 3 technical replicates. Data shown as the mean ± standard deviation.

The results of the FACS analysis and qRT-PCR experiments showed that the miRNA mimic can be successfully transfected into the cells. To investigate if this miRNA is able to affect expression of the CD69-firefly luciferase, HEK293 cells were transfected with 25 ng pMirTarget-CD69, 1 ng phRG-TK Renilla and 75 nM, 150 nM or 300 nM of microRNA mimic. Cells were cultured for 48 hours in medium before being lysed and fluorescence measured. Figure 5.12 shows the Firefly to Renilla luciferase ratio. The cells transfected with the *C. elegans* mimic had similar fluorescence values to cells transfected only with the Firefly and Renilla plasmids (data not shown). Cells transfected with the Hco-miR-5352 mimic at all concentrations showed significantly lower ratios compared to cells transfected with the *C. elegans* negative control miRNA mimic (P-value ≤ 0.001, Mann-Whitney test). In the experiment shown in Figure 5.12, 75 nM of miR-5352 mimic showed an average 50% decrease in the ratio, while 150 nM and 300 nM mimic resulted in a 60% decrease. It was noted that the highest concentration of control *C. elegans* miRNA mimic reduced the luciferase ratio compared to the lower levels, although this decrease was not significant (P-value 0.062, Mann-

Whitney test). From these data, a significant decrease in the luciferase ratios was observed in the cells with Hco-miR-5352 present, indicating that Hco-miR-5352 miRNA mimic is able to interact with the pMirTarget-CD69 and cause its down-regulation.

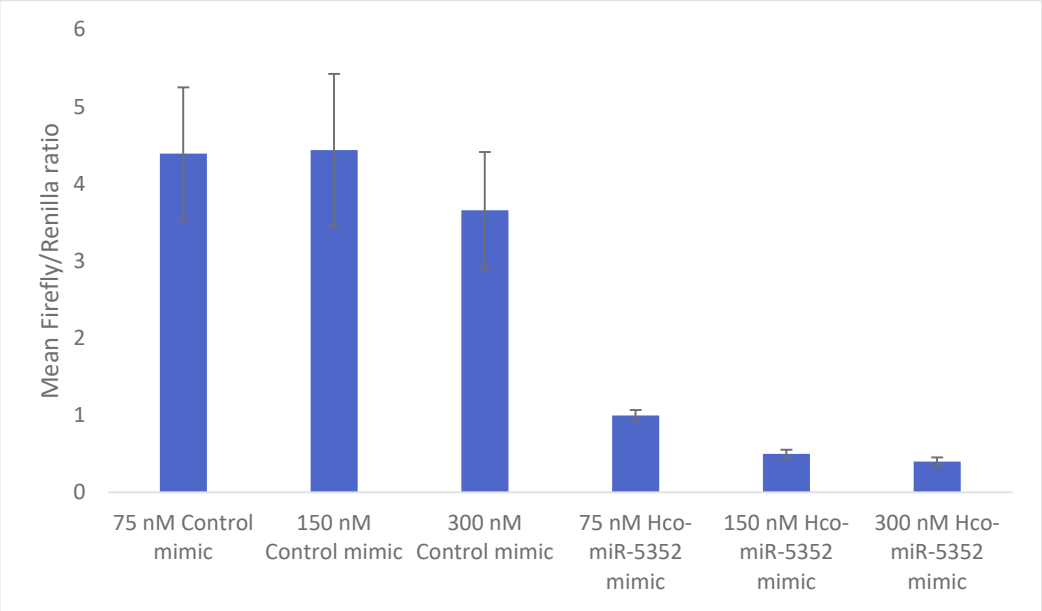


Figure 5.12 Hco-miR-5352 mimic down-regulates the expression of CD69 in a dual luciferase assay. HEK293 cells normalised to non-transfected cells. Cells were co-transfected with 25 ng pMirTarget-CD69, 1 ng phRG-TK Renilla and 75 nM, 150 nM or 300 nM of microRNA mimic Hco-miR-5352 or the *C. elegans* negative control miRNA. Each combination was replicated six times and the experiment was repeated three times. 48 hours after transfection, cell extracts were prepared to measure CD69 expression, which is expressed as ratio of Firefly luciferase/Renilla luciferase. Graph shows the mean ratio +/- standard deviation of a single representative experiment.

To confirm this finding, the CD69 3' UTR was scrambled to disrupt the predicted Hco-miR-5352 target sites whilst keeping the same distribution of nucleotides. Target prediction showed that the mutated CD69 3' UTR did not possess any binding sites for either Hco-miR-5352 or Cel-miR-67. HEK293 cells were transfected with 25 ng of the mutated pMirTarget-CD69 or the wild-type pMirTarget-CD69, 1 ng phRG-TK Renilla and 150 nM of microRNA mimic. Cells were cultured for 48 hours in medium before being lysed and fluorescence measured. Figure 5.13 shows the result of a single representative experiment. The graph shows that 150 mM Hco-miR-5352 resulted in an average decrease of 72% compared to the no-mimic cells (p-value ≤ 0.01, Mann-Whitney test). However, the Cel-miR-67 mimic was also able to reduce the Firefly/Renilla ratio by 48% compared to the no mimic cells (p-value ≤ 0.01, Mann-Whitney test). The

decrease in Firefly/Renilla ratio between the Cel-miR-67 and Hco-miR-5352 mimic transfected cells is also significant (p-value ≤ 0.01, Mann-Whitney test). In the parallel transfection with the wild-type pMirTarget-CD69, Cel-miR-67 controll mimic did not cause any significant change in the Firefly/Renilla ratio (p-value = 0.07, Mann-Whitney test) and showed a similar percentage reduction as Figure 5.12. This experiment is difficult to interpret as, surprisingly, transfection of the Hco-miR-5352 mimic significantly reduced the Firefly/Renilla ratio as did the control Cel-miR-67 mimic when using the mutated CD69 3' UTR.

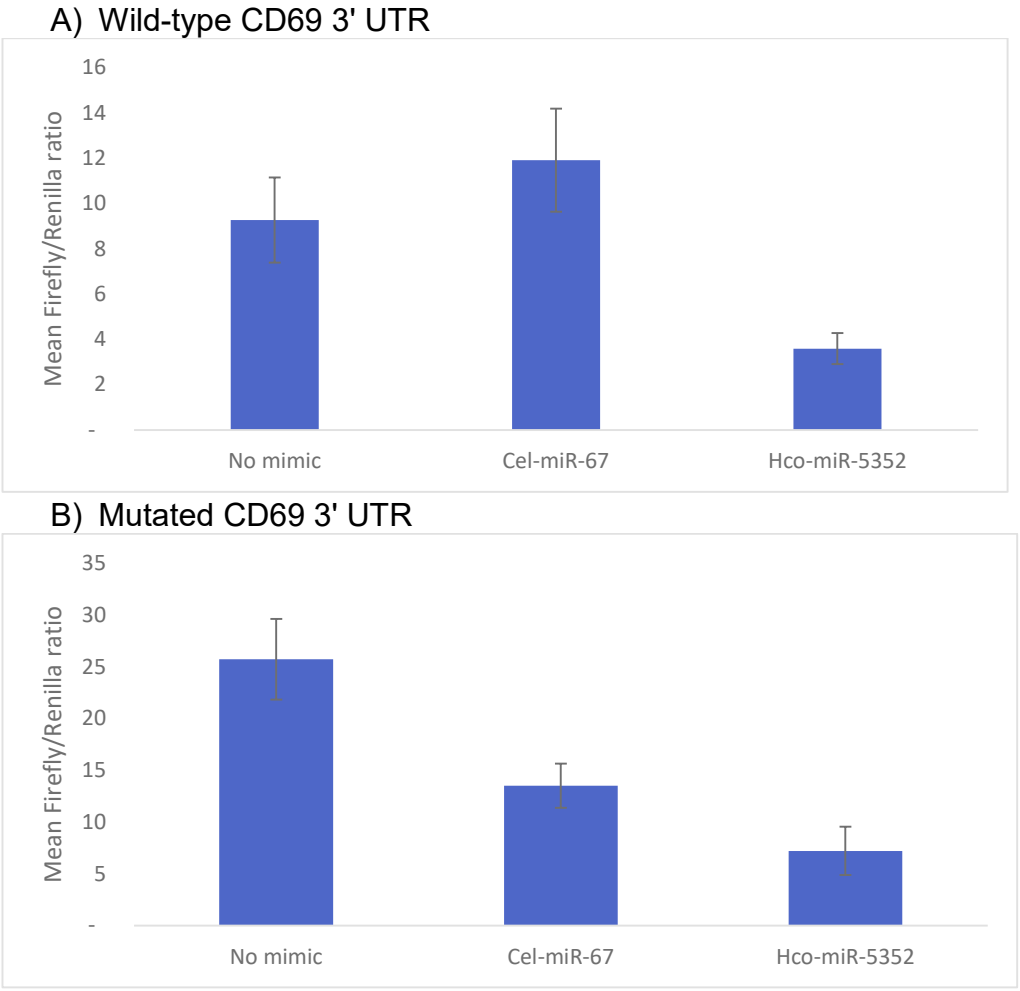


Figure 5.13 Hco-miR-5352 mimic interacts with the CD69 3' UTR. HEK293 cells were co-transfected with 25 ng pMirTarget-CD69, 1 ng phRG-TK Renilla and 150 nM of microRNA mimic Hco-miR-5352 or the *C. elegans* negative control miRNA. A) Cells were transfected with the wild-type CD69 3' UTR. B) Cells transfected with the mutated CD69 3' UTR. Each combination was replicated six times and the experiment was repeated three times. 48 hours after transfection, cell extracts were prepared to measure CD69 expression, which is expressed as ratio of Firefly luciferase/Renilla luciferase. Graph shows the mean ratio +/- standard deviation of a single representative experiment.

5.2.4.2 The effect of Hco-miR-5352 on CD69 induction in a Jurkat T cell line.

The dual luciferase assay showed that Hco-miR-5352 is able to downregulate the CD69 3′ UTR in a synthetic luciferase reporter construct. To investigate the effect of miR-5352 in a more biological system, a human T cell line (Jurkat) was used. This cell line is able to express CD69 on the cell surface in response to stimulation of the cells with phorbol 12-myristate 13-acetate (PMA).

The expression of CD69 on Jurkat cells cultured with differing concentrations of PMA from 2.5 ng/ml to 20 ng/ml was first investigated. Cells were exposed to PMA for an hour and then left for 24 hours before staining with anti-CD69 antibody or the isotype control antibody. They were then analysed using FACS. Figure 5.14 shows the FACS plot for four different concentrations of PMA with a DMSO control. The lowest two concentrations of PMA had negligible staining compared to the control, rising to 36% at 10 ng/ml and 80% at 20 ng/ml. 20 ng/ml of PMA was used in all further experiments in order to optimize CD69 expression on as many cells as possible.

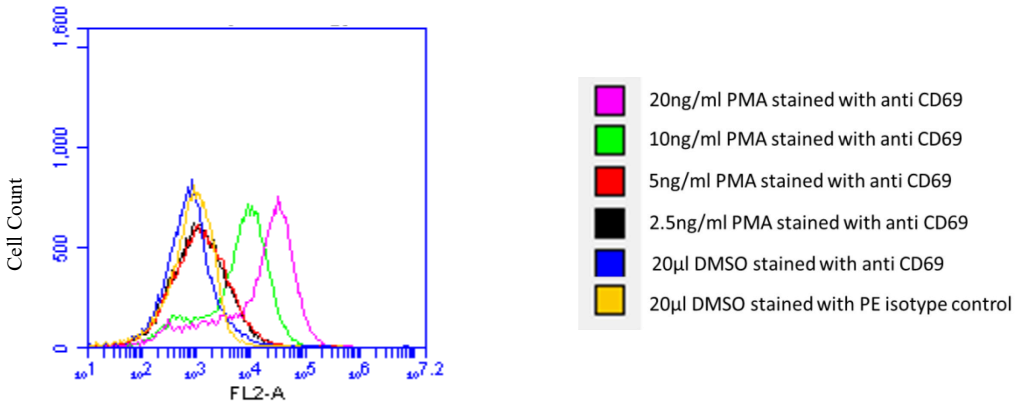


Figure 5.14 CD69 expression on Jurkat cells stimulated with varying concentrations of PMA. PMA was dissolved in DMSO and diluted to the stated concentrations in RPMI. DMSO controls were used to assess the effect of the DMSO. Cells were analysed by FACS following staining with anti-CD69 antibody or isotype control antibody. A shift to the right indicates positive expression of CD69.

The Amaxa™ nucleofector kit was chosen to transfect the cells as the kit has protocols designed to transfect Jurkat T cells. The manual suggested two different nucleofection programmes, (programme 1 – high viability, programme 5 – high expression). To assess the effectiveness of the protocol, Jurkat cells were transfected using both programmes using the pmaxGFP control plasmid provided in the Amaxa™ kit. Figure 5.15 shows the number of live fluorescent cells with both programmes. From the overlay histogram, 74% of the live cells transfected using programme 5 were fluorescent, whereas only 40% of cells were fluorescent using programme 1.

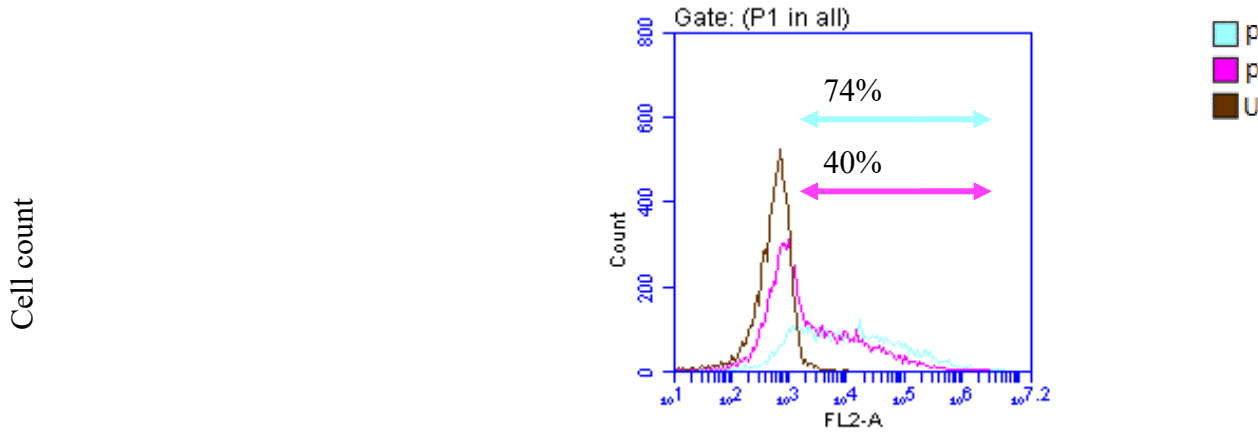


Figure 5.15 Overlay histogram plot showing the percentage of live cells that are fluorescent after the Jurkat cells were transfected with the pMax plasmid provided with the Amaxa™ transfection kit using either Programme 1 for higher cell survival or Programme 5 for high expression.

However, programme 1 showed higher cell survival following nucleofection (data not shown) and was selected for all further experiments. To confirm that microRNA mimics could be successfully transfected into Jurkat cells using the nucleofection method, siGLO was used as a transfection marker. Figure 5.16 shows a comparison of fluorescence between the untransfected and siGLO transfected cells. 98% of live cells were successfully transfected with siGLO (Figure 5.16). Transfection of cells with siGLO appears to give a higher viability than in cells transfected with plasmid (compare Figure 5.15 and Figure 5.16).

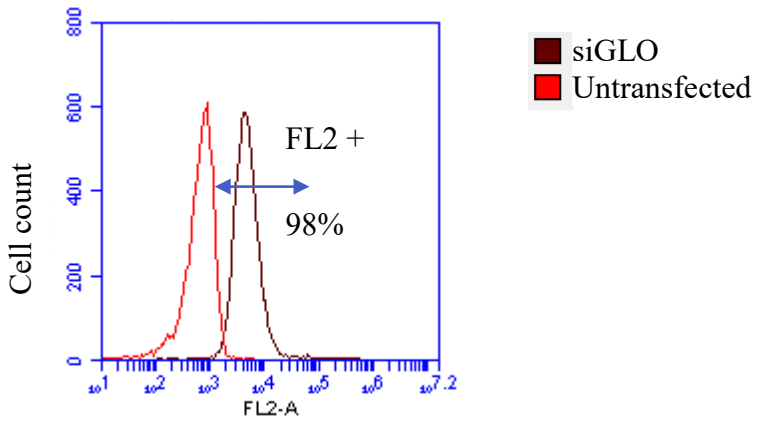


Figure 5.16 Overlay histogram plot showing the number of live cells that are fluorescent. Jurkat cells were transfected with 300 nM of siGLO transfection indicator. Cells were transfected with siGLO using the Amaxa™ nucleofection kit then processed for FACS analysis 24 hours. The red line shows the background fluorescence from untransfected cells.

The FACS analysis confirmed that miRNA mimics can be successfully transfected into Jurkat cells very efficiently. To examine the effect of Hco-miR-5352 on the expression of CD69, Jurkat cells were transfected with Hco-miR-5352 mimic for 24 hours, then stimulated with 20 ng/ml of PMA. Figure 5.17 shows the percentage of live cells and the percentage of live cells that were positive for CD69. Approximately 60% of live cells expressed CD69 following stimulation with PMA. However, when miR-5352 mimic was transfected into the cells, there was no significant difference in CD69 staining (P-value = 0.96 CD69 expression +/- miR-5352). These results suggest that the miRNA mimic was not able to interact with CD69 induced on Jurkat cells and downregulate its expression. Figure 5.17 also shows that cell survival after transfection was low, ranging from 50% to 60%. Cell survival in the untransfected group was higher, averaging 70%.

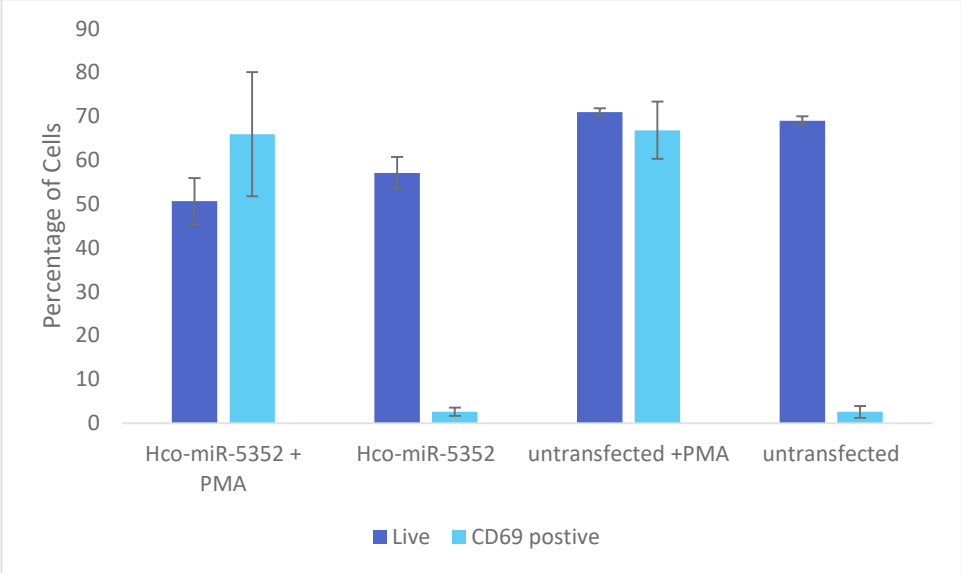


Figure 5.17 miR-5352 mimic does not affect CD69 expression in Jurkat cells following PMA stimulation. Cells were transfected with Hco-miR-5352 mimic or left untransfected for 24 hours and then either stimulated with PMA for 24 hours or left unstimulated before FACS analysis. Graph shows the percentage of live cells and the percentage of live cells that stained positive for CD69. The mean of three biological replicates is shown ± standard deviation.

5.2.4.3 Transfecting Jurkat cells with Lipofectamine®.

The nucleofection method did not show any significant difference between the number of live cells that were CD69 positive between the untransfected and Hco-miR-5352 transfected cells. However, Jurkat cell survival following nucleofection was relatively low as only 50% of the cells were alive during the FACS analysis. As Lipofectamine® had been extremely efficient in transfecting HEK293 cells, attempts were also made to transfect Jurkat cells using the Lipofectamine® reagent.

To test the transfection efficiency of Lipofectamine®, Jurkat T cells were transfected with either 150 or 300 mM of the transfection indicator siGLO using Lipofectamine® LTX. Figure 5.18 shows the average results of three replicate experiments. Cell survival using lipofectamine was higher compared to nucleofection. Doubling the amount of siGLO used in the transfection increased the percentage of siGLO positive cells from 60% to 88% (P-value < 0.0001, Mann-Whitney). However, increasing the amount of siGLO lowered the number of live cells from 86% to 72% (P-value < 0.0001, Mann-Whitney).

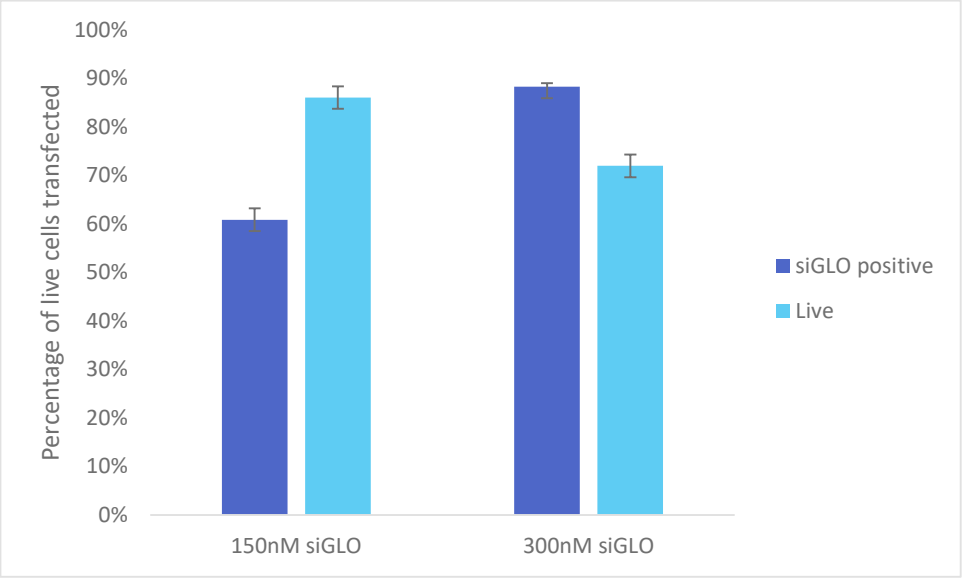


Figure 5.18 The efficiency of transfection of Jurkat T cells using lipofectamine. Cells were transfected with either 150 nM or 300 nM of siGLO transfection indicator and left for 48 hours before FACS analysis. Graph shows the percentage of live cells and the percentage of live cells that were positive for siGLO. The mean of three biological replicates is shown ± standard deviation.

From Figure 5.18, increasing the amount of siGLO being transfected into the cells had a significant effect on cell survival and on the number of cells being transfected. To control for this effect, a negative control mimic, *C. elegans* Cel-miR-67 was used. In these experiments, Jurkat cells were transfected with either the negative control mimic or Hco-miR-5352 and left for 24 hours. The cells were then stimulated with 20 ng/ml of PMA for an hour. Cells were taken at 8 and at 24 hours for FACS analysis. At 8 hours, less than 40% of the cells in any of the conditions were CD69 positive. The number of CD69 positive cells at 24 hours was far higher, nearer to 90%. The data shown in Figure 5.19 show that there was no effect of the miR-5352 mimic on CD69 expression at either 150 or 300 nM (for both, p-value > 0.05, Mann-Whitney).

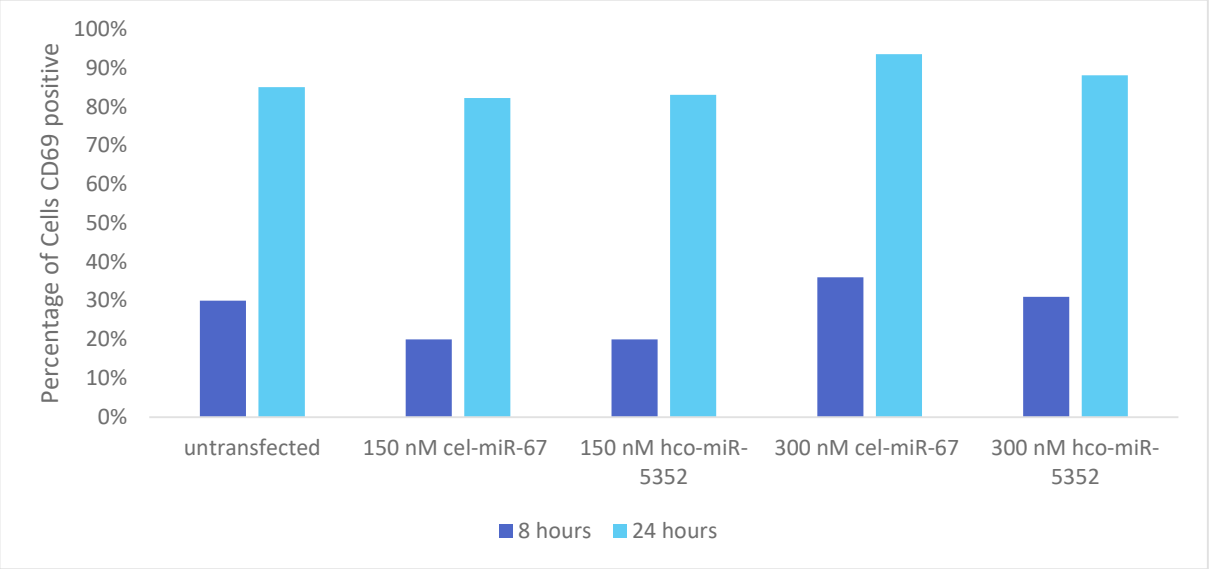


Figure 5.19 Hco-miR-5352 does not affect PMA-induced CD69 expression following transfection of Jurkat T cells. Cells were transfected with 150 or 300 nM of Hco-miR-5352 mimic or the negative control Cel-miR-67 mimic and were analysed at 8 and 24 hours after PMA stimulation. Data shown is from a single experiment.

In a final attempt, Jurkat cells were transfected with 300 nM of Hco-miR-5352 or the negative control Cel-miR-67. This experiment was repeated three times. Cells were stimulated with PMA 24 hours after transfection and analysed by FACS 24 hours after stimulation. While a higher percentage of cells transfected with Cel-miR-67 were CD69 positive than the untransfected control, the difference was not statistically significant (p-value = 0.88, Mann-Whitney) (Figure 5.20). The difference between the Cel-miR-67 and Hco-miR-5352 was also not statistically significant (p-value = 0.31, Mann-Whitney) (Figure 5.20), suggesting that the presence of the mimic had no effect on the expression of CD69 by Jurkat cells when induced by PMA.

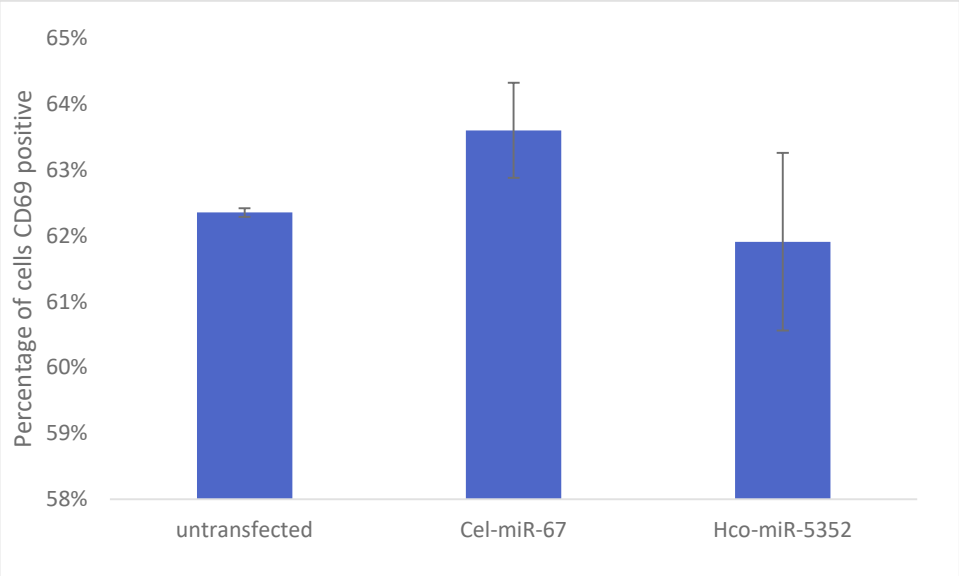


Figure 5.20 Hco-miR-5352 does not affect CD69 expression on Jurkat T cells. Cells were transfected with 300 nM of Hco-miR-5352 or the negative control Cel-miR-67 and analysed at 24 hours after PMA stimulation at 20 ng/ml. The mean of three experiments is shown ± standard deviation.

5.3 Discussion

This chapter presents an attempt to elucidate the target of Hco-miR-5352 through the use of bioinformatics and experimental techniques. The presence of miRNAs in the ES implies that these miRNAs may have a function outside of the adult *H. contortus*. Two potential targets exist: the sheep host or other *H. contortus* adults. If the miRNA targets sheep cells, the miRNA would have to be able to enter the cells and should be detectable. Three miRNAs, Hco-miR-5352, Hco-miR-5895 and Hco-miR-5960 were selected that were highly abundant in the ES library and did not have any homologues in the most current version of the sheep genome (Oar v3.1) (Jiang *et al.*, 2014). From qRT-PCR experiments, Hco-miR-5352 and Hco-miR-5895 could be detected from both abomasal and lymph node tissues from sheep infected with *H. contortus*. However, Hco-miR-5960 was detected in both tissue types and was amplified from the uninfected tissue as well. Whilst the initial search of the sheep genome identified no homologues, further BLAST searches of the NCBI nucleotide collection database (nr) limited to *Ovis aries* identified 100 sequences that have between 50 to 65% query coverage with 100% identity to the Hco-miR-5960 sequence. All of the sequences identified are predicted mRNA sequences and cover a range of different proteins from von Willebrand factor C domain containing proteins to serine/threonine kinase 11. The large number of predicted mRNA sequences with similarity to Hco-miR-5960 suggests that there may be non-specific amplification of a sequence similar to the *H. contortus* miR-5960 sequence. However, BLAST searches with the Hco-miR-5352 and Hco-miR-5895 sequences also identified mRNA sequences with 50 to 65% query coverage and 100% identity but were not detected in the uninfected sheep, suggesting that the Hco-miR-5960 may not be originating from the sheep. In addition, there are large differences in the abundance of miRNAs between sheep. This is especially noticeable between the two infected abomasal tissues when looking at Hco-miR-5895 expression, although this difference is not as large in the lymph node samples. This difference may be due to the level of infection in the different sheep or the low number of biological replicates. If the miRNA levels are reflecting the infection severity, it may be possible to utilise miRNAs as a biomarker of *H. contortus* infection severity.

On the other hand, the qRT-PCR showed that Hco-miR-5352 and miR-5895 are present in infected sheep tissue compared to the uninfected sheep and at a level that the miRNA can be amplified. The presence of *H. contortus* specific miRNAs in the sheep tissue may indicate that these particular miRNAs have a specific role within sheep tissue. Since the discovery that miRNAs can be released from mammalian cells and are stable in the body fluids, there has been increased interest in miRNAs as possible non-invasive biomarkers. Parasite specific miRNAs from *S. mansoni* have been found in serum from mice and humans infected with the parasite (Hoy *et al.*, 2014). In addition, parasite miRNAs have also been identified in plasma or nodule fluid of animals infected with filarial parasites (Britton *et al.*, 2015). For gastro-intestinal (GI) nematodes such as *Heligmosomoides polygyrus,* parasite miRNAs were not found in the circulation and our initial studies support this observation (Winter *et al.,* unpublished data) and suggest that instead GI nematodes may release miRNAs into the local gut environment (this chapter). Detailed analysis of *H. polygyrus* ES products showed that adult worms secreted exosomes containing miRNAs. These nematode exosomes were able to suppress T –helper 2-type responses in mice and Dusp1, a regulator of MAPK signalling, was found to be repressed by nematode miRNAs in a luciferase assay (Buck *et al.*, 2014). This phenomenon is not just limited to parasites. Zhang analysed global miRNA levels in a range of mammals and found there is selective uptake of certain miRNAs from the consumption of rice (Zhang *et al.*, 2012). One particular miRNA, mi168a is present in human and animal serum and was shown to mediate expression of the liver gene LDLRAP1 (Zhang *et al.*, 2012). This is not limited to plants as RNAs from a range of sources including bacteria and fungi have been found in human plasma (Wang *et al.*, 2012b), however, other studies have disputed these findings (Snow *et al.*, 2013).

Although the presence of miRNAs originating from an external source is well established, the function of these miRNAs is still unknown. Computational methods are important in identifying mRNA targets of miRNAs. As validating a potential miRNA target in the laboratory is costly and time consuming, computational approaches can assist in narrowing down potential target genes. Hco-miR-5352 is conserved in selected parasitic nematodes and the nematodes that possess Hco-miR-5352 parasitise a wide range of hosts suggesting that the targets are likely to be well conserved across different organisms. To identify

targets, a miRNA target prediction program, Targetscan, was used to predict mammalian targets for these miRNAs. Targetscan contains a database of 17,840 3′ UTR sequences from 84 different species and is able to take into account conservation of miRNA binding sites.

Targetscan predicted 500 potential mammalian targets of Hco-miR-5352, which is the maximum number of targets produced by the programme. To narrow down the potential targets, mRNAs in common between the miRNAs of the cluster were examined. Since miRNA clusters are co-expressed, members of the Hco-miR-5352 cluster could be acting together to co-ordinately control a particular pathway or target the same genes (Hausser and Zavolan, 2014; Xu and Wong, 2008).

ADAMTS5, CNOT6L and ZNF148 are predicted to be targeted by three of the four miRNAs in the cluster. No gene was targeted by all four miRNAs. However, the three targets had relatively weak energy scores associated with the interaction compared to the top hit of Hco-miR-5352, CD69, suggesting that such interactions were unlikely. The IPA analysis identified a number of pathways based on the Targetscan targets, however none of the pathways had a statistically significant p-value, although the Wnt/β-catenin signalling pathway was closest to significance. Targetscan puts CD69 at the top of the list of 500 targets for Hco-miR-5352, suggesting that it is one of the most likely target genes. The binding sites in CD69 were further confirmed using three additional programs; Probability of Target Accessibility (PITA), miRanda and RNAhybrid (data not shown).

To demonstrate an interaction between Hco-miR-5352 and CD69, mammalian HEK293 cells were transfected with the miRNA and a dual luciferase assay system was used to readout the potential interaction. Initial experiments used two different miRNA expression plasmids, each expressing the whole Hco-miR-5352 cluster. Both miRNA plasmids contain a fluorescent GFP protein, but this signal could not be detected in HEK293 cells under a fluorescent microscope in the presence of the cloned miRNA sequence. Furthermore, qRT-PCR could not detect a significant level of Hco-miR-5352 expression in the transfected cells. It is possible that the plasmids are not being transfected into the cells, although

this seems unlikely as the empty plasmids can be easily visualised by UV microscopy. Alternatively, it is possible that they are not able to drive the expression of the miRNA.

From the same qRT-PCR experiment, a Hco-miR-5352 miRNA mimic could be detected in the transfected cells. These mimics are chemically modified double-stranded RNA molecules intended to mimic endogenous miRNAs. Once introduced into the system, the mimics are thought to enter the RISC and bind to the mRNA target. In the dual luciferase assay system, the Hco-miR-5352 mimic was able to decrease the level of the Firefly luciferase reporter gene cloned downstream of the CD69 3' UTR. In contrast, the Cel-miR-67 negative control mimic did not reduce Firefly luciferase levels, providing evidence of a specific interaction between Hco-miR-5352 and the CD69 3' UTR. Additional studies were carried out to further confirm this, using a mutated form of the CD69 3′ UTR which ought to disrupt the binding of miR-5352. However, the expression of the mutated CD69 plasmid was significantly affected by the presence of the Hco-miR-5352 mimic as well as the negative control *C. elegans* mimic, although no binding sites were predicted (Figure 5.13), complicating interpretation of the data and suggesting some non-specific interaction. The reason for this apparent interaction is not known.

Further experiments sought to confirm the possible interaction of miR-5352 and CD69 in a more biological scenario. Jurkat T cells were chosen due to their ability to express a high level of CD69 after stimulation with PMA. Transfection of these cells proved to be problematic. The nucleofection system initially chosen was designed to transfect Jurkat cells with 90% efficiency with 90% cell viability (http://bio.lonza.com). In experiments conducted with the miRNA mimic as well as the pmaxGFP® Vector provided in the kit, cell viability was less than 50%. Although the protocol had been optimised for Jurkat cells, it is likely that electroporation is detrimental to the cells. A transient transfection method, the lipofectamine method applied to the HEK293 cells, was used instead. The transient transfection achieved much higher cell viability, with over 60% cell survival depending on the amount of mimic being transfected, and over 70% of the cells containing the mimic.

Transfecting Jurkat cells with Hco-miR-5352 did not influence the level of CD69 expression compared to the untransfected cells or the Cel-miR-67 control transfected cells. Previous experiments have shown that PMA stimulation can activate T cells and elicit expression of CD69 on the cell surface (Brocardo *et al.*, 2001; Vance *et al.*, 1997; Wiskocil *et al.*, 1985). Jurkat cells were stimulated with a range of PMA concentrations to determine the optimum concentration required to obtain the highest level of CD69 expression and cell survival. 20 ng/mL of PMA was chosen as it gave a high level of CD69 expression and did not significantly impact cell viability. Jurkat cells transfected with control or Hco-miR-5352 mimic and then stimulated with PMA did not show significant difference in the level of CD69 expression.

PMA is a very strong non-biological stimulus (Barr *et al.*, 2007) and may potentially overpower the effect of the mimic. The assay used also examined the level of the protein and not the actual target of the mimic, the mRNA. Examining the number of CD69 transcripts could give a better indication as to whether or not the mimic is targeting CD69 mRNA, as the level of mimic is finite, whilst the number of cells and the number of CD69 mRNAs will increase throughout the experiment. A shorter time point was chosen to investigate if the Hco-miR-5352 had an impact when the level of CD69 expression is lower. At 8 hours, around 40% of the cells were CD69 positive, however the level of CD69 expression did not differ significantly following transfection with the control or the Hco-miR-5352 mimic. A more physiological stimulus, CD3/CD28 was also tested (Creson *et al.*, 1999), however, preliminary experiments using CD3/CD28 to stimulate Jurkat cells had a survival rate averaging 50%, of which, approximately 10% were CD69 positive (data not shown). Analysis of the CD69 3' UTR by Targetscan demonstrated that it was a possible target of several mammalian miRNAs including miR-17, miR-21, miR-25, miR-130/301 and miR-181. In the CD69 3' UTR, the binding sites for Hco-miR-5352 overlap significantly with miR-25, suggesting that miR-5352 might be competing for binding with miR-25.

In conclusion, studies carried out in this chapter sought to identify potential mammalian targets of the miR-5352 cluster using bioinformatics approaches. A single mRNA, CD69, was selected for additional study and while some evidence

was produced of an interaction between miR-5352 and the 3′ UTR of CD69 this could not be confirmed under all experimental conditions. The mRNA transcriptome of cells transfected with either the control miRNA mimic or the Hco-miR-5352 mimic is examined in more detail in Chapter 6.

6 Analysis of mRNA changes in MODE-K cells following transfection with miR-5352

6.1 Introduction.

The initial target prediction performed using Targetscan predicted 500 unique mammalian genes as potential targets of Hco-miR-5352 (see Chapter 5). Even considering only those sites that have an exact seed sequence match (8mers) still resulted in 166 unique genes. Whilst it is unlikely that miR-5352 targets all 500 of these genes, miRNAs have been shown to be able to target multiple genes. For example, in prostate cancer, miR-15a and miR-16 were shown to simultaneously target BCL2, CCND1 and WNT3A (Bonci *et al.*, 2008). Whilst any one miRNA can potentially target many genes, other studies have shown that miRNAs tend to target functionally related genes or components of the same pathway (Hausser and Zavolan, 2014; Kim *et al.*, 2011).

The presence of selective miRNAs in the ES and exosomes from *H. contortus* L4 and adult parasites, as well as in sheep tissue, suggested that these miRNAs may be involved in immunomodulation. Hco-miR-5352 was demonstrated to interact with the 3' UTR of CD69 in the dual luciferase assay (Chapter 5). However, there was no formal proof of that interaction in a more physiological system, although there are several reasons that could explain these data (see Chapter 5 for discussion). Since only CD69 was investigated in the previous experiments, it is unknown if Hco-miR-5352 might have any effect on other immune related genes.

A common method of identifying miRNA targets is to investigate gene expression following miRNA transfection. This has been performed in human cells where transfection of miR-124, a miRNA localised in the brain, was able to shift the mRNA expression profile of transfected cells towards that of the brain, whilst the muscle miRNA, miR-1 was able to shift it towards the muscle (Lim *et al.*, 2005). Whilst miRNAs are able to silence their targets by blocking translation, they are also able to cause mRNA degradation (Doench and Sharp, 2004; Hu and Coller, 2012). The resulting transcript changes can then be monitored using microarrays or by RT-PCR or RNA-Sequencing.

To further understand the effects of Hco-miR-5352 on global gene expression, experiments described in this chapter were carried out using a miR-5352 mimic transfected into a mouse gut epithelial cell line, called MODE-K. These cells originate from the epithelium of the mouse small intestine (Vidal *et al.*, 1993) and were chosen due to the large amount of data available on mouse cells compared to sheep cells . Two different types of experiment were carried out. For both, total RNA was obtained from cells transfected with either the Hco-miR-5352 miRNA mimic or from a control mimic based upon Cel-miR-67, which shows minimal identity to mouse miRNA sequences. In the first approach, a Qiagen RT^2 Profiler PCR array was used to examine the difference in expression patterns of a number of important immune cell surface markers as well as markers of innate and adaptive immune responses. In the second approach, RNA-Seq was carried out as this allows the interrogation of a large number of differential genes and does not rely on probes or primers. RNA-Seq is also not limited to existing genomic sequences and can take into account sequence variations.

6.2 Results

6.2.1 RT^2 Profiler

In this experiment, RNA from MODE-K cells transfected with Hco-miR-5352 mimic or Cel-miR-67 control mimic was isolated and these cell samples used to probe two separate mouse arrays; Mouse Cell Surface Markers and Mouse Innate and Adaptive Immune Responses as described in Chapter 2.24. Arrays were probed with RNA from a single sample of Hco-mir-5352 or control mimic-transfected cells.

Table 6.1 shows a summary of the C_T values obtained from the study presented as given by the RT^2 programme. Genes with a C_T value >35 were discarded from subsequent analysis because the software considers these genes to be undetectable.

Table 6.1 Summary of gene C_T values from both PCR arrays tested.

	Mouse cell surface markers	Mouse Innate and Adaptive immune response genes
Number of genes being tested	84	84
Number of genes that had a C_T value of above 35 or none detected in both cell samples	40	52
Number of genes with a C_T > 30 in cell sample and C_T < 30 in the other	3	2
Number of genes with the average C_T > 30 in both cell samples	24	20
Number of genes with C_T < 30 in both cell samples	17	10

For both PCR arrays, a large number of the genes, nearly half of the genes being tested, had C_T values which were higher than 35 in both the MODE-K cells transfected with Hco-miR-5352 mimic and with the negative control Cel-miR-67 mimic making these values erroneous. These were discarded from the analysis. A

much smaller group of genes (3 and 2 respectively from cell surface markers and innate and adaptive immune response) had a high C_T value in one cell sample type (low expression) and a low C_T value (high expression) in another; this group of genes were considered in the analysis. Genes with C_T values above 30 means that the expression of these genes is low in both control and test samples and C_T values < 30 suggest high expression. Out of the genes with at least one C_T value below the threshold of 35 (44 in mouse cell surface markers and 32 in mouse innate and adaptive immune response), the majority of genes are down-regulated in the cells transfected with the Hco-miR-5352 mimic compared to the control mimic transfected cells, although there are a few mRNAs that are up-regulated as well.

6.2.2 RT² Profiler™ PCR Array Mouse Cell Surface Markers (PAMM-055Z)

Table 6.2 Quality control check for the mouse cell surface markers array for cells transfected with Hco-miR-5352 mimic and the control Cel-miR-67 mimic.

Test Performed	Test Result
1. PCR Array Reproducibility	All Samples Passed
2. RT Efficiency	All Samples Passed
3. Genomic DNA Contamination	All Samples Passed

From both the Hco-miR-5352 mimic and Cel-miR-67 control mimic transfected cell arrays, control wells passed data quality checks (Table 6.2). Figure 6.1 shows the normalised expression of Hco-miR-5352 transfected MODE-K cells and the control samples (Cel-miR-67 transfected MODE-K cells). The majority of the genes are down-regulated and no genes are significantly up-regulated in miR-5352 transfected cells (Figure 6.1).

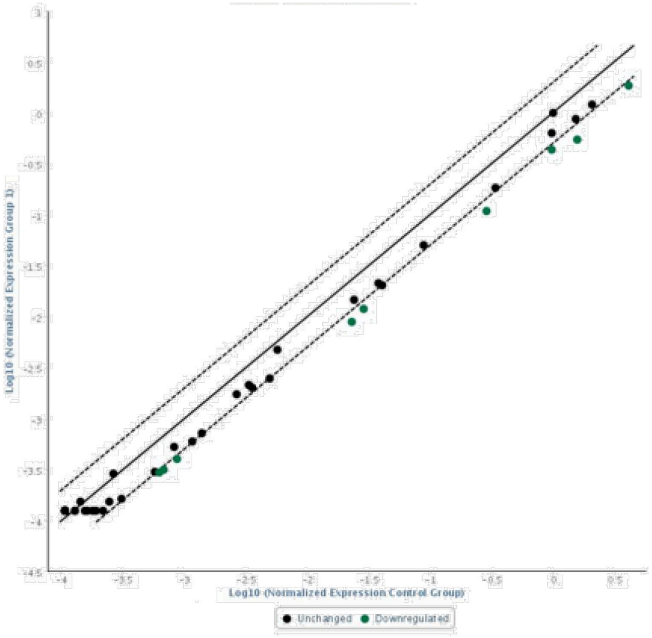


Figure 6.1 Graph showing normalised gene expression from the RT² profiler array. The solid lines indicate equal expression in both groups, the dashed lines above and below indicate the two-fold change. Dots above the lines indicate up-regulation (higher expression in miR-5352 mimic transfected cells), dots below indicate down-regulation (lower expression in miR-5352 mimic transfected cells). No genes were highly up-regulated although several are slightly up-regulated. Eleven genes were down-regulated with a fold change greater than or equal to 2.0 in miR-5352 transfected cells. Two genes have a value of 2 and are not coloured in green.

Only two genes, Fcer2a and S100a8, showed a very low level of up-regulation of (1.06 and 1.07) respectively. Eleven genes have a fold change greater than or equal to 2 (Table 6.3).

Table 6.3 Genes with fold changes ≥ 2.0 in miR-5352 transfected cells.

Symbol	Fold Change	Gene name	Gene function
Actb	-2.2191	actin, beta	
Ncam1	-2.5847	Neural cell adhesion molecule	Natural Killer (NK) Cell Surface Markers
Myh9	-2.639	Myosin heavy polypeptide 9	Smooth Muscle Cell Surface Markers
Itga3	-2.4116	Integrin alpha 3	Mature B-cell Surface Markers
Col1a1	-2.8679	Collagen, type I, alpha I	Fibroblast (Stromal Cell) Surface Markers
Chst10	-2.1735	Carbohydrate sulfotransferase	B-Cell Surface Markers
Cd63	-2.2346	CD63 antigen	Monocyte & Macrophage Cell Surface Markers
Cd276	-2	CD276 antigen	T-Cell Surface Markers
Cd247	-2	CD247 antigen	T-Cell Surface Markers
Cd1d1	-2.114	CD1d1 antigen	Epithelial Cell Surface Markers
Cd160	-2.1585	CD160 antigen	T-Cell Surface Markers

Of the eleven down-regulated genes, eight encoded immune cell surface markers. The previously predicted cell surface marker CD69 was also present on the array, however no signal was detected in either the Hco-miR-5352 mimic transfected cells or with the control mimic, suggesting that CD69 levels were too low to be detected. None of the down-regulated genes are predicted targets of Hco-miR-5352 using Targetscan Custom seed match programme. To investigate in more detail if these differentially expressed genes are predicted targets using other programmes, the 3' UTR sequences for each of the genes were obtained and binding sites for Hco-miR-5352 were predicted using PITA with default settings. From the eleven genes, only eight binding sites across four genes were predicted by PITA (Table 6.4). However only one gene, Col1a1 had a binding site with a ΔΔG of less than the threshold (ΔΔG > -10), suggesting that apart from Col1a1, these genes are unlikely to be directly targeted by Hco-miR-5352. However, this gene set is still showing a large down regulation between the Hco-miR-5352 mimic transfected cells compared to the control mimic transfected cells and, since they do not appear to direct targets of miR-5352, an upstream gene in the pathway might be targeted.

Table 6.4 PITA predicted binding sites for Hco-miR-5352 on the 3' UTR of interest. The more negative the value of ΔΔG, the stronger the binding of the miRNA to the site.

Gene	Position	Seed	dGduplex	dGopen	ΔΔG
Col1a1	725	08:00:00	-25.65	-9.91	-15.73
Myh9	566	08:01:00	-14.93	-6.19	-8.73
Cd276	417	08:01:00	-19.85	-13.88	-5.96
Cd1d1	414	08:01:01	-13.5	-7.63	-5.86
Col1a1	1406	08:01:00	-12.9	-8.54	-4.35
Cd1d1	206	08:01:01	-11.7	-8.06	-3.63
Cd1d1	311	08:01:01	-13.6	-10.13	-3.46
Myh9	1229	08:01:01	-12.8	-9.77	-3.02

6.2.3 RT² Profiler™ PCR Array Mouse Innate and Adaptive Immune Responses (PAMM-052Z)

The Qiagen data quality checks show that the two Mouse Innate and Adaptive Immune Response arrays were of good quality (Table 6.5).

Table 6.5 Quality control check for the mouse innate and adaptive immune response array for cells transfected with Hco-miR-5352 mimic and the control Cel-miR-67 mimic.

Test Performed	Test Result
1. PCR Array Reproducibility	All Samples Passed
2. RT Efficiency	All Samples Passed
3. Genomic DNA Contamination	All Samples Passed

From Figure 6.2 six genes had lower expression in the Hco-miR-5352 mimic transfected cells relative to the Cel-miR-67 control mimic transfected cells, with a fold change of greater than 2 (Table 6.6). Interestingly, only one gene, CD80 had higher expression in the Hco-miR-5352 transfected cells. CD80 is a co-stimulatory molecule for T cells, while each of the significantly down-regulated genes have important roles in the induction or development of the immune response. A further 9 genes were higher in Hco-miR-5352 mimic transfected cells and 16 genes had lower expression in Hco-miR-5352 mimic transfected cells with fold change between 1.5 and 2.

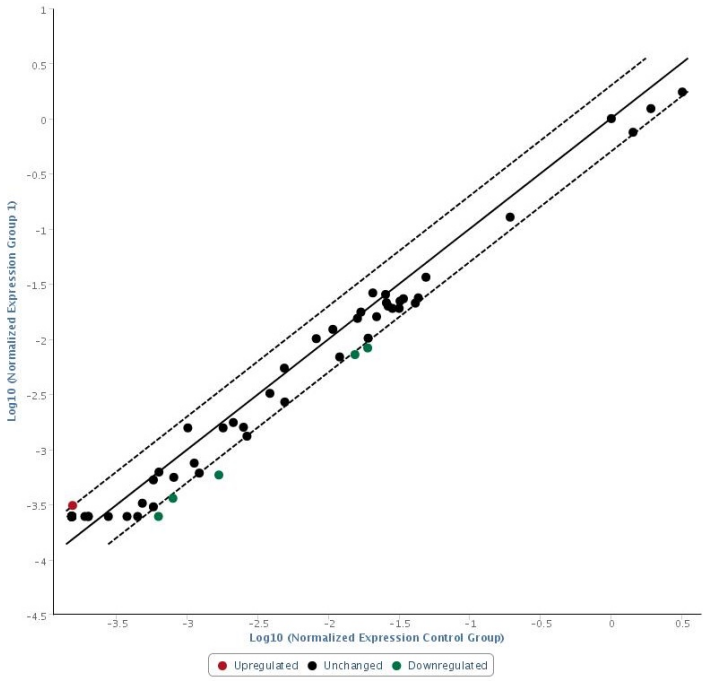


Figure 6.2 Graph showing the normalised gene expression from the RT^2 profiler. The solid line indicates a line of equal expression in both groups, the dashed lines above and below indicate the two-fold change. Dots above the line indicate up-regulation, dots below indicate down-regulation. Five genes are down-regulated and one gene up-regulated with a fold change greater than two. Tlr2 has a fold change of 2 and was not coloured green

Table 6.6 Genes with fold changes ≥ 2 are shown

Symbol	Fold Change	Gene name	Gene function
CD80	2.01	CD80 antigen	Th1 Markers & Immune Response
Foxp3	-2.53	Foxhead box P3	Treg Markers
Tlr2	-2	Toll like receptor 2	Pattern Recognition Receptors (PPRs)
Tlr4	-2.114	Toll like receptor 4	Pattern Recognition Receptors (PPRs)
Myd88	-2.1886	Myeloid differentiation primary response gene 88	Defence Response to Bacteria
Ifnar1	-2.2658	Interferon (alpha and beta) receptor 1	Defence Response to Viruses
Tlr6	-2.8481	Toll like receptor 6	Pattern Recognition Receptors (PPRs)

microRNA target prediction was conducted using PITA on the 3' UTR sequences for the seven genes whose expression is significantly changed. CD80 did not have any predicted miRNA binding sites for miR-5352, however, sites were predicted for five of the other genes with two sites predicted on Myd88. All of these sites have an eight seed match although the ΔΔG scores are still higher than the threshold of -10 (Table 6.7).

Table 6.7 PITA predicted binding sites for Hco-miR-5352 on the 3' UTR of interest. The more negative the value of ΔΔG, the stronger the binding of the miRNA to the site.

Gene	Position	Seed	dGduplex	dGopen	ΔΔG
Myd88	855	08:01:01	-13.3	-5.6	-7.69
Foxp3	620	08:00:01	-20.2	-13.47	-6.72
Myd88	210	08:01:00	-13.31	-7.03	-6.27
Tlr4	1162	08:01:01	-14.81	-8.78	-6.02
Ifnar1	693	08:01:00	-18.14	-12.19	-5.94

From the RT² Profiler™ PCR Array experiments, a number of genes were identified that had significant expression differences between the Hco-miR-5352 mimic transfected cells and the control mimic transfected cells, indicating that the Hco-miR-5352 mimic is having an impact on gene expression in the MODE-K cells. However, none of these genes show good binding sites for Hco-miR-5352,

suggesting that whilst these genes may be altered in the presence of the mimic, they are not likely to be the direct targets of Hco-miR-5352. The analysis is limited to the genes on the array and the low number of genes showing differences means that analysis using the Qiagen IPA tool will not be effective.

6.2.4 Global gene analysis of MODE-K cells transfected with miR-5352.

The data from the RT^2 profiling arrays identified a number of immune related genes that were down-regulated in MODE-K cells transfected with the Hco-miR-5352 mimic compared to cells transfected with the control mimic, suggesting that the Hco-miR-5352 mimic is causing certain genes to be down-regulated. However, miRNA target prediction did not produce sites of high binding energy implying that these genes may not be direct targets. IPA canonical pathway or network analysis is able to determine possible upstream regulators or genes with direct or indirect relationships to the down-regulated genes, but no expression data was available for these using the RT profiler approach.

A solution would be to examine the expression of all the genes in the MODE-K cells by RNA-Sequencing following transfection with miR-5352. Cells were transfected exactly as described for the RT^2 Profiler array and RNA extracted and sent for sequencing (see Chapter 2.23). In total, 23,776 mouse genes were identified, of which 467 (2%) showed significant differences (q-value less than 0.05) in expression between miR-5352 and control miRNA transfections. 348 genes had higher levels of expression in the cells transfected with the control mimic compared to the Hco-miR-5352 (and therefore could potentially be down-regulated by miR-5352), while 119 genes had higher expression in the Hco-miR-5352 transfected cells compared to the control (see Table 6.8 for a summary).

Table 6.8 Summary of genes from the RNA-Seq data

Number of genes identified	23,776
Number of genes differentially expressed	467 (2%)
Number of differentially expressed genes decreased in cells transfected with Hco-miR-5352	348
Number of differentially expressed genes increased in cells transfected with Hco-miR-5352	119
Number of differentially expressed genes with Hco-miR-5352 binding sites	46 44-down-regulated 2-up-regulated

The top 100 most significantly differently expressed genes that are up or down-regulated are shown in appendix Table 8.3 and Table 8.4. The gene that had the greatest decrease in expression in the presence of Hco-miR-5352 was the purinergic receptor P2RY12, with a log2(fold change) of -4.7. The list of down-regulated genes also included a number of chemokine receptors, CD molecules and components of the complement pathway.

Conversely, of the 100 most significantly up-regulated genes, solute carrier family 6 (SLC6A1) was the most highly up-regulated with a fold change of 5.56. A further nine solute carrier proteins are present in this list. No solute carrier proteins were detected to be significantly down-regulated. Angiotensinogen is also present and shows a fold change of 5.3. Of the up-regulated genes, a few genes, such as the chemokine ligand 10 are immune related, however the large majority are not.

To determine if any of these significantly altered genes have predicted binding sites for Hco-miR-5352, the gene list was compared to the list of potential target genes originally obtained from Targetscan. Targetscan searches for the presence of 7 to 8 bases that match the seed region of the mature miRNA and that also possess an adenosine at the very 3' terminus of the miRNA complementary site, as this “A anchor” has been shown to increase target prediction specificity (Lewis *et al.*, 2005).

In total, a statistically significant number of genes (46) that have significant expression differences between the transfected cell types also have predicted binding sites for Hco-miR-5352 (p-value of 7.94^{-19}). The majority of these genes have decreased expression in the Hco-miR-5352 transfected cells, as may be expected if the genes are targeted by this microRNA (Table 6.9). Only two genes that were predicted targets of Hco-miR-5352, phosphatidic acid phosphatase type 2B (PPAP2B) and four joined box 1 (FJX1), showed higher expression in the Hco-miR-5352 transfected cells (Table 6.9). PPAP2B is a member of the phosphatidic acid phosphatase (PAP) family which was found to be up-regulated in type 2 dendritic cells although its function in DCs is unclear (Hata *et al.*, 2009). FJX1is a transmembrane glycoprotein that may play key roles in cancer

related pathways and has been shown to be overexpressed in various tumours (Chai *et al.*, 2015) and correlates with poor survival in colorectal cancer patients (Al-Greene *et al.*, 2013).

GO terms analysis of the 44 genes that are down-regulated in the Hco-miR-5352 transfected cells and have predicted binding sites for Hco-miR-5352 was performed using pantherdb (www.pantherdb.org). The analysis showed that the majority of genes were involved in the regulation of cellular and biological processes. No terms were identified that are specifically related to immunology. The location of these genes also appear to be widespread across the cell, with most being intracellular and in organelles (Appendix Table 8.2). The 44 genes that are down-regulated and also have miR-5352 binding sites account for only 12 % of the total number of down-regulated genes. However, GO term analysis of the 348 genes that are significantly down-regulated in the presence of Hco-miR-5352 (44 genes that have binding sites for Hco-miR-5352 and 308 that do not have binding sites), shows that 89 are associated with the immune system (Table 6.10). Out of the 35 biological process GO terms, the vast majority have functions related to various parts of the immune system and its regulation.

Table 6.9 List of genes that have significantly different expression between the Hco-miR-5352 and Cel-miR-67 transfected cells that also have miR-5352 binding sites on the 3' UTR as predicted by Targetscan. Table is sorted by number and the type of sites. 8mer stands for an exact match to positions 2-8 of the mature miRNA followed by an "A", 7mer-m8: An exact match to positions 2-8. 7mer-1A: An exact match to positions 2-7 followed by an "A". (Lewis *et al.*, 2005) **Genes in black indicate lower expression in Hco-miR-5352 transfected cells. Genes in green indicate genes with higher expression in Hco-miR-5352 transfected cells.**

Target gene	Gene name	total	8mer	7mer-m8	7mer-1A
USP47	ubiquitin specific peptidase 47	3	1	1	1
LHFPL2	lipoma HMGIC fusion partner-like 2	2	1	0	1
RGS17	regulator of G-protein signaling 17	2	1	0	1
APPL1	adaptor protein, phosphotyrosine interaction, PH domain and leucine zipper containing 1	1	1	0	0
EIF5	eukaryotic translation initiation factor 5	1	1	0	0
FAM126B	family with sequence similarity 126, member B	1	1	0	0
HAND2	heart and neural crest derivatives expressed 2	1	1	0	0
JMY	junction-mediating and regulatory protein	1	1	0	0
LPGAT1	lysophosphatidylglycerol acyltransferase 1	1	1	0	0
MYO1B	myosin IB	1	1	0	0
PHF15	PHD finger protein 15	1	1	0	0
PJA2	praja 2, RING-H2 motif containing	1	1	0	0
PPAP2B	phosphatidic acid phosphatase type 2B	1	1	0	0
PRKCE	protein kinase C, epsilon	1	1	0	0
REV3L	REV3-like, catalytic subunit of DNA polymerase zeta (yeast)	1	1	0	0
SOCS5	suppressor of cytokine signaling 5	1	1	0	0
TWF1	twinfilin, actin-binding protein, homolog 1 (Drosophila)	1	1	0	0
USP28	ubiquitin specific peptidase 28	1	1	0	0
BTG2	BTG family, member 2	2	0	0	2
AFF4	AF4/FMR2 family, member 4	1	0	1	0
ANKRD50	ankyrin repeat domain 50	1	0	1	0
ARNT	aryl hydrocarbon receptor nuclear translocator	1	0	1	0
FJX1	four jointed box 1 (Drosophila)	1	0	1	0

GOSR1	golgi SNAP receptor complex member 1	1	0	1	0
H3F3B	H3 histone, family 3B (H3.3B)	1	0	1	0
LPP	LIM domain containing preferred translocation partner in lipoma	1	0	1	0
MARCH6	membrane-associated ring finger (C3HC4) 6	1	0	1	0
MED15	mediator complex subunit 15	1	0	1	0
SOCS6	suppressor of cytokine signaling 6	1	0	1	0
SUB1	SUB1 homolog (S. cerevisiae)	1	0	1	0
ADAM10	ADAM metallopeptidase domain 10	1	0	0	1
ARPC5	actin related protein 2/3 complex, subunit 5, 16kDa	1	0	0	1
BNC2	basonuclin 2	1	0	0	1
CEBPA	CCAAT/enhancer binding protein (C/EBP), alpha	1	0	0	1
DOCK11	dedicator of cytokinesis 11	1	0	0	1
DUSP10	dual specificity phosphatase 10	1	0	0	1
ITM2B	integral membrane protein 2B	1	0	0	1
MLF2	myeloid leukemia factor 2	1	0	0	1
PHTF2	putative homeodomain transcription factor 2	1	0	0	1
PLEKHB2	pleckstrin homology domain containing, family B (evectins) member 2	1	0	0	1
PTEN	phosphatase and tensin homolog (mutated in multiple advanced cancers 1)	1	0	0	1
RAPH1	Ras association (RalGDS/AF-6) and pleckstrin homology domains 1	1	0	0	1
RBPMS2	RNA binding protein with multiple splicing 2	1	0	0	1
UBE2W	ubiquitin-conjugating enzyme E2W (putative)	1	0	0	1
VGLL3	vestigial like 3 (Drosophila)	1	0	0	1
VPS26A	vacuolar protein sorting 26 homolog A (S. pombe)	1	0	0	1

Table 6.10 GO term analysis of the 348 genes that have lower expression in the Hco-miR-5352 transfected cells compared to the Cel-miR-67 transfected cells.

Biological Process	Number of genes
GO:0002376 immune system process	89
GO:0048584 positive regulation of response to stimulus	82
GO:0048870 cell motility	65
GO:0051674 localization of cell	65
GO:0002682 regulation of immune system process	54
GO:0006955 immune response	51
GO:0006952 defense response	51
GO:0002684 positive regulation of immune system process	41
GO:0045321 Leucocyte activation	41
GO:0002520 immune system development	38
GO:0016337 single organismal cell-cell adhesion	37
GO:0048534 hematopoietic or lymphoid organ development	36
GO:0002252 immune effector process	36
GO:0030097 haemopoiesis	34
GO:0006954 inflammatory response	33
GO:0007010 cytoskeleton organization	31
GO:0050776 regulation of immune response	30
GO:0046649 lymphocyte activation	30
GO:0034097 response to cytokine	28
GO:0022604 regulation of cell morphogenesis	28
GO:0034109 homotypic cell-cell adhesion	28
GO:0031347 regulation of defense response	26
GO:0045087 innate immune response	24
GO:0007159 Leucocyte cell-cell adhesion	24
GO:0042110 T cell activation	23
GO:0070486 Leucocyte aggregation	23
GO:0070489 T cell aggregation	23
GO:0071593 lymphocyte aggregation	23
GO:0002521 Leucocyte differentiation	22
GO:0050778 positive regulation of immune response	22
GO:0002683 negative regulation of immune system process	22
GO:0030036 actin cytoskeleton organization	22
GO:0030099 myeloid cell differentiation	21
GO:0050900 Leucocyte migration	21
GO:0060627 regulation of vesicle-mediated transport	21

Molecular Function	Number of genes
GO:0042803 protein homodimerization activity	24
GO:0042802 identical protein binding	42
GO:0003779 actin binding	20
GO:0004175 endopeptidase activity	10
Cellular Component	**Number of genes**
GO:0016023 cytoplasmic membrane-bounded vesicle	30
GO:0015629 actin cytoskeleton	17
GO:0005912 adherens junction	17
GO:0070161 anchoring junction	17
GO:0005924 cell-substrate adherens junction	15
GO:0005925 focal adhesion	15
GO:0030055 cell-substrate junction	15
GO:0005581 collagen trimer	6
GO:0030139 endocytic vesicle	5
GO:0009925 basal plasma membrane	4
GO:0045178 basal part of cell	4

Table 6.11 shows the list of GO terms for the 119 genes that have higher expression in the Hco-miR-5352 transfected cells compared to the Cel-miR-67 cells. The GO terms that were identified from the genes that are decreased in the presence of Hco-miR-5352 are very different from the GO terms of the genes that are increased. The genes that are decreased are predominantly related to immune responses, whereas the genes that are increased are more likely to be associated with the response to chemicals and regulation of gene expression.

Table 6.11 GO term analysis of the genes that have higher expression in the Hco-miR-5352 transfected cells compared to the Cel-miR-67 transfected cells.

Biological Process	Number of genes
GO:0042221 response to chemical	119
GO:0010467 gene expression	104
GO:0010468 regulation of gene expression	100
GO:0010033 response to organic substance	94
GO:0008219 cell death	89
GO:0016265 death	89
GO:0009892 negative regulation of metabolic process	74
GO:0031324 negative regulation of cellular metabolic process	64

`

GO:0022607 cellular component assembly	62
GO:0010605 negative regulation of macromolecule metabolic process	61
GO:0044281 small molecule metabolic process	56
GO:1901700 response to oxygen-containing compound	53
GO:0048878 chemical homeostasis	43
GO:0051248 negative regulation of protein metabolic process	42
GO:0007417 central nervous system development	37
GO:1901701 cellular response to oxygen-containing compound	37
GO:0009628 response to abiotic stimulus	35
GO:0032269 negative regulation of cellular protein metabolic process	34
GO:0019725 cellular homeostasis	33
GO:0055082 cellular chemical homeostasis	32
GO:0010942 positive regulation of cell death	31
GO:0010629 negative regulation of gene expression	31
GO:0043068 positive regulation of programmed cell death	29
GO:0010563 negative regulation of phosphorus metabolic process	27
GO:0045936 negative regulation of phosphate metabolic process	27
GO:0006820 anion transport	27
GO:1901698 response to nitrogen compound	26
GO:0006873 cellular ion homeostasis	26
GO:0061061 muscle structure development	24
GO:0010243 response to organonitrogen compound	24
GO:0071396 cellular response to lipid	24
GO:0097193 intrinsic apoptotic signaling pathway	23
GO:0070997 neuron death	20
GO:0042326 negative regulation of phosphorylation	20
GO:0010608 posttranscriptional regulation of gene expression	20
GO:0031400 negative regulation of protein modification process	20
Molecular Function	**Number of genes**
GO:0005215 transporter activity	33
GO:0022892 substrate-specific transporter activity	29
GO:0022857 transmembrane transporter activity	25
GO:0015075 ion transmembrane transporter activity	24
GO:0022891 substrate-specific transmembrane transporter activity	24
Cellular Component	**Number of genes**
GO:0042995 cell projection	62
GO:0097458 neuron part	45
GO:0043005 neuron projection	36
GO:0030425 dendrite	17

`

6.2.5 IPA analysis in epithelial cells

6.2.5.1 Canonical Pathways

IPA was used to investigate the relationship between differential gene expression in the presence of Hco-miR-5352 or the control mimic. Initial analysis was focused on epithelial cells as the MODE-K cell line is an intestinal epithelial cell line. An excel spreadsheet containing the gene IDs, the log2(fold change) values, the p-values and the q-values (false discovery rate) were uploaded into the IPA programme. Of the 23,776 genes that were identified by RNA-Seq, 20,000 were identified in the IPA knowledgebase.

IPA is able to generate a huge amount of data from RNA-Seq results. It was therefore necessary to filter the data to obtain the most likely pathways being affected by Hco-miR-5352. To this end, the data was first filtered for only genes expressed in epithelial cells. This reduced the input list to 1414 genes. In general, for a significantly differentially expressed gene, the q-value (false discovery rate) is considered to be less than or equal to 0.05. As the experiment used multiple comparisons, the false discovery rate was used instead of the p-value to limit the number of false positives in the experiment. Figure 6.3 shows a frequency histogram of the q-values of the epithelial genes. 46 genes fulfil the q-value requirement with the majority of genes having a q-value of 1, which is too few for IPA analysis. However, as the cell line used is an epithelial cell line, and *Haemonchus* specific miRNAs have been detected in sheep mucosal tissue, these epithelial genes could be very interesting.

`

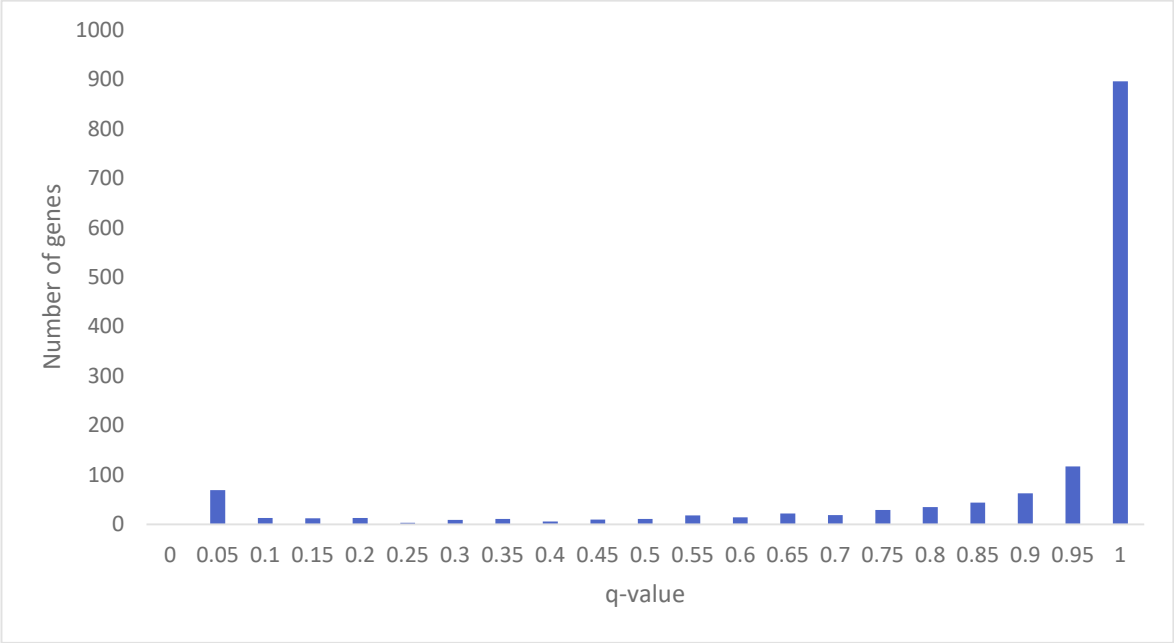


Figure 6.3 Frequency histogram of the q-values for the 1414 epithelial cell genes

To this end, all 1414 genes were used in an IPA Core Analysis. IPA is also able to predict whether or not a pathway is activated or inhibited (z-score). This is done by observing which genes are increased or decreased and comparing this with the expected direction from the literature. A z-score of ≥ 2 is considered significant. The pathway is predicted to be activated if the z-score is ≥ 2 or inhibited if the z-score ≤ -2.

A total of 382 pathways were identified and of these, 47 were considered significant by IPA (p-value < 0.05). The 20 most significant pathways are shown in Table 6.12 and are ordered by their z-score, none of which had a positive value (activated).

Table 6.12 The 20 most significant pathways identified from the epithelial cell analysis. P-value calculated using a right-tailed Fisher Exact Test. Ratio indicates the number of genes from the input list that are present in the pathway. A ratio of one indicates all the input list genes are present, a score of 0 indicates none are present. Z-score is a prediction of whether the pathway is activated or inhibited. NaN indicates z-scores were not available.

Ingenuity Canonical Pathways	p-value	Ratio	z-score
Colorectal Cancer Metastasis Signalling	1.70E-02	0.91	-2.66
NF-κB Signalling	1.32E-02	0.92	-2.57
Role of NFAT in Cardiac Hypertrophy	4.27E-03	0.96	-2.36
Leucocyte Extravasation Signalling	8.13E-03	0.95	-2.29
p38 MAPK Signalling	1.66E-02	0.97	-2.27
Cardiac Hypertrophy Signalling	1.74E-02	0.95	-2.27
IL-1 Signalling	1.45E-02	1	-2.07
Induction of Apoptosis by HIV1	1.45E-02	1	-1.71
Type I Diabetes Mellitus Signalling	2.95E-03	0.98	-1.46
Pancreatic Adenocarcinoma Signalling	4.17E-03	0.98	-0.78
Adipogenesis pathway	2.00E-04	1	NaN
Molecular Mechanisms of Cancer	1.38E-03	0.93	NaN
Tight Junction Signalling	1.70E-03	1	NaN
Axonal Guidance Signalling	3.24E-03	0.93	NaN
Breast Cancer Regulation by Stathmin1	7.94E-03	0.96	NaN
Small Cell Lung Cancer Signalling	8.13E-03	1	NaN
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis	1.00E-02	0.92	NaN
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	1.41E-02	0.9	NaN
RAR Activation	1.74E-02	0.92	NaN
Chronic Myeloid Leukaemia Signalling	1.95E-02	0.97	NaN

Discussed below are three pathways (NF-κB Signalling, Role of NFAT in Cardiac Hypertrophy, Leucocyte Extravasation Signalling) that were selected based on the following criteria: they had a negative z-score (pathway is inhibited), the pathways contained genes that are predicted Hco-miR-5352 targets and may be involved in modulation of the host response against *H. contortus*.

The NF-κB canonical pathway contains seven predicted Hco-miR-5352 targets, none of which show significant expression differences in the RNA-Seq data, while other genes in this pathway are differentially expressed (Figure 6.4). The z-score is -2.6 showing an overall inhibition of the pathway. The NF-κB protein complex controls the transcription of DNA, cytokine production and cell survival.

`

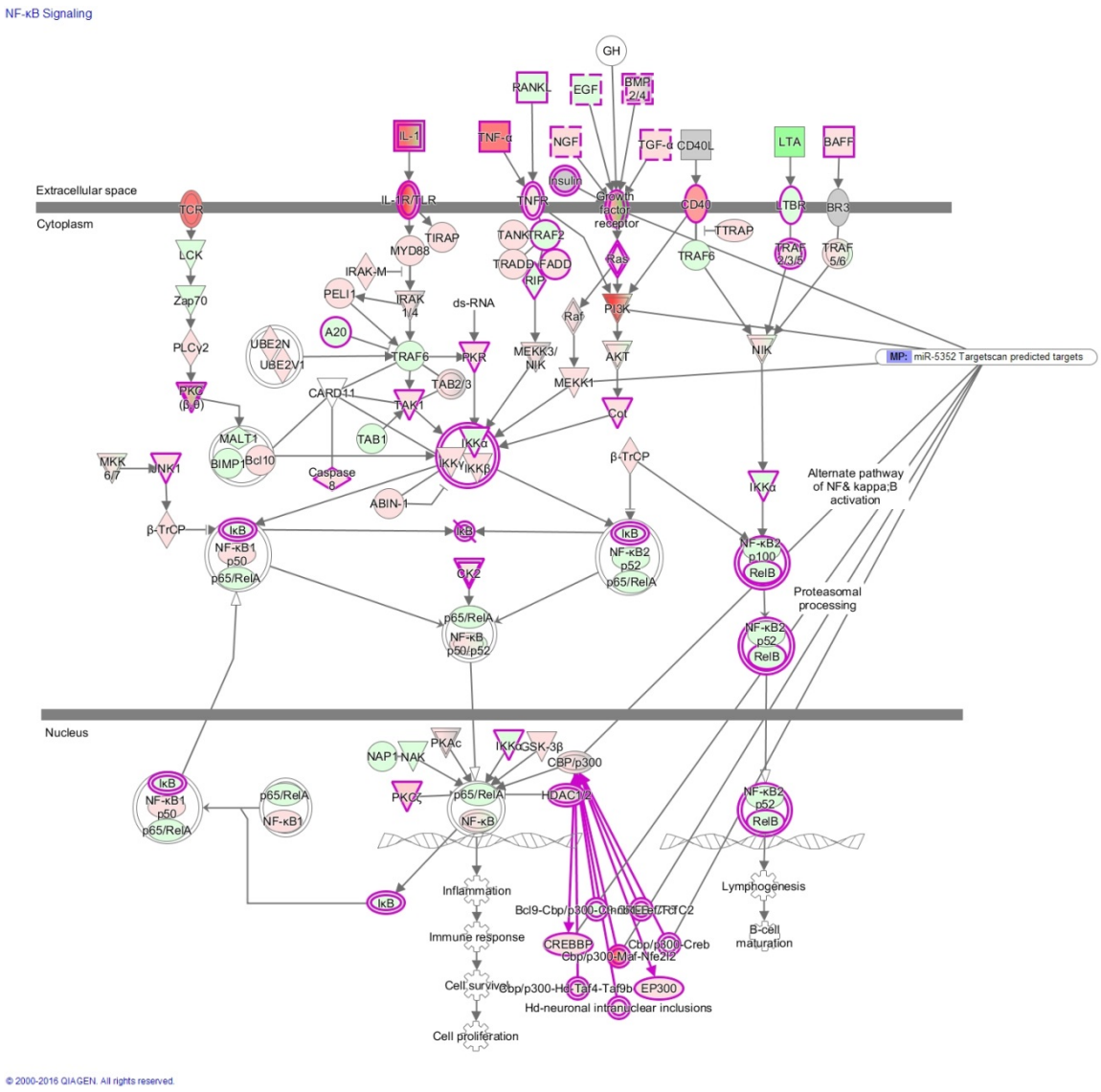


Figure 6.4 Canonical pathway of NF-κB showing the genes which are predicted targets of miR-5352 by Targetscan and those targets that have significant differential expression. Genes highlighted in red indicate decreased expression when Hco-miR-5352 is present. Genes in green indicate increased expression.

The second pathway analysed is “Role of NFAT (Nuclear factor of activated T-cells) in Cardiac Hypertrophy” (Figure 6.5). This pathway contains seven predicted targets of Hco-miR-5352. In both Figure 6.5 and Figure 6.6, the significantly differentially expressed predicted target is PKC (protein kinase C). The pathway also has a Z-score of -2.4 showing that it is predicted to be down-regulated. The NFAT protein family are a group of transcription factors important in inducing expression of cytokine genes in T cells (Daniel *et al.*, 2014).

`

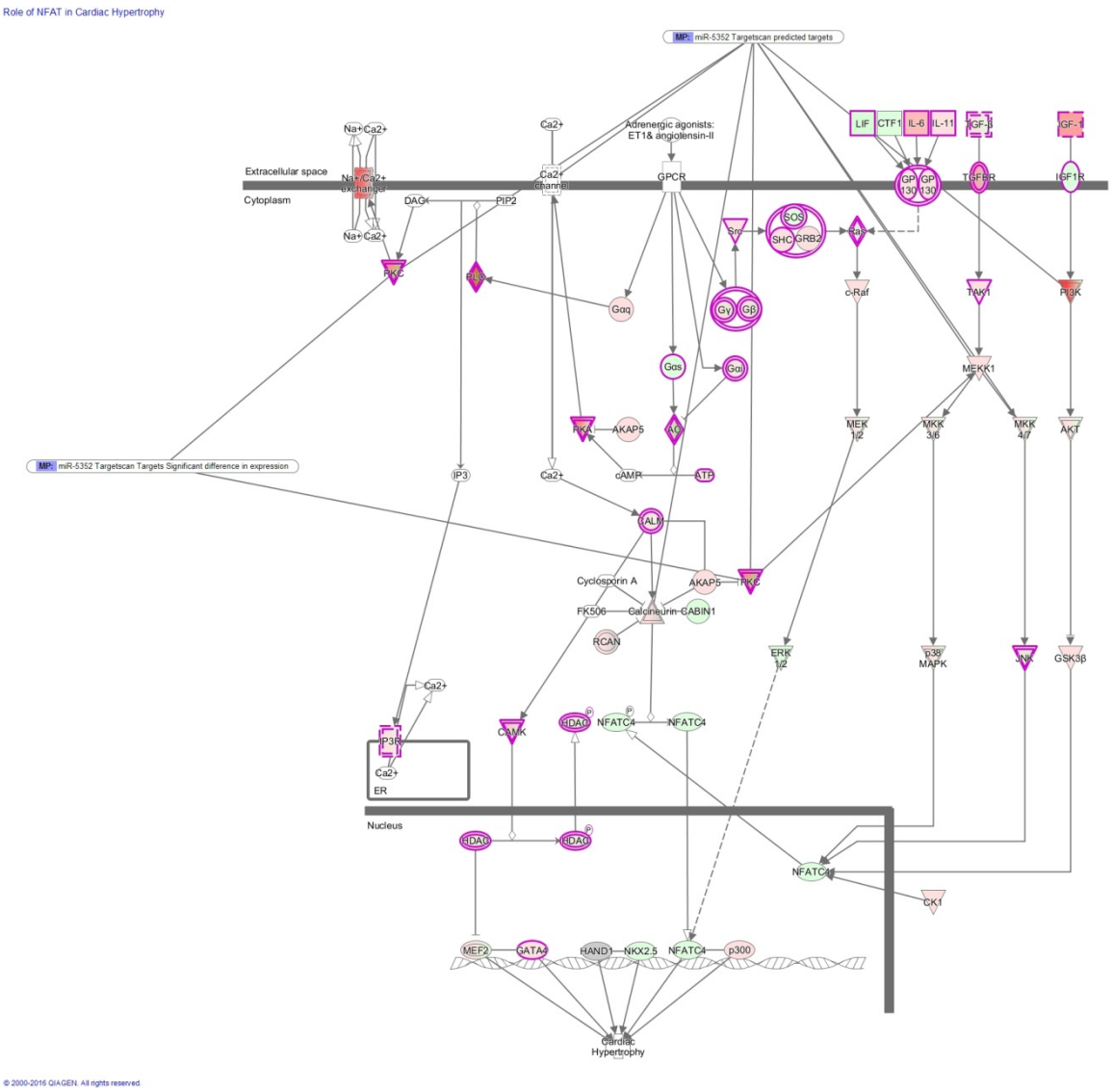


Figure 6.5 Canonical pathway of Role of NFAT in Cardiac Hypertrophy showing the genes which are predicted targets of miR-5352 by Targetscan and those targets that have significantly differential expression. Genes highlighted in red indicate decreased expression when Hco-miR-5352 is present. Genes in green indicate increased expression.

The canonical pathway shown in Figure 6.6 is the "Leucocyte extravasation signalling pathway". The pathway contains ten predicted targets of miR-5352 using the Targetscan algorithm and of those, one (PKC, protein kinase C) is differentially expressed according to the q-value. This gene appears twice in the pathway and is the same gene present in Figure 6.5. The pathway also has a predicted z-score of -2.3, which represents a significant downregulation when Hco-miR-5352 is present. This pathway is the process by which leucocytes migrate from the blood to the tissue during inflammation.

`

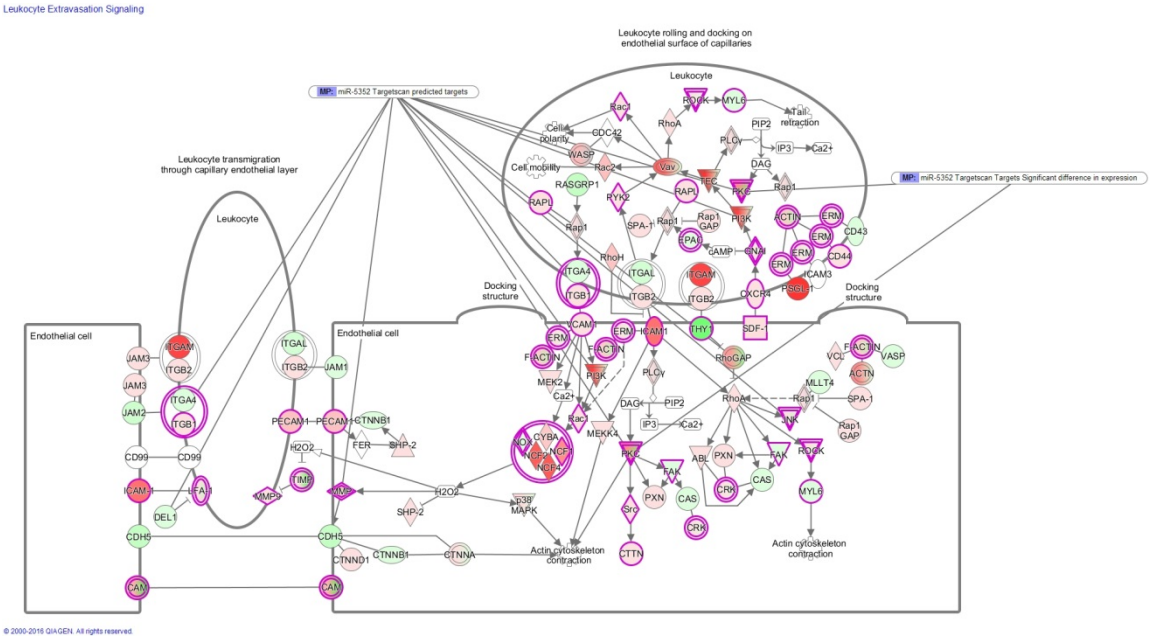


Figure 6.6 Canonical pathway of Leucocyte Extravasation Signalling showing the genes that are predicted targets of miR-5352 by Targetscan and those targets that have significant differential expression (Table 6.9). Genes highlighted in red indicate decreased expression when Hco-miR-5352 is present. Genes in green indicate increased expression.

6.2.5.2 Predicted upstream regulators of identified genes

IPA is also able to predict any upstream regulators, such as proteins or miRNAs that target the genes identified in the RNA-Seq data. The programme is then able to predict whether these upstream molecules are either activated or inhibited based on the correlation observed between the genes that are increased or decreased and the prediction expression pattern from the literature.

A total of nine molecules were identified in the upstream analysis, three of which (ARNTL, GHRL and GAS6) are predicted to be activated, and two (TLR3 and CD40) predicted to be inhibited (Table 6.13). Three more (RB1, IL17C and IGF1) have reasonable predicted z-scores although they do not pass the threshold of ≥ 2 or ≤ 2. The final gene, FOS is predicted to target 49 of the 1414 genes, whereas the rest are predicted to target between four to nine genes in total. A z-score was unavailable for FOS as the literature only provides evidence of regulation, but not the direction. However, FOS was the only upstream regulator predicted to target a significant number of genes (see p-values in Table 6.13).

`

Table 6.13 Predicted upstream regulators of the 1414 epithelial cell genes.

Upstream Regulator	Molecule Type	Predicted Activation State	Activation z-score	Target molecules in dataset
FOS	transcription regulator			ABCC1,ACOX1,AFF4,AKR1C3,AQP3,ATP2C1,CA2,CAMK2D,CAT,CD44,CSF3,DLG1,DSG3,ELOVL3,EPS8,FOS,FRZB,GATA4,HDAC6,HMOX1,HSD17B10,IRS2,ITGA6,ITGB1,KRT13,LAMA3,LRP8,Macf1,MET,MLH1,MMP3,MXD1,ODC1,PDPN,PFN1,PTGS2,PTH1R,PTTG1,RAPGEF3,RBP1,RIPK1,SDC1,SFRP2,SMAD7,TGFB1,WNT1,WNT10B,WNT4,WNT6
IL17C	cytokine		-1.890	CCL20,IL17C,IL1A,IL1B,IL36RN,IL6,LCN2,S100A8,TNF
IGF1	growth factor		-1.974	ASGR2,CAMP,IGF1,LCN2,TF
TLR3	transmembrane receptor	Inhibited	-2.219	Ccl2,IL1B,IL6,OAS1,TNF
CD40	transmembrane receptor	Inhibited	-2.401	CCL2,FASLG,FOS,IL15,IL6,JUN
RB1	transcription regulator		1.969	BCL2,CDKN1A,CLDN3,IGF1,MYC,PARP1
ARNTL	transcription regulator	Activated	2.000	ACACA,FASN,GPAM,SREBF1
GHRL	growth factor	Activated	2.228	ACACA,FASN,GPAM,PPARG,SREBF1
GAS6	growth factor	Activated	2.236	IL1B,IL6,SOCS1,SOCS3,TNF

FOS or c-Fos is a proto-oncogene and a member of the Fos gene family along with FOSB, FOSL1 and FOSL2. Fos has been shown to be silenced by miR-155 in dendritic cells and is critical for dendritic cell maturation and function (Dunand-Sauthier *et al.*, 2011). Apart from FOS, none of the other upstream regulators have a significant p-value of overlap.

Two cell surface receptors were identified in this analysis, CD40 and TLR-3. CD40 is an interesting target. This gene has the lowest z-score of -2.4 based on

`

six genes which were all decreased to different degrees. CD40 is a co-stimulatory molecule expressed on antigen presenting cells (APC) and is required for T cell activation (Grewal and Flavell, 1996). The second cell surface molecule that was decreased is Toll-like receptor (TLR) 3. TLR were also identified in the RT^2 profiler array as genes of interest and will be discussed in more detail below. Two other regulators were identified in the analysis, IL-17C and IGF1, however both had a z-score higher than -2.0.

6.2.5.3 Networks

IPA is also able to generate a network analysis. The network analysis attempts to identify connections between the input genes through indirect or direct relationships irrespective of canonical pathways. The IPA network analysis identified a large number of networks with a variety of predicted functions (Table 6.14). The networks were selected based on number of Hco-miR-5352 predicted targets and potential to be involved in modulation of host immunity. Only the network titled “Inflammatory Disease, Tissue Development, Cellular Function and Maintenance” networks appears to be a promising target of Hco-miR-5352 as detailed below.

`

Table 6.14 Top ten networks identified by IPA. The score indicates the likelihood of the genes in a network being found together by chance. A score of 2 or higher indicates a 99% chance of not being generated by chance.

Top Diseases and Functions	Score	Molecules in Network
Cancer, Organismal Injury and Abnormalities, Cell Death and Survival	3	ADM,ALB,ALDH2,ATP2B1,BTG2,CCND2,CEBPB,CFB,CLEC11A,CRHR1,DSC3,ETS1,GCH1,GSTM2,IGFBP2,IKBKE,IL33,IL4R,IL7R,JARID2,KRT6B,LAMC2,NCK1,NEDD9,NGF,OAS2,OSM,OSMR,PHGDH,PPARD,RELB,TM4SF1,TNF,TNFAIP2,WNT7A
Inflammatory Disease, Tissue Development, Cellular Function and Maintenance	3	ABCC1,ABCC2,AHR,ARG1,AXL,CEBPA,CISH,CYP2B6,CYP2C8,CYP2C9,FN1,GAS6,GSTA1,HNF4A,IGFBP4,IL6,IL1B,IL6R,IL6ST,KDR,LBP,MERTK,NOS2,NR1I2,NRP1,S100A10,SAA1,SEMA3A,SLC10A1,SLC10A2,SOCS2,SOCS3,TGFB1,TLR3,TYRO3
Cell-To-Cell Signaling and Interaction, Cellular Movement, Hematological System Development and Function	3	CAMP,CASP1,Ccl2,CCL20,CSF2,CXCL2,CXCL3,CXCL10,DLX3,EGFR,FADD,FAS,FASLG,ICAM1,IFNGR1,IFNGR2,IL18,IL1A,IL1R1,IL1RL2,IL1RN,IRF1,IRF7,MMP1,NFKBIA,PRKCA,PRL,SOCS1,STAT1,TAC1,TLR2,TNFAIP3,VDR,VEGFA,YES1
Cell Cycle, Connective Tissue Development and Function, Cellular Development	3	ACTL6A,CCND1,CCNE1,CDK1,CDK2,CDKN1A,CDKN1B,CDKN2B,DICER1,E2F1,EGF,ENSA,ERBB2,HES1,HSPA5,IGF1,Ins1,ITGA6,ITGB1,ITGB4,KRT14,MYC,NFATC2,NOTCH1,PCNA,PPARG,PRKCD,PRKCH,PRKD3,RB1,RBPJ,RND3,TP53,TP63,WNT4
Tissue Development, Cellular Growth and Proliferation, Embryonic Development	3	CASP3,CCL2,CD40,CD44,CDKN2A,CHUK,CREB1,CTBP1,DLG1,DSG1,ESR2,FOS,GJB1,HBEGF,HGF,JUN,JUNB,JUND,KRT8,LCN2,MET,MKI67,MXD1,PSCA,PTGS2,RAC1,RARA,RARG,SDC1,SMAD2,SMAD3,SMAD4,SMAD7,TERT,TGFA
Cell Death and Survival, Cancer, Organismal Injury and Abnormalities	3	ABCB1,Abcb1b,AGER,APP,AQP8,AR,CD46,CLDN1,CLDN3,CLDN11,DSC2,DSG3,DSP,ELMO1,FABP5,HDAC1,IL12B,IL17C,IL17F,IL23A,IL36RN,MSI1,MTA2,OCLN,PTGDS,ROCK2,RPS6,S100A8,SCEL,SERPINB2,SRD5A1,TGFB2,TGFB3,TJP1,VCAM1
Lipid Metabolism, Small Molecule Biochemistry,	3	ABCB4,ACADM,ACADVL,adenosine triphosphate,APOB,APOE,ATP1A1,B2M,CASP8,CAT,CD74,CD1D,CES1,CPT

`

Energy Production		1A,FABP1,FASN,GJB2,GSN,HMGCS1,IL10,INS,LDLR,LPIN1,MTTP,NCF1,PCK1,PPARA,PPARGC1A,PRKCZ,SCARB1,SIRT1,SLC2A1,SREBF1,THRSP,TNFSF10
Embryonic Development, Organismal Development, Cellular Development	3	ACVR1,ACVR2A,AURKB,BMP2,BMP4,BMP6,BMPR2,BMPR1A,DLL1,E2F4,E2F5,EDN1,EDNRB,ERBB3,FGF1,FGF2,FGFR1,FGFR2,FGFR4,Hamp/Hamp2,ID1,KIT,KITLG,LRP2,MITF,NF1,NOTCH3,NRG1,PA2G4,PKD2,RBL2,SHC1,SRC,TCF3,TF
Cell Death and Survival, Cancer, Organismal Injury and Abnormalities	3	ADRB2,AQP3,ASGR2,ATF2,CYP2A6 (includes others),CYP2E1,ELF3,ENPEP,EZR,FOXA1,FOXA2,FOXJ1,GATA4,GATA6,GFER,GLI1,HNF1A,HOXA5,HSD17B2,JAG2,KLF4,KRT19,MNX1,NKX2-1,NR1D1,RARB,SERPINA1,SLC9A3R1,SOX2,SSTR1,SULF2,TFF1,TFRC,TTR,VTN
Cardiovascular System Development and Function, Cellular Movement, Organismal Development	3	AGT,AGTR1,BIRC2,CDH3,CHRM4,COL4A1,CSTA,GADD45A,HIF1A,ILK,INHBB,ITGAV,LAMA3,LEF1,MAP3K7,MAPK8,MMP3,MMP7,MMP9,PDGFB,PLA2G2A,PLAU,PTK2,RAB8B,RIPK1,SERPINE1,SNAI1,SPARC,THBS1,THRB,TIMP1,TNFRSF1B,TRAF2,VIM,XIAP

The network “Inflammatory Disease, Tissue Development, Cellular Function and Maintenance” contains several genes involved in inflammatory disease. From Targetscan, two of the molecules in the network, AXL and CEBPA (CCAAT/enhancer-binding protein alpha), are predicted targets of Hco-miR-5352, with the latter molecule showing a significant decrease in expression in the presence of Hco-miR-5352. The network is focused around IL-6 and IL-1, both of which show a slight (but insignificant) decrease in expression in the presence of miR-5352. Interesting, IL-6 has previously been implicated in resistance in *H. contortus* infection (Estrada-Reyes *et al.*, 2015).

`

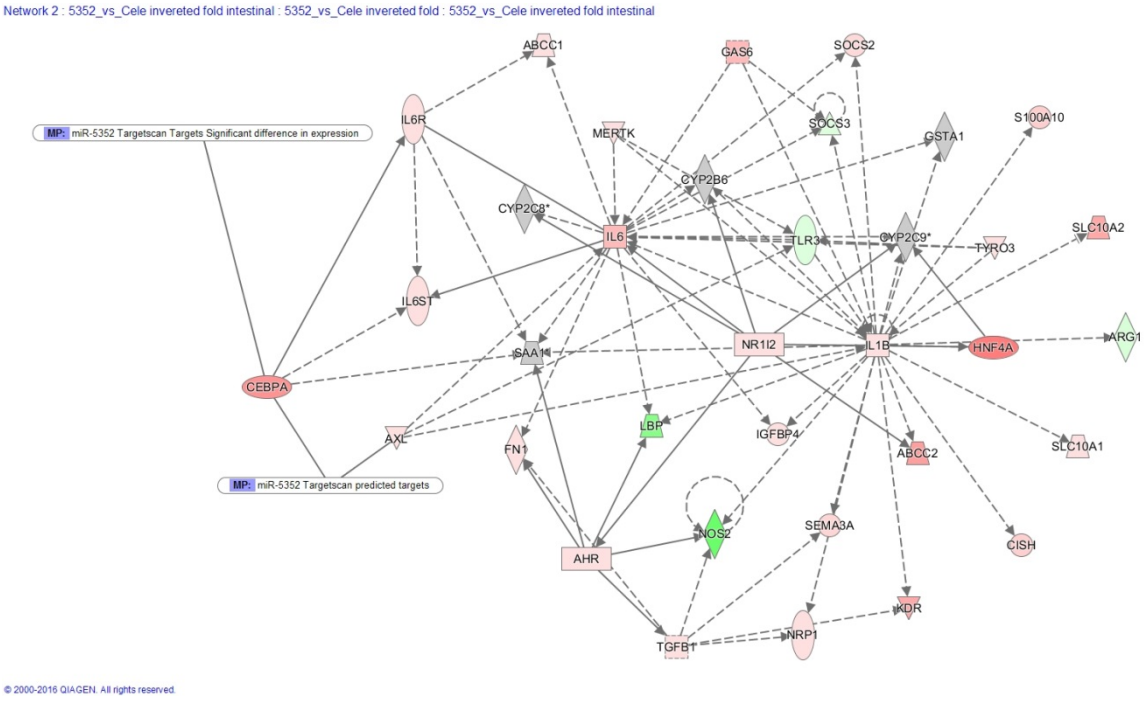


Figure 6.7 IPA network titled "Inflammatory Disease, Tissue Development, Cellular Function and Maintenance". Molecules coloured in red indicate decreased expression in Hco-miR-5352 transfected cells, whereas molecules in green indicate increased expression.

`

6.2.6 IPA analysis in all cell types

The previous section focused on IPA analysis specifically of epithelial cells, as the cell line used was an intestinal epithelial cell line. However, of the 23,000 genes that were identified by RNA-Seq, only 1414 of these genes are present in the epithelial cell list and of these, only 46 are classified as significantly differentially expressed. This could be due in part to the lack of information available on various genes, as certain interactions may have been performed in a specific cell line, although the results may also be applicable to other cells or tissues. In Chapter 5.2.1, the expression of *Haemonchus*-specific miRNAs was analysed in the abomasal mucosa and abomasal lymph nodes. For Hco-miR-5352, the miRNA was identified in both tissues, suggesting that the miRNA may not be limited to just epithelial cells and may be present in other cells or tissues.

To this end, the analysis was re-run using genes from all the tissues and cell lines that are present in the IPA database. 20,764 IDs were mapped out of the 23,776 genes identified by RNA-Seq (a similar number as mapped to the epithelial cell gene list). Of these, 17,772 genes were identified in the IPA database. This number was too large for IPA, so a subset of genes with a false discovery rate (q-value) of less than 0.05 was selected, leaving 447 genes (i.e. the most significantly differentially expressed genes. The 100 genes with the lowest q-values are shown in Table 8.5).

6.2.6.1 Canonical pathways

The 447 genes with significantly altered expression were then analysed by IPA. Of the canonical pathways identified, a large number of these pathways appear to be specific to immune cells, such as B cell receptor signalling, production of nitric oxide and reactive oxygen in macrophages and dendritic cell maturation. Table 6.15 shows a list of the 20 most significant pathways identified. Pathways shown below ("Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes", "Role of NFAT in Regulation of the Immune Response" and "Leucocyte Extravasation Signalling") were selected because they had a negative z-score, the pathways contained genes that are predicted Hco-miR-5352 targets and may be involved in modulation of the host response against *H. contortus*. In addition to these three pathways, there are a number of other pathways, all

`

with significant p-values but low z-scores such as “IL-1 signalling”, “IL-10 signalling”, “macrophages”, “the complement system”, “dendritic cells” and “B cells”. Some of these pathways are down-regulated and could be utilised by the parasite to reduce immune responses even though the z-score is higher than the threshold.

Table 6.15 The 20 most significant IPA canonical pathways identified. P-value calculated using a right-tailed Fisher Exact Test. Ratio indicates the number of genes from the input list that are present in the pathway. A ratio of one indicates all the input list genes are present, a score of 0 indicates none are present. Z-score is a prediction of whether the pathway is activated or inhibited. NaN indicates z-scores were not available.

Ingenuity Canonical Pathways	p-value	Ratio	z-score
Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes	2.69E-05	0.1	-3.162
Role of NFAT in Regulation of the Immune Response	8.13E-04	0.06	-2.333
Dendritic Cell Maturation	8.51E-04	0.06	-1.414
Leucocyte Extravasation Signalling	2.57E-03	0.05	-1.414
ATM Signalling	1.07E-03	0.1	-1.342
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	3.47E-04	0.06	-1.155
Complement System	8.71E-04	0.13	-1
Integrin Signalling	2.95E-03	0.05	-1
IL-1 Signalling	2.63E-03	0.07	-0.816
Ephrin Receptor Signalling	2.82E-03	0.06	-0.816
B Cell Receptor Signalling	2.45E-04	0.07	-0.632
Protein Kinase A Signalling	2.57E-03	0.04	0
PI3K Signalling in B Lymphocytes	1.26E-05	0.09	0.302
STAT3 Pathway	1.05E-04	0.11	0.707
p53 Signalling	7.24E-04	0.08	0.816
IL-10 Signalling	5.50E-04	0.1	NaN
Hepatic Fibrosis / Hepatic Stellate Cell Activation	1.29E-03	0.06	NaN
phagosome formation	1.45E-03	0.07	NaN
Role of Tissue Factor in Cancer	1.74E-03	0.07	NaN
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	2.88E-03	0.05	NaN

The first pathway identified with a significant z-score and which contained Hco-miR-5352 predicted targets is “Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes” (Figure 6.8). This is an extremely interesting pathway as it is predicted to be inhibited in the Hco-miR-5352 transfected cell, with a z-score of -3.2 and a p-value of 0.007. Macrophages and monocytes are an important component of the host response against GI nematodes and are known to be targeted by parasites (reviewed by McNeilly & Nisbet 2014).

`

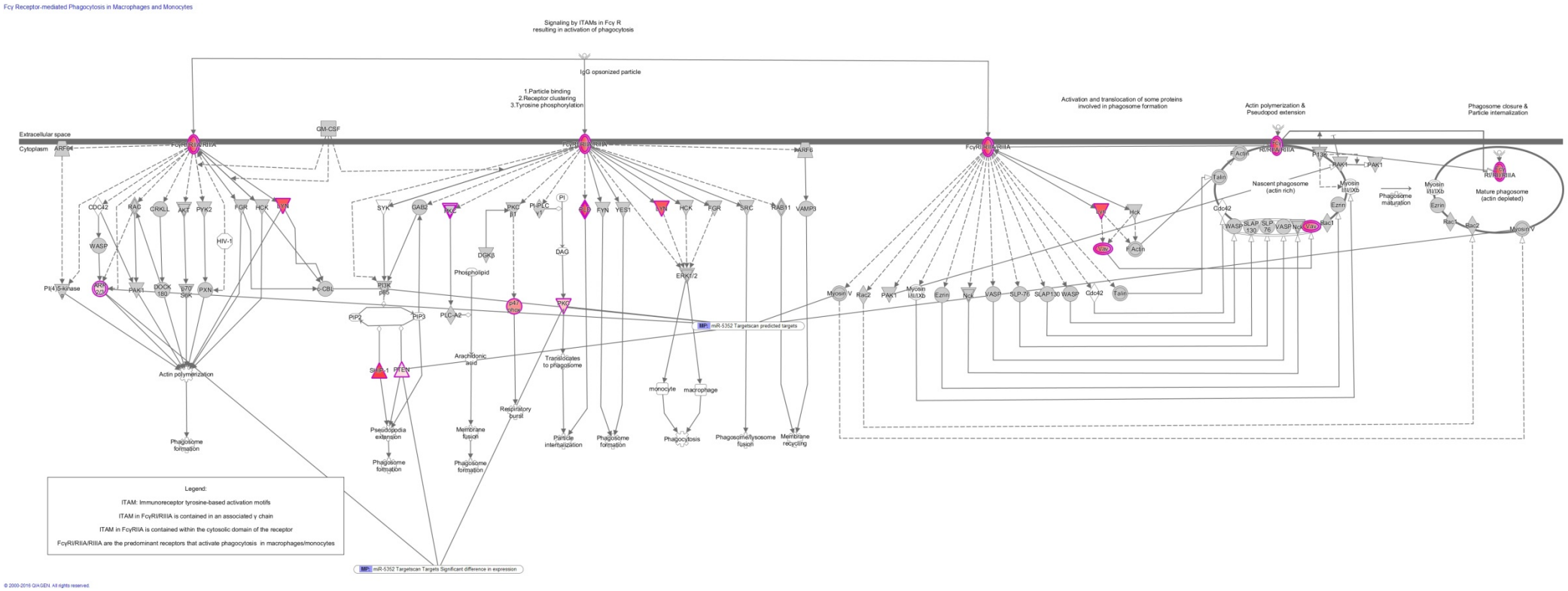


Figure 6.8 Canonical pathway of Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes

The second pathway identified was “Role of NFAT in the regulation of the immune response” (Figure 6.9). This pathway has some similarities with the NFAT pathway identified in the epithelial cell specific analysis (see section 6.2.5.1). This pathway has a z-score of -2.3 indicating inhibition. The pathway contains three genes that are predicted targets of Hco-miR-5352 and 11 genes that are significantly differentially expressed, although no gene falls into both categories. This pathway also contains NFATc4 which was identified in section 6.2.5.1. as well as c-Fos which was identified in 6.2.5.2.

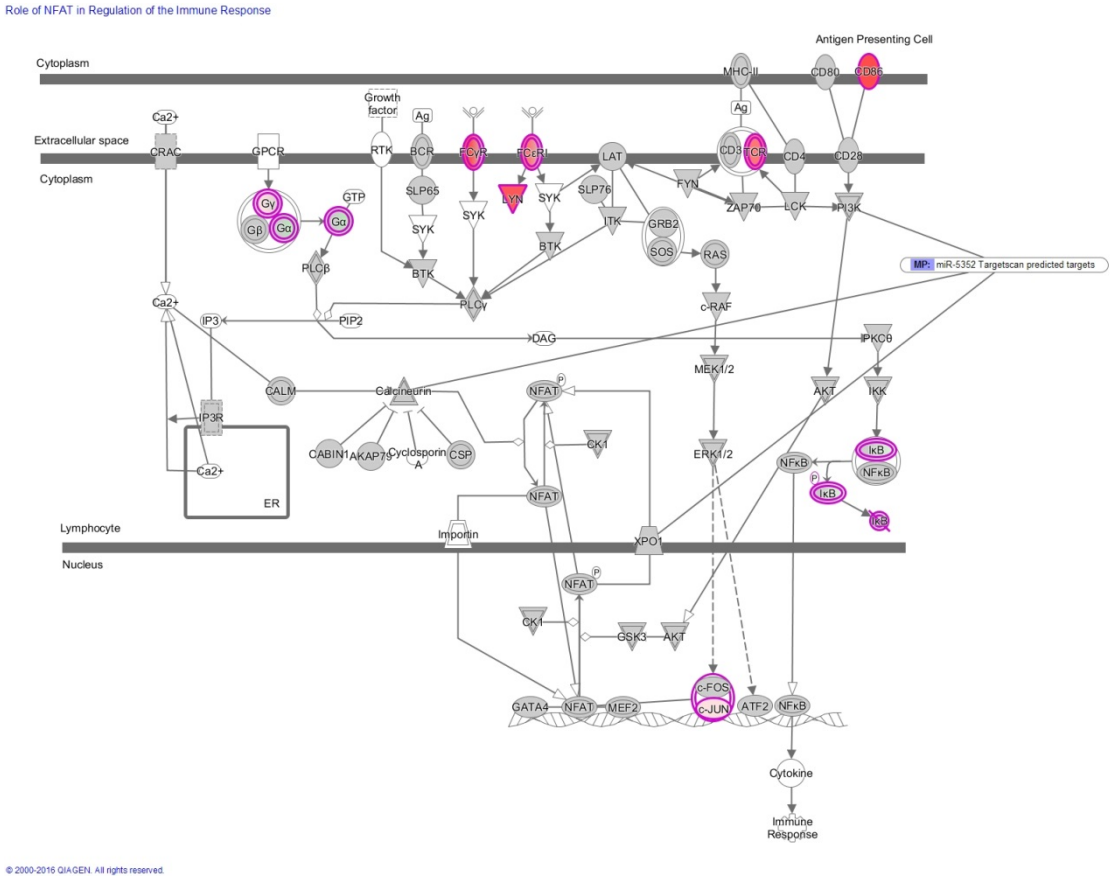


Figure 6.9 Canonical pathway of the role of NFAT in the regulation of the immune response showing the genes which are predicted targets of miR-5352 by Targetscan and those targets that have significant differential expression. Genes highlighted in red indicate decreased expression when Hco-miR-5352 is present. Genes in green indicate increased expression.

Leucocyte extravasation was again identified as a pathway. This pathway is identical to that shown previously in Figure 6.6. With the q-value cut off, only 11 genes of the 92 shown in Figure 6.10 have significantly differential expression and all are decreased. Based on these 11 genes, the z-score is calculated as -1.4, compared to a z-score of -2.3 in Figure 6.6, and is below the threshold.

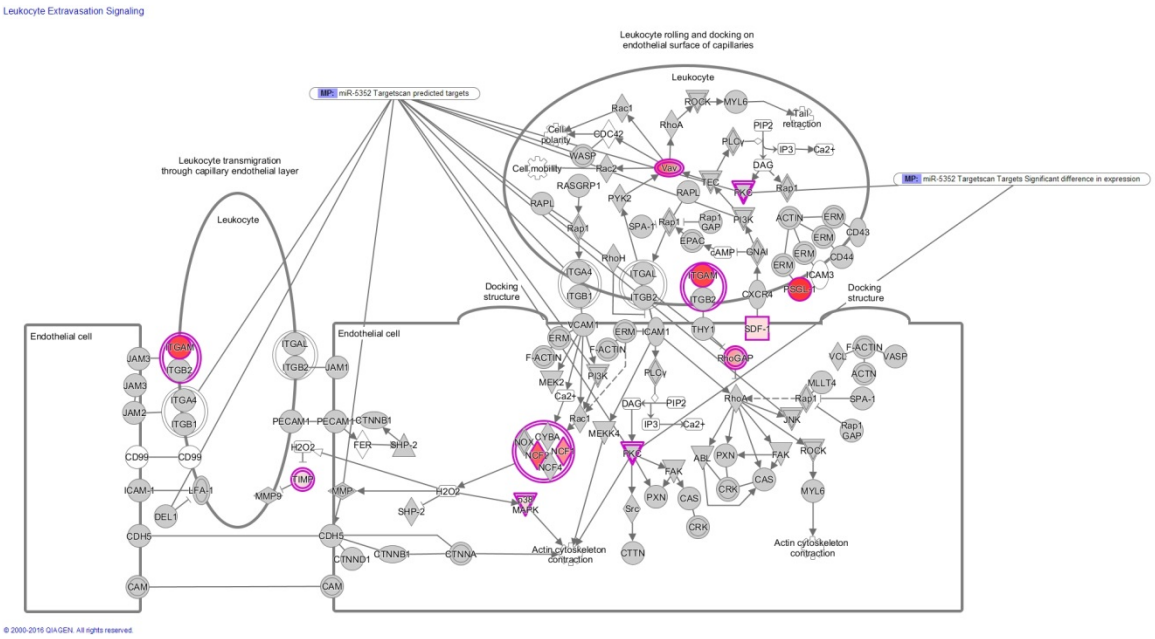


Figure 6.10 Canonical pathway of Leucocyte extravasation showing the genes which are predicted targets of miR-5352 by Targetscan and those targets that have significant differential expression. Genes highlighted in red indicate decreased expression when Hco-miR-5352 is present.

6.2.6.2 Predicted upstream regulators of identified genes

Upstream regulator analysis of the 449 significantly differentially expressed genes identified 1,679 molecules, of which 52 were predicted to be inhibited and 30 activated. This list of regulators includes miRNAs as well as a variety of proteins and chemicals. When focused on miRNAs only, 22 miRNAs were predicted upstream regulators, with one miRNA being predicted to be inhibited and four miRNAs that are activated. The other miRNAs all have z scores of less than 2.0 (Table 6.16).

Table 6.16 predicted miRNA upstream regulators that are predicted to be activated or inhibited.

Upstream Regulator	Predicted Activation State	Activation z-score	Target molecules in dataset
miR-223	Inhibited	-2.65	CD180, CD33, CSF1, HPGDS, Mt2, NDRG2, SGPL1
miR-122-5p (miRNAs w/seed GGAGUGU)		-1.28	ANK2,ATP1A2,SLC7A11,TMEM50B
miR-25		-0.81	BBC3,CDKN1A,PTEN,REV3L,TGFBR1
miR-17		-0.34	CDKN1A,FIGF,PTEN,S1PR1,TGFBR1
let-7		0.05	BTG2,CCNB1,CDK8,E2F6,IGF2BP1,TGFBR1
miR-221-3p (and other miRNAs w/seed GCUACAU)		0.10	BBC3,DDIT4,PTEN,TIMP3
miR-30c-5p (and other miRNAs w/seed GUAAACA)		0.45	IFRD1,JUN,MYO10,SLC7A11,SYPL1
miR-10		1.06	BAK1,CD86,FGFR3,GJA1,RPS6KA1,SGPL1,TGFBR1
let-7a-5p (and other miRNAs w/seed GAGGUAG)		1.21	EIF3J,FAM105A,IFRD1,IGF2BP1,MTRR,SPCS3,SYPL1,TGFBR1
miR-155-5p (miRNAs w/seed UAAUGCU)		1.48	CSF1R,CYP51A1,DPP7,INPP5D,MAFB,MYO10,PLXND1,SCAMP1,SYPL1,TWF1
miR-21		1.55	AIF1,BTG2,CCL3L3,CD180,CDKN1A,CLU,COL3A1,CSF3R,CXCL10,FCGR1A,FCGR2A,PTEN,SOCS6,TIMP3
miR-21-5p (and other miRNAs w/seed AGCUUAU)		1.64	BTG2,CDKN1A,PTEN,SOCS5,TIMP3
miR-291a-3p (and other miRNAs w/seed		1.81	ANKRD52,BAZ1A,CDKN1A,PRKACB,SEPT2,VPS26A

`

AAGUGCU)			
miR-29b-3p (and other miRNAs w/seed AGCACCA)		1.93	COL3A1,FBN1,LOXL2,PTEN
miR-486-5p (and other miRNAs w/seed CCUGUAC)		1.97	ADARB1,PIK3AP1,PTEN,TWF1
miR-130		1.98	GTF2H1,IRF8,MAFB,TGFBR1
miR-486		1.98	ADARB1,PIK3AP1,PTEN,TWF1
miR-1	Activated	2.19	EML4,GJA1,H3F3A/H3F3B,SNAI2,TWF1
miR-125b-5p (and other miRNAs w/seed CCCUGAG)	Activated	2.76	BAK1,C10orf54,CCR5,H3F3A/H3F3B,JUN,LIPA,MAPK14,SGPL1
miR-1-3p (and other miRNAs w/seed GGAAUGU)	Activated	3.40	ANKIB1,CTSC,EML4,FERMT2,GJA1,H3F3A/H3F3B,HAND2,MMD,PLEKHB2,TIMP3,TWF1,UNC93B1
miR-124-3p (and other miRNAs w/seed AAGGCAC)	Activated	3.71	CAV1,Ccl2,CD164,CEBPA,CHST14,CYP1B1,IFRD2,MAN2A1,MAPK14,MYO10,NEK6,SNAI2,SWAP70,SYPL1,TEAD1,TNFRSF21,ZNF367

miR-223 is the only miRNA that is predicted to be inhibited and has been shown to be dysregulated in leukaemia and lymphomas (reviewed by Haneklaus *et al.*, 2013) and suppresses the pro-inflammatory activation of macrophages. miR-223 does not show any sequence similarity to miR-5352.

miR-25 shows a significant degree of overlap with Hco-miR-5352 (see Chapter 5.2.3) and was identified by IPA. However, the miRNA has a z-score of -0.81 which indicates a slight inhibition based on the expression of BBC3, CDKN1A, PTEN, REV3L, TGFBR1, although this is below the threshold. Of these genes, only TGFBR1 and REV3L are also predicted targets of Hco-miR-5352.

`

6.2.6.3 Networks

IPA identified 25 different networks from the list of input genes and the ten best scoring networks are shown in Table 6.17. Two networks (Infectious Diseases, Cellular Development, Connective Tissue Development and Function; and Cancer, Cell Death and Survival, Organismal Injury and Abnormalities) were identified to be of interest and are discussed below.

Table 6.17 Top ten networks identified by IPA. The score indicates the likelihood of the genes in a network being found together by chance. A score of 2 or higher indicates a 99% chance of not being generated by chance. Genes highlighted in red or green indicate down-regulated or up-regulated genes respectively, from the input list

Top Diseases and Functions	Score	Focus Molecules	Molecules in Network
Infectious Diseases, Cellular Development, Connective Tissue Development and Function	42	27	ADAM10,ADD3,ADRB,Angiotensin II receptor type 1,ATF4,ATF5,BAK1,C10orf54,CD74,CDK2AP1,CHAC1,CSPG5,CXCL10,EML4,Fcrls,IFN Beta,Ifnar,IL10RA,LGMN,MFSD4,miR-143,Na+,K+ -ATPase,NDRG2,NEK6,Notch,PARVG,PI3K (family),PTN,RDH11,SASH1,Siglech,SLC7A11,SLCO2B1,Vegf,WNT7B
Cancer, Organismal Injury and Abnormalities, Gastro-intestinal Disease	38	25	ALT,ARHGAP28,ATP7A,C1q,C1QA,C1QB,C1QC,CFH,CTSC,ELL2,Fcgr2,GLRX,IFN type 1,IgG,IgG1,IgG4,Igh (family),INF2,ITM2B,KIAA0196,MAN2A1,MLKL,Mt1,Mt2,NFkB (complex),P-TEFb,PAK1IP1,RIOK3,ST3GAL1,STK40,TRIB3,TRIM16,UNC93B1,VPS26A,ZFAND6
Cell Morphology, DNA Replication, Recombination, and Repair, Nucleic Acid Metabolism	32	23	ADAMTS14,ANK2,ARHGAP4,ATP1A2,ATP1B2,BGN,COL3A1,Collagen Alpha1,Collagen type I,Collagen type II,Collagen type III,Collagen(s),Cpla2,CSF3R,CTSS,ENC1,ENTPD1,Eotaxin,ERK1/2,FBN1,Fc gamma receptor,FERMT2,Fibrin,FIGF,ITGAV,LAIR1,MMD,OLFM1,PCOLCE2,PKC alpha/beta,SEMA7A,SGPP1,SYK/ZAP,TIMP2,WISP1
Cancer, Cell Death and Survival, Organismal Injury and Abnormalities	31	23	Alpha tubulin,ASB10,CCNB1,Ck2,Cyclin E,DUSP10,GADD45B,H3F3A/H3F3B,Hdac,HISTONE,HPGDS,IFRD1,LEPROT,LOXL2,miR-148,Ngp,Pdgf (complex),PGAM5,phosphatase,PPP1CB,PPTC7,PTEN,PTPN2,Ras,Rb,Sapk,SAT1,SETD7,SIRPA,SNAI2,Sos,TCF,TINAGL1,VAV1,VGLL3
Cellular Development, Cellular Growth and Proliferation, Haematological System Development and Function	30	21	AIF1,Akt,APC/APC2,CD37,CD53,CD180,COL6A2,COL6A3,collagen,Collagen type VI,Cyclin A,DAAM1,ERRFI1,Fascin,Foxo,IGF2BP1,Igg2,Igg3,IgG2b,LAG3,LCP1,LY86,miR-27,miR-192,N-Cadherin,PIK3AP1,PIM3,PROS1,Pvr,Raf,S6K1,SOCS6,TNFRSF21,Tpm4,TSH
Lipid Metabolism,	28	20	ANGPTL2,ANGPTL4,ANKH,APOE,BTG2,Cd3

`

Small Molecule Biochemistry, Connective Tissue Development and Function			3,ERK,GOT,GOT1,HDL,HDL-cholesterol,L-type Calcium Channel,LDL-cholesterol,LIPA,LOXL4,LTBP1,LTBP2,Mac1,MAFB,MED15,PLIN2,PRKAA,RASAL3,SAA,SERCA,SGMS1,Smad2/3,SORT1,Sphk,TBCEL,TEAD1,TMSB4,TTR,VLDL,VLDL-cholesterol
Hereditary Disorder, Immunological Disease, Inflammatory Disease	28	20	7S NGF,ARMC1,Atf,Beta Tubulin,CNTFR,Creb,CYP51A1,DDIT4,DDOST,Dynamin,FES,Fibrinogen,GJA1,Gm-csf,GNAO1,GPS1,KIFAP3,Mapk,NCF1,NCF2,NGF,NPHP1,p70 S6k,Ppp2c,PRKAC,PRKACB,Rap1,RGS17,RPS6KA1,Rsk,SEMA3A,SLC30A4,SLC39A12,sphingomyelinase,unspecific monooxygenase
Hair and Skin Development and Function, Organ Morphology, Cancer	28	20	26s Proteasome,ADCY,CDKN1A,Cg,DPY30,EIF5,FJX1,GOSR1,GOSR2,Histone h3,Hsp70,Hsp90,Ikb,IKK (complex),IL1,Immunoglobulin,Insulin,KCTD12,LMAN1,MAP1LC3,MEGF10,miR-142,NFKBIE,PRKCE,Proinsulin,RNA polymerase II,S100A10,SCAMP1,SEPP1,SPCS3,TCN2,TGFBR1,Ubiquitin,XBP1,ZBTB18
Energy Production, Lipid Metabolism, Small Molecule Biochemistry	27	20	14-3-3,AChR,ADH5,AFF4,Ahr-aryl hydrocarbon-Arnt,alcohol group acceptor phosphotransferase,ALDOC,ARNT,Cbp/p300,CDA,CDK8,CDK16,CYP1B1,EAF1,FYTTD1,GTF2H1,Holo RNA polymerase II,IgG2a,Importin alpha,ISCA1,KPNA3,mediator,Mitochondrial complex 1,NDUFA5,Nos,OAZ1,Pka,PLEKHB2,PRKX,PTPase,SAE1,Smad,Tgf beta,TRIOBP,TSPAN7
Immunological Disease, Inflammatory Disease, Inflammatory Response	26	19	ADIPOR1,AMMECR1,CCP110,CPNE8,CYB5B,EBP,ELAVL1,ENOPH1,FAM102B,FAM105A,GALNT4,GSTP1,HMGB2,IER3IP1,Krtap14,LPGAT1,MARCH6,MOV10,MTRR,OSBPL2,PHTF2,PIGN,PJA2,PPFIBP1,RAPH1,RIOK3,SFXN3,SLC25A44,TNFRSF21,TSEN54,VKORC1L1,WDR55,XPO7,XXYLT1,ZDHHC7

The first network identified (shown in Figure 6.11) is centred around CXCL10 and VEGF with an IL10RA offshoot that is extremely down-regulated in the presence of miR-5352, based on the RNA-Seq data. VEGF (Vascular endothelial growth factor) is produced by many cell types including tumour cells, macrophages and platelets (Sunderkötter *et al.*, 1994; Verheul *et al.*, 1997). CXCL10 (C-X-C motif chemokine 10) is a cytokine that is secreted in response to IFN-γ and is a chemo-attractant for macrophages, T and dendritic cells as well as promoting T cell adhesion to endothelial cells (Angiolillo *et al.*, 1995; Dufour *et al.*, 2002). IL10RA (interleukin 10 receptor, alpha) signaling is a key component in regulating mucosal immune homeostasis in mice and humans. Loss of this receptor impairs the differentiation and function of T regulatory cells. In the current study, levels

`

of VEGF were unchanged in the presence of miR-5352, while CXCL10 was increased and IL10RA was significantly decreased in expression level.

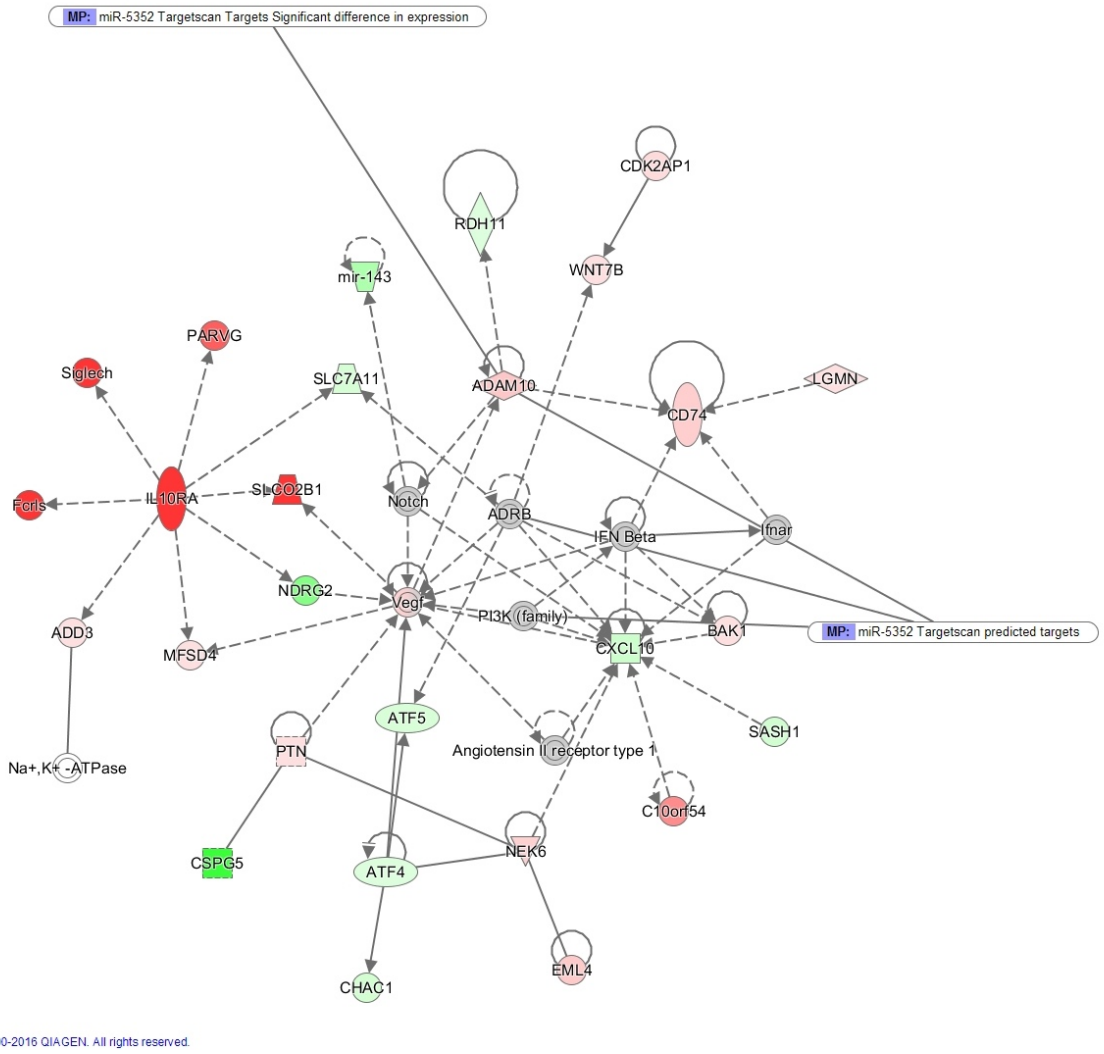


Figure 6.11 Network titled Infectious Diseases, Cellular Development, Connective Tissue Development and Function. Molecules coloured in red indicate decreased expression in Hco-miR-5352 transfected cells, whereas molecules in green indicate increased expression

Figure 6.12 shows clustering around Pten (Phosphatase and tensin homolog). This pathway is included because it contains 7 genes that are miR-5352 predicted targets (the highest number of all the networks analyzed). Pten de-phosphorylates PtdIns (3,4,5)P_3 to produce PtdIns(4,5)P_2 which terminates PI3K signaling. PTEN is commonly associated with cancers (Soond *et al.*, 2012) and is also important in T regulatory cells as it functions to repress Th1 cells and follicular helper T cells. Deficiency of Pten in Tregs resulted in excessive IFN-γ production, Th1 responses and mice spontaneously developed lymphoid hyperplasia (Shrestha *et al.*, 2015).

`

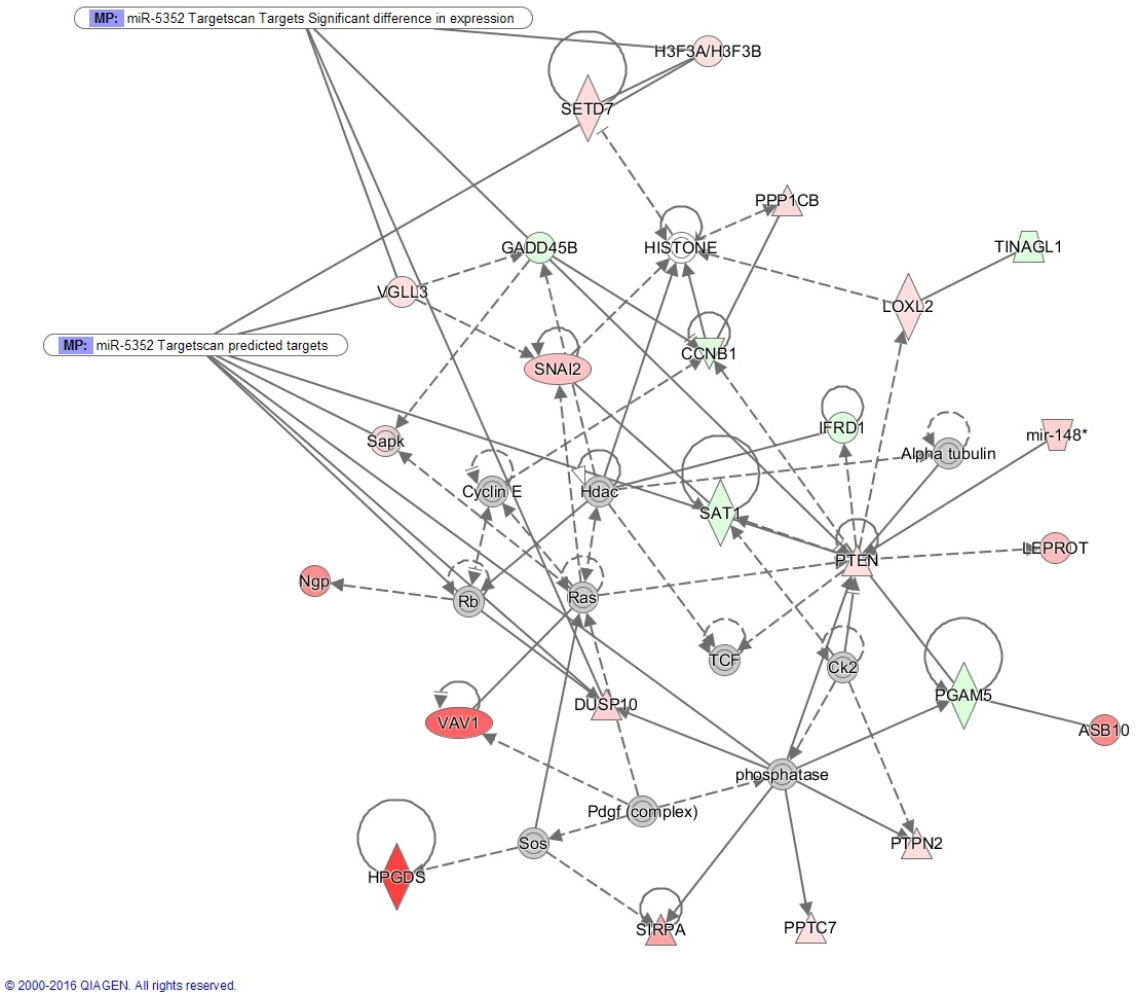


Figure 6.12 Network titled Cancer, Cell Death and Survival, Organismal Injury and Abnormalities. Molecules coloured in red indicate decreased expression in Hco-miR-5352 transfected cells, whereas molecules in green indicate increased expression

`

6.3 Discussion

Previous experiments using Jurkat cells and HEK293 cells were focused on analysing the effect of Hco-miR-5352 on CD69 expression; however, these experiments gave no indication as to the global effect of Hco-miR-5352 within the cell. In this chapter, gene expression was focused on MODE-K cells, a mouse intestinal epithelial cell line. To do this, two approaches were used. The first approach used RT² Profiler™ PCR arrays and focused on selected immune associated genes and the second approach used RNA-Seq to examine a much larger number of genes. MODE-K are mouse intestinal epithelial cells immortalised using SV40 (Vidal *et al.*, 1993). Further studies demonstrated that MODE-K cells were able to modulate antigen presentation to T cells and can bind IL-10, which inhibited the IFN-γ induced expression of MHC II, suggesting that these cells are a good model for intestinal epithelial cells.

The RT² Profiler™ PCR array identified a number of interesting genes from the Mouse Innate and Adaptive Immune Responses. For example, four members of the TLR signalling pathway, MyD88 and TLR 2, 4 and 6, were all significantly decreased when Hco-miR-5352 is present. The array also contained TLR 1, 3, 5, 7, 8 and 9. TLR 1, 3 and 5 showed fold changes of -1, indicating an insignificant decrease in expression whilst TLRs 7, 8 and 9 were undetectable in both Hco-miR-5352 and Cel-miR-67 transfected cells. Remarkably, the grouping follows the location of the TLRs as TLR 1, 2, 4, 5 and 6 are located on the cell surface whilst TLR 3, 7, 8 and 9 are localized to the endosomal/lysosomal compartment (Wu *et al.*, 2010). Of the endosomal/lysosomal localized TLRs, only TLR3 was detectable and it is also the only TLR that does not activate NFκB through MyD88 (Kawai and Akira, 2006). TLR3 instead induces production of type I interferons (Kawai and Akira, 2006). It is unclear why TLR 2, 4 and 6 are significantly decreased whilst the other cell surface TLRs are not. Although all five cell surface TLRs all signal through MyD88, each TLR has different ligands and TLR 2, 4 and 6 are not specific to any particular cell type.

Targeting of TLRs by GI nematodes has been shown in the literature. Work on ES-62 from *Acanthocheilonema vitae* showed that it interacts with TLR4 to inhibit degranulation and release of inflammatory mediators (Harnett *et al.*, 2010). In genetically resistant sheep, TLR2, 4 and 9 were abundantly expressed when

`

challenged with *H. contortus* or *Trichostrongylus colubriformis* (Ingham *et al.*, 2008).

Another interesting gene that was identified was CD80. This was the only gene that had a significant positive fold change from both arrays. CD80 along with CD86 provides co-stimulation to T cells by interacting with CD28 and CTLA4. Increasing CD80 expression would lead to activation of T cells. The increased expression of CD80 goes against the hypothesis that Hco-miR-5352 is targeting the host immune response as upregulation of this molecule would be detrimental to the nematode. Soluble egg antigen from schistosome parasites was shown to suppress DC activation and expression of CD80, MHC class II and CD86 (Kane *et al.*, 2004). However, human monocyte-derived DCs stimulated with *H. contortus* ES, showed no change in CD80 expression, although co-stimulation with ES and LPS showed slight, but insignificant decreases in expression (Rehman *et al.*, 2015).

Foxp3 (forkhead p3) was identified as significantly decreased from the array data. Foxp3 is transcription factor important in the development and function of regulatory T cells (Treg) (Fontenot *et al.*, 2003; Hori *et al.*, 2003). In general, induction of Foxp3 leads to increased Treg cells, which can reduce the inflammatory immune response. *H. polygyrus* infection in mice results in increased number of T cells expressing Foxp3 in the colon (Hang *et al.*, 2013). *H. polygyrus* and *T. circumcincta* secretions were also shown to induce de novo Foxp3 expression in Foxp3- splenocytes in mice (Grainger *et al.*, 2010). The data from the Hco-miR-5352 infected cells show a decrease in Foxp3 mRNA, but how that might relate to the presence of Tregs in the abomasum of infected sheep would require further study.

The second RT2 profiler was ‘cell surface markers’ and again some potentially interesting immune-related genes came out of this analysis. For example, of the eleven genes identified as significantly decreased, five are cell differentiation (CD) genes (CD63, CD276, CD247, CD1d1 and CD160). These five CD genes do not appear to share anything in common; they are present on different cell types, have different functions and have not been associated with GI nematode infections. CD63 is a tetraspanin protein frequently found in extracellular vesicles and used as a marker for EV (Lo Cicero *et al.*, 2015; Yáñez-Mó *et al.*,

`

2015b) although the abundance of CD63 was variable depending on the origin of the EV (Yoshioka *et al.*, 2013). It has also been shown that depletion of CD63 increases activation of CD4+ T cells, although this appears to be due to the increased production of exosomes when CD63 is depleted. In addition, CD63 is required for mast cell degranulation (Kraft *et al.*, 2013; Petersen *et al.*, 2011).

Interestingly, in *H. polygyrus*, exosomes were enriched for homologues of tetraspanin although it is unclear if these tetraspanins included CD63. Searches of the CD63 protein sequence in the NCBI nr database with BLAST identified a tetraspanin domain containing protein in *H. contortus* as well as in homologues in *T. suis*, *A. ceylanicum*, *A. suum B. malayi, C. elegans, C. remanei, C. briggsae L. loa* and *W. bancrofti*. Most of the hits had a query cover of around 98% although the percentage identity was low, between 29 to 38%.

The main issue with the array method is the small number of genes included in the analysis. The array is limited to 96 wells including the controls, and although this can be increased to 384 wells, it is still far fewer than the number of genes that can be analysed using the RNA-Seq method. This limitation means that when a gene of interest is identified, no data is available for other genes that are up or downstream, making it difficult to determine the overall effect of decreasing the expression of that one particular gene. On the other hand, the RT² Profiler™ may be more useful if a particular pathway of interest is found. If the RT² Profiler™ plate is populated with genes from a particular pathway, it would be relatively simple to determine the expression of those genes from many different conditions, such as differing concentrations of Hco-miR-5352, or gene expression over time.

In contrast, the RNA-Seq approach was able to identify 23,776 genes in the mouse cell, of which 467 showed significant differences in expression when transfected with Hco-miR-5352 relative to control miRNA mimic. The significantly down-regulated genes in the presence of Hco-miR-5352 appear to be predominantly associated with the immune response, whereas the up-regulated genes do not. The most significantly decreased gene was the purinergic receptor P2RY12. In response to tissue damage, extracellular nucleotides signal through purinergic receptors to activate a range of cells including mast cells (Gallucci and Matzinger, 2001). Enzymes secreted by *T.*

`

spiralis target the extracellular nucleotides to prevent the activation of mast cells and promote survival (Afferson *et al.*, 2012).

In comparison, the most significantly up-regulated gene, SLC6A1, is one of ten solute carrier genes that are up-regulated. These solute carrier genes are membrane transport proteins and are mostly located in the cell membrane, although the SLC25 family are located in the mitochondria (Palmieri, 2013). The reason for the up-regulation of these SLC proteins is unknown. The fourth most up-regulated gene is angiotensinogen. Angiotensinogen is produced from the liver and is cleaved by renin to create angiotensin I which is then targeted by the angiotensin converting enzyme (ACE) to form angiotensin II. Angiotensin II is a hormone that causes vasoconstriction, water and salt retention and increases blood pressure (Weir and Dzau, 1999). Although angiotensin has not been reported to be involved in the host response to *H. contortus* infection, it is possible that an increase in blood pressure within the gut may be advantageous for blood-feeding parasites such as *H. contortus*. *C. elegans* is known to possess an ACE-like protein which has been shown to be essential for larval development and morphogenesis (Brooks *et al.*, 2003), although no homologues have been found in parasitic nematodes (Page *et al.*, 2014).

With such a large number of genes, IPA was used in an attempt to make sense of the data by initially filtering the dataset for epithelial specific genes and then by looking at all cell types and applying various criteria to limit the number of genes. From the filtered genes, IPA was used to identify canonical pathways and networks that might be of interest. With this approach, all the genes that are affected by Hco-miR-5352, either directly or indirectly, can be identified and the network of the interactions can be determined using IPA.

The analysis using IPA generated a huge amount of data and identified a myriad of canonical pathways and networks involved in very different biological processes. It was therefore challenging to determine which of the pathways and networks are the most likely candidates. IPA was consequently filtered based on the hypothesis that the targets of Hco-miR-5352 are host immune response genes. This method was not able to identify the specific target of Hco-miR-5352, however, through use of the literature, we can theorise on how down-regulation of a certain pathway may benefit the parasite.

`

From IPA, the Leucocyte Extravasation Signalling canonical pathway was identified from the analysis of “epithelial cells” and from “all cell” types, both of which produced a negative z-score. Interestingly, when including all the genes in the pathway, the z-score was -2.3, whilst analysing only the significant differentially expressed genes gave a z-score of -1.4, which is below the threshold of significance. This result shows the difference between the two methodologies. It is unknown if Hco-miR-5352 will cause a significant decrease in the gene expression of the whole pathway. This pathway is involved in capturing leucocytes from the blood, tethering to the endothelial cells, following which, they migrate to the site of inflammation. Targeting of this pathway could reduce the migration of leucocyte to the site of infection. In the abomasal nematode *Ostertagia ostertagi*, transcriptomic analysis of the abomasum from cattle identified the leucocyte extravasation pathway as significantly induced during infection (Li *et al.*, 2010). In rats challenged with *N. brasiliensis*, the leucocyte extravasation signalling pathway was identified as significantly overrepresented in the intestinal muscosa (Athanasiadou *et al.*, 2011). Whilst this pathway appears to be significantly induced during nematode infection, the present analysis suggests it may be targeted by Hco-miR-5352.

Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes was identified from the “all cell” data analysis. Fcγ receptors are a class of Fc receptor that bind to IgG attached to foreign bodies and results in phagocytosis. The role of the Fcγ receptor in nematode infections is not clear. ES-62 from *A. vitae,* mentioned previously, is able to inhibit another class of Fc receptor, FcεRI, mediated immune responses. A possible role for the Fcγ receptor was suggested in mucosal mast cell responses against *Strongyloides venezuelensis* where Fcγ knockout mice had no difference in mast cell responses, but deletion of Fcγ reduced mast cell degranulation resulting in significantly higher faecal egg counts, worm burden and delayed worm expulsion (Onah *et al.*, 2000).

Another interesting pathway is the two NFAT (factor of activated T-cells) pathways. From the epithelial cell analysis, the pathway “The role of NFAT in cardiac hypertrophy” was identified, while “The role of NFAT in the regulation of the immune response” was identified in the “all cell” analysis. The two pathways show some similarity and both have a z-score less than -2 suggesting

`

that the pathway is down-regulated in the presence of Hco-miR-5352. NFAT is important in the function of T regulatory cells and T cell anergy, as well as having an impact on Th1 and Th2 differentiation, although the role of NFAT in immune cells has been focused on NFATc1-3 (reviewed by Daniel *et al.*, 2014). One noteworthy function of NFAT is that in mast cells, high NFAT activity is required for NF-α and IL-13 promoter induction. This coupled with the Fcγ pathway indicates that Hco-miR-5352 may be targeting mast cells.

One gene family, ADAM, was identified in several different analyses. Three ADAM genes were identified in the RNA-Seq dataset; ADAMTS14, ADAM32 and ADAM10. All three are significantly decreased in the presence of Hco-miR-5352, although only ADAM10 is a predicted target of Hco-miR-5352. ADAM10 is also in the Ephrin Receptor Signalling pathway, one of the 20 most significant IPA canonical pathways as well as in the highest scoring IPA network in the "all cell" analysis. ADAMTS14 is in the third highest scoring IPA network. ADAM10 is a sheddase and functions to cleave membrane proteins and substrates of ADAM10 include CX3CL1, CXCL16, IL-6, TNFα and VE-cadherin (Dreymueller *et al.*, 2012). These substrates of ADAM10 as well as substrates of ADAM17 are involved in the activation of inflammation in endothelial cells and the subsequent recruitment of leucocytes. In cattle infected with *O. ostertagi,* there is substantial hyperplasia in the abomasum, which may to be due to increased ADAM10 and ADAM17 in the gastric mucosa (Joh *et al.*, 2005; Mihi *et al.*, 2013).

Interestingly, from WormBase ParaSite, a homologue of ADAM10 can also be identified in a large number of different nematodes including two homologues in *H. contortus*. Of the two homologues, only one HCOI00696800 had a predicted binding site for Hco-miR-5352 using the Miranda prediction program. *H. contortus* also had one homologue of ADAMTSL5, ADAM11 and ADAM17 (ftp://ftp.wormbase.org/pub/wormbase/releases/WS247/species/o_volvulus/PRJEB513/annotation/o_volvulus.PRJEB513.WS247.orthologs.txt). The function of these ADAM homologues are unlikely to be inflammatory related, however, ADAM10 also has roles in neurodevelopment and has also been shown to regulate Notch signalling in mice (Hartmann *et al.*, 2002; Tsai *et al.*, 2014; Yang *et al.*, 2006).

`

Differences were also detected between the RT² Profiler™ PCR array and the RNA-Seq. For example, CD80, was the only gene on the array to show an increase in expression following transfection with miR-5352, relative to control transfection (Cel-miR-67). However, the RNA sequencing data showed a reduction in CD80 expression in miR-5352 transfected cells, although the difference between groups was not significant. IL-6 also showed a large fold change from the PCR arrays, however in the RNA-Sequencing gave values below the threshold for the Hco-miR-5352 transfected and Cel-miR-67 transfected cells.

In conclusion, this chapter sought to identify the genes that are differentially expressed when exposed to Hco-miR-5352 using a PCR array and then by RNA-Seq. A large number of genes were identified using the RNA-Seq method and were subsequently filtered using Qiagen's Ingenuity Pathway Analysis to identify pathways and networks of particular interest.

`

7 General Discussion

This project has focused primarily on analysing the possible functions of the *H. contortus* miRNA, Hco-miR-5352. This miRNA was initially found in *H. contortus* and was of interest due to its location in a miRNA cluster and its identification in the parasitic nematode *A. suum* but not in free-living *Caenorhabditis* species. Analysis in this project expanded our knowledge of this miRNA cluster, identifying it in ten different parasitic helminths all of which parasitise the GI tract. In addition, the miRNA cluster showed significant conservation in the order and orientation of the miRNAs present in the respective parasite nematode genomes. This conservation, coupled with the expression pattern of the cluster in adult *H. contortus*, presence in the excretory/secretory products and in the abomasal tissues of infected sheep, suggests possible functions for this miRNA cluster in adult worms and in host-parasite interactions.

Bioinformatic analysis was initially used to identify possible targets of the cluster in the *H. contortus* genome. However, target identification is hampered by the quality of the *H. contortus* genome as well as the availability of target prediction programmes. The genome is not yet fully annotated, making it difficult to identify genes of interest. To characterise the genes identified, homologues of *H. contortus* genes were identified in the closely related nematode, *C. elegans*. Previous analysis of the *H. contortus* and *C. elegans* genomes showed that over 80% of the orthologues that are present on the same *H. contortus* scaffold appear on the same *C. elegans* chromosome, although the order of genes is less well conserved (Laing *et al.*, 2013).

Target prediction programmes all have limitations based on the various features used to identify target genes. For example, there is evidence that non-conserved sites are functional, therefore programmes that focus on conservation would be unlikely to identify these sites (Farh *et al.*, 2005; Witkos *et al.*, 2011) and a high predicted free energy does not guarantee that an interaction occurs (Watanabe *et al.*, 2007). The analysis was also limited by the availability of programmes. To improve the power of the bioinformatic prediction, the results of multiple target prediction programmes were combined with available transcriptomic data for the mRNAs, which allowed the identification of miRNA-mRNA that show inverse expression data. However, not all interactions may produce a significant

`

decrease in expression. A number of programmes have been developed such as miRanda-mirSVR (Betel *et al.*, 2010) and DIANA-microT-CDS (Maragkakis *et al.*, 2009; Paraskevopoulou *et al.*, 2013) that utilise machine learning approaches. Machine learning techniques use a set of known miRNA-mRNA interactions to develop a model of miRNA targets. The model is then able to determine the power of certain features and predict further targets. These programmes were not used in the analysis as the prediction parameters cannot be adjusted and the results are pre-computed and do not accept custom miRNA or UTR sequences. It would be interesting to train these programmes using known *C. elegans* miRNA-mRNA interactions, and then apply this model to *H. contortus*.

The ES of parasitic helminths is known to be a source of host immunomodulatory molecules. Studies aimed at identifying *H. contortus* specific miRNAs in the ES using RNA-Seq identified a large number of RNA sequences in the ES, including a significant number of miRNAs. Comparing the ten most abundant miRNAs from the ES and in whole worm libraries showed significant differences, suggesting that there is selective release of miRNAs in the ES. Comparing the miRNAs identified in the ES of the parasitic nematode *H. polygyrus* (HES) with that of *H. contortus* ES showed that a number of miRNAs, such as miR-100, miR-71, let-7, lin-4 and miR-60, are present in both libraries and are also widely conserved not just in nematodes (Buck *et al.*, 2014). The ranking of these miRNAs in the two libraries varied significantly, for example, miR-71 is ranked second in HES, whilst it is ranked 18 in *H. contortus*.

miRNAs have been identified in the ES of a number of parasitic helminths and studies show that parasite miRNAs in serum could be used to distinguish between infected and uninfected individuals, although there is not necessarily a correlation between miRNA copy numbers and the intensity of infection (Tritten *et al.*, 2014). Most studies have focused on filarial nematodes and schistosome parasites (Hoy *et al.*, 2014; Quintana *et al.*, 2015; Sotillo *et al.*, 2016; Wang *et al.*, 2015; Zamanian *et al.*, 2015). Interestingly, these helminths all reside or pass through the blood stream of the host. For GI tract nematodes, there is no evidence to date that miRNAs are released into the bloodstream: parasite specific miRNAs have not been detected in the serum of mice infected with *H. polygyrus* (Buck *et al.*, 2014) and preliminary experiments were unable to

`

amplify *H. contortus* miRNAs from serum of infected sheep (Winter *et al.*, unpublished data). As adult *H. polygyrus* and *H. contortus* reside primarily in the GI tract, the miRNAs may not gain access into the circulation. However, *H. contortus* does secrete a range of parasite specific miRNAs in the excretory-secretory products, at least *in vitro,* and some of these can be detected in gut epithelial and lymph node tissues from the site of infection. One possibility is that GI nematode miRNAs could be present in the faeces. Recent work has shown that miRNAs can be detected in the faeces of humans and can be used as a non-invasive biomarker for colorectal cancer (Yau *et al.*, 2015).

In this thesis, it was shown that *H. contortus* miRNAs can be detected in the ES and that selected miRNAs are present within extracellular vesicle-like structures (EVs) present in the ES. However, it is not clear whether the vesicles observed in *H. contortus* ES are exosomes, as no further characterisation was carried out to enable their precise identification. Further studies could include immunoblotting or mass spectrometry on isolated vesicles to identify proteins that are known to be present in exosomes, such as Alix, CD9, CD63 and Heat shock protein 8 (Mathivanan and Simpson, 2009).

Exosomes from cultured mouse and human cells have been shown to contain miRNAs (Valadi *et al.*, 2007), and have been demonstrated to transfer miRNAs between cells (Chen *et al.*, 2012). These structures are impermeable to RNAses, suggesting a possible explanation for the stability of miRNAs in bodily fluids (Koga *et al.*, 2011). The analysis of the *H. contortus* EV-enriched fraction versus the EV-depleted fraction suggested that not all of the miRNAs may be encapsulated in EVs. The EV-depleted fraction had around 30% lower reads compared to the EV-enriched fraction, however, there were still approximately 8 million reads in the adult and L4 EV-depleted ES. In the mammalian cell lines HepG2, A549, T98, BSEA2B and human primary fibroblasts, miRNAs were found outside of the exosomes and the cells also released a number of RNA-binding proteins that may protect these miRNAs from degradation (Wang *et al.*, 2010d). Further studies suggest that these extracellular miRNAs are most likely associated with argonaute proteins, which are highly stable (Arroyo *et al.*, 2011; Turchinovich *et al.*, 2011). In contrast to the data from *H. contortus* ES, it has been suggested that the majority of extracellular miRNAs from plasma/serum

`

and cell culture media are not contained within EVs (Arroyo *et al.*, 2011; Turchinovich *et al.*, 2011). Exosomes purified from plasma, seminal fluid, dendritic cells, mast cells, and ovarian cancer cells showed, on average, less than one miRNA per exosome suggesting that the exosomes may not carry biologically significant numbers of miRNAs (Chevillet *et al.*, 2014).

The presence of miRNAs in the ES products correlates with data from other nematodes, where miRNAs have been shown to be present in extracellular vesicles or exosomes (Buck *et al.*, 2014; Nowacki *et al.*, 2015; Quintana *et al.*, 2015). In *H. contortus,* many miRNAs showed approximately equal abundance in both EV-depleted and EV-enriched libraries. The purpose of enclosing certain miRNAs in EVs but not others, is unclear. It is possible that this separation is due to functional differences. Exosomes have been identified in a range of protozoan, fungal and helminth infections and all appear to target various components of the immune system to aid parasite/pathogen survival within the host (reviewed by Coakley *et al.*, 2015), although whether all of these exosomes contain miRNAs and what their function may be is unknown. An attractive hypothesis is that miRNAs present in the EV are targeting host cells, as exosomes can fuse with the plasma membrane and deliver a cargo to the recipient cell (Raposo and Stoorvogel, 2013b). This could include microRNAs (Buck *et al.*, 2014) or proteins capable of modulating immune responses.

In addition to miRNAs, a large number of other RNA species were identified in the ES libraries. Whilst the function and origin of the rRNAs and mRNAs in the ES is still unknown, there is evidence from *T. cruzi* and *S. mansoni* that EVs contain a class of tRNA-derived-small RNAs which were able to modify expression of certain host genes (Garcia-Silva *et al.*, 2014a, b; Nowacki *et al.*, 2015).

As the ES is presumably released within the abomasum of infected animals, it is also possible that miRNAs could target other *H. contortus* nematodes, to promote survival, for example. Transfer of miRNAs between individuals of the same species has been demonstrated in bovines. Microvesicles containing RNA can be isolated from milk. These miRNAs were shown to target various immune genes, were not affected by the pH of the GI tract and could be detected in cells cultured with the microvesicles, suggesting a possible role in supporting the development of the calf (Hata *et al.*, 2010). It has been shown using

`

fluorescently-labelled siRNA that developing L4 larvae of *H. contortus* can ingest RNA, at least in vitro (Britton *et al.*, 2015). As the miR-5352 sequence is extremely similar between all the GI nematodes with available genome information, it is possible that released miRNAs may be a form of cross-species communication, such as in sheep with mixed nematode infections.

While the identification of possible *H. contortus* targets of miRNAs did not yield many convincing targets, the characteristics of the miR-5352 cluster suggested that the targets of the miRNAs might be expressed in host cells. Carrying out target gene predication on mammalian mRNAs suggested CD69 as a possible target. The targeting of host genes by miRNAs from an infectious body has been shown in a number of viruses such as *Herpesvirus saimiri*, which was able to modulate the expression of cell cycle regulators of marmoset T cells (Guo *et al.*, 2015). miRNAs from the Epstein–Barr virus have also been shown to target host genes to prevent apoptosis through the pro-apoptotic gene p53, as well as *BCL2L11* (BIM) (Marquitz *et al.*, 2011; Riley *et al.*, 2012; Skalsky *et al.*, 2012).

To investigate the possible interactions of a parasite miRNA with host genes, Targetscan was used to identify genes with binding sites for Hco-miR-5352 that are conserved across different species. Conservation was considered important as miR-5352 is present in a number of different parasitic nematodes, all of which have different hosts, suggesting that the targets may also be conserved. CD69 was the top predicted target for Hco-miR-5352. CD69 is an interesting target gene as it plays an essential role in regulating the inflammatory process in the intestine (Radulovic and Niess, 2015) and has been shown to be targeted by endogenous miRNAs. The rhesus macaque version of CD69 mRNA was predicted by Targetscan to be targeted by miR-92a and subsequent dual luciferase assays showed that the miR-92a mimic reduced CD69 expression by 50% (Kumar *et al.*, 2016). Luciferase reporter assays also showed that miR-181a is able to repress the expression of CD69 (Neilson *et al.*, 2007). In the gut, most T cells are reported to express CD69 compared to circulating T cells (Testi *et al.*, 1994) and CD69 is also up-regulated on T cells at the site of infection (Mori, 2011; Pignatti *et al.*, 2002). The high level of CD69 expression in intestinal T cells has been shown to be a result of the intestinal microflora and the constant exposure to antigen, as reducing the microflora resulted in decreased CD69 expression

`

(Jeong *et al.*, 2012; Radulovic *et al.*, 2012). T cells within the GI tract play a role in eliminating infections as well as regulating the immune response. In several different diseases, such as asthma (Martín *et al.*, 2010) and inflammatory bowel disease (Radulovic *et al.*, 2013), the absence of CD69 exacerbated disease severity. For example, $CD69^{-/-}$ mice infected with the intracellular pathogen *Listeria monocytogenes* were more susceptible to infection. The mice showed increased levels of Th1 cytokines, cell death and tissue damage suggesting an uncontrolled inflammatory response and reduced number of pathogen specific T cells (Vega-Ramos *et al.*, 2010). Activation of CD69 is also reported to induce TGF-β expression, suppress production of pro-inflammatory cytokines IL-17 and IFN-γ (Sancho *et al.*, 2003) and induce Foxp3+ regulatory T cells (Radulovic *et al.*, 2012). CD69 up-regulation leads to ERK activation and maintains high expression of membrane-bound TGF-β on T cells (Han *et al.*, 2009). There are no data available on the potential role of CD69 in *H. contortus* infection, but targeting of the CD69 gene by Hco-miR-5352 could perhaps benefit *H. contortus* by reducing specific T cell responses and inflammation at the site of infection. Further work will be required to investigate the precise cell type that expresses CD69 in the ovine abomasum or GI tract.

A dual luciferase assay was utilised and showed that co-expression of a Hco-miR-5352 mimic in transfected HEK cells resulted in a significant decrease in expression of CD69. However, this interaction was not confirmed in the more physiological Jurkat T cell after PMA stimulation. It is unknown what type of cells Hco-miR-5352 targets, or if the differences between mouse and sheep cells are important for the function of miR-5352, although analysis of the CD69 gene in both species suggests that the binding site is well conserved. From the Targetscan prediction, it is also known that miR-25 is predicted to target CD69 and shows significant binding site overlap with miR-5352. It is therefore possible that there may be competition for binding on the CD69 3' UTR between Hco-miR-5352 and other mammalian miRNAs.

Subsequent experiments attempted to identify Hco-miR-5352 targets from a larger number of genes with the RT^2 Profiler Arrays and then with a whole transcriptome approach. The most interesting target pathway identified by the RT^2 Profiler Array was the TLR signalling pathway as four genes, MyD88, TLR 2, 4

`

and 6 were all significantly decreased in the presence of Hco-miR-5352, with TLR 1, 3 and 5 also showing a decrease, although this was below the significance threshold. Whilst these genes are not predicted to be direct targets of Hco-miR-5352, TLRs are important in the immune response against GI nematodes (Harnett *et al.*, 2010; Ingham *et al.*, 2008) and their decrease could influence host susceptibility to *H. contortus*.

A whole transcriptome approach to identify Hco-miR-5352 targets utilised RNA-Seq of MODEK cells transfected with Hco-miR-5352 mimic. This method identified a huge list of differentially regulated genes; significantly, a large proportion of down-regulated genes in the presence of Hco-miR-5352 were immune related while genes that were up-regulated appeared predominantly to be solute carriers. Unfortunately, the method is unable to pinpoint which of the genes are the direct targets of Hco-miR-5352. Analysis by Targetscan of genes that were predicted to contain miR-5352 binding sites did not show any significant enrichment of GO terms, nor did they share any features in common, perhaps signifying the limitation of current target prediction programmes. Ingenuity Pathway Analysis and Gene ontology were used to attempt to identify canonical pathways that may be targeted by Hco-miR-5352. With the large number of pathways identified by IPA, it was necessary to limit the analysis and because many of the down-regulated genes had immune function, pathways that are involved in the immune response were the focus of subsequent analysis.

IPA however, cannot definitively identify a pathway or network that is targeted by Hco-miR-5352, although it was able to identify several targets for further study. In particular, the leucocyte extravasation pathway was identified both in the “epithelial cell” and “all cell” analysis and may be important in resistance against *H. contortus* infection. The leucocyte extravasation pathway is the process by which leucocytes are attracted to the site of infection and migrate from the blood to the tissue during inflammation (Middleton *et al.*, 2002). Down-regulation of this pathway would reduce the number of leucocytes present at the site of infection and reduce the immune response to *H. contortus* infection. The expression of a number of genes/pathways associated with mast cells were also altered in the presence of Hco-miR-5352. Mucosal mast cells have been shown to have an important function in expulsion of parasites including *H.*

`

contortus and decreasing mast cell activity could benefit survival of *H. contortus* (Meeusen *et al.*, 2005; Pérez *et al.*, 2003). However, it cannot be ruled out that Hco-miR-5352 may not be targeting the host immune system. Other methods, such as High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP) may be more useful (Chi *et al.*, 2009). HITS-CLIP utilises ultraviolet radiation to crosslink RNA to RNA binding proteins followed by immunoprecipitation of the argonaute-RNA complexes. The proteins are then degraded and RNA extracted and sequenced (Zisoulis *et al.*, 2010). This approach would produce a list of miRNAs and mRNAs that interact within the cell. However, this method still requires bioinformatic analysis to identify interactions between the miRNA and mRNAs.

Perhaps the most interesting aspect of these ES miRNAs is their presence in sheep tissue and the possible transfer of miRNAs between the nematode and the host with the subsequent targeting of host mRNAs. There has been a large amount of work showing that miRNAs can be a form of intercellular communication, as referred to above. In addition to intracellular communication, evidence of cross-species modulation by microRNAs has been demonstrated. Recent studies by Zhang *et al.* showed that plant miRNAs could be detected in the blood and tissues of humans and other plant eating mammals. Further studies suggest that these plant miRNAs may be able to regulate gene expression in mammals (Zhang *et al.*, 2012) although this has been disputed (Dickinson *et al.*, 2013). It has also been found that in individuals with sickle-cell anaemia, the sickle-cells have a dysregulated miRNA profile which may be associated with protection against *Plasmodium* infection. miR-451 and let-7i, which are highly enriched in sickle-cell erythrocytes were able to reduce *Plasmodium falciparum* growth. However, *Plasmodium* does not contain Dicer or Argonaute homologues and thus do not possess RISC; here RNA interference is through the miRNAs fusing with the *Plasmodium* mRNA and impairing ribosome loading (LaMonte *et al.*, 2012). In the nematode *H. polygyrus,* parasite specific miRNAs were shown to be able to reduce expression of certain host genes using a dual luciferase assay (Buck *et al.*, 2014).

In addition, host miRNAs have also been suggested to play a role in regulating the gut microbiota. Faecal miRNAs from mice, some of which were identified in

`

EVs, were shown to enter cultured bacteria and reduce growth. Mice lacking Dicer had reduced miRNA levels and more diverse gut flora. Transplanting of faecal miRNAs restored diversity to wild type levels (Liu *et al.*, 2016). The data suggest that the gut flora is susceptible to miRNA-mediated regulation, and raises the possibility that the miRNAs released by *H. contortus* could also be targeting the gut flora.

In summary, this project has characterised the mir-5352 cluster from *H. contortus* and shown that it is highly conserved within GI nematodes. *H. contortus* was demonstrated to release a select number of miRNAs in the ES products, either in extracellular vesicles or in the supernatant. Some of these miRNAs were detected in sheep tissues collected from the site of infection. Analysis of the effect of Hco-miR-5352 on a mouse intestinal epithelial cell suggests that Hco-miR-5352 could be involved in modulating the host immune system although further studies will be required to elucidate the direct targets of Hco-miR-5352 and its contributions to *H. contortus* survival *in vivo*

`

8 Appendix

8.1 Conservation of 5352

Table 8.1 Nematode species that contain the miR-5352 cluster.

Nematode	Organism	Site of infection	Source	Reference if published
Haemonchus contortus	sheep	Abomasum	http://parasite.wormbase.org/Haemonchus_contortus_prjeb506/Info/Index/ http://parasite.wormbase.org/Haemonchus_contortus_prjna205202/Info/Index/	(Laing *et al.*, 2013) (Schwarz *et al.*, 2013)
***Ascaris* suum**	Pigs	Small intestine	http://parasite.wormbase.org/Ascaris_suum_prjna62057/Info/Index/ http://parasite.wormbase.org/Ascaris_suum_prjna62057/Info/Index/	(Wang *et al.*, 2012a) (Jex *et al.*, 2011)
Oesophagostomum dentatum	Pigs	Large intestine, caecum	http://parasite.wormbase.org/Oesophagostomum_dentatum_prjna72579/Info/Index/	
Nippostrongylus brasiliensis	Rats	Small intestine	http://parasite.wormbase.org/Nippostrongylus_brasiliensis_prjeb511/Info/Index/	
Necator Americanus	Humans	intestine	http://parasite.wormbase.org/Necator_americanus_prjna72135/Info/Index/	(Tang *et al.*, 2014)
Strongylus vulgaris	Equine	Large intestine	http://parasite.wormbase.org	

`

			/Strongylus_vulgaris_prjeb531/Info/Index/	
Heligmosomoides bakeri/ (polygyrus)	Woodmice and small rodents	Duodenum and small intestine	http://parasite.wormbase.org/Heligmosomoides_bakeri_prjeb1203/Info/Index/	
Cylicostephanus goldi	Horses	Large intestine	http://parasite.wormbase.org/Cylicostephanus_goldi_prjeb498/Info/Index/	
Ancylostoma duodenale	Humans	Intestine	http://parasite.wormbase.org/Ancylostoma_duodenale_prjna72581/Info/Index/	
Ancylostoma ceylaniucm	Humans and hamsters	Intestine	http://parasite.wormbase.org/Ancylostoma_ceylanicum_prjna72583/Info/Index/	

`

8.2 Folding of Dviv

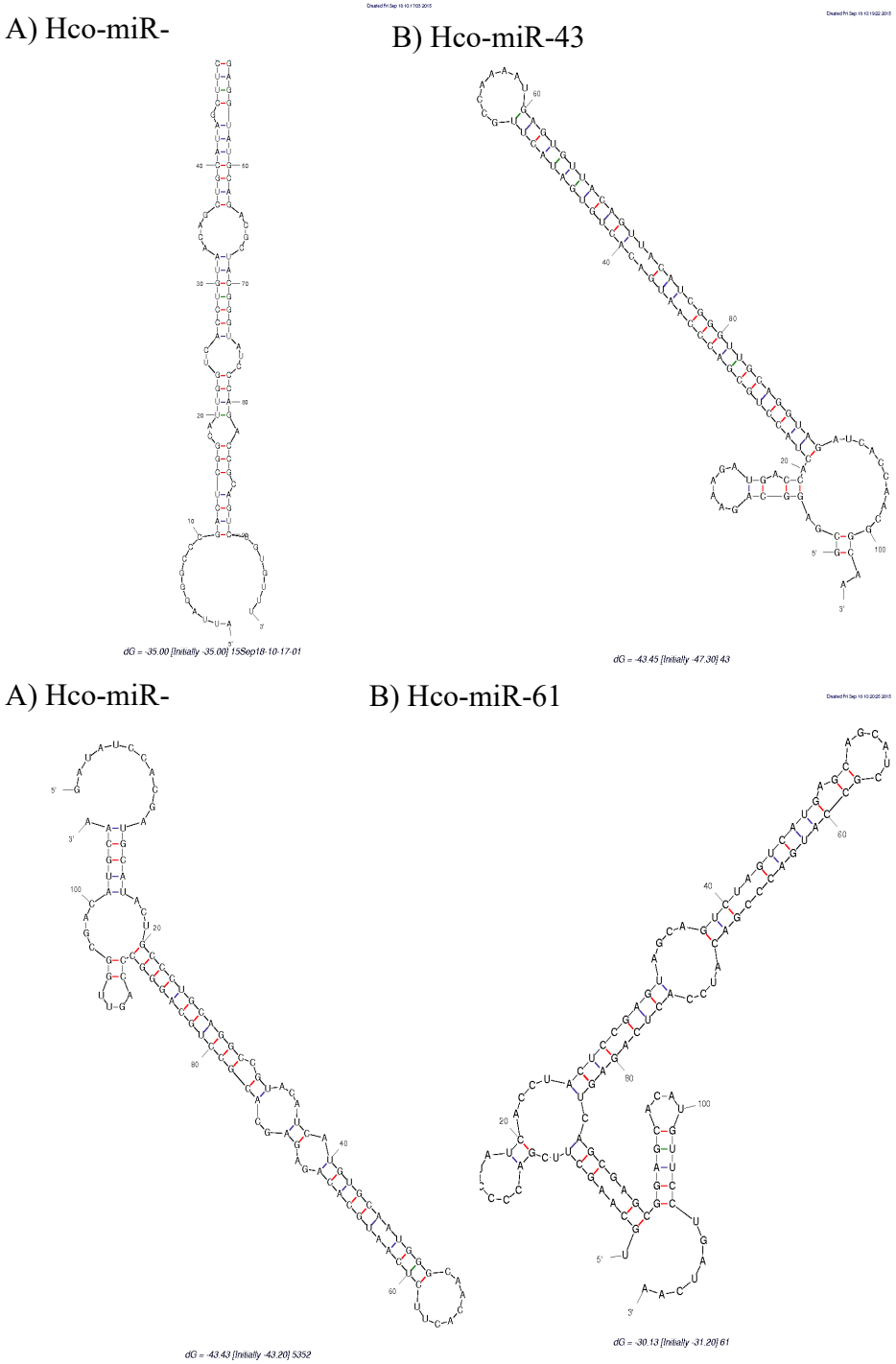


Figure 8.1 RNA folding of miR-5352 homologues in *D. viviparus* using mfold (http://unafold.rna.albany.edu/?q=mfold).

`

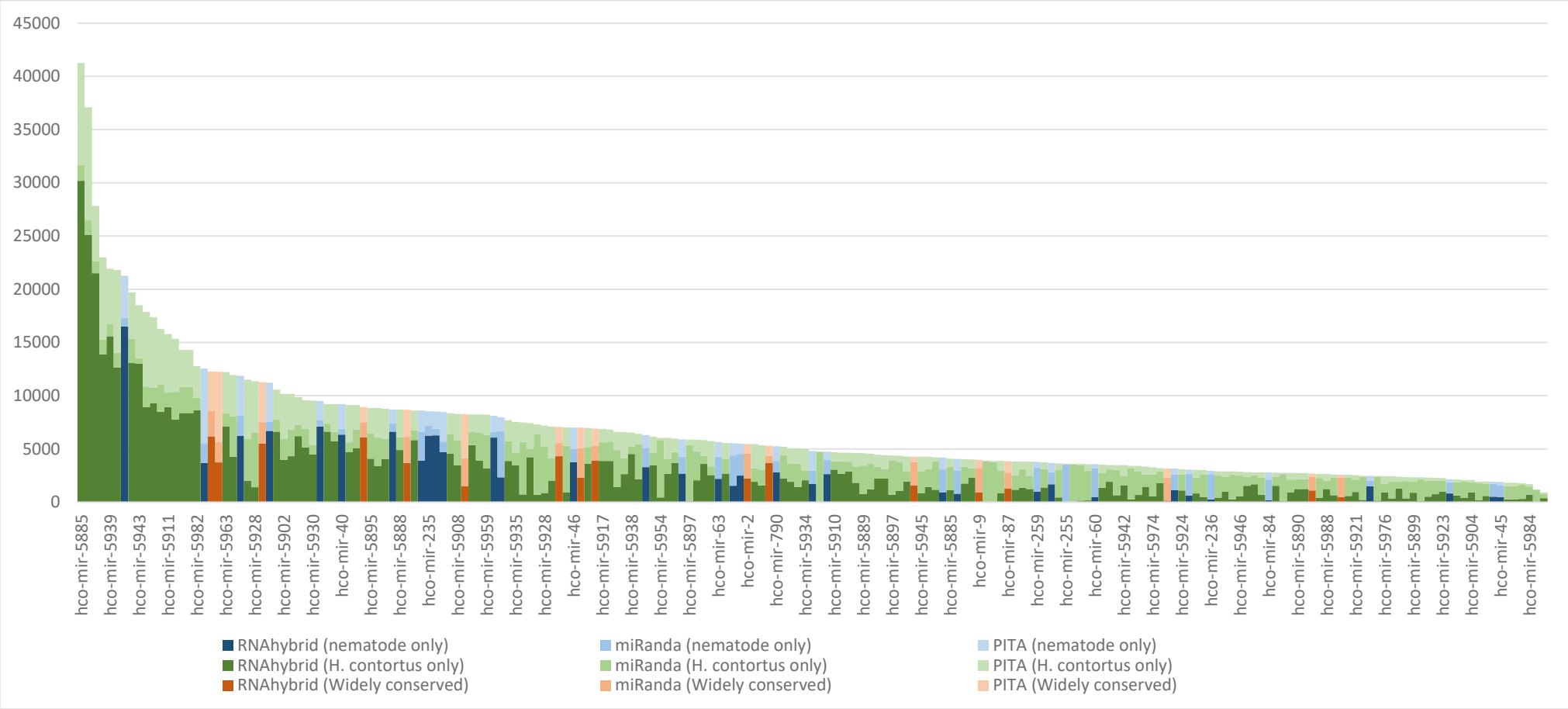


Figure 8.2 Number of targets identified for each miRNA using RNAhybrid, miRanda and PITA. miRNAs coloured in blue indicate they that are nematode specific. Green indicate miRNAs that are unique to *H. contortus*. Red indicates miRNAs that are widely conserved.

`

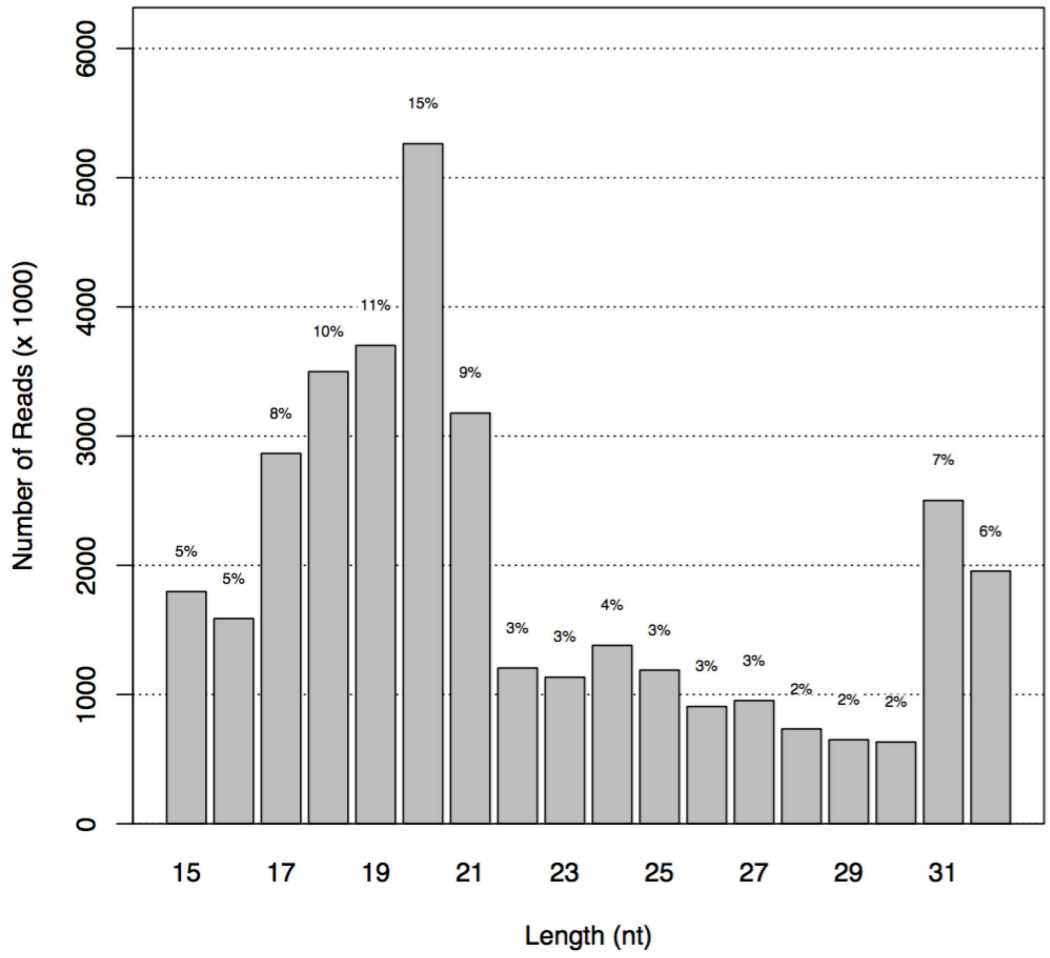


Figure 8.3 Length distribution of the mappable reads in all the ES libraries.

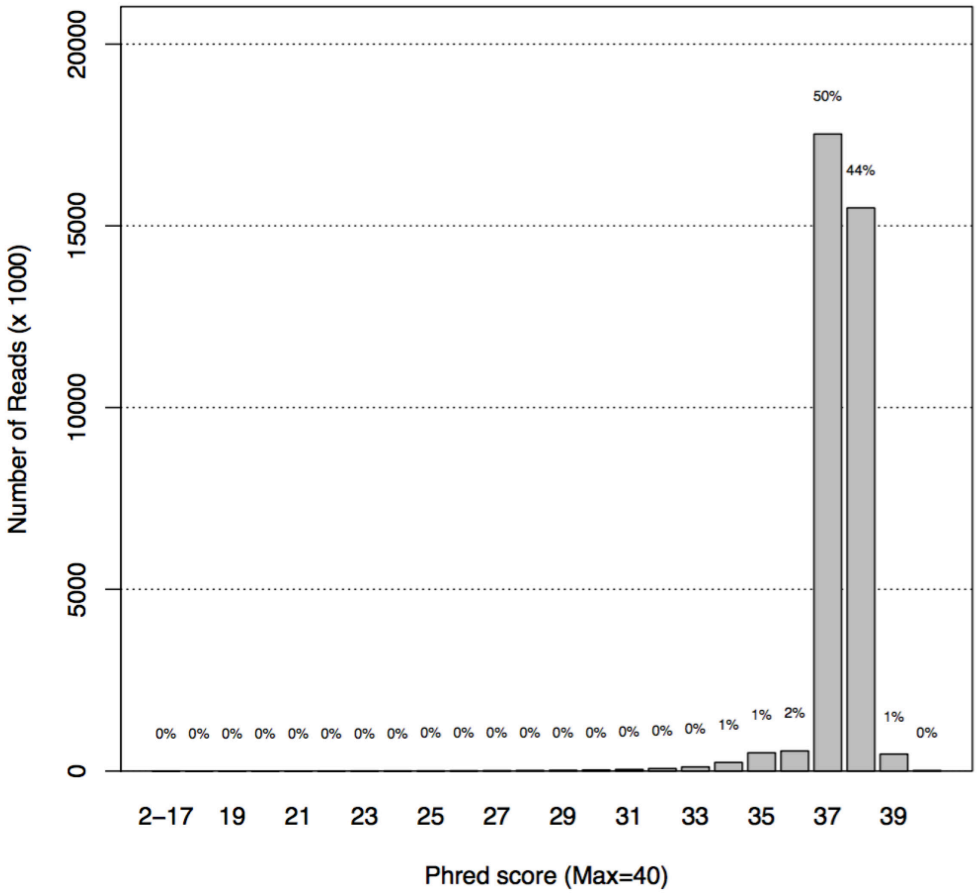


Figure 8.4 Histogram of the average phred score per base of mappable reads for all ES libraries. Phred score larger than 30 stands for probability of incorrect base calls less than 1 in 1,000 in one sequencing read

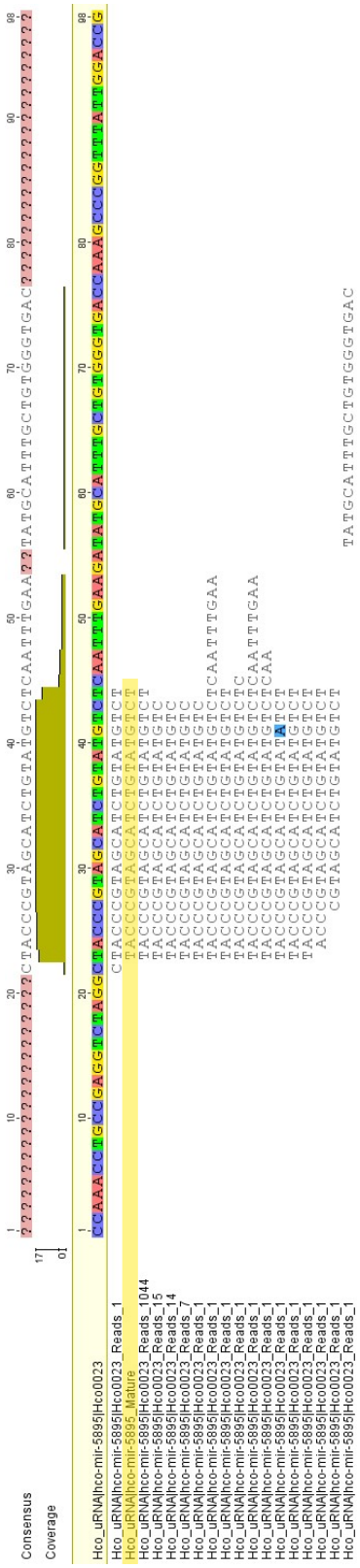


Figure 8.5 Alignment of *Hco-miR-5895* to its precursor. Entry in yellow is the predicted mature sequence.

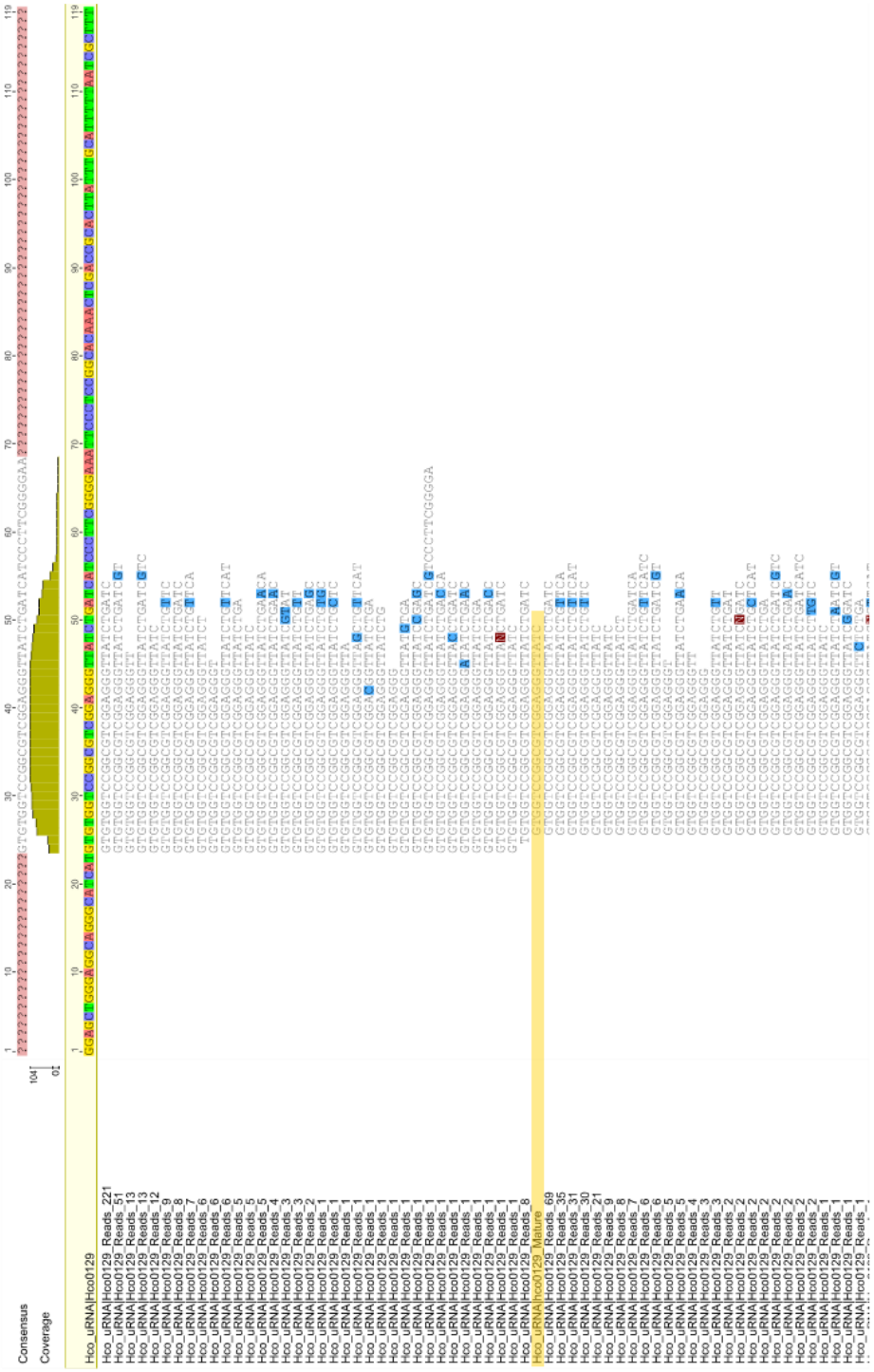


Figure 8.6 Alignment of all Hco0129 reads to its precursor. Entry highlighted in yellow is the predicted mature sequence. Sequence was not submitted to miRBase as it did not fold according to CID-miRNA.

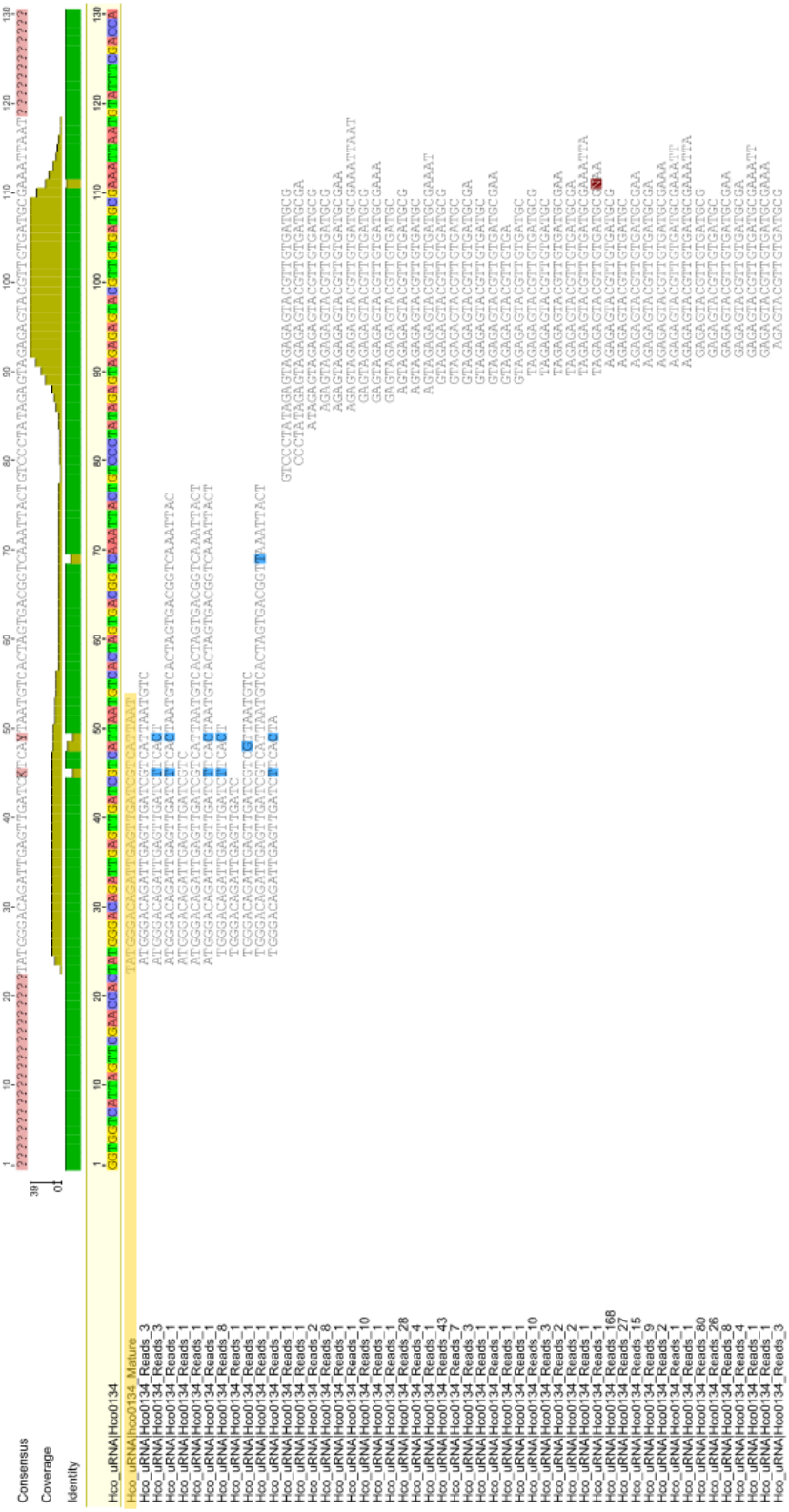


Figure 8.7 Alignment of Hco0134 to its precursor. Sequence was rejected from the list of miRNAs due to less than 10 reads in the adult library. Entry highlighted in yellow is the predicted mature miRNA sequence. Nucleotides highlighted in blue indicate mismatches with the precursor sequence.

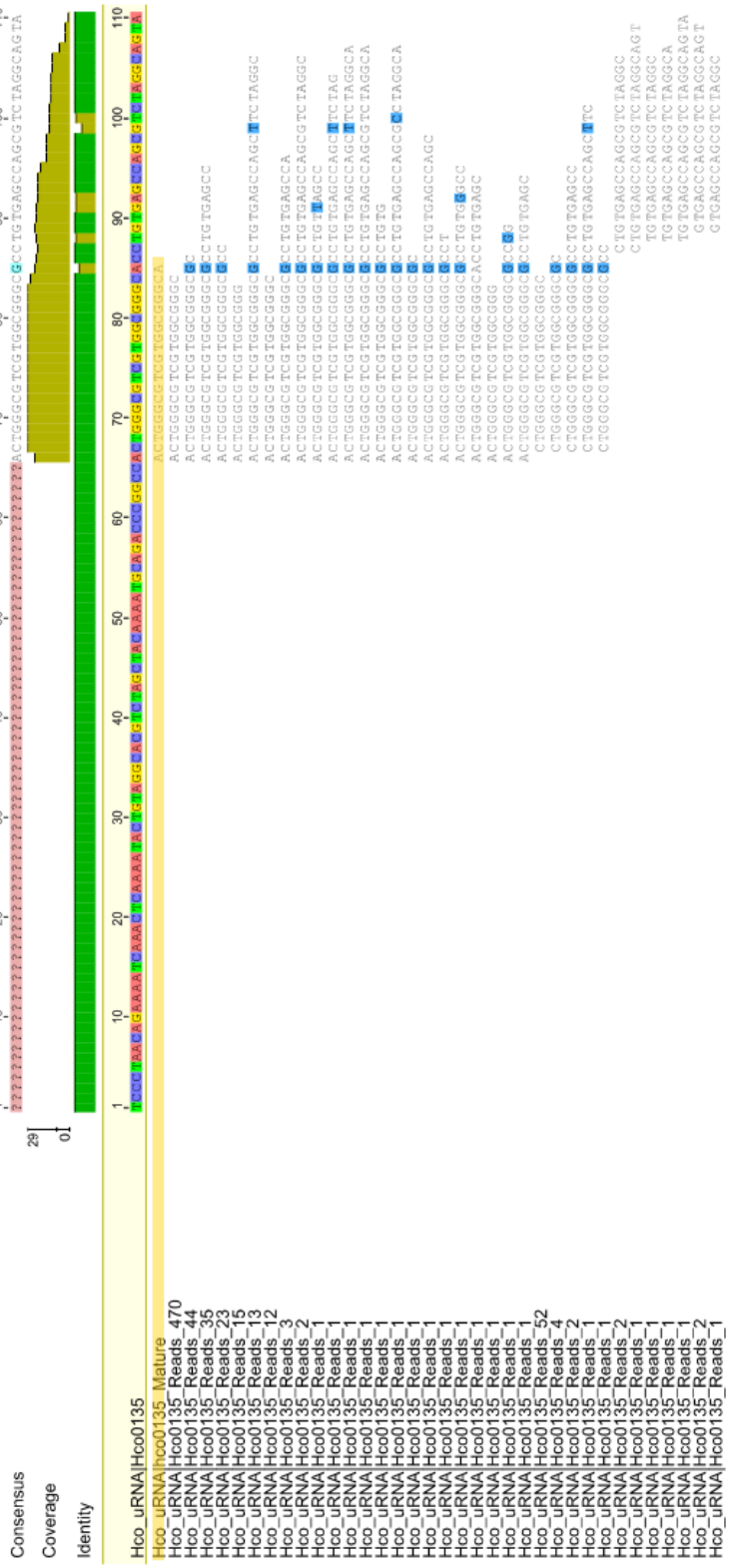


Figure 8.8 Alignment of Hco0135 to its precursor sequence. miRBase programs could not fold the sequence properly and was rejected. Entry in yellow is the predicted mature sequence. The majority of reads align closely to the mature sequence. Nucleotides highlighted in blue indicate mismatches with the precursor sequence.

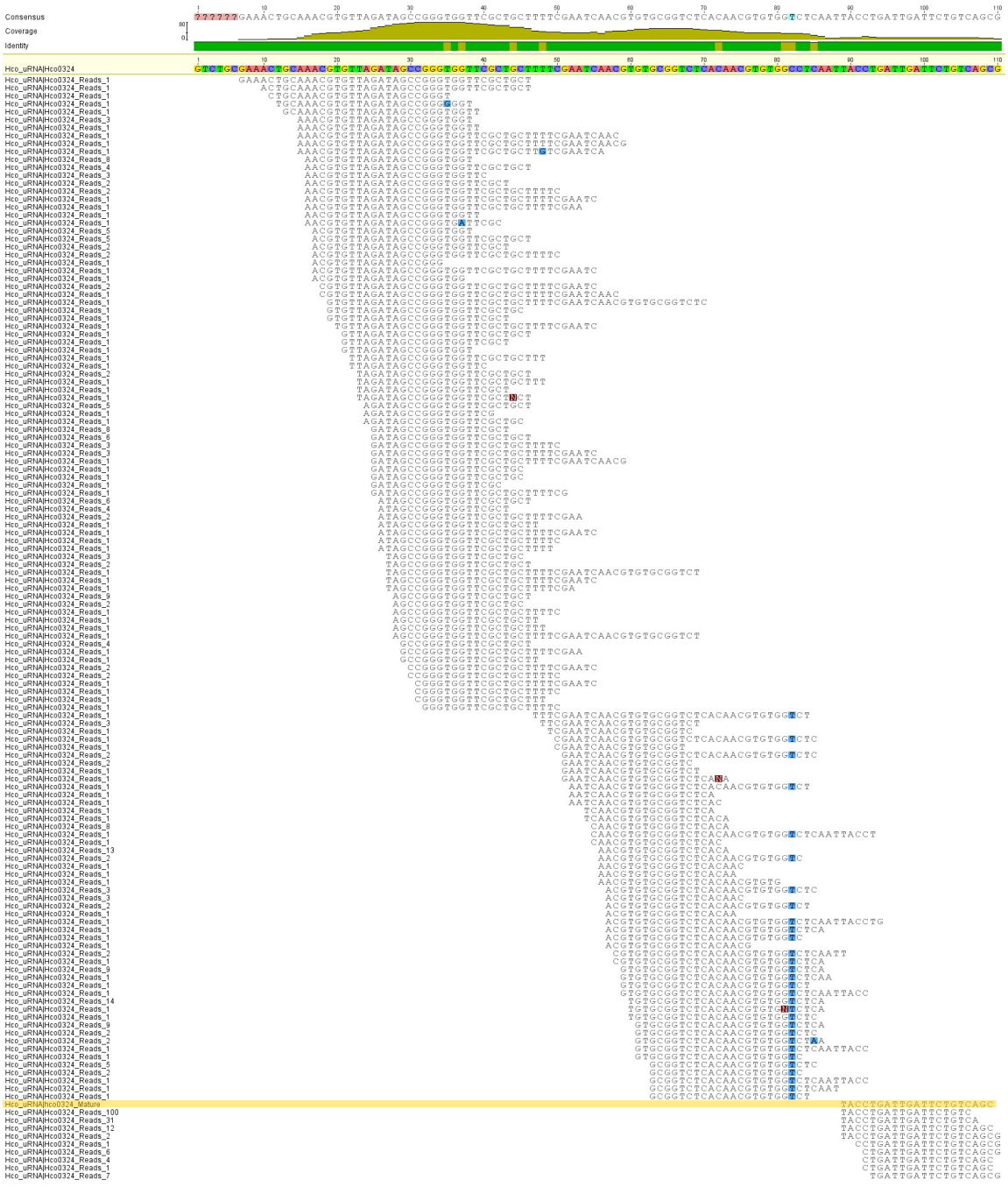


Figure 8.9 Alignment of Hco0324 to its precursor sequence. This sequence was rejected due to similarity with a rRNA. Entry in yellow is the predicted mature sequence. The reads align rather evenly across the miRNA precursor, which is commonly seen in rRNA sequence alignments. There is an enrichment of a Hco0324 mature-like sequence which is three nucleotides shorter. Nucleotides highlighted in blue indicate mismatches with the precursor sequence. Nucleotides highlighted in brown are unknown.

Figure 8.10 Alignment of Hco0325 to its precursor. Sequence was not included in the original miRNA list due to similarity with tRNA sequences. Entry in yellow is the predicted mature sequence. Nucleotides highlighted in blue indicate mismatches with the precursor sequence. Nucleotides highlighted in brown are unknown.

8.3 Genes in Innate and Adaptive Immune Responses RT² Profiler PCR Array

Innate Immunity

Pattern Recognition Receptors (PPRs): Ddx58 (RIG-

`

I), Nlrp3, Nod1 (Card4), Nod2, Tlr1, Tlr2, Tlr3, Tlr4,Tlr5, Tlr6, Tlr7, Tlr8, Tlr9.
Cytokines: Ccl12 (MCP-5, Scya12), Ccl5 (RANTES), Csf2 (GMCSF), Cxcl10 (INP10), Ifna2, Ifnb1, Il18, Il1a,Il1b, Il2, Tnf.
Other Innate Immunity Genes: Apcs, C3, C5ar1 (Gpr77), Casp1 (ICE), Cd14, Cd4, Cd40 (Tnfrsf5), Cd40lg(Tnfsf5), Cd8a, Crp, H2-Q10, H2-T23, Il1r1, Irak1, Irf3, Irf7, Itgam, Ly96 (MD-2), Lyz2, Mapk1 (Erk2), Mapk8(JNK1), Mbl2, Mpo, Mx1, Myd88, Nfkb1, Nfkbia (Iκbα, Mad3), Stat1, Ticam1 (TRIF), Traf6.

Adaptive Immunity
Th1 Markers & Immune Response: Ccr5, Cd80, Cxcr3, Ifng, Il18, Il23a, Slc11a1, Stat4, Tbx21, Tlr4, Tlr6.
Th2 Markers & Immune Response: Ccr4, Ccr8, Cd86, Gata3, Ifnb1, Il10, Il13, Il18, Il4, Il5, Il6, Nod2, Stat6.
Th17 Markers: Ccr6, Il17a, Rorc, Stat3.
Treg Markers: Ccr4, Ccr8, Foxp3, Il10.
T-Cell Activation: Cd80, Cd86, Icam1, Ifng, Il23a, Il6, Slc11a1.
Cytokines: Ccl12 (MCP-5, Scya12), Ccl5 (RANTES), Csf2 (GMCSF), Cxcl10 (INP10), Ifna2, Ifng, Il10, Il13,Il17a, Il18, Il2, Il23a, Il4, Il5, Il6, Tnf.
Other Adaptive Immunity Genes: Cd4, Cd40 (Tnfrsf5), Cd40lg (Tnfsf5), Cd8a, Crp, Fasl (Tnfsf6), H2-Q10,Ifnar1, Ifngr1, Il1b, Il1r1, Irf3, Irf7, Itgam, Jak2, Mapk8 (JNK1), Mbl2, Mx1, Nfkb1, Rag1, Stat1.

Humoral Immunity
C3, C5ar1 (Gpr77), Ccl12 (MCP-5, Scya12), Ccr6, Crp, Ifnb1, Ifng, Il6, Mbl2, Nod2, Tnf.

Inflammatory Response
Apcs, C3, Ccl5 (RANTES), Crp, Foxp3, Il1a, Il1b, Il4, Il6, Mbl2, Stat3, Tnf.

Defense Response to Bacteria
C5ar1 (Gpr77), Ifnb1, Ifng, Il23a, Il6, Lyz2, Mbl2, Myd88, Nod1 (Card4), Nod2, Slc11a1, Tlr1, Tlr3, Tlr4, Tlr6,Tlr9, Tnf.

Defense Response to Viruses
Cd4, Cd40 (Tnfrsf5), Cd86, Cd8a, Cxcl10 (INP10), Ddx58 (RIG-I), H2-Q10, Ifnar1, Ifnb1, Il23a, Il6, Irf3,Nlrp3, Ticam1 (TRIF), Tlr3, Tlr7, Tlr8, Tyk2.

8.4 Genes in Cell Surface Markers RT2 Profiler PCR Array

B-Cell Surface Markers
Activated B-Cells: Cd28, Cd38, Cd69, Cd70 (Tnfsf7), Cd80, Cd83, Cd86, Dpp4, Fcer2a, Il2ra (CD25), Tnfrsf8.
Mature B-

`

Cells: Cd19, Cd22, Cd24a, Cd37, Cd40 (Tnfrsf5), Cd72, Cd74, Cd79a, Cd79b, Cr2, Il1r2, Itga2,Itga3, Ms4a1, St6gal1.
Other B-Cell Surface Markers: Cd200r1, Chst10, H2-Q10, Nt5e.

T-Cell Surface Markers
Cytotoxic T-Cells: Cd8a, Cd8b1, Ctla4.
Helper T-Cells: Cd4.
Activated T-Cells: Alcam, Cd2, Cd276, Cd38, Cd40lg (Tnfsf5), Cd69, Cd70 (Tnfsf7), Cd83, Cd96, Ctla4,Dpp4, Il12rb1, Il2ra (CD25), Itga1, Tnfrsf4 (Ox40), Tnfrsf8.
Other T-Cell Surface Markers: Cd160, Cd247 (Cd3z), Cd28, Cd37, Cd3d, Cd3g, Cd5, Cd6, Cd7, Fas(Tnfrsf6), Klrd1, Nt5e, St6gal1.

Natural Killer (NK) Cell Surface Markers
Cd2, Cd244, Cd247 (Cd3z), Cd7, Cd96, Chst10, Il12rb1, Klrc1, Klrd1, Ncam1.

Monocyte & Macrophage Cell Surface Markers
Activated Macrophages: Cd69, Eng (Evi-1), Fcer2a, Il2ra (CD25).
Other Monocyte & Macrophage Cell Surface Markers: C5ar1 (Gpr77), Cd163, Cd34, Cd40 (Tnfrsf5), Cd63,Cd70 (Tnfsf7), Cd74, Cd86, Chst10, Csf1r, Dpp4, Fcgr1, Icam2, Il1r2, Itga1, Itga2, S100a8, Tnfrsf8.

Endothelial Cell Surface Markers
Eng (Evi-1), Icam2, Nos3 (eNOS), Pecam1, Selp, Tek (Tie-2), Vcam1, Vwf.

Smooth Muscle Cell Surface Markers
Myh10, Myh9, Myocd.

Dendritic Cell Surface Markers
Cd209a, Cd40 (Tnfrsf5), Cd83, Cd86, Cr2, Fcer2a.

Mast Cell Surface Markers
C5ar1 (Gpr77), Fcer1a, Fcer2a, Tpsab1.

Fibroblast (Stromal Cell) Surface Markers
Alcam, Col1a1, Col1a2.

Epithelial Cell Surface Markers
Cd1d1, Cd1d2, Epcam, Krt18, Krt5, Krt8.

Adipocyte Surface Markers
Retn.

`

Table 8.2 GO terms of the 44 genes that have decreased expression in the Hco-miR-5352 transfected cell compared to the control and have predicted Hco-miR-5352 binding sites according to Targetscan. PHF15 and H3F3B were not found in the pantherdb database so were not included in the analysis.

GO biological process complete	Number of genes
biological_process (GO:0008150)	42
cellular process (GO:0009987)	34
single-organism process (GO:0044699)	31
single-organism cellular process (GO:0044763)	29
biological regulation (GO:0065007)	28
regulation of cellular process (GO:0050794)	28
regulation of biological process (GO:0050789)	28
organic substance metabolic process (GO:0071704)	25
primary metabolic process (GO:0044238)	25
cellular metabolic process (GO:0044237)	25
metabolic process (GO:0008152)	25
cellular macromolecule metabolic process (GO:0044260)	24
macromolecule metabolic process (GO:0043170)	24
regulation of metabolic process (GO:0019222)	23
regulation of primary metabolic process (GO:0080090)	21
regulation of macromolecule metabolic process (GO:0060255)	21
regulation of cellular metabolic process (GO:0031323)	21
response to stimulus (GO:0050896)	20
GO molecular function complete	Number of genes
molecular_function (GO:0003674)	42
binding (GO:0005488)	33
protein binding (GO:0005515)	27
organic cyclic compound binding (GO:0097159)	14
heterocyclic compound binding (GO:1901363)	14
ion binding (GO:0043167)	13
catalytic activity (GO:0003824)	13
enzyme binding (GO:0019899)	10
GO cellular component complete	Number of genes
cellular_component (GO:0005575)	42
cell (GO:0005623)	37
cell part (GO:0044464)	37
intracellular (GO:0005622)	36

`

intracellular part (GO:0044424)	35
organelle (GO:0043226)	33
intracellular organelle (GO:0043229)	32
membrane-bounded organelle (GO:0043227)	32
intracellular membrane-bounded organelle (GO:0043231)	30
cytoplasm (GO:0005737)	29
nucleus (GO:0005634)	22
intracellular organelle part (GO:0044446)	21
organelle part (GO:0044422)	21
membrane (GO:0016020)	20

Table 8.3 100 genes with the largest fold change decrease in Hco-miR-5352 transfected cells compared to the Cel-miR-67 transfected cells. Genes filtered for Exp False (q-value less than 0.05)

Log2(Fold Change)	Exp False Discovery Rate (q-value)	Symbol	Entrez Gene Name	Location	Type(s)
-4.658	0.00254	P2RY12	purinergic receptor P2Y, G-protein coupled, 12	Plasma Membrane	G-protein coupled receptor
-4.219	0.00254	Fcrls	Fc receptor-like S, scavenger receptor	Plasma Membrane	other
-3.872	0.00254	Siglech	sialic acid binding Ig-like lectin H	Plasma Membrane	other
-3.859	0.00254	CX3CR1	chemokine (C-X3-C motif) receptor 1	Plasma Membrane	G-protein coupled receptor
-3.843	0.00254	SLCO2B1	solute carrier organic anion transporter family member 2B1	Plasma Membrane	transporter
-3.796	0.00254	PLXDC2	plexin domain containing 2	Extracellular Space	other
-3.711	0.00254	SELPLG	selectin P ligand	Plasma Membrane	other
-3.707	0.0429	CD37	CD37 molecule	Plasma Membrane	other
-3.424	0.00254	SALL1	spalt-like transcription factor 1	Nucleus	transcription regulator
-3.362	0.00254	TNFRSF21	tumor necrosis factor receptor superfamily member 21	Plasma Membrane	transmembrane receptor
-3.345	0.00254	CSF1R	colony stimulating factor 1 receptor	Plasma Membrane	kinase
-3.335	0.00254	SLC7A8	solute carrier family 7 (amino acid transporter light chain, L system), member 8	Plasma Membrane	transporter
-2.998	0.00254	LAPTM5	lysosomal protein transmembrane 5	Plasma Membrane	other
-2.978	0.00254	C1QC	complement component 1, q subcomponent, C chain	Extracellular Space	other
-2.931	0.00254	Trem2	triggering receptor expressed on myeloid cells 2	Other	other
-2.919	0.00254	C1QB	complement component 1, q subcomponent, B chain	Extracellular Space	other
-2.917	0.00254	MAFB	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog B	Nucleus	transcription regulator
-2.909	0.00254	PLD4	phospholipase D family member 4	Extracellular Space	enzyme

`

-2.873	0.00254	IL10RA	interleukin 10 receptor subunit alpha	Plasma Membrane	transmembrane receptor
-2.81	0.00254	CD180	CD180 molecule	Plasma Membrane	other
-2.796	0.00254	KCTD12	potassium channel tetramerization domain containing 12	Plasma Membrane	ion channel
-2.79	0.00254	Cd33	CD33 antigen	Plasma Membrane	other
-2.75	0.00254	C1QA	complement component 1, q subcomponent, A chain	Extracellular Space	other
-2.731	0.00254	CD53	CD53 molecule	Plasma Membrane	other
-2.657	0.00254	LAG3	lymphocyte-activation gene 3	Plasma Membrane	transmembrane receptor
-2.634	0.00254	CCR5	chemokine (C-C motif) receptor 5 (gene/pseudogene)	Plasma Membrane	G-protein coupled receptor
-2.559	0.00254	LY86	lymphocyte antigen 86	Plasma Membrane	other
-2.534	0.00254	INPP5D	inositol polyphosphate-5-phosphatase D	Cytoplasm	phosphatase
-2.509	0.00254	HPGDS	hematopoietic prostaglandin D synthase	Cytoplasm	enzyme
-2.508	0.00254	LAIR1	leucocyte-associated immunoglobulin-like receptor 1	Plasma Membrane	transmembrane receptor
-2.48	0.0242	PIK3AP1	phosphoinositide-3-kinase adaptor protein 1	Cytoplasm	other
-2.457	0.00254	Ppfia4	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 4	Plasma Membrane	other
-2.448	0.00254	ITGAM	integrin subunit alpha M	Plasma Membrane	transmembrane receptor
-2.422	0.00254	ENTPD1	ectonucleoside triphosphate diphosphohydrolase 1	Plasma Membrane	enzyme
-2.366	0.00254	CYTH4	cytohesin 4	Cytoplasm	other
-2.358	0.00254	NCKAP1L	NCK associated protein 1 like	Plasma Membrane	other
-2.348	0.0132	CD86	CD86 molecule	Plasma Membrane	transmembrane receptor
-2.333	0.00254	IRF8	interferon regulatory factor 8	Nucleus	transcription regulator
-2.295	0.00819	CCL3L3	chemokine (C-C motif) ligand 3-like 3	Extracellular Space	cytokine
-2.284	0.00254	FCGR2A	Fc fragment of IgG receptor IIa	Plasma Membrane	transmembrane receptor
-2.198	0.00254	LYN	LYN proto-oncogene, Src family tyrosine kinase	Cytoplasm	kinase
-2.137	0.0175	NCF2	neutrophil cytosolic factor 2	Cytoplasm	enzyme
-2.103	0.0216	PARVG	parvin gamma	Cytoplasm	other
-2.084	0.00984	P2RY6	pyrimidinergic receptor P2Y, G-protein coupled, 6	Plasma Membrane	G-protein coupled receptor
-2.078	0.00254	TF	transferrin	Extracellular Space	transporter
-2.07	0.00254	RASAL3	RAS protein activator like 3	Cytoplasm	other
-2.041	0.00254	FCGR1A	Fc fragment of IgG receptor Ia	Plasma Membrane	transmembrane receptor
-2.036	0.00254	VAV1	vav guanine nucleotide exchange factor 1	Nucleus	transcription regulator
-2.035	0.0322	AIF1	allograft inflammatory factor 1	Nucleus	other
-1.982	0.00254	UNC93B1	unc-93 homolog B1 (C. elegans)	Cytoplasm	transporter
-1.9	0.0175	ADAP2	ArfGAP with dual PH domains 2	Cytoplasm	other
-1.866	0.00254	CFH	complement factor H	Extracellular Space	other
-1.808	0.00254	CTSS	cathepsin S	Cytoplasm	peptidase
-1.769	0.00254	TYROBP	TYRO protein tyrosine kinase binding protein	Plasma Membrane	transmembrane receptor
-1.767	0.00984	FCER1G	Fc fragment of IgE receptor Ig	Plasma Membrane	transmembrane receptor
-1.75	0.00464	FAM105A	family with sequence	Other	other

`

			similarity 105 member A		
-1.742	0.00464	HAVCR2	hepatitis A virus cellular receptor 2	Plasma Membrane	other
-1.729	0.03	SPINT1	serine peptidase inhibitor, Kunitz type 1	Extracellular Space	other
-1.714	0.00254	RTN4RL1	reticulon 4 receptor-like 1	Plasma Membrane	other
-1.707	0.00819	CTSH	cathepsin H	Cytoplasm	peptidase
-1.699	0.00254	ELMO1	engulfment and cell motility 1	Cytoplasm	other
-1.673	0.0288	CCDC88B	coiled-coil domain containing 88B	Nucleus	enzyme
-1.657	0.0311	LTF	lactotransferrin	Extracellular Space	peptidase
-1.637	0.00254	CSF3R	colony stimulating factor 3 receptor	Plasma Membrane	transmembr ane receptor
-1.633	0.0189	FCGR2B	Fc fragment of IgG receptor IIb	Plasma Membrane	transmembr ane receptor
-1.622	0.00254	LPCAT2	lysophosphatidylcholine acyltransferase 2	Cytoplasm	enzyme
-1.621	0.00819	CTSC	cathepsin C	Cytoplasm	peptidase
-1.572	0.00254	LCP1	lymphocyte cytosolic protein 1 (L-plastin)	Cytoplasm	other
-1.554	0.00254	MPEG1	macrophage expressed 1	Cytoplasm	other
-1.539	0.0378	ASB10	ankyrin repeat and SOCS box containing 10	Other	other
-1.515	0.00254	OCA2	OCA2 melanosomal transmembrane protein	Plasma Membrane	transporter
-1.509	0.00254	HMHA1	histocompatibility (minor) HA-1	Cytoplasm	transporter
-1.506	0.045	Ngp	neutrophilic granule protein	Extracellular Space	other
-1.462	0.0262	NCF1	neutrophil cytosolic factor 1	Cytoplasm	enzyme
-1.449	0.00254	CORO1A	coronin 1A	Cytoplasm	other
-1.403	0.0242	CEBPA	CCAAT/enhancer binding protein alpha	Nucleus	transcription regulator
-1.338	0.00254	FGD2	FYVE, RhoGEF and PH domain containing 2	Cytoplasm	other
-1.327	0.0132	ARHGAP4	Rho GTPase activating protein 4	Cytoplasm	other
-1.325	0.0148	Krtap14	keratin associated protein 14	Other	other
-1.324	0.0274	FES	FES proto-oncogene, tyrosine kinase	Cytoplasm	kinase
-1.288	0.00254	RPS6KA1	ribosomal protein S6 kinase, 90kDa, polypeptide 1	Cytoplasm	kinase
-1.288	0.00254	SEPP1	selenoprotein P, plasma, 1	Extracellular Space	other
-1.247	0.00254	HEXB	hexosaminidase B (beta polypeptide)	Cytoplasm	enzyme
-1.228	0.00254	TMEM119	transmembrane protein 119	Cytoplasm	other
-1.224	0.0366	RAB3IL1	RAB3A interacting protein (rabin3)-like 1	Other	other
-1.181	0.00254	SIRPA	signal-regulatory protein alpha	Plasma Membrane	phosphatase
-1.171	0.00254	TGFBR1	transforming growth factor beta receptor I	Plasma Membrane	kinase
-1.123	0.0346	B630019K0 6Rik	novel protein similar to F-box and leucine-rich repeat protein 17 (Fbxl17)	Other	other
-1.038	0.00254	ENOPH1	enolase-phosphatase 1	Cytoplasm	enzyme
-0.998	0.00254	TMEM50B	transmembrane protein 50B	Plasma Membrane	other
-0.983	0.0175	BTG2	BTG family member 2	Nucleus	transcription regulator
-0.955	0.00254	VPS26A	VPS26 retromer complex component A	Cytoplasm	transporter
-0.935	0.00254	PAK1IP1	PAK1 interacting protein 1	Nucleus	other
-0.897	0.00254	LPGAT1	lysophosphatidylglycerol acyltransferase 1	Cytoplasm	other
-0.893	0.00254	SCOC	short coiled-coil protein	Cytoplasm	other
-0.888	0.00254	LEPROT	leptin receptor overlapping transcript	Plasma Membrane	other
-0.875	0.0408	TCN2	transcobalamin 2	Extracellular	transporter

`

				Space	
-0.843	0.00254	PHTF2	putative homeodomain transcription factor 2	Cytoplasm	other
-0.843	0.0189	PAG1	phosphoprotein membrane anchor with glycosphingolipid microdomains 1	Plasma Membrane	other

Table 8.4 100 genes with the largest fold change increase in Hco-miR-5352 transfected cells compared to the Cel-miR-67 transfected cells. Genes filtered for Exp False (q-value less than 0.05)

Log2(Fold Change)	False Discovery Rate (q-value)	Symbol	Entrez Gene Name	Location	Type(s)
5.56	0.00254	SLC6A1	solute carrier family 6 (neurotransmitter transporter), member 1	Plasma Membrane	transporter
5.553	0.00254	SLC6A11	solu3te carrier family 6 (neurotransmitter transporter), member 11	Plasma Membrane	transporter
5.382	0.00254	GPR37L1	G protein-coupled receptor 37 like 1	Plasma Membrane	G-protein coupled receptor
5.313	0.0429	AGT	angiotensinogen	Extracellular Space	growth factor
5.021	0.00254	ATP1B2	ATPase, Na+/K+ transporting, beta 2 polypeptide	Plasma Membrane	transporter
4.947	0.00254	SLC1A2	solute carrier family 1 (glial high affinity glutamate transporter), member 2	Plasma Membrane	transporter
4.843	0.00254	SPARCL1	SPARC like 1	Extracellular Space	other
4.821	0.00254	BCAN	brevican	Extracellular Space	other
4.68	0.00254	ATP1A2	ATPase, Na+/K+ transporting, alpha 2 polypeptide	Plasma Membrane	transporter
4.444	0.00254	FAM107A	family with sequence similarity 107 member A	Nucleus	other
4.253	0.00254	CLU	clusterin	Cytoplasm	other
4.025	0.00254	NTRK2	neurotrophic tyrosine kinase, receptor, type 2	Plasma Membrane	kinase
3.967	0.0132	SLC39A12	solute carrier family 39 (zinc transporter), member 12	Plasma Membrane	transporter
3.963	0.0408	SLC7A10	solute carrier family 7 (neutral amino acid transporter light chain, asc system), member 10	Plasma Membrane	transporter
3.508	0.00254	TTYH1	tweety family member 1	Plasma Membrane	ion channel
3.123	0.045	APLP1	amyloid beta (A4) precursor-like protein 1	Extracellular Space	other
3.035	0.00254	CSPG5	chondroitin sulfate proteoglycan 5	Extracellular Space	growth factor
2.985	0.00464	CNTFR	ciliary neurotrophic factor receptor	Plasma Membrane	transmembrane receptor
2.769	0.0116	SLC25A18	solute carrier family 25 (glutamate carrier), member 18	Cytoplasm	transporter
2.265	0.00653	ALDH1L1	aldehyde dehydrogenase 1 family member L1	Cytoplasm	enzyme
1.979	0.00254	Cyp2d22	cytochrome P450, family 2, subfamily d, polypeptide 22	Cytoplasm	enzyme
1.961	0.00254	SLC1A3	solute carrier family 1 (glial high affinity glutamate transporter), member 3	Plasma Membrane	transporter
1.863	0.00254	NDRG2	NDRG family member 2	Cytoplasm	other
1.637	0.0288	FJX1	four jointed box 1	Extracellular Space	other

`

1.602	0.00254	TTR	transthyretin	Extracellular Space	transporter
1.528	0.00464	mir-142	microRNA 142	Cytoplasm	microRNA
1.524	0.0189	mir-146	microRNA 146a	Cytoplasm	microRNA
1.407	0.0311	mir-27	microRNA 27a	Cytoplasm	microRNA
1.389	0.00254	TSPAN7	tetraspanin 7	Plasma Membrane	other
1.351	0.00819	GNAO1	guanine nucleotide binding protein (G protein), alpha activating activity polypeptide O	Plasma Membrane	enzyme
1.347	0.00254	FGFR3	fibroblast growth factor receptor 3	Plasma Membrane	kinase
1.336	0.00653	GPNMB	glycoprotein (transmembrane) nmb	Plasma Membrane	enzyme
1.28	0.00254	ENPP2	ectonucleotide pyrophosphatase/phosphodiesterase 2	Plasma Membrane	enzyme
1.253	0.00254	mir-143	microRNA 143	Cytoplasm	microRNA
1.177	0.00254	mir-10	microRNA 100	Other	microRNA
1.086	0.00254	KRT80	keratin 80, type II	Cytoplasm	other
1.086	0.00254	S1PR1	sphingosine-1-phosphate receptor 1	Plasma Membrane	G-protein coupled receptor
1.081	0.0162	ANK2	ankyrin 2, neuronal	Plasma Membrane	other
0.987	0.00254	ALDOC	aldolase, fructose-bisphosphate C	Cytoplasm	enzyme
0.985	0.00254	ALS2CL	ALS2 C-terminal like	Cytoplasm	other
0.945	0.00819	Oasl2	2'-5' oligoadenylate synthetase-like 2	Other	enzyme
0.823	0.00254	CKB	creatine kinase, brain	Cytoplasm	kinase
0.783	0.00653	CXCL10	chemokine (C-X-C motif) ligand 10	Extracellular Space	cytokine
0.755	0.0485	BBC3	BCL2 binding component 3	Cytoplasm	other
0.725	0.00254	mir-192	microRNA 192	Cytoplasm	microRNA
0.702	0.00653	MEGF10	multiple EGF-like-domains 10	Plasma Membrane	other
0.69	0.0242	SASH1	SAM and SH3 domain containing 1	Extracellular Space	other
0.675	0.00254	CHAC1	ChaC glutathione-specific gamma-glutamylcyclotransferase 1	Cytoplasm	other
0.667	0.00254	TRIB3	tribbles pseudokinase 3	Nucleus	kinase
0.636	0.0229	LAMC2	laminin subunit gamma 2	Extracellular Space	other
0.626	0.00254	SLC7A11	solute carrier family 7 (anionic amino acid transporter light chain, xc- system), member 11	Plasma Membrane	transporter
0.604	0.00254	UAP1L1	UDP-N-acetylglucosamine pyrophosphorylase 1 like 1	Other	other
0.596	0.00254	NUPR1	nuclear protein 1, transcriptional regulator	Nucleus	transcription regulator
0.595	0.00254	ATF5	activating transcription factor 5	Nucleus	transcription regulator
0.594	0.00254	ZBTB18	zinc finger and BTB domain containing 18	Nucleus	transcription regulator
0.59	0.00254	DVL3	dishevelled segment polarity protein 3	Cytoplasm	other
0.583	0.00254	CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)	Nucleus	kinase
0.58	0.0189	ATP6V0E2	ATPase, H+ transporting V0 subunit e2	Cytoplasm	enzyme
0.559	0.00254	PGAM5	PGAM family member 5, serine/threonine protein phosphatase, mitochondrial	Cytoplasm	enzyme
0.547	0.00653	Tsc22d3	TSC22 domain family, member 3	Nucleus	other

`

0.545	0.00254	DDIT4	DNA damage inducible transcript 4	Cytoplasm	other
0.545	0.0162	Zfp568	zinc finger protein 568	Other	other
0.544	0.00254	PDGFB	platelet derived growth factor subunit B	Extracellular Space	growth factor
0.543	0.00254	LOXL4	lysyl oxidase like 4	Extracellular Space	enzyme
0.524	0.00254	Mt1	metallothionein 1	Cytoplasm	other
0.519	0.046	NFKBIE	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon	Nucleus	transcription regulator
0.509	0.0148	CTNS	cystinosin, lysosomal cystine transporter	Cytoplasm	transporter
0.509	0.0262	F3	coagulation factor III, tissue factor	Plasma Membrane	transmembrane receptor
0.507	0.00254	Plagl2	PLAG1 like zinc finger 2	Nucleus	transcription regulator
0.5	0.00254	Mt2	metallothionein 2	Other	other
0.492	0.00254	APOE	apolipoprotein E	Extracellular Space	transporter
0.488	0.0262	OLFM1	olfactomedin 1	Cytoplasm	other
0.484	0.00254	PCDH1	protocadherin 1	Plasma Membrane	other
0.476	0.0175	STK40	serine/threonine kinase 40	Cytoplasm	kinase
0.472	0.0162	GADD45B	growth arrest and DNA damage inducible beta	Cytoplasm	other
0.461	0.00254	SLC27A1	solute carrier family 27 (fatty acid transporter), member 1	Plasma Membrane	transporter
0.457	0.00254	TINAGL1	tubulointerstitial nephritis antigen-like 1	Extracellular Space	transporter
0.446	0.00464	USP31	ubiquitin specific peptidase 31	Nucleus	peptidase
0.442	0.00984	IGF2BP1	insulin like growth factor 2 mRNA binding protein 1	Cytoplasm	translation regulator
0.441	0.00464	RDH11	retinol dehydrogenase 11 (all-trans/9-cis/11-cis)	Cytoplasm	enzyme
0.44	0.00254	FLCN	folliculin	Cytoplasm	other
0.439	0.00254	DPP7	dipeptidyl-peptidase 7	Cytoplasm	peptidase
0.438	0.0408	ASPA	aspartoacylase	Cytoplasm	enzyme
0.432	0.00254	ZNF367	zinc finger protein 367	Nucleus	transcription regulator
0.425	0.00819	SLC25A44	solute carrier family 25 member 44	Cytoplasm	transporter
0.423	0.045	CRY2	cryptochrome circadian clock 2	Nucleus	enzyme
0.422	0.00464	SNX8	sorting nexin 8	Cytoplasm	transporter
0.42	0.0216	SCRN1	secernin 1	Cytoplasm	other
0.404	0.00254	CCNB1	cyclin B1	Cytoplasm	kinase
0.403	0.00254	ATF4	activating transcription factor 4	Nucleus	transcription regulator
0.403	0.0253	CA13	carbonic anhydrase XIII	Cytoplasm	enzyme
0.402	0.0132	CDK8	cyclin-dependent kinase 8	Nucleus	kinase
0.399	0.0262	MLKL	mixed lineage kinase domain-like	Cytoplasm	kinase
0.397	0.0378	TBCEL	tubulin folding cofactor E-like	Other	other
0.394	0.00819	RALB	v-ral simian leukemia viral oncogene homolog B	Cytoplasm	enzyme
0.392	0.045	TRIM16	tripartite motif containing 16	Cytoplasm	transcription regulator
0.388	0.0229	ANGPTL4	angiopoietin like 4	Extracellular Space	other
0.387	0.00254	PLIN2	perilipin 2	Plasma Membrane	other
0.383	0.0346	Pvr	poliovirus receptor	Plasma Membrane	other

`

Table 8.5 100 most significant genes according to their false discovery rate. Genes in green indicate increased expression whilst genes coloured red indicated decreased expression in the presence of Hco-miR-5352.

Log2(Fold Change)	False Discovery Rate (q-value)	Symbol	Entrez Gene Name	Location	Type(s)
5.56	0.003	SLC6A1	solute carrier family 6 member 1	Plasma Membrane	transporter
5.553	0.003	SLC6A11	solute carrier family 6 member 11	Plasma Membrane	transporter
5.382	0.003	GPR37L1	G protein-coupled receptor 37 like 1	Plasma Membrane	G-protein coupled receptor
5.021	0.003	ATP1B2	ATPase Na+/K+ transporting subunit beta 2	Plasma Membrane	transporter
4.947	0.003	SLC1A2	solute carrier family 1 member 2	Plasma Membrane	transporter
4.843	0.003	SPARCL1	SPARC like 1	Extracellular Space	other
4.821	0.003	BCAN	brevican	Extracellular Space	other
4.68	0.003	ATP1A2	ATPase Na+/K+ transporting subunit alpha 2	Plasma Membrane	transporter
-4.658	0.003	P2RY12	purinergic receptor P2Y12	Plasma Membrane	G-protein coupled receptor
4.444	0.003	FAM107A	family with sequence similarity 107 member A	Nucleus	other
4.253	0.003	CLU	clusterin	Cytoplasm	other
-4.219	0.003	Fcrls	Fc receptor-like S, scavenger receptor	Plasma Membrane	other
4.025	0.003	NTRK2	neurotrophic receptor tyrosine kinase 2	Plasma Membrane	kinase
-3.872	0.003	Siglech	sialic acid binding Ig-like lectin H	Plasma Membrane	other
-3.859	0.003	CX3CR1	C-X3-C motif chemokine receptor 1	Plasma Membrane	G-protein coupled receptor
-3.843	0.003	SLCO2B1	solute carrier organic anion transporter family member 2B1	Plasma Membrane	transporter
-3.796	0.003	PLXDC2	plexin domain containing 2	Extracellular Space	other
-3.711	0.003	SELPLG	selectin P ligand	Plasma Membrane	other
3.508	0.003	TTYH1	tweety family member 1	Plasma Membrane	ion channel
-3.424	0.003	SALL1	spalt-like transcription factor 1	Nucleus	transcription regulator
-3.362	0.003	TNFRSF21	tumor necrosis factor receptor superfamily member 21	Plasma Membrane	transmembrane receptor
-3.345	0.003	CSF1R	colony stimulating factor 1 receptor	Plasma Membrane	kinase
-3.335	0.003	SLC7A8	solute carrier family 7 member 8	Plasma Membrane	transporter
3.035	0.003	CSPG5	chondroitin sulfate proteoglycan 5	Extracellular Space	growth factor
-2.998	0.003	LAPTM5	lysosomal protein transmembrane 5	Plasma Membrane	other
-2.978	0.003	C1QC	complement component 1, q subcomponent, C chain	Extracellular Space	other
-2.931	0.003	Trem2	triggering receptor expressed on myeloid cells 2	Other	other
-2.919	0.003	C1QB	complement component 1, q subcomponent, B chain	Extracellular Space	other
-2.917	0.003	MAFB	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog B	Nucleus	transcription regulator
-2.909	0.003	PLD4	phospholipase D family member 4	Extracellular Space	enzyme
-2.873	0.003	IL10RA	interleukin 10 receptor subunit alpha	Plasma Membrane	transmembrane receptor
-2.81	0.003	CD180	CD180 molecule	Plasma	other

				Membrane	
-2.796	0.003	KCTD12	potassium channel tetramerization domain containing 12	Plasma Membrane	ion channel
-2.79	0.003	Cd33	CD33 antigen	Plasma Membrane	other
-2.75	0.003	C1QA	complement component 1, q subcomponent, A chain	Extracellular Space	peptidase
-2.731	0.003	CD53	CD53 molecule	Plasma Membrane	other
-2.657	0.003	LAG3	lymphocyte activating 3	Plasma Membrane	transmembrane receptor
-2.634	0.003	CCR5	C-C motif chemokine receptor 5 (gene/pseudogene)	Plasma Membrane	G-protein coupled receptor
-2.559	0.003	LY86	lymphocyte antigen 86	Plasma Membrane	other
-2.534	0.003	INPP5D	inositol polyphosphate-5-phosphatase D	Cytoplasm	phosphatase
-2.509	0.003	HPGDS	hematopoietic prostaglandin D synthase	Cytoplasm	enzyme
-2.508	0.003	LAIR1	leucocyte associated immunoglobulin like receptor 1	Plasma Membrane	transmembrane receptor
-2.457	0.003	Ppfia4	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 4	Plasma Membrane	other
-2.448	0.003	ITGAM	integrin subunit alpha M	Plasma Membrane	transmembrane receptor
-2.422	0.003	ENTPD1	ectonucleoside triphosphate diphosphohydrolase 1	Plasma Membrane	enzyme
-2.366	0.003	CYTH4	cytohesin 4	Cytoplasm	other
-2.358	0.003	NCKAP1L	NCK associated protein 1 like	Plasma Membrane	other
-2.333	0.003	IRF8	interferon regulatory factor 8	Nucleus	transcription regulator
-2.284	0.003	FCGR2A	Fc fragment of IgG receptor IIa	Plasma Membrane	transmembrane receptor
-2.198	0.003	LYN	LYN proto-oncogene, Src family tyrosine kinase	Cytoplasm	kinase
-2.078	0.003	TF	transferrin	Extracellular Space	transporter
-2.07	0.003	RASAL3	RAS protein activator like 3	Cytoplasm	other
-2.041	0.003	FCGR1A	Fc fragment of IgG receptor Ia	Plasma Membrane	transmembrane receptor
-2.036	0.003	VAV1	vav guanine nucleotide exchange factor 1	Nucleus	transcription regulator
-1.982	0.003	UNC93B1	unc-93 homolog B1 (C. elegans)	Cytoplasm	transporter
1.979	0.003	Cyp2d22	cytochrome P450, family 2, subfamily d, polypeptide 22	Cytoplasm	enzyme
1.961	0.003	SLC1A3	solute carrier family 1 member 3	Plasma Membrane	transporter
-1.866	0.003	CFH	complement factor H	Extracellular Space	other
1.863	0.003	NDRG2	NDRG family member 2	Cytoplasm	other
-1.808	0.003	CTSS	cathepsin S	Cytoplasm	peptidase
-1.769	0.003	TYROBP	TYRO protein tyrosine kinase binding protein	Plasma Membrane	transmembrane receptor
-1.714	0.003	RTN4RL1	reticulon 4 receptor like 1	Plasma Membrane	other
-1.699	0.003	ELMO1	engulfment and cell motility 1	Cytoplasm	other
-1.637	0.003	CSF3R	colony stimulating factor 3 receptor	Plasma Membrane	transmembrane receptor
-1.622	0.003	LPCAT2	lysophosphatidylcholine acyltransferase 2	Cytoplasm	enzyme
1.602	0.003	TTR	transthyretin	Extracellular Space	transporter
-1.572	0.003	LCP1	lymphocyte cytosolic protein 1	Cytoplasm	other
-1.554	0.003	MPEG1	macrophage expressed 1	Cytoplasm	other
-1.515	0.003	OCA2	OCA2 melanosomal transmembrane protein	Plasma Membrane	transporter
-1.509	0.003	HMHA1	histocompatibility (minor) HA-1	Cytoplasm	transporter
-1.449	0.003	CORO1A	coronin 1A	Cytoplasm	other

1.389	0.003	TSPAN7	tetraspanin 7	Plasma Membrane	other
1.347	0.003	FGFR3	fibroblast growth factor receptor 3	Plasma Membrane	kinase
-1.338	0.003	FGD2	FYVE, RhoGEF and PH domain containing 2	Cytoplasm	other
-1.288	0.003	RPS6KA1	ribosomal protein S6 kinase A1	Cytoplasm	kinase
-1.288	0.003	SEPP1	selenoprotein P, plasma, 1	Extracellular Space	other
1.28	0.003	ENPP2	ectonucleotide pyrophosphatase/phosphodiesterase 2	Plasma Membrane	enzyme
1.253	0.003	mir-143	microRNA 143	Cytoplasm	microRNA
-1.247	0.003	HEXB	hexosaminidase subunit beta	Cytoplasm	enzyme
-1.228	0.003	TMEM119	transmembrane protein 119	Cytoplasm	other
-1.181	0.003	SIRPA	signal regulatory protein alpha	Plasma Membrane	phosphatase
1.177	0.003	mir-10	microRNA 100	Other	microRNA
-1.171	0.003	TGFBR1	transforming growth factor beta receptor 1	Plasma Membrane	kinase
1.086	0.003	KRT80	keratin 80	Cytoplasm	other
1.086	0.003	S1PR1	sphingosine-1-phosphate receptor 1	Plasma Membrane	G-protein coupled receptor
-1.038	0.003	ENOPH1	enolase-phosphatase 1	Cytoplasm	enzyme
-0.998	0.003	TMEM50B	transmembrane protein 50B	Plasma Membrane	other
0.987	0.003	ALDOC	aldolase, fructose-bisphosphate C	Cytoplasm	enzyme
0.985	0.003	ALS2CL	ALS2 C-terminal like	Cytoplasm	other
-0.955	0.003	VPS26A	VPS26 retromer complex component A	Cytoplasm	transporter
-0.935	0.003	PAK1IP1	PAK1 interacting protein 1	Nucleus	other
-0.897	0.003	LPGAT1	lysophosphatidylglycerol acyltransferase 1	Cytoplasm	other
-0.893	0.003	SCOC	short coiled-coil protein	Cytoplasm	other
-0.888	0.003	LEPROT	leptin receptor overlapping transcript	Plasma Membrane	other
-0.843	0.003	PHTF2	putative homeodomain transcription factor 2	Cytoplasm	other
0.823	0.003	CKB	creatine kinase B	Cytoplasm	kinase
-0.817	0.003	ANLN	anillin actin binding protein	Cytoplasm	other
-0.809	0.003	RBPMS2	RNA binding protein with multiple splicing 2	Other	other
-0.79	0.003	SNAI2	snail family zinc finger 2	Nucleus	transcription regulator
-0.774	0.003	REV3L	REV3 like, DNA directed polymerase zeta catalytic subunit	Nucleus	enzyme

9 Bibliography

Aboobaker AA and **Blaxter ML (2003)** Use of RNA interference to investigate gene function in the human filarial nematode parasite Brugia malayi. *Mol. Biochem. Parasitol.* 129(1): 41–51.

Afferson HC, **Eleftheriou E**, **Selkirk ME** and **Gounaris K (2012)** Trichinella spiralis secreted enzymes regulate nucleotide-induced mast cell activation and release of mouse mast cell protease 1. *Infect. Immun.* 80(11): 3761–7.

Ai J, Zhang R, Li Y, Pu J, Lu Y, Jiao J, Li K, Yu B, Li Z & other authors (2010) Circulating microRNA-1 as a potential novel biomarker for acute myocardial infarction. *Biochem. Biophys. Res. Commun.* 391(1): 73–7.

Al-Greene NT, Means AL, Lu P, Jiang A, Schmidt CR, Chakravarthy AB, Merchant NB, Washington MK, Zhang B & other authors (2013) Four jointed box 1 promotes angiogenesis and is associated with poor patient survival in colorectal carcinoma. *PLoS One*. 8(7): e69660.

Allen JE and **Maizels RM (2011)** Diversity and dialogue in immunity to helminths. *Nat. Rev. Immunol.* 11(6): 375–88.

Altuvia Y, **Landgraf P**, **Lithwick G**, **Elefant N**, **Pfeffer S**, **Aravin A**, **Brownstein MJ**, **Tuschl T** and **Margalit H (2005)** Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res.* 33(8): 2697–2706.

Amarante AFT, **Bricarello PA**, **Huntley JF**, **Mazzolin LP** and **Gomes JC (2005)** Relationship of abomasal histology and parasite-specific immunoglobulin A with the resistance to Haemonchus contortus infection in three breeds of sheep. *Vet. Parasitol.* 128(1–2): 99–107.

Anderson P and **Ivanov P (2014)** tRNA fragments in human health and disease. *FEBS Lett.* 588(23): 4297–304.

Angiolillo AL, **Sgadari C**, **Taub DD**, **Liao F**, **Farber JM**, **Maheshwari S**, **Kleinman HK**, **Reaman GH** and **Tosato G (1995)** Human interferon-inducible protein 10 is a potent inhibitor of angiogenesis in vivo. *J. Exp. Med.* 182(1): 155–62.

Anthony RM, **Urban JF**, **Alem F**, **Hamed HA**, **Rozo CT**, **Boucher J-L**, **Van Rooijen N** and **Gause WC (2006)** Memory T(H)2 cells induce alternatively activated macrophages to mediate protection against nematode parasites. *Nat. Med.* 12(8): 955–60. editor & translator, Nature Publishing Group.

Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF, Mitchell PS, Bennett CF, Pogosova-Agadjanyan EL & other authors (2011) Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. U. S. A.* 108(12): 5003–5008. editor & translator, National Academy of Sciences.

`

Athanasiadou S, **Jones LA**, **Burgess STG**, **Kyriazakis I**, **Pemberton AD**, **Houdijk JGM** and **Huntley JF (2011)** Genome-wide transcriptomic analysis of intestinal tissue to assess the impact of nutrition and a secondary nematode challenge in lactating rats. *PLoS One*. 6(6): e20771. editor & translator, Public Library of Science.

Bai S, **Nasser MW**, **Wang B**, **Hsu S-H**, **Datta J**, **Kutay H**, **Yadav A**, **Nuovo G**, **Kumar P** and **Ghoshal K (2009)** MicroRNA-122 inhibits tumorigenic properties of hepatocellular carcinoma cells and sensitizes these cells to sorafenib. *J. Biol. Chem.* 284(46): 32015–27.

Baietti MF, Zhang Z, Mortier E, Melchior A, Degeest G, Geeraerts A, Ivarsson Y, Depoortere F, Coomans C & other authors (2012) Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. *Nat. Cell Biol.* 14(7): 677–85. editor & translator, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.

Baker RL, **Mwamachi DM**, **Audho JO**, **Aduda EO** and **Thorpe W (1998)** Resistance of Galla and Small East African goats in the sub-humid tropics to gastrointestinal nematode infections and the peri-parturient rise in faecal egg counts. *Vet. Parasitol.* 79(1): 53–64.

Bakker N, **Vervelde L**, **Kanobana K**, **Knox DP**, **Cornelissen AWCA**, **Vries E d.** and **Yatsuda AP (2004)** Vaccination against the nematode Haemonchus contortus with a thiol-binding fraction from the excretory/secretory products (ES). *Vaccine*. 22(5–6): 618–628.

Balic A, **Bowles V.** and **Meeusen EN. (2000)** Cellular profiles in the abomasal mucosa and lymph node during primary infection with Haemonchus contortus in sheep. *Vet. Immunol. Immunopathol.* 75(1–2): 109–120.

Balic A, **Cunningham CP** and **Meeusen ENT (2006)** Eosinophil interactions with Haemonchus contortus larvae in the ovine gastrointestinal tract. *Parasite Immunol.* 28(3): 107–15.

Bandiera S, **Pfeffer S**, **Baumert TF** and **Zeisel MB (2015)** miR-122--a key factor and therapeutic target in liver disease. *J. Hepatol.* 62(2): 448–57.

Barr TA, **Brown S**, **Ryan G**, **Zhao J** and **Gray D (2007)** TLR-mediated stimulation of APC: Distinct cytokine responses of B cells and dendritic cells. *Eur. J. Immunol.* 37(11): 3040–53.

Bartel DPD (2009) MicroRNAs: Target Recognition and Regulatory Functions. *Cell.* 136(2): 215–233.

Baskerville S and **Bartel DP (2005)** Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*. 11(3): 241–247.

Beasley AM, **Kahn LP** and **Windon RG (2010)** The periparturient relaxation of immunity in Merino ewes infected with Trichostrongylus colubriformis: parasitological and immunological responses. *Vet. Parasitol.* 168(1–2): 60–70.

Beech RN, **Wolstenholme AJ**, **Neveu C** and **Dent JA (2010)** Nematode parasite genes: what's in a name? *Trends Parasitol.* 26(7): 334–340.

Belkaid Y and **Oldenhove G (2008)** Tuning microenvironments: induction of regulatory T

`

cells by dendritic cells. *Immunity*. 29(3): 362–71.

Berezikov E, **Guryev V**, **van de Belt J**, **Wienholds E**, **Plasterk RHA** and **Cuppen E (2005)** Phylogenetic shadowing and computational identification of human microRNA genes. *Cell*. 120(1): 21–24.

Bernal D, **de la Rubia JE**, **Carrasco-Abad AM**, **Toledo R**, **Mas-Coma S** and **Marcilla A (2004)** Identification of enolase as a plasminogen-binding protein in excretory-secretory products of Fasciola hepatica. *FEBS Lett*. 563(1–3): 203–6.

Bernal D, **Trelis M**, **Montaner S**, **Cantalapiedra F**, **Galiano A**, **Hackenberg M** and **Marcilla A (2014)** Surface analysis of Dicrocoelium dendriticum. The molecular characterization of exosomes reveals the presence of miRNAs. *J. Proteomics*. 105232–41.

Betel D, **Koppal A**, **Agius P**, **Sander C** and **Leslie C (2010)** Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol*. 11(8): R90. editor & translator, BioMed Central.

Bethony JM, **Loukas A**, **Hotez PJ** and **Knox DP (2006)** Vaccines against blood-feeding nematodes of humans and livestock. *Parasitology*. 133 Suppl(S2): S63-79. editor & translator, Cambridge University Press.

Bi Y, **Liu G** and **Yang R (2007)** Th17 cell induction and immune regulatory effects. *J. Cell. Physiol*. 211(2): 273–8.

Blackwell NM and **Else KJ (2001)** B cells and antibodies are required for resistance to the parasitic gastrointestinal nematode Trichuris muris. *Infect. Immun*. 69(6): 3860–8.

Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, Vierstraete A, Vanfleteren JR, Mackey LY, Dorris M & other authors (1998) A molecular evolutionary framework for the phylum Nematoda. *Nature*. 392(6671): 71–5.

Blitz NM and **Gibbs HC (1972)** Studies on the arrested development of Haemonchus contortus in sheep—II. Termination of arrested development and the spring rise phenomenon. *Int. J. Parasitol*. 2(1): 13–22.

Bonci D, Coppola V, Musumeci M, Addario A, Giuffrida R, Memeo L, D'Urso L, Pagliuca A, Biffoni M & other authors (2008) The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. *Nat. Med*. 14(11): 1271–7. editor & translator, Nature Publishing Group.

Boulin T, **Fauvin A**, **Charvet CL**, **Cortet J**, **Cabaret J**, **Bessereau J-L** and **Neveu C (2011)** Functional reconstitution of Haemonchus contortus acetylcholine receptors in Xenopus oocytes provides mechanistic insights into levamisole resistance. *Br. J. Pharmacol*. 164(5): 1421–32. editor & translator, Wiley-Blackwell.

Boulin T, **Gielen M**, **Richmond JE**, **Williams DC**, **Paoletti P** and **Bessereau J-L (2008)** Eight genes are required for functional reconstitution of the Caenorhabditis elegans levamisole-sensitive acetylcholine receptor. *Proc. Natl. Acad. Sci. U. S. A.* 105(47): 18590–5.

`

Bowdridge SA, **Zajac AM** and **Notter DR (2015)** St. Croix sheep produce a rapid and greater cellular immune response contributing to reduced establishment of Haemonchus contortus. *Vet. Parasitol.* 208(3–4): 204–10.

Brennecke J, **Stark A**, **Russell RB** and **Cohen SM (2005)** Principles of microRNA-target recognition. *PLoS Biol.* 3(3): e85. editor & translator, Public Library of Science.

Britton C, **Winter AD**, **Gillan V** and **Devaney E (2014)** microRNAs of parasitic helminths - Identification, characterization and potential as drug targets. *Int. J. Parasitol. Drugs drug Resist.* 4(2): 85–94. editor & translator, Elsevier.

Britton C, **Winter AD**, **Marks ND**, **Gu H**, **McNeilly TN**, **Gillan V** and **Devaney E (2015)** Application of small RNA technology for improved control of parasitic helminths. *Vet. Parasitol.* 212(1–2): 47–53.

Brocardo MG, **Schillaci R**, **Galeano A**, **Radrizzani M**, **White V**, **Guerrico AG**, **Santa-Coloma TA** and **Roldan A (2001)** Early effects of insulin-like growth factor-1 in activated human T lymphocytes. *J. Leukoc. Biol.* 70(2): 297–305.

Brooks DR, **Appleford PJ**, **Murray L** and **Isaac RE (2003)** An essential role in molting and morphogenesis of Caenorhabditis elegans for ACN-1, a novel member of the angiotensin-converting enzyme family that lacks a metallopeptidase active site. *J. Biol. Chem.* 278(52): 52340–6.

Brown A, **Burleigh JM**, **Billett EE** and **Pritchard DI (1995)** An initial characterization of the proteolytic enzymes secreted by the adult stage of the human hookworm Necator americanus. *Parasitology.* 110 (Pt 5(5): 555–563. editor & translator, Cambridge University Press.

Buck AH and **Blaxter M (2013)** Functional diversification of Argonautes in nematodes: an expanding universe. *Biochem. Soc. Trans.* 41(4): 881–6.

Buck AH, Coakley G, Simbari F, McSorley HJ, Quintana JF, Le Bihan T, Kumar S, Abreu-Goodger C, Lear M & other authors (2014) Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity. *Nat. Commun.* 55488. editor & translator, Nature Publishing Group.

Burgess CGS, **Bartley Y**, **Redman E**, **Skuce PJ**, **Nath M**, **Whitelaw F**, **Tait A**, **Gilleard JS** and **Jackson F (2012)** A survey of the trichostrongylid nematode species present on UK sheep farms and associated anthelmintic control practices. *Vet. Parasitol.* 189(2–4): 299–307. editor & translator, Elsevier B.V.

Bush EW and **van Rooij E (2014)** miR-25 in heart failure. *Circ. Res.* 115(7): 610–2.

Cai M-Y, **Luo R-Z**, **Li Y-H**, **Dong P**, **Zhang Z-L**, **Zhou F-J**, **Chen J-W**, **Yun J-P**, **Zhang CZ-Y** and **Cao Y (2012)** High-expression of ZBP-89 correlates with distal metastasis and poor prognosis of patients in clear cell renal cell carcinoma. *Biochem. Biophys. Res. Commun.* 426(4): 636–42.

Cai P, **Piao X**, **Hao L**, **Liu S**, **Hou N**, **Wang H** and **Chen Q (2013)** A deep analysis of the small

`

non-coding RNA population in Schistosoma japonicum eggs. *PLoS One*. 8(5): e64003. editor & translator, Public Library of Science.

Cai P, **Gobert GN** and **McManus DP (2015a)** MicroRNAs in Parasitic Helminthiases: Current Status and Future Perspectives. *Trends Parasitol*.

Cai P, **Gobert GN**, **You H**, **Duke M** and **McManus DP (2015b)** Circulating miRNAs: Potential Novel Biomarkers for Hepatopathology Progression and Diagnosis of Schistosomiasis Japonica in Two Murine Models. *PLoS Negl. Trop. Dis.* 9(7): e0003965. editor & translator, Public Library of Science.

Cao G, Huang B, Liu Z, Zhang J, Xu H, Xia W, Li J, Li S, Chen L & other authors (2010) Intronic miR-301 feedback regulates its host gene, ska2, in A549 cells by targeting MEOX2 to affect ERK/CREB pathways. *Biochem. Biophys. Res. Commun.* 396(4): 978–82.

Cassada RC and **Russell RL (1975)** The dauerlarva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans. *Dev. Biol.* 46(2): 326–42.

Chai SJ, **Yap YY**, **Foo YC**, **Yap LF**, **Ponniah S**, **Teo SH**, **Cheong SC**, **Patel V** and **Lim KP (2015)** Identification of Four-Jointed Box 1 (FJX1)-Specific Peptides for Immunotherapy of Nasopharyngeal Carcinoma. *PLoS One*. 10(11): e0130464.

Chaudhry U, **Redman EM**, **Raman M** and **Gilleard JS (2015)** Genetic evidence for the spread of a benzimidazole resistance mutation across southern India from a single origin in the parasitic nematode Haemonchus contortus. *Int. J. Parasitol.* 45(11): 721–8.

Cheeseman CL, **Delany NS**, **Woods DJ** and **Wolstenholme AJ (2001)** High-affinity ivermectin binding to recombinant subunits of the Haemonchus contortus glutamate-gated chloride channel. *Mol. Biochem. Parasitol.* 114(2): 161–168.

Chen X, **Liang H**, **Zhang J**, **Zen K** and **Zhang C-Y (2012)** Horizontal transfer of microRNAs: molecular mechanisms and clinical applications. *Protein Cell*. 3(1): 28–37.

Chen Y, Shen A, Rider PJ, Yu Y, Wu K, Mu Y, Hao Q, Liu Y, Gong H & other authors (2011) A liver-specific microRNA binds to a highly conserved RNA sequence of hepatitis B virus and negatively regulates viral gene expression and replication. *FASEB J.* 25(12): 4511–21.

Chen Z and **Duan X (2011)** Ribosomal RNA depletion for massively parallel bacterial RNA-sequencing applications. *Methods Mol. Biol.* 73393–103.

Cheng G, **Luo R**, **Hu C**, **Cao J** and **Jin Y (2013)** Deep sequencing-based identification of pathogen-specific microRNAs in the plasma of rabbits infected with Schistosoma japonicum. *Parasitology*. 140(14): 1751–61. editor & translator, Cambridge University Press.

Chertov O, **Ueda H**, **Xu LL**, **Tani K**, **Murphy WJ**, **Wang JM**, **Howard OM**, **Sayers TJ** and **Oppenheim JJ (1997)** Identification of human neutrophil-derived cathepsin G and azurocidin/CAP37 as chemoattractants for mononuclear cells and neutrophils. *J. Exp. Med.* 186(5): 739–47.

`

Chevillet JR, Kang Q, Ruf IK, Briggs HA, Vojtech LN, Hughes SM, Cheng HH, Arroyo JD, Meredith EK & other authors (2014) Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc. Natl. Acad. Sci. U. S. A.* 111(41): 14888–93. editor & translator, National Academy of Sciences.

Chi SW, **Zang JB**, **Mele A** and **Darnell RB (2009)** Argonaute HITS-CLIP decodes microRNA|[ndash]|mRNA interaction maps. *Nature, Publ. online 17 June 2009; | doi10.1038/10.1038/nature08170.* 460(7254): 479. editor & translator, Nature Publishing Group.

Chiejina SN, **Behnke JM** and **Fakae BB (2015)** Haemonchotolerance in West African Dwarf goats: contribution to sustainable, anthelmintics-free helminth control in traditionally managed Nigerian dwarf goats. *Parasite*. 227. editor & translator, EDP Sciences.

Lo Cicero A, **Stahl PD** and **Raposo G (2015)** Extracellular vesicles shuffling intercellular messages: for good or for bad. *Curr. Opin. Cell Biol.* 3569–77.

Cliffe LJ and **Grencis RK (2004)** The Trichuris muris system: a paradigm of resistance and susceptibility to intestinal nematode infection. *Adv. Parasitol.* 57255–307.

Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibé B, Bouix J, Caiment F, Elsen J-M & other authors (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat. Genet.* 38(7): 813–8.

Coakley G, Maizels RM, Buck AH, Murray CJ, al. et, Lustigman S, al. et, Prichard RK, al. et & other authors (2015) Exosomes and Other Extracellular Vesicles: The New Communicators in Parasite Infections. *Trends Parasitol.* 31(10): 477–89. editor & translator, Elsevier.

Connan RM (1968) Studies on the Worm Populations in the Alimentary Tract of Breeding Ewes. *J. Helminthol.* 42(1–2): 9–28. editor & translator, Cambridge University Press.

Conrad KD, **Giering F**, **Erfurth C**, **Neumann A**, **Fehr C**, **Meister G** and **Niepmann M (2013)** MicroRNA-122 dependent binding of Ago2 protein to hepatitis C virus RNA is associated with enhanced RNA stability and translation stimulation. *PLoS One.* 8(2): e56272. editor & translator, Public Library of Science.

Coutinho LL, **Matukumalli LK**, **Sonstegard TS**, **Van Tassell CP**, **Gasbarre LC**, **Capuco A V** and **Smith TPL (2007)** Discovery and profiling of bovine microRNAs from immune-related and embryonic tissues. *Physiol. Genomics.* 29(1): 35–43.

Cowie RH (2013) Biology, systematics, life cycle, and distribution of Angiostrongylus cantonensis, the cause of rat lungworm disease. *Hawaii. J. Med. Public Health.* 72(6 Suppl 2): 6–9.

Creson JR, **Lin AA**, **Li Q**, **Broad DF**, **Roberts MR** and **Anderson SJ (1999)** The Mode and Duration of Anti-CD28 Costimulation Determine Resistance to Infection by Macrophage-Tropic Strains of Human Immunodeficiency Virus Type 1 In Vitro. *J. Virol.* 73(11): 9337–9347.

`

D'Elia R, **Behnke JM**, **Bradley JE** and **Else KJ (2009)** Regulatory T cells: a role in the control of helminth-driven intestinal pathology and worm survival. *J. Immunol.* 182(4): 2340–8. editor & translator, American Association of Immunologists.

Dalley BK and **Golomb M (1992)** Gene expression in the Caenorhabditis elegans dauer larva: developmental regulation of Hsp90 and other genes. *Dev. Biol.* 151(1): 80–90.

Dalton J. and **Mulcahy G (2001)** Parasite vaccines — a reality? *Vet. Parasitol.* 98(1–3): 149–167.

Dalton JP, **Brindley PJ**, **Knox DP**, **Brady CP**, **Hotez PJ**, **Donnelly S**, **O'Neill SM**, **Mulcahy G** and **Loukas A (2003)** Helminth vaccines: from mining genomic information for vaccine targets to systems used for protein expression. *Int. J. Parasitol.* 33(5–6): 621–640.

Daniel C, **Gerlach K**, **Väth M**, **Neurath MF** and **Weigmann B (2014)** Nuclear factor of activated T cells - a transcription factor family as critical regulator in lung and colon cancer. *Int. J. cancer.* 134(8): 1767–75.

Davies SJ and **McKerrow JH (2003)** Developmental plasticity in schistosomes and other helminths. *Int. J. Parasitol.* 33(11): 1277–84.

Denli AM, **Tops BBJ**, **Plasterk RH a**, **Ketting RF** and **Hannon GJ (2004)** Processing of primary microRNAs by the Microprocessor complex. *Nature.* 432(7014): 231–235.

Dent J a, **Smith MM**, **Vassilatis DK** and **Avery L (2000)** The genetics of ivermectin resistance in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* 97(6): 2674–2679.

Dent JA, **Davis MW** and **Avery L (1997)** avr-15 encodes a chloride channel subunit that mediates inhibitory glutamatergic neurotransmission and ivermectin sensitivity in Caenorhabditis elegans. *EMBO J.* 16(19): 5867–79.

Dickinson B, **Zhang Y**, **Petrick JS**, **Heck G**, **Ivashuta S** and **Marshall WS (2013)** Lack of detectable oral bioavailability of plant microRNAs after feeding in mice. *Nat. Biotechnol.* 31(11): 965–967. editor & translator, Nature Publishing Group.

Diebold SS (2008) Determination of T-cell fate by dendritic cells. *Immunol. Cell Biol.* 86(5): 389–97. editor & translator, Australasian Society for Immunology Inc.

Doench JG and **Sharp PA (2004)** Specificity of microRNA target selection in translational repression. *Genes Dev.* 18(5): 504–11.

Doench JG, **Petersen CP** and **Sharp P a. (2003)** siRNAs can function as miRNAs. *Genes Dev.* 17(4): 438–442. editor & translator, Cold Spring Harbor Laboratory Press.

Dreymueller D, **Pruessmeyer J**, **Groth E** and **Ludwig A (2012)** The role of ADAM-mediated shedding in vascular biology. *Eur. J. Cell Biol.* 91(6): 472–485.

Drudge J, **Jr SL** and **Wyant Z (1957)** Strain variation in the response of sheep nematodes to the action of phenothiazine. II. Studies on pure infections of *Haemonchus contortus*. *Am. J. ….* 18(67): 317–325.

Dufour JH, **Dziejman M**, **Liu MT**, **Leung JH**, **Lane TE** and **Luster AD (2002)** IFN- -Inducible

`

Protein 10 (IP-10; CXCL10)-Deficient Mice Reveal a Role for IP-10 in Effector T Cell Generation and Trafficking. *J. Immunol.* 168(7): 3195–3204. editor & translator, American Association of Immunologists.

Dunand-Sauthier I, Santiago-Raber M-L, Capponi L, Vejnar CE, Schaad O, Irla M, Seguín-Estévez Q, Descombes P, Zdobnov EM & other authors (2011) Silencing of c-Fos expression by microRNA-155 is critical for dendritic cell maturation and function. *Blood.* 117(17): 4490–500.

Entrican G, **Wattegedera SR** and **Griffiths DJ (2015)** Exploiting ovine immunology to improve the relevance of biomedical models. *Mol. Immunol.* 66(1): 68–77.

Estrada-Reyes ZM, **López-Reyes AG**, **Lagunas-Martínez A**, **Ramírez-Vargas G**, **Olazarán-Jenkins S**, **Hernández-Romano J**, **Mendoza-de-Gives P** and **López-Arellano ME (2015)** Relative expression analysis of IL-5 and IL-6 genes in tropical sheep breed Pelibuey infected with Haemonchus contortus. *Parasite Immunol.* 37(9): 446–52.

Etheridge A, **Lee I**, **Hood L**, **Galas D** and **Wang K (2011)** Extracellular microRNA: a new source of biomarkers. *Mutat. Res.* 717(1–2): 85–90.

Falzon LC, **Menzies PI**, **Shakya KP**, **Jones-Bitton A**, **Vanleeuwen J**, **Avula J**, **Jansen JT** and **Peregrine AS (2013)** A longitudinal study on the effect of lambing season on the periparturient egg rise in Ontario sheep flocks. *Prev. Vet. Med.* 110(3–4): 467–80.

Faraoni I, **Antonetti FR**, **Cardone J** and **Bonmassar E (2009)** miR-155 gene: a typical multifunctional microRNA. *Biochim. Biophys. Acta.* 1792(6): 497–505.

Farh KK-H, **Grimson A**, **Jan C**, **Lewis BP**, **Johnston WK**, **Lim LP**, **Burge CB** and **Bartel DP (2005)** The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science.* 310(5755): 1817–21. editor & translator, American Association for the Advancement of Science.

Fetterer RH and **Rhoads ML (1997)** The in vitro uptake and incorporation of hemoglobin by adult Haemonchus-contortus. *Vet. Parasitol.* 69(1–2): 77–87.

Finney CAM, **Taylor MD**, **Wilson MS** and **Maizels RM (2007)** Expansion and activation of CD4(+)CD25(+) regulatory T cells in Heligmosomoides polygyrus infection. *Eur. J. Immunol.* 37(7): 1874–86.

Fire A, **Xu S**, **Montgomery MK**, **Kostas SA**, **Driver SE** and **Mello CC (1998)** Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature.* 391(6669): 806–11.

Fishelson Z (1995) Novel mechanisms of immune evasion by Schistosoma mansoni. *Memo□rias do Inst. Oswaldo Cruz.* 90(2): 289–292. editor & translator, Fundação Oswaldo Cruz.

Fontenot JD, **Gavin MA** and **Rudensky AY (2003)** Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat. Immunol.* 4(4): 330–336. editor & translator, Nature Publishing Group.

`

Forrester SG, **Hamdan FF**, **Prichard RK** and **Beech RN (1999)** Cloning, sequencing, and developmental expression levels of a novel glutamate-gated chloride channel homologue in the parasitic nematode Haemonchus contortus. *Biochem. Biophys. Res. Commun.* 254(3): 529–34.

Forrester SG, **Prichard RK** and **Beech RN (2002)** A glutamate-gated chloride channel subunit from Haemonchus contortus:: Expression in a mammalian cell line, ligand binding, and modulation of anthelmintic binding by glutamate. *Biochem. Pharmacol.* 63(6): 1061–1068.

Frankel LB, **Christoffersen NR**, **Jacobsen A**, **Lindow M**, **Krogh A** and **Lund AH (2008)** Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. *J. Biol. Chem.* 283(2): 1026–33.

Freudenstein-Dan A, **Gold D** and **Fishelson Z (2003)** Killing of schistosomes by elastase and hydrogen peroxide: implications for leukocyte-mediated schistosome killing. *J. Parasitol.* 89(6): 1129–35.

Fu Y, Wang W, Tong J, Pan Q, Long Y, Qian W, Hou X, Fu Y, Wang W & other authors (2009) Th17: A New Participant in Gut Dysfunction in Mice Infected with *Trichinella spiralis*. *Mediators Inflamm.* 20091–7. editor & translator, Hindawi Publishing Corporation.

Gallucci S and **Matzinger P (2001)** Danger signals: SOS to the immune system. *Curr. Opin. Immunol.* 13(1): 114–119.

Gantier MP, **Sadler AJ** and **Williams BRG (2007)** Fine-tuning of the innate immune response by microRNAs. *Immunol. Cell Biol.* 85(6): 458–62.

Gao XH, Liu QZ, Chang W, Xu XD, Du Y, Han Y, Liu Y, Yu ZQ, Zuo ZG & other authors (2013) Expression of ZNF148 in different developing stages of colorectal cancer and its prognostic value: A large Chinese study based on tissue microarray. *Cancer.* 119(12): 2212–2222.

Garcia-Silva MR, **Cabrera-Cabrera F**, **das Neves RFC**, **Souto-Padrón T**, **de Souza W** and **Cayota A (2014a)** Gene expression changes induced by Trypanosoma cruzi shed microvesicles in mammalian host cells: relevance of tRNA-derived halves. *Biomed Res. Int.* 2014305239. editor & translator, Hindawi Publishing Corporation.

Garcia-Silva MR, Cura das Neves RF, Cabrera-Cabrera F, Sanguinetti J, Medeiros LC, Robello C, Naya H, Fernandez-Calero T, Souto-Padron T & other authors (2014b) Extracellular vesicles shed by Trypanosoma cruzi are linked to small RNA pathways, life cycle regulation, and susceptibility to infection of mammalian cells. *Parasitol. Res.* 113(1): 285–304. editor & translator, Springer Berlin Heidelberg.

Geldhof P, **Meyvis Y**, **Vercruysse J** and **Claerebout E (2008)** Vaccine testing of a recombinant activation-associated secreted protein (ASP1) from Ostertagia ostertagi. *Parasite Immunol.* 30(1): 57–60.

`

Geldhof P, **Claerebout E**, **Knox D**, **Vercauteren I**, **Looszova a.** and **Vercruysse J (2002)** Vaccination of calves against ostertagia ostertagi with cysteine proteinase enriched protein fractions. *Parasite Immunol*. 24(5): 263–270.

Geldhof P, **Vercauteren I**, **Vercruysse J**, **Knox DP**, **Van Den Broeck W** and **Claerebout E (2004)** Validation of the protective *Ostertagia ostertagi* ES-thiol antigens with different adjuvantia. *Parasite Immunol*. 26(1): 37–43.

Geldhof P, **Murray L**, **Couthier A**, **Gilleard JS**, **McLauchlan G**, **Knox DP** and **Britton C (2006)** Testing the efficacy of RNA interference in Haemonchus contortus. *Int. J. Parasitol*. 36(7): 801–10.

Ghendler Y, **Parizade M**, **Arnon R**, **McKerrow JH** and **Fishelson Z (1996)** Schistosoma mansoni: evidence for a 28-kDa membrane-anchored protease on schistosomula. *Exp. Parasitol*. 83(1): 73–82.

Ghosh R, **Andersen EC**, **Shapiro JA**, **Gerke JP** and **Kruglyak L (2012)** Natural variation in a chloride channel subunit confers avermectin resistance in C. elegans. *Science*. 335(6068): 574–8.

Giannoukos G, Ciulla DM, Huang K, Haas BJ, Izard J, Levin JZ, Livny J, Earl AM, Gevers D & other authors (2012) Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. *Genome Biol*. 13(3): R23.

Gill HS, **Altmann K**, **Cross ML** and **Husband AJ (2000)** Induction of T helper 1- and T helper 2-type immune responses during Haemonchus contortus infection in sheep. *Immunology*. 99(3): 458–63.

Gillan V and **Devaney E (2005)** Regulatory T cells modulate Th2 responses induced by Brugia pahangi third-stage larvae. *Infect. Immun.* 73(7): 4034–42.

Gilleard JS (2013) Haemonchus contortus as a paradigm and model to study anthelmintic drug resistance. *Parasitology*. 140(12): 1506–22. editor & translator, Cambridge University Press.

Gilleard JS (2004) The use of Caenorhabditis elegans in parasitic nematode research. *Parasitology*. 128(Supplement S1): S49–S70.

Gironella M, Seux M, Xie M-J, Cano C, Tomasini R, Gommeaux J, Garcia S, Nowak J, Yeung ML & other authors (2007) Tumor protein 53-induced nuclear protein 1 expression is repressed by miR-155, and its restoration inhibits pancreatic tumor development. *Proc. Natl. Acad. Sci. U. S. A.* 104(41): 16170–5.

Glasson SS, Askew R, Sheppard B, Carito B, Blanchet T, Ma H-L, Flannery CR, Peluso D, Kanki K & other authors (2005) Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis. *Nature*. 434(7033): 644–8.

Gomez-Escobar N, **Lewis E** and **Maizels RM (1998)** A novel member of the transforming growth factor-beta (TGF-beta) superfamily from the filarial nematodes Brugia malayi and B. pahangi. *Exp. Parasitol*. 88(3): 200–9.

`

Gomez-Escobar N, **Gregory WF** and **Maizels RM (2000)** Identification of tgh-2, a filarial nematode homolog of Caenorhabditis elegans daf-7 and human transforming growth factor beta, expressed in microfilarial and adult stages of Brugia malayi. *Infect. Immun.* 68(11): 6402–10.

Gossner AG, **Venturina VM**, **Shaw DJ**, **Pemberton JM** and **Hopkins J (2012)** Relationship between susceptibility of Blackface sheep to Teladorsagia circumcincta infection and an inflammatory mucosal T cell response. *Vet. Res.* 43(1): 26. editor & translator, BioMed Central.

Grainger JR, Smith KA, Hewitson JP, McSorley HJ, Harcus Y, Filbey KJ, Finney CAMM, Greenwood EJDD, Knox DP & other authors (2010) Helminth secretions induce de novo T cell Foxp3 expression and regulatory function through the TGF-β pathway. *J. Exp. Med.* 207(11): 2331–2341.

Grencis RK, **Humphreys NE** and **Bancroft AJ (2014)** Immunity to gastrointestinal nematodes: mechanisms and myths. *Immunol. Rev.* 260(1): 183–205.

Grewal IS and **Flavell RA (1996)** The Role of CD40 Ligand in Costimulation and T-Cell Activation. *Immunol. Rev.* 153(1): 85–106.

Guo YE, **Oei T** and **Steitz JA (2015)** Herpesvirus saimiri MicroRNAs Preferentially Target Host Cell Cycle Regulators. *J. Virol.* 89(21): 10901–11. editor & translator, American Society for Microbiology (ASM).

Halachmi N and **Lev Z (2002)** The Sec1 Family: A Novel Family of Proteins Involved in Synaptic Transmission and General Secretion. *J. Neurochem.* 66(3): 889–897.

Han J, **Lee Y**, **Yeom KH**, **Kim YK**, **Jin H** and **Kim VN (2004)** The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* 18(24): 3016–3027.

Han Y, **Guo Q**, **Zhang M**, **Chen Z** and **Cao X (2009)** CD69+CD4+CD25- T Cells, a New Subset of Regulatory T Cells, Suppress T Cell Proliferation through Membrane-Bound TGF- 1. *J. Immunol.* 182(1): 111–120.

Haneklaus M, **Gerlic M**, **O'Neill LAJ** and **Masters SL (2013)** miR-223: infection, inflammation and cancer. *J. Intern. Med.* 274(3): 215–26.

Hang L, **Blum AM**, **Setiawan T**, **Urban JP**, **Stoyanoff KM** and **Weinstock J V (2013)** Heligmosomoides polygyrus bakeri infection activates colonic Foxp3+ T cells enhancing their capacity to prevent colitis. *J. Immunol.* 191(4): 1927–34. editor & translator, NIH Public Access.

Hansen EP, **Kringel H**, **Williams AR** and **Nejsum P (2015)** Secretion of RNA-Containing Extracellular Vesicles by the Porcine Whipworm, Trichuris suis. *J. Parasitol.* 101(3): 336–40. editor & translator, American Society of Parasitologists.

Harnett MM, **Melendez AJ** and **Harnett W (2010)** The therapeutic potential of the filarial nematode-derived immunodulator, ES-62 in inflammatory disease. *Clin. Exp. Immunol.* 159(3): 256–67.

`

Harrington LE, **Hatton RD**, **Mangan PR**, **Turner H**, **Murphy TL**, **Murphy KM** and **Weaver CT (2005)** Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nat. Immunol.* 6(11): 1123–32.

Harrington LE, **Mangan PR** and **Weaver CT (2006)** Expanding the effector CD4 T-cell repertoire: the Th17 lineage. *Curr. Opin. Immunol.* 18(3): 349–56.

Hartmann D, Strooper B de, Serneels L, Craessaerts K, Herreman A, Annaert W, Umans L, Lübke T, Illert AL & other authors (2002) The disintegrin/metalloprotease ADAM 10 is essential for Notch signalling but not for alpha-secretase activity in fibroblasts. *Hum. Mol. Genet.* 11(21): 2615–2624. editor & translator, Oxford University Press.

Hartmann S and **Lucius R (2003)** Modulation of host immune responses by nematode cystatins. *Int. J. Parasitol.* 33(11): 1291–1302.

Hata M, **Takahara S**, **Tsuzaki H**, **Ishii Y**, **Nakata K**, **Akagawa KS** and **Satoh K (2009)** Expression of Th2-skewed pathology mediators in monocyte-derived type 2 of dendritic cells (DC2). *Immunol. Lett.* 126(1–2): 29–36.

Hata T, **Murakami K**, **Nakatani H**, **Yamamoto Y**, **Matsuda T** and **Aoki N (2010)** *Isolation of bovine milk-derived microvesicles carrying mRNAs and microRNAs. Biochem. Biophys. Res. Commun.*

Hatfield SD, **Shcherbata HR**, **Fischer KA**, **Nakahara K**, **Carthew RW** and **Ruohola-Baker H (2005)** Stem cell division is regulated by the microRNA pathway. *Nature.* 435(7044): 974–8.

Hausser J and **Zavolan M (2014)** Identification and consequences of miRNA-target interactions - beyond repression of gene expression. *Nat. Rev. Genet.* 15(9): 599–612. editor & translator, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.

Hayes J, **Peruzzi PP** and **Lawler S (2014)** MicroRNAs in cancer: biomarkers, functions and therapy. *Trends Mol. Med.* 20(8): 460–9. editor & translator, Elsevier.

He X, **Sai X**, **Chen C**, **Zhang Y**, **Xu X**, **Zhang D** and **Pan W (2013)** Host serum miR-223 is a potential new biomarker for Schistosoma japonicum infection and the response to chemotherapy. *Parasit. Vectors.* 6272.

Herbert DR, Hölscher C, Mohrs M, Arendse B, Schwegmann A, Radwanska M, Leeto M, Kirsch R, Hall P & other authors (2004) Alternative Macrophage Activation Is Essential for Survival during Schistosomiasis and Downmodulates T Helper 1 Responses and Immunopathology. *Immunity.* 20(5): 623–635. editor & translator, Elsevier.

Hernando G, **Bergé I**, **Rayes D** and **Bouzat C (2012)** Contribution of subunits to Caenorhabditis elegans levamisole-sensitive nicotinic receptor function. *Mol. Pharmacol.* 82(3): 550–60.

Hewitson JP, **Harcus YM**, **Curwen RS**, **Dowle AA**, **Atmadja AK**, **Ashton PD**, **Wilson A** and **Maizels RM (2008)** The secretome of the filarial parasite, Brugia malayi: proteomic

`

profile of adult excretory-secretory products. *Mol. Biochem. Parasitol.* 160(1): 8–21.

Hewitson JP, **Grainger JR** and **Maizels RM (2009)** Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. *Mol. Biochem. Parasitol.* 167(1): 1–11.

Höglund J, **Gustafsson K**, **Ljungström B-L**, **Engström A**, **Donnan A** and **Skuce P (2009)** Anthelmintic resistance in Swedish sheep flocks based on a comparison of the results from the faecal egg count reduction test and resistant allele frequencies of the beta-tubulin gene. *Vet. Parasitol.* 161(1–2): 60–8.

Holden-Dye L and **Walker RJ (2007)** Anthelmintic drugs. *WormBook*. 1–13. editor & translator, WormBook.

Holmes PH (1985) Pathogenesis of trichostrongylosis. *Vet. Parasitol.* 18(2): 89–101.

Hori S, Nomura T, Sakaguchi S, Parijs L Van, Abbas AK, Sakaguchi S, Maloy KJ, Powrie F, Coutinho A & other authors (2003) Control of regulatory T cell development by the transcription factor Foxp3. *Science*. 299(5609): 1057–61. editor & translator, American Association for the Advancement of Science.

Hoy AM, Lundie RJ, Ivens A, Quintana JF, Nausch N, Forster T, Jones F, Kabatereine NB, Dunne DW & other authors (2014) Parasite-derived microRNAs in host serum as novel biomarkers of helminth infection. *PLoS Negl. Trop. Dis.* 8(2): e2701. editor & translator, Public Library of Science.

Hsu S-H, Wang B, Kota J, Yu J, Costinean S, Kutay H, Yu L, Bai S, La Perle K & other authors (2012) Essential metabolic, anti-inflammatory, and anti-tumorigenic functions of miR-122 in liver. *J. Clin. Invest.* 122(8): 2871–83. editor & translator, American Society for Clinical Investigation.

Hu G, **Drescher KM** and **Chen X-M (2012)** Exosomal miRNAs: Biological Properties and Therapeutic Potential. *Front. Genet.* 356. editor & translator, Frontiers Media SA.

Hu W and **Coller J (2012)** What comes first: translational repression or mRNA degradation? The deepening mystery of microRNA function. *Cell Res.* 22(9): 1322–4. editor & translator, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences.

Huang Z, **Huang D**, **Ni S**, **Peng Z**, **Sheng W** and **Du X (2010)** Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. *Int. J. Cancer*. 127(1): 118–26.

Huntley JF, **Redmond J**, **Welfare W**, **Brennan G**, **Jackson F**, **Kooyman F** and **Vervelde L (2001)** Studies on the immunoglobulin E responses to Teladorsagia circumcincta in sheep: purification of a major high molecular weight allergen. *Parasite Immunol.* 23(5): 227–235.

Hussein AS, **Kichenin K** and **Selkirk ME (2002)** Suppression of secreted acetylcholinesterase expression in Nippostrongylus brasiliensis by RNA interference. *Mol. Biochem. Parasitol.* 122(1): 91–94.

`

Hutvágner G, **McLachlan J**, **Pasquinelli AE**, **Bálint É**, **Tuschl T** and **Zamore PD (2001)** A Cellular Function for the RNA-Interference Enzyme Dicer in the Maturation of the let-7 Small Temporal RNA. *Science (80-.).* 293(5531): 834–838.

Ingham A, **Reverter A**, **Windon R**, **Hunt P** and **Menzies M (2008)** Gastrointestinal nematode challenge induces some conserved gene expression changes in the gut mucosa of genetically resistant sheep. *Int. J. Parasitol.* 38(3–4): 431–42.

Islam MK, **Miyoshi T**, **Yamada M** and **Tsuji N (2005)** Pyrophosphatase of the roundworm Ascaris suum plays an essential role in the worm's molting and development. *Infect. Immun.* 73(4): 1995–2004.

Issa Z, **Grant WN**, **Stasiuk S** and **Shoemaker CB (2005)** Development of methods for RNA interference in the sheep gastrointestinal parasite, Trichostrongylus colubriformis. *Int. J. Parasitol.* 35(9): 935–40.

Jangra RK, **Yi M** and **Lemon SM (2010)** Regulation of hepatitis C virus translation and infectious virus production by the microRNA miR-122. *J. Virol.* 84(13): 6615–25.

Janssen HLA, Reesink HW, Lawitz EJ, Zeuzem S, Rodriguez-Torres M, Patel K, van der Meer AJ, Patick AK, Chen A & other authors (2013) Treatment of HCV infection by targeting microRNA. *N. Engl. J. Med.* 368(18): 1685–94. editor & translator, Massachusetts Medical Society.

Jarrett WF, **Jennings FW**, **Mcintyre WI**, **Mulligan W** and **Urquhart GM (1958)** Irradiated helminth larvae in vaccination. *Proc. R. Soc. Med.* 51(9): 743–4.

Jarrett WF, **Jennings FW**, **Mcintyre WI**, **Mulligan W**, **Thomas BA** and **Urquhart GM (1959)** Immunological studies on Dictyocaulus viviparus infection: the immunity resulting from experimental infection. *Immunology*. 2252–61.

Jarrett WF, **Jennings FW**, **Mcintyre WI**, **Mulligan W** and **Urquhart GM (1960)** Immunological studies on Dictyocaulus viviparus infection; immunity produced by the administration of irradiated larvae. *Immunology*. 3145–51.

Jeong SP, **Kang J-A** and **Park S-G (2012)** Intestinal intraepithelial TCRγδ+ T cells are activated by normal commensal bacteria. *J. Microbiol.* 50(5): 837–841. editor & translator, The Microbiological Society of Korea.

Jex AR, Liu S, Li B, Young ND, Hall RS, Li Y, Yang L, Zeng N, Xu X & other authors (2011) Ascaris suum draft genome. *Nature*. 479(7374): 529–33.

Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, Wu C, Muzny DM, Li Y & other authors (2014) The sheep genome illuminates biology of the rumen and lipid metabolism. *Science*. 344(6188): 1168–73.

Joh T, **Kataoka H**, **Tanida S**, **Watanabe K**, **Ohshima T**, **Sasaki M**, **Nakao H**, **Ohhara H**, **Higashiyama S** and **Itoh M (2005)** Helicobacter pylori-Stimulated Interleukin-8 (IL-8) Promotes Cell Proliferation Through Transactivation of Epidermal Growth Factor Receptor (EGFR) by Disintegrin and Metalloproteinase (ADAM) Activation. *Dig. Dis. Sci.*

`

50(11): 2081–2089. editor & translator, Kluwer Academic Publishers-Plenum Publishers.

Jones AK, **Buckingham SD** and **Sattelle DB (2005)** Chemistry-to-gene screens in Caenorhabditis elegans. *Nat. Rev. Drug Discov.* 4(4): 321–30. editor & translator, Nature Publishing Group.

Jopling CL, **Yi M**, **Lancaster AM**, **Lemon SM** and **Sarnow P (2005)** Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science*. 309(5740): 1577–81. editor & translator, American Association for the Advancement of Science.

Kamal SM and **El Sayed Khalifa K (2006)** Immune modulation by helminthic infections: worms and viral infections. *Parasite Immunol*. 28(10): 483–496.

Kaminsky R (2003) Drug resistance in nematodes: a paper tiger or a real problem? *Curr. Opin. Infect. Dis.*

Kaminsky R, Gauvry N, Schorderet Weber S, Skripsky T, Bouvier J, Wenger a., Schroeder F, Desaules Y, Hotz R & other authors (2008a) Identification of the amino-acetonitrile derivative monepantel (AAD 1566) as a new anthelmintic drug development candidate. *Parasitol. Res.* 103(4): 931–939.

Kaminsky R, Ducray P, Jung M, Clover R, Rufener L, Bouvier J, Weber SS, Wenger A, Wieland-Berghausen S & other authors (2008b) A new class of anthelmintics effective against drug-resistant nematodes. *Nature*. 452(7184): 176–180.

Kane CM, **Cervi L**, **Sun J**, **McKee AS**, **Masek KS**, **Shapira S**, **Hunter CA** and **Pearce EJ (2004)** Helminth Antigens Modulate TLR-Initiated Dendritic Cell Activation. *J. Immunol.* 173(12): 7454–7461. editor & translator, American Association of Immunologists.

Kaplan RM (2004) Drug resistance in nematodes of veterinary importance: a status report. *Trends Parasitol*. 20(10): 477–81.

Karanu FN, **Rurangirwa FR**, **McGuire TC** and **Jasmer DP (1993)** Haemonchus contortus: identification of proteases with diverse characteristics in adult worm excretory-secretory products. *Exp. Parasitol*. 77(3): 362–71.

Kawai T and **Akira S (2006)** TLR signaling. *Cell Death Differ*. 13(5): 816–825. editor & translator, Nature Publishing Group.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S & other authors (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28(12): 1647–9.

Khan AA, **Betel D**, **Miller ML**, **Sander C**, **Leslie CS** and **Marks DS (2009)** Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. *Nat. Biotechnol*. 27(6): 549–55.

Khvorova A, **Reynolds A** and **Jayasena SD (2003)** Functional siRNAs and miRNAs exhibit strand bias. *Cell*. 115(2): 209–216.

`

Kim NH, Kim HS, Kim N-G, Lee I, Choi H-S, Li X-Y, Kang SE, Cha SY, Ryu JK & other authors (2011) p53 and microRNA-34 are suppressors of canonical Wnt signaling. *Sci. Signal*. 4(197): ra71.

Kim VN (2005) MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol*. 6(5): 376–385. editor & translator, Nature Publishing Group.

Kincaid RP and **Sullivan CS (2012)** Virus-Encoded microRNAs: An Overview and a Look to the Future. *PLoS Pathog*. 8(12): e1003018 (T. C. Hobman, Ed.). editor & translator, Public Library of Science.

Klass M and **Hirsh D (1976)** Non-ageing developmental variant of Caenorhabditis elegans. *Nature*. 260(5551): 523–5.

Klaver EJ, **Kuijk LM**, **Laan LC**, **Kringel H**, **van Vliet SJ**, **Bouma G**, **Cummings RD**, **Kraal G** and **van Die I (2013)** Trichuris suis-induced modulation of human dendritic cell function is glycan-mediated. *Int. J. Parasitol*. 43(3–4): 191–200.

Knox DP (2007) Proteinase inhibitors and helminth parasite infection. *Parasite Immunol*. 29(2): 57–71.

Knox DP, **Smith SK** and **Smith WD (1999)** Immunization with an affinity purified protein extract from the adult parasite protects lambs against infection with Haemonchus contortus. *Parasite Immunol*. 21(4): 201–10.

Knox D (2011) Proteases in blood-feeding nematodes and their potential as vaccine candidates. *Adv. Exp. Med. Biol*. 712155–76.

Knox DP, **Redmond DL**, **Newlands GF**, **Skuce PJ**, **Pettit D** and **Smith WD (2003)** The nature and prospects for gut membrane proteins as vaccine candidates for Haemonchus contortus and other ruminant trichostrongyloids. *Int. J. Parasitol*. 33(11): 1129–1137.

Köberle V, Pleli T, Schmithals C, Augusto Alonso E, Haupenthal J, Bönig H, Peveling-Oberhag J, Biondi RM, Zeuzem S & other authors (2013) Differential stability of cell-free circulating microRNAs: implications for their utilization as biomarkers. *PLoS One*. 8(9): e75184. editor & translator, Public Library of Science.

Koga Y, **Yasunaga M**, **Moriya Y**, **Akasu T**, **Fujita S**, **Yamamoto S** and **Matsumura Y (2011)** Exosome can prevent RNase from degrading microRNA in feces. *J. Gastrointest. Oncol*. 2(4): 215–22. editor & translator, AME Publications.

Komai-Koma M, **Xu D**, **Li Y**, **McKenzie ANJ**, **McInnes IB** and **Liew FY (2007)** IL-33 is a chemoattractant for human Th2 cells. *Eur. J. Immunol*. 37(10): 2779–86.

Kong YW, Cannell IG, de Moor CH, Hill K, Garside PG, Hamilton TL, Meijer HA, Dobbyn HC, Stoneley M & other authors (2008) The mechanism of micro-RNA-mediated translation repression is determined by the promoter of the target gene. *Proc. Natl. Acad. Sci*. 105(26): 8866–8871.

Kooyman F, **Schallig H**, **Van Leeuwen MA**, **MacKellar A**, **Huntley J**, **Cornelissen A** and **Vervelde L (2000)** Protection in lambs vaccinated with Haemonchus contortus antigens

`

is age related, and correlates with IgE rather than IgG1 antibody. *Parasite Immunol*. 22(1): 13–20.

Korn T, **Bettelli E**, **Gao W**, **Awasthi A**, **Jäger A**, **Strom TB**, **Oukka M** and **Kuchroo VK (2007)** IL-21 initiates an alternative pathway to induce proinflammatory T(H)17 cells. *Nature*. 448(7152): 484–7.

Korn T, **Bettelli E**, **Oukka M** and **Kuchroo VK (2009)** IL-17 and Th17 Cells. *Annu. Rev. Immunol*. 27485–517. editor & translator, Annual Reviews.

Kotze AC and **Bagnall NH (2006)** RNA interference in Haemonchus contortus: suppression of beta-tubulin gene expression in L3, L4 and adult worms in vitro. *Mol. Biochem. Parasitol*. 145(1): 101–10.

Kraft S, **Jouvin M-H**, **Kulkarni N**, **Kissing S**, **Morgan ES**, **Dvorak AM**, **Schroder B**, **Saftig P** and **Kinet J-P (2013)** The Tetraspanin CD63 Is Required for Efficient IgE-Mediated Mast Cell Degranulation and Anaphylaxis. *J. Immunol*. 191(6): 2871–2878. editor & translator, American Association of Immunologists.

Kuhn DE, **Martin MM**, **Feldman DS**, **Terry A V.**, **Nuovo GJ** and **Elton TS (2008)** Experimental validation of miRNA targets. *Methods*. 44(1): 47–54.

Kumar V, **Torben W**, **Kenway CS**, **Schiro FR** and **Mohan M (2016)** Longitudinal Examination of the Intestinal Lamina Propria Cellular Compartment of Simian Immunodeficiency Virus-Infected Rhesus Macaques Provides Broader and Deeper Insights into the Link between Aberrant MicroRNA Expression and Persistent Immune Activati. *J. Virol*. . 90(10): 5003–5019.

Kwa MS, **Kooyman FN**, **Boersema JH** and **Roos MH (1993)** Effect of selection for benzimidazole resistance in Haemonchus contortus on beta-tubulin isotype 1 and isotype 2 genes. *Biochem. Biophys. Res. Commun*. 191(2): 413–9.

Kwa MS, **Veenstra JG** and **Roos MH (1994)** Benzimidazole resistance in Haemonchus contortus is correlated with a conserved mutation at amino acid 200 in beta-tubulin isotype 1. *Mol. Biochem. Parasitol*. 63(2): 299–303.

Kwa MS, **Veenstra JG**, **Van Dijk M** and **Roos MH (1995)** Beta-tubulin genes from the parasitic nematode Haemonchus contortus modulate drug resistance in Caenorhabditis elegans. *J. Mol. Biol*. 246(4): 500–10.

de la Torre-Escudero E, **Manzano-Román R**, **Pérez-Sánchez R**, **Siles-Lucas M** and **Oleaga A (2010)** Cloning and characterization of a plasminogen-binding surface-associated enolase from Schistosoma bovis. *Vet. Parasitol*. 173(1–2): 76–84.

Lacroux C, Nguyen THC, Andreoletti O, Prevot F, Grisez C, Bergeaud J-P, Gruner L, Brunel J-C, Francois D & other authors (2006) Haemonchus contortus (Nematoda: Trichostrongylidae) infection in lambs elicits an unequivocal Th2 immune response. *Vet. Res*. 37(4): 607–22. editor & translator, EDP Sciences.

Laing R, Hunt M, Protasio A V, Saunders G, Mungall K, Laing S, Jackson F, Quail M, Beech

`

R & other authors (2011) Annotation of Two Large Contiguous Regions from the <italic>Haemonchus contortus</italic> Genome Using RNA-seq and Comparative Analysis with <italic>Caenorhabditis elegans</italic>. *PLoS One*. 6(8): e23216. editor & translator, Public Library of Science.

Laing R, Kikuchi T, Martinelli A, Tsai I, Beech R, Redman E, Holroyd N, Bartley D, Beasley H & other authors (2013) The genome and transcriptome of Haemonchus contortus, a key model parasite for drug and vaccine discovery. *Genome Biol*. 14(8): R88. editor & translator, BioMed Central Ltd.

Lambertz U, Oviedo Ovando ME, Vasconcelos EJR, Unrau PJP, Myler PJ, Reiner NE, Zandbergen G, Klinger M, Mueller A & other authors (2015) Small RNAs derived from tRNAs and rRNAs are highly enriched in exosomes from both old and new world Leishmania providing evidence for conserved exosomal RNA Packaging 16(1): 151. editor & translator, BioMed Central.

LaMonte G, Philip N, Reardon J, Lacsina JR, Majoros W, Chapman L, Thornburg CD, Telen MJ, Ohler U & other authors (2012) Translocation of sickle cell erythrocyte microRNAs into Plasmodium falciparum inhibits parasite translation and contributes to malaria resistance. *Cell Host Microbe*. 12(2): 187–99. editor & translator, NIH Public Access.

Landthaler M, **Yalcin A** and **Tuschl T (2004)** The Human DiGeorge Syndrome Critical Region Gene 8 and Its *D. melanogaster* Homolog Are Required for miRNA Biogenesis. *Curr. Biol.* 14(23): 2162–2167.

Lanford RE, **Hildebrandt-Eriksen ES**, **Petri A**, **Persson R**, **Lindow M**, **Munk ME**, **Kauppinen S** and **Ørum H (2010)** Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science*. 327(5962): 198–201.

Larsen PL (1993) Aging and resistance to oxidative damage in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* 90(19): 8905–9.

Lee LW, **Zhang S**, **Etheridge A**, **Ma L**, **Martin D**, **Galas D** and **Wang K (2010)** Complexity of the microRNA repertoire revealed by next-generation sequencing. *RNA*. 16(11): 2170–80.

Lee RC, **Feinbaum RL** and **Ambros V (1993a)** The C . elegans Heterochronic Gene lin-4 Encodes Small RNAs with Antisense Complementarity to & II-14. *Cell*. 75(5): 843–854.

Lee RC, **Feinbaum RL** and **Ambros V (1993b)** The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*. 75(5): 843–854.

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O & other authors (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature*. 425(6956): 415–419.

Lee Y, **Kim M**, **Han J**, **Yeom K-H**, **Lee S**, **Baek SH** and **Kim VN (2004)** MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 23(20): 4051–4060.

`

Leipe J, **Grunke M**, **Dechant C**, **Reindl C**, **Kerzendorf U**, **Schulze-Koops H** and **Skapenko A (2010)** Role of Th17 cells in human autoimmune arthritis. *Arthritis Rheum.* 62(10): 2876–85.

Lendner M, **Doligalska M**, **Lucius R** and **Hartmann S (2008)** Attempts to establish RNA interference in the parasitic nematode Heligmosomoides polygyrus. *Mol. Biochem. Parasitol.* 161(1): 21–31.

Lewis BP, **Shih I**, **Jones-Rhoades MW**, **Bartel DP** and **Burge CB (2003)** Prediction of mammalian microRNA targets. *Cell.* 115(7): 787–798.

Lewis BP, **Burge CB** and **Bartel DP (2005)** Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell.* 120(1): 15–20.

Li J, **Katiyar SK** and **Edlind TD (1996)** Site-directed mutagenesis of Saccharomyces cerevisiae beta-tubulin: interaction between residue 167 and benzimidazole compounds. *FEBS Lett.* 385(1–2): 7–10.

Li RW, **Hou Y**, **Li C** and **Gasbarre LC (2010)** Localized complement activation in the development of protective immunity against Ostertagia ostertagi infections in cattle. *Vet. Parasitol.* 174(3–4): 247–56.

Li Y-P, **Gottwein JM**, **Scheel TK**, **Jensen TB** and **Bukh J (2011)** MicroRNA-122 antagonism against hepatitis C virus genotypes 1-6 and reduced efficacy by host RNA insertion or mutations in the HCV 5' UTR. *Proc. Natl. Acad. Sci. U. S. A.* 108(12): 4991–6.

Li Y, **Masaki T**, **Yamane D**, **McGivern DR** and **Lemon SM (2013)** Competing and noncompeting activities of miR-122 and the 5' exonuclease Xrn1 in regulation of hepatitis C virus replication. *Proc. Natl. Acad. Sci. U. S. A.* 110(5): 1881–6.

Li Z and **Rana TM (2014)** Therapeutic targeting of microRNAs: current status and future challenges. *Nat. Rev. Drug Discov.* 13(8): 622–638. editor & translator, Nature Publishing Group.

Liang G, **Zhu Y**, **Sun B**, **Shao Y**, **Jing A**, **Wang J** and **Xiao Z (2014)** Assessing the survival of exogenous plant microRNA in mice. *Food Sci. Nutr.* 2(4): 380–8.

Lim LP, **Lau NC**, **Garrett-Engele P**, **Grimson A**, **Schelter JM**, **Castle J**, **Bartel DP**, **Linsley PS** and **Johnson JM (2005)** Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature.* 433(7027): 769–773.

Little PR, **Hodges A**, **Watson TG**, **Seed JA** and **Maeder SJ (2010)** Field efficacy and safety of an oral formulation of the novel combination anthelmintic, derquantel-abamectin, in sheep in New Zealand. *N. Z. Vet. J.* 58(3): 121–9.

Liu Q, **Liu Z**, **Rozo CT**, **Hamed HA**, **Alem F**, **Urban JF** and **Gause WC (2007)** The role of B cells in the development of CD4 effector T cells during a polarized Th2 immune response. *J. Immunol.* 179(6): 3821–30.

Liu S, da Cunha AP, Rezende RMM, Cialic R, Wei Z, Bry L, Comstock LEE, Gandhi R,

`

Weiner HLL & other authors (2016) The Host Shapes the Gut Microbiota via Fecal MicroRNA. *Cell Host Microbe*. 19(1): 32–43.

Lodish H, **Berk A**, **Zipursky SL**, **Matsudaira P**, **Baltimore D** and **Darnell J (2000)** *Processing of rRNA and tRNA*. editor & translator, W. H. Freeman.

López-Cabrera M, **Santis AG**, **Fernández-Ruiz E**, **Blacher R**, **Esch F**, **Sánchez-Mateos P** and **Sánchez-Madrid F (1993)** Molecular cloning, expression, and chromosomal localization of the human earliest lymphocyte activation antigen AIM/CD69, a new member of the C-type animal lectin superfamily of signal-transmitting receptors. *J. Exp. Med*. 178(2): 537–47.

Lu D, Davis MPA, Abreu-Goodger C, Wang W, Campos LS, Siede J, Vigorito E, Skarnes WC, Dunham I & other authors (2012) MiR-25 regulates Wwp2 and Fbxw7 and promotes reprogramming of mouse fibroblast cells to iPSCs. *PLoS One*. 7(8): e40938. editor & translator, Public Library of Science.

Lund E, **Güttinger S**, **Calado A**, **Dahlberg JE** and **Kutay U (2004)** Nuclear export of microRNA precursors. *Science*. 303(5654): 95–98.

Luo W, **Friedman MS**, **Shedden K**, **Hankenson KD** and **Woolf PJ (2009)** GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*. 10(1): 161. editor & translator, BioMed Central.

Lytle JR, **Yario TA** and **Steitz JA (2007)** Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. *Proc. Natl. Acad. Sci. U. S. A.* 104(23): 9667–72.

Mack GS (2007) MicroRNA gets down to business. *Nat. Biotechnol*. 25(6): 631–8. editor & translator, Nature Publishing Group.

Madden KB, **Yeung KA**, **Zhao A**, **Gause WC**, **Finkelman FD**, **Katona IM**, **Urban JF** and **Shea-Donohue T (2004)** Enteric Nematodes Induce Stereotypic STAT6-Dependent Alterations in Intestinal Epithelial Cell Function. *J. Immunol*. 172(9): 5616–5621. editor & translator, American Association of Immunologists.

Maizels RM and **Balic A (2004)** Resistance to helminth infection: the case for interleukin-5-dependent mechanisms. *J. Infect. Dis*. 190(3): 427–9.

Mangan PR, **Harrington LE**, **O'Quinn DB**, **Helms WS**, **Bullard DC**, **Elson CO**, **Hatton RD**, **Wahl SM**, **Schoeb TR** and **Weaver CT (2006)** Transforming growth factor-beta induces development of the T(H)17 lineage. *Nature*. 441(7090): 231–234.

Manoury B, **Gregory WF**, **Maizels RM** and **Watts C (2001)** Bm-CPI-2, a cystatin homolog secreted by the filarial parasite Brugia malayi, inhibits class II MHC-restricted antigen processing. *Curr. Biol*. 11(6): 447–51. editor & translator, Elsevier.

Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E & other authors (2009) DIANA-microT web server: elucidating microRNA functions through target prediction. *Nucleic Acids Res.*

`

37(Web Server issue): W273-6. editor & translator, Oxford University Press.

Marquitz AR, **Mathur A**, **Nam CS** and **Raab-Traub N (2011)** The Epstein-Barr Virus BART microRNAs target the pro-apoptotic protein Bim. *Virology*. 412(2): 392–400. editor & translator, NIH Public Access.

Martín P, Gómez M, Lamana A, Marín AM, Cortés JR, Ramírez-Huesca M, Barreiro O, López-Romero P, Gutiérrez-Vázquez C & other authors (2010) The leukocyte activation antigen CD69 limits allergic asthma and skin contact hypersensitivity. *J. Allergy Clin. Immunol*. 126(2): 355–365.e3.

Martin RJ and **Pennington AJ (1989)** A patch-clamp study of effects of dihydroavermectin on Ascaris muscle. *Br. J. Pharmacol*. 98(3): 747–756.

Martinez-Sanchez A and **Murphy CL (2013)** MicroRNA Target Identification-Experimental Approaches. *Biology (Basel)*. 2(1): 189–205. editor & translator, Multidisciplinary Digital Publishing Institute (MDPI).

Mathivanan S and **Simpson RJ (2009)** ExoCarta: A compendium of exosomal proteins and RNA. *Proteomics*. 9(21): 4997–5000. editor & translator, WILEY-VCH Verlag.

McCoy KD, Stoel M, Stettler R, Merky P, Fink K, Senn BM, Schaer C, Massacand J, Odermatt B & other authors (2008) Polyclonal and specific antibodies mediate protective immunity against enteric helminth infection. *Cell Host Microbe*. 4(4): 362–73. editor & translator, Elsevier.

McEwen JM and **Kaplan JM (2008)** UNC-18 promotes both the anterograde trafficking and synaptic function of syntaxin. *Mol. Biol. Cell*. 19(9): 3836–46.

McKerrow JH, **Caffrey C**, **Kelly B**, **Loke P** and **Sajid M (2006)** Proteases in parasitic diseases. *Annu. Rev. Pathol*. 1497–536. editor & translator, Annual Reviews.

McNeilly TN, **Devaney E** and **Matthews JB (2009)** Teladorsagia circumcincta in the sheep abomasum: defining the role of dendritic cells in T cell regulation and protective immunity. *Parasite Immunol*. 31(7): 347–56.

McNeilly TN and **Nisbet AJ (2014)** Immune modulation by helminth parasites of ruminants: implications for vaccine development and host immune competence. *Parasite*. 2151.

McSorley HJ, **Grainger JR**, **Harcus Y**, **Murray J**, **Nisbet AJ**, **Knox DP** and **Maizels RM (2010)** daf-7-related TGF-beta homologues from Trichostrongyloid nematodes show contrasting life-cycle expression patterns. *Parasitology*. 137(1): 159–71. editor & translator, Cambridge University Press.

Mearns H, **Horsnell WGC**, **Hoving JC**, **Dewals B**, **Cutler AJ**, **Kirstein F**, **Myburgh E**, **Arendse B** and **Brombacher F (2008)** Interleukin-4-promoted T helper 2 responses enhance Nippostrongylus brasiliensis-induced pulmonary pathology. *Infect. Immun.* 76(12): 5535–42.

Mederos AE, **Ramos Z** and **Banchero GE (2014)** First report of monepantel Haemonchus contortus resistance on sheep farms in Uruguay. *Parasit. Vectors*. 7(1): 598. editor &

`

translator, BioMed Central.

Meeusen ENT, **Balic A** and **Bowles V (2005)** Cells, cytokines and other molecules associated with rejection of gastrointestinal nematode parasites. *Vet. Immunol. Immunopathol.* 108(1): 121–125.

Meng F, **Henson R**, **Wehbe-Janek H**, **Ghoshal K**, **Jacob ST** and **Patel T (2007)** MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. *Gastroenterology*. 133(2): 647–58. editor & translator, Elsevier.

Middleton J, Patterson AAM, Gardner L, Schmutz C, Ashton BA, Rot A, Hub E, Middleton J, Baggiolini M & other authors (2002) Leukocyte extravasation: chemokine transport and presentation by the endothelium. *Blood*. 100(12): 3853–60. editor & translator, American Society of Hematology.

Mihi B, **Van Meulder F**, **Rinaldi M**, **Van Coppernolle S**, **Chiers K**, **Van den Broeck W**, **Goddeeris B**, **Vercruysse J**, **Claerebout E** and **Geldhof P (2013)** Analysis of cell hyperplasia and parietal cell dysfunction induced by Ostertagia ostertagi infection. *Vet. Res.* 44(1): 121. editor & translator, BioMed Central.

Miller JE and **Horohov DW (2006)** Immunological aspects of nematode parasite control in sheep. *J. Anim. Sci.* 84 SupplE124-32.

Min B, Prout M, Hu-Li J, Zhu J, Jankovic D, Morgan ES, Urban JF, Dvorak AM, Finkelman FD & other authors (2004) Basophils produce IL-4 and accumulate in tissues after infection with a Th2-inducing parasite. *J. Exp. Med.* 200(4): 507–17.

Min H and **Yoon S (2010)** Got target? Computational methods for microRNA target prediction and their extension. *Exp. Mol. Med.* 42(4): 233–44. editor & translator, Nature Publishing Group.

Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC & other authors (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. U. S. A.* 105(30): 10513–10518.

Mittal S, **Aslam A**, **Doidge R**, **Medica R** and **Winkler GS (2011)** The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence. *Mol. Biol. Cell*. 22(6): 748–58.

Molehin AJ, **Gobert GN** and **McManus DP (2012)** Serine protease inhibitors of parasitic helminths. *Parasitology*. 139(6): 681–95. editor & translator, Cambridge University Press.

Mori N (2011) Induction of CD69 expression by *cag* PAI-positive *Helicobacter pylori* infection. *World J. Gastroenterol.* 17(32): 3691.

Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu A-L, Zhao Y, McDonald H, Zeng T & other authors (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells.

`

Genome Res. 18(4): 610–21.

Morlan JD, **Qu K** and **Sinicropi D V (2012)** Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. *PLoS One*. 7(8): e42882. editor & translator, Public Library of Science.

Mountford AP, **Anderson S** and **Wilson RA (1996)** Induction of Th1 cell-mediated protective immunity to Schistosoma mansoni by co-administration of larval antigens and IL-12 as an adjuvant. *J. Immunol*. 156(12): 4739–45. editor & translator, American Association of Immunologists.

Mourelatos Z, **Dostie J**, **Paushkin S**, **Sharma A**, **Charroux B**, **Abel L**, **Rappsilber J**, **Mann M** and **Dreyfuss G (2002)** miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev*. 16(6): 720–728.

Munn EA, **Smith TS**, **Graham M**, **Greenwood CA**, **Tavernor AS** and **Coetzee G (1993)** Vaccination of merino lambs against haemonchosis with membrane-associated proteins from the adult parasite. *Parasitology*. 106 (Pt 163–6.

Murakami Y, **Aly HH**, **Tajima A**, **Inoue I** and **Shimotohno K (2009)** Regulation of the hepatitis C virus genome replication by miR-199a. *J. Hepatol*. 50(3): 453–60.

Murphy L, **Eckersall PD**, **Bishop SC**, **Pettit JJ**, **Huntley JF**, **Burchmore R** and **Stear MJ (2010)** Genetic variation among lambs in peripheral IgE activity against the larval stages of Teladorsagia circumcincta. *Parasitology*. 137(8): 1249–60. editor & translator, Cambridge University Press.

Murri S, **Knubben-Schweizer G**, **Torgerson P** and **Hertzberg H (2014)** Frequency of eprinomectin resistance in gastrointestinal nematodes of goats in canton Berne, Switzerland. *Vet. Parasitol*. 203(1–2): 114–9.

Naguibneva I, **Ameyar-Zazoua M**, **Polesskaya A**, **Ait-Si-Ali S**, **Groisman R**, **Souidi M**, **Cuvellier S** and **Harel-Bellan A (2006)** The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. *Nat. Cell Biol*. 8(3): 278–84.

Nakae S, **Iwakura Y**, **Suto H** and **Galli SJ (2007)** Phenotypic differences between Th1 and Th17 cells and negative regulation of Th1 cell differentiation by IL-17. *J. Leukoc. Biol*. 81(5): 1258–68.

Nasheri N, **Singaravelu R**, **Goodmurphy M**, **Lyn RK** and **Pezacki JP (2011)** Competing roles of microRNA-122 recognition elements in hepatitis C virus RNA. *Virology*. 410(2): 336–44.

Neilson JR, **Zheng GXY**, **Burge CB** and **Sharp PA (2007)** Dynamic regulation of miRNA expression in ordered stages of cellular development. *Genes Dev*. 21(5): 578–89. editor & translator, Cold Spring Harbor Laboratory Press.

Neveu C, **Charvet CL**, **Fauvin A**, **Cortet J**, **Beech RN** and **Cabaret J (2010)** Genetic diversity of levamisole receptor subunits in parasitic nematode species and

`

abbreviated transcripts associated with resistance. *Pharmacogenet. Genomics*. 20(7): 414–25.

Newton S. and **Munn E. (1999)** The Development of Vaccines against Gastrointestinal Nematode Parasites, Particularly Haemonchus contortus. *Parasitol. Today*. 15(3): 116–122. editor & translator, Elsevier.

Nieuwhof GJ and **Bishop SC (2005)** Costs of the major endemic diseases of sheep in Great Britain and the potential benefits of reduction in disease impact. *Anim. Sci*. 81(1): 23–29.

Nisbet AJ, **Bell NE V**, **McNeilly TN**, **Knox DP**, **Maizels RM**, **Meikle LI**, **Wildblood LA** and **Matthews JB (2010)** A macrophage migration inhibitory factor-like tautomerase from Teladorsagia circumcincta (Nematoda: Strongylida). *Parasite Immunol*. 32(7): 503–11.

Nisbet AJ, **McNeilly TN**, **Wildblood LA**, **Morrison AA**, **Bartley DJ**, **Bartley Y**, **Longhi C**, **McKendrick IJ**, **Palarea-Albaladejo J** and **Matthews JB (2013)** Successful immunization against a parasitic nematode by vaccination with recombinant proteins. *Vaccine*. 31(37): 4017–23.

Novina CD and **Sharp PA (2004)** The RNAi revolution. *Nature*. 430(6996): 161–4. editor & translator, Nature Publishing Group.

Nowacki FC, **Swain MT**, **Klychnikov OI**, **Niazi U**, **Ivens A**, **Quintana JF**, **Hensbergen PJ**, **Hokke CH**, **Buck AH** and **Hoffmann KF (2015, October 5)** Protein and small non-coding RNA-enriched extracellular vesicles are released by the pathogenic blood fluke Schistosoma mansoni. *J. Extracell. Vesicles*.

Nygren MK, Tekle C, Ingebrigtsen VA, Mäkelä R, Krohn M, Aure MR, Nunes-Xavier CE, Perälä M, Tramm T & other authors (2014) Identifying microRNAs regulating B7-H3 in breast cancer: the clinical impact of microRNA-29c. *Br. J. Cancer*. 110(8): 2072–80. editor & translator, Cancer Research UK.

O'Sullivan BM and **Donald AD (1973)** Responses to infection with Haemonchus contortus and Trichostrongylus colubriformis in ewes of different reproductive status. *Int. J. Parasitol*. 3(4): 521–530.

O'Sullivan BM and **Donald AD (1970)** A field study of nematode parasite populations in the lactating ewe. *Parasitology*. 61(2): 301–15.

Okano M, **Satoskar AR**, **Nishizaki K**, **Abe M** and **Harn DA (1999)** Induction of Th2 responses and IgE is largely due to carbohydrates functioning as adjuvants on Schistosoma mansoni egg antigens. *J. Immunol*. 163(12): 6712–7.

Onah DN, **Uchiyama F**, **Nagakui Y**, **Ono M**, **Takai T** and **Nawa Y (2000)** Mucosal defense against gastrointestinal nematodes: Responses of mucosal mast cells and mouse mast cell protease 1 during primary Strongyloides venezuelensis infection in FcR??-knockout mice. *Infect. Immun.* 68(9): 4968–4971. editor & translator, American Society for Microbiology (ASM).

`

Page AP, **Stepek G**, **Winter AD** and **Pertab D (2014)** Enzymology of the nematode cuticle: A potential drug target? *Int. J. Parasitol. Drugs Drug Resist.* 4(2): 133–141.

Palmieri F (2013) The mitochondrial transporter family SLC25: Identification, properties and physiopathology. *Mol. Aspects Med.* 34(2): 465–484.

Paraskevopoulou MD, **Georgakilas G**, **Kostoulas N**, **Vlachos IS**, **Vergoulis T**, **Reczko M**, **Filippidis C**, **Dalamagas T** and **Hatzigeorgiou AG (2013)** DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Res.* 41(Web Server issue): W169-73. editor & translator, Oxford University Press.

Park H, Li Z, Yang XO, Chang SH, Nurieva R, Wang YY-H, Wang YY-H, Hood L, Zhu Z & other authors (2005) A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nat. Immunol.* 6(11): 1133–41.

Parker JS and **Barford D (2006)** Argonaute: A scaffold for the function of short regulatory RNAs. *Trends Biochem. Sci.* 31(11): 622–630.

Parker JS, **Roe SM** and **Barford D (2005)** Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature.* 434(7033): 663–666.

Pauley KM, **Cha S** and **Chan EKL** MicroRNA in autoimmunity and autoimmune diseases. *J. Autoimmun.* 32(3–4): 189–94.

Pearce EJ and **MacDonald AS (2002)** The immunobiology of schistosomiasis. *Nat. Rev. Immunol.* 2(7): 499–511.

Pérez J, **Garcı́a PM**, **Hernández S**, **Mozos E**, **Cámara S** and **Martı́nez-Moreno A (2003)** Experimental haemonchosis in goats: effects of single and multiple infections in the host response. *Vet. Parasitol.* 111(4): 333–342.

Pernthaner A, **Stasiuk SJ**, **Roberts JM** and **Sutherland IA (2012)** The response of monocyte derived dendritic cells following exposure to a nematode larval carbohydrate antigen. *Vet. Immunol. Immunopathol.* 148(3–4): 284–92.

Peters L and **Meister G (2007)** Argonaute Proteins: Mediators of RNA Silencing. *Mol. Cell.* 26(5): 611–623. editor & translator, Elsevier.

Petersen SH, **Odintsova E**, **Haigh TA**, **Rickinson AB**, **Taylor GS** and **Berditchevski F (2011)** The role of tetraspanin CD63 in antigen presentation via MHC class II. *Eur. J. Immunol.* 41(9): 2556–2561. editor & translator, WILEY-VCH Verlag.

Peterson SM, **Thompson J a.**, **Ufkin ML**, **Sathyanarayana P**, **Liaw L** and **Congdon CB (2014)** Common features of microRNA target prediction tools. *Front. Genet.* 5(FEB): 1–10. editor & translator, Frontiers.

Pettit JJ, **Jackson F**, **Rocchi M** and **Huntley JF (2005)** The relationship between responsiveness against gastrointestinal nematodes in lambs and the numbers of circulating IgE-bearing cells. *Vet. Parasitol.* 134(1–2): 131–9.

Pignatti P, **Perfetti L**, **Galdi E**, **Pozzi V**, **Bossi A**, **Biale C** and **Moscato G (2002)** Increased

`

CD69 expression on peripheral blood eosinophils after specific inhalation challenge. *Allergy*. 57(5): 411–416. editor & translator, Munksgaard International Publishers.

Pillai R (2005) MicroRNA function: Multiple mechanisms for a tiny RNA? *Rna*. 11(12): 1753–1761.

Pleass RJ, **Kusel JR** and **Woof JM (2000)** Cleavage of human IgE mediated by Schistosoma mansoni. *Int. Arch. Allergy Immunol*. 121(3): 194–204. editor & translator, Karger Publishers.

Prichard RK (1990) Anthelmintic resistance in nematodes: Extent, recent understanding and future directions for control and research. *Int. J. Parasitol*. 20(4): 515–523.

Prichard R (2001) Genetic variability following selection of Haemonchus contortus with anthelmintics. *Trends Parasitol*. 17(9): 445–453.

Prieto-Lafuente L, **Gregory WF**, **Allen JE** and **Maizels RM (2009)** MIF homologues from a filarial nematode parasite synergize with IL-4 to induce alternative activation of host macrophages. *J. Leukoc. Biol*. 85(5): 844–54.

Puttachary S, **Trailovic SM**, **Robertson AP**, **Thompson DP**, **Woods DJ** and **Martin RJ (2013)** Derquantel and abamectin: effects and interactions on isolated tissues of *Ascaris suum*. *Mol. Biochem. Parasitol*. 188(2): 79–86.

Quintana JF, Makepeace BL, Babayan SA, Ivens A, Pfarr KM, Blaxter M, Debrah A, Wanji S, Ngangyung HF & other authors (2015) Extracellular Onchocerca-derived small RNAs in host nodules and blood. *Parasit. Vectors*. 8(1): 58. editor & translator, BioMed Central Ltd.

Radulovic K, **Manta C**, **Rossini V**, **Holzmann K**, **Kestler HA**, **Wegenka UM**, **Nakayama T** and **Niess JH (2012)** CD69 Regulates Type I IFN-Induced Tolerogenic Signals to Mucosal CD4 T Cells That Attenuate Their Colitogenic Potential. *J. Immunol*. 188(4): 2001–2013.

Radulovic K and **Niess JH (2015)** CD69 Is the Crucial Regulator of Intestinal Inflammation: A New Target Molecule for IBD Treatment? *J. Immunol. Res*. 20151–12.

Radulovic K, **Rossini V**, **Manta C**, **Holzmann K**, **Kestler HA** and **Niess JH (2013)** The Early Activation Marker CD69 Regulates the Expression of Chemokines and CD4 T Cell Accumulation in Intestine. *PLoS One*. 8(6): e65413 (B. Siegmund, Ed.).

Raposo G and **Stoorvogel W (2013a)** Extracellular vesicles: Exosomes, microvesicles, and friends. *J. Cell Biol*. 200(4): .

Raposo G and **Stoorvogel W (2013b)** Extracellular vesicles: Exosomes, microvesicles, and friends. *J. Cell Biol*. 200(4): 373–383. editor & translator, Rockefeller University Press.

Rathore DK, **Suchitra S**, **Saini M**, **Singh BP** and **Joshi P (2006)** Identification of a 66 kDa Haemonchus contortus excretory/secretory antigen that inhibits host monocytes. *Vet. Parasitol*. 138(3–4): 291–300.

Redman E, **Packard E**, **Grillo V**, **Smith J**, **Jackson F** and **Gilleard JS (2008)** Microsatellite

`

analysis reveals marked genetic differentiation between Haemonchus contortus laboratory isolates and provides a rapid system of genetic fingerprinting. *Int. J. Parasitol.* 38(1): 111–22.

Redmond DL and **Knox DP (2006)** Further protection studies using recombinant forms of Haemonchus contortus cysteine proteinases. *Parasite Immunol.* 28(5): 213–9.

Redmond DL and **Knox DP (2004)** Protection studies in sheep using affinity-purified and recombinant cysteine proteinases of adult Haemonchus contortus. *Vaccine.* 22(31–32): 4252–61.

Rehman ZU, **Knight JS**, **Koolaard J**, **Simpson H V** and **Pernthaner A (2015)** Immunomodulatory effects of adult Haemonchus contortus excretory/secretory products on human monocyte-derived dendritic cells. *Parasite Immunol.* 37(12): 657–69.

Reiner SL (2005) Epigenetic control in the immune response. *Hum. Mol. Genet.* 14 Spec No(suppl_1): R41-6.

Reinhart BJ, **Slack FJ**, **Basson M**, **Pasquinelli a E**, **Bettinger JC**, **Rougvie a E**, **Horvitz HR** and **Ruvkun G (2000)** The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature.* 403(6772): 901–906.

Ren Z and **Ambros VR (2015)** Caenorhabditis elegans microRNAs of the let-7 family act in innate immune response circuits and confer robust developmental timing against pathogen stress. *Proc. Natl. Acad. Sci. U. S. A.* 112(18): E2366-75.

Reszka N, **Rijsewijk FAM**, **Zelnik V**, **Moskwa B** and **Bieńkowska-Szewczyk K (2007)** Haemonchus contortus: characterization of the baculovirus expressed form of aminopeptidase H11. *Exp. Parasitol.* 117(2): 208–13.

Reyes JL and **Terrazas LI (2007)** The divergent roles of alternatively activated macrophages in helminthic infections. *Parasite Immunol.* 29(12): 609–19.

Reynolds LA, **Filbey KJ** and **Maizels RM (2012)** Immunity to the model intestinal helminth parasite Heligmosomoides polygyrus. *Semin. Immunopathol.*

El Ridi R and **Tallima H (2012)** Adjuvant selection for vaccination against murine schistosomiasis. *Scand. J. Immunol.* 76(6): 552–8.

Riley KJ, **Rabinowitz GS**, **Yario TA**, **Luna JM**, **Darnell RB** and **Steitz JA (2012)** EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *EMBO J.* 31(9): 2207–21. editor & translator, European Molecular Biology Organization.

Roberts B, **Antonopoulos A**, **Haslam SM**, **Dicker AJ**, **McNeilly TN**, **Johnston SL**, **Dell A**, **Knox DP** and **Britton C (2013)** Novel expression of Haemonchus contortus vaccine candidate aminopeptidase H11 using the free-living nematode Caenorhabditis elegans. *Vet. Res.* 44(1): 111. editor & translator, BioMed Central Ltd.

Rodriguez A, **Griffiths-Jones S**, **Ashurst JL** and **Bradley A (2004)** Identification of mammalian microRNA host genes and transcription units. *Genome Res.* 14(10 A): 1902–

`

10.

Roeber F, Jex AAR, Gasser RBRRB, McLeod R, Sackett D, Holmes P, Gilleard J, Kaplan R, Wolstenholme A & other authors (2013) Impact of gastrointestinal parasitic nematodes of sheep, and the role of advanced molecular tools for exploring epidemiology and drug resistance - an Australian perspective. *Parasit. Vectors*. 6(1): 153. editor & translator, BioMed Central.

Rook GAW (2009) Review series on helminths, immune modulation and the hygiene hypothesis: the broader implications of the hygiene hypothesis. *Immunology*. 126(1): 3–11.

Rottiers V and **Näär AM (2012)** MicroRNAs in metabolism and metabolic disorders. *Nat. Rev. Mol. Cell Biol.* 13(4): 239–50. editor & translator, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.

Rutherford K, **Parkhill J**, **Crook J**, **Horsnell T**, **Rice P**, **Rajandream M-A** and **Barrell B (2000)** Artemis: sequence visualization and annotation. *Bioinformatics*. 16(10): 944–945.

Rutitzky LI, **Lopes da Rosa JR** and **Stadecker MJ (2005)** Severe CD4 T cell-mediated immunopathology in murine schistosomiasis is dependent on IL-12p40 and correlates with high levels of IL-17. *J. Immunol.* 175(6): 3920–6.

Sahoo S, **Murugavel S**, **Devi IK**, **Vedamurthy G V**, **Gupta SC**, **Singh BP** and **Joshi P (2013)** Glyceraldehyde-3-phosphate dehydrogenase of the parasitic nematode Haemonchus contortus binds to complement C3 and inhibits its activity. *Parasite Immunol.* 35(12): 457–67.

Samarasinghe B, **Knox DP** and **Britton C (2011)** Factors affecting susceptibility to RNA interference in Haemonchus contortus and in vivo silencing of an H11 aminopeptidase gene. *Int. J. Parasitol.* 41(1): 51–9.

Sánchez-Mateos P and **Sánchez-Madrid F (1991)** Structure-function relationship and immunochemical mapping of external and intracellular antigenic sites on the lymphocyte activation inducer molecule, AIM/CD69. *Eur. J. Immunol.* 21(10): 2317–25.

Sancho D, **Gómez M**, **Viedma F**, **Esplugues E**, **Gordón-Alonso M**, **Angeles García-López M**, **de la Fuente H**, **Martínez-A C**, **Lauzurica P** and **Sánchez-Madrid F (2003)** CD69 downregulates autoimmune reactivity through active transforming growth factor-β production in collagen-induced arthritis. *J. Clin. Invest.* 112(6): 872–882.

Sargison ND, **Jackson F**, **Bartley DJ**, **Wilson DJ**, **Stenhouse LJ** and **Penny CD (2007)** Observations on the emergence of multiple anthelmintic resistance in sheep flocks in the south-east of Scotland. *Vet. Parasitol.* 145(1–2): 65–76.

Sarkies P, Selkirk ME, Jones JT, Blok V, Boothby T, Goldstein B, Hanelt B, Ardila-Garcia A, Fast NM & other authors (2015) Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. *PLoS Biol.* 13(2):

`

e1002061. editor & translator, Public Library of Science.

Sayed D and **Abdellatif M (2011)** MicroRNAs in Development and Disease. *Physiol. Rev.* 91(3): 827–887.

Schallig HD f. h., **Van Leeuwen MA w.** and **Cornelissen AW c. a. (1997)** Protective immunity induced by vaccination with two Haemonchus contortus excretory secretory proteins in sheep. *Parasite Immunol.* 19(10): 447–453.

Schierack P, **Lucius R**, **Sonnenburg B**, **Schilling K** and **Hartmann S (2003)** Parasite-specific immunomodulatory functions of filarial cystatin. *Infect. Immun.* 71(5): 2422–9.

Schmeier S, **MacPherson CR**, **Essack M**, **Kaur M**, **Schaefer U**, **Suzuki H**, **Hayashizaki Y** and **Bajic VB (2009)** Deciphering the transcriptional circuitry of microRNA genes expressed during human monocytic differentiation. *BMC Genomics.* 10(1): 595.

Schmitz J, Owyang A, Oldham E, Song Y, Murphy E, McClanahan TK, Zurawski G, Moshrefi M, Qin J & other authors (2005) IL-33, an interleukin-1-like cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokines. *Immunity.* 23(5): 479–90. editor & translator, Elsevier.

Schwarz EM, Korhonen PK, Campbell BE, Young ND, Jex AR, Jabbar A, Hall RS, Mondal A, Howe AC & other authors (2013) The genome and developmental transcriptome of the strongylid nematode Haemonchus contortus. *Genome Biol.* 14(8): R89.

Segura M, **Su Z**, **Piccirillo C** and **Stevenson MM (2007)** Impairment of dendritic cell function by excretory-secretory products: a potential mechanism for nematode-induced immunosuppression. *Eur. J. Immunol.* 37(7): 1887–904.

Selbach M, **Schwanhäusser B**, **Thierfelder N**, **Fang Z**, **Khanin R** and **Rajewsky N (2008)** Widespread changes in protein synthesis induced by microRNAs. *Nature.* 455(7209): 58–63.

Shakya KP, **Miller JE** and **Horohov DW (2009)** A Th2 type of immune response is associated with increased resistance to Haemonchus contortus in naturally infected Gulf Coast Native lambs. *Vet. Parasitol.* 163(1–2): 57–66.

Shi Z, **Montgomery TA**, **Qi Y** and **Ruvkun G (2013)** High-throughput sequencing reveals extraordinary fluidity of miRNA, piRNA, and siRNA pathways in nematodes. *Genome Res.* 23(3): 497–508.

Shrestha S, **Yang K**, **Guy C**, **Vogel P**, **Neale G** and **Chi H (2015)** Treg cells require the phosphatase PTEN to restrain TH1 and TFH cell responses. *Nat. Immunol.* 16(2): 178–87.

Skalsky RL, Corcoran DL, Gottwein E, Frank CL, Kang D, Hafner M, Nusbaum JD, Feederle R, Delecluse H-J & other authors (2012) The viral and cellular microRNA targetome in lymphoblastoid cell lines. *PLoS Pathog.* 8(1): e1002484. editor & translator, Public Library of Science.

Skuce PJ, **Stewart EM**, **Smith WD** and **Knox DP (1999)** Cloning and characterization of glutamate dehydrogenase (GDH) from the gut of Haemonchus contortus. *Parasitology.*

`

118 (Pt 3297–304.

Skuce PJ, **Newlands GF.**, **Stewart EM**, **Pettit D**, **Smith SK**, **Smith WD** and **Knox DP (2001)** Cloning and characterisation of thrombospondin, a novel multidomain glycoprotein found in association with a host protective gut extract from Haemonchus contortus. *Mol. Biochem. Parasitol.* 117(2): 241–244.

Smith SK, **Pettit D**, **Newlands GFJ**, **Redmond DL**, **Skuce PJ**, **Knox DP** and **Smith WD (1999)** Further immunization and biochemical studies with a protective antigen complex from the microvillar membrane of the intestine of Haemonchus contortus. *Parasite Immunol.* 21(4): 187–199.

Smith WD and **Zarlenga DS (2006)** Developments and hurdles in generating vaccines for controlling helminth parasites of grazing ruminants. *Vet. Parasitol.* 139(4): 347–359.

Smith WD, **Newlands GFJ**, **Smith SK**, **Pettit D** and **Skuce PJ (2003)** Metalloendopeptidases from the intestinal brush border of Haemonchus contortus as protective antigens for sheep. *Parasite Immunol.* 25(6): 313–323.

Smith WD (1999) Prospects for vaccines of helminth parasites of grazing ruminants. *Int. J. Parasitol.* 29(1): 17–24.

Smith WD and **Angus KW (1980)** Haemonchus contortus: attempts to immunise lambs with irradiated larvae. *Res. Vet. Sci.* 29(1): 45–50.

Smith WD, **Smith SK** and **Murray JM (1994)** Protection studies with integral membrane fractions of Haemonchus contortus. *Parasite Immunol.* 16(5): 231–41.

Snow JW, **Hale AE**, **Isaacs SK**, **Baggish AL** and **Chan SY (2013)** Ineffective delivery of diet-derived microRNAs to recipient animal organisms. *RNA Biol.* 10(7): 1107–16. editor & translator, Landes Bioscience.

Song J-J and **Joshua-Tor L (2006)** Argonaute and RNA — getting into the groove. *Curr. Opin. Struct. Biol.* 16(1): 5–11.

Soond DR, **Garçon F**, **Patton DT**, **Rolf J**, **Turner M**, **Scudamore C**, **Garden OA** and **Okkenhaug K (2012)** Pten loss in CD4 T cells enhances their helper function but does not lead to autoimmunity or lymphoma. *J. Immunol.* 188(12): 5935–43.

Sotillo J, **Pearson M**, **Potriquet J**, **Becker L**, **Pickering D**, **Mulvenna J** and **Loukas A (2016)** *Extracellular vesicles secreted by Schistosoma mansoni contain protein vaccine candidates. Int. J. Parasitol.*

Stanke M, **Keller O**, **Gunduz I**, **Hayes A**, **Waack S** and **Morgenstern B (2006)** AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 34(Web Server issue): W435-9.

Stanley P and **Stein PE (2003)** BmSPN2, a serpin secreted by the filarial nematode Brugia malayi, does not inhibit human neutrophil proteinases but plays a noninhibitory role. *Biochemistry.* 42(20): 6241–8.

`

Stanton H, Rogerson FM, East CJ, Golub SB, Lawlor KE, Meeker CT, Little CB, Last K, Farmer PJ & other authors (2005) ADAMTS5 is the major aggrecanase in mouse cartilage in vivo and in vitro. *Nature*. 434(7033): 648–52.

Starega-Roslan J, **Krol J**, **Koscianska E**, **Kozlowski P**, **Szlachcic WJ**, **Sobczak K** and **Krzyzosiak WJ (2011)** Structural basis of microRNA length variety. *Nucleic Acids Res.* 39(1): 257–68.

Steinhauer J and **Kalderon D (2005)** The RNA-binding protein Squid is required for the establishment of anteroposterior polarity in the Drosophila oocyte. *Development*. 132(24): 5515–25. editor & translator, The Company of Biologists Ltd.

Suarez VH, **Cristel SL** and **Busetti MR (2009)** Epidemiology and effects of gastrointestinal nematode infection on milk productions of dairy ewes. *Parasite*. 16(2): 141–7.

Suchitra S, **Anbu KA**, **Rathore DK**, **Mahawar M**, **Singh BP** and **Joshi P (2008)** Haemonchus contortus calreticulin binds to C-reactive protein of its host, a novel survival strategy of the parasite. *Parasite Immunol.* 30(6–7): 371–4.

Suchitra S and **Joshi P (2005)** Characterization of Haemonchus contortus calreticulin suggests its role in feeding and immune evasion by the parasite. *Biochim. Biophys. Acta*. 1722(3): 293–303.

Sunderkötter C, **Steinbrink K**, **Goebeler M**, **Bhardwaj R** and **Sorg C (1994)** Macrophages and angiogenesis. *J. Leukoc. Biol.* 55(3): 410–22.

Tabara H (1998) REVERSE GENETICS:RNAi in C. elegans: Soaking in the Genome Sequence. *Science (80-.).* 282(5388): 430–431. editor & translator, American Association for the Advancement of Science.

Tang Q and **Bluestone JA (2008)** The Foxp3+ regulatory T cell: a jack of all trades, master of regulation. *Nat. Immunol.* 9(3): 239–44.

Tang YT, Gao X, Rosa BA, Abubucker S, Hallsworth-Pepin K, Martin J, Tyagi R, Heizer E, Zhang X & other authors (2014) Genome of the human hookworm Necator americanus. *Nat. Genet.* 46(3): 261–9.

Testi R, **D'Ambrosio D**, **De Maria R** and **Santoni A (1994)** The CD69 receptor: a multipurpose cell-surface trigger for hematopoietic cells. *Immunol. Today*. 15(10): 479–83.

Théry C, **Boussac M**, **Véron P**, **Ricciardi-Castagnoli P**, **Raposo G**, **Garin J** and **Amigorena S (2001)** Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles. *J. Immunol.* 166(12): 7309–18.

Thomson DW, **Bracken CP** and **Goodall GJ (2011)** Experimental strategies for microRNA target identification. *Nucleic Acids Res.* 39(16): 6845–53. editor & translator, Oxford University Press.

Timmons L and **Fire A (1998)** Specific interference by ingested dsRNA. *Nature*. 395(6705): 854. editor & translator, Macmillan Magazines Ltd.

`

Trapnell C, **Pachter L** and **Salzberg SL (2009)** TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 25(9): 1105–11.

Trapnell C, **Hendrickson DG**, **Sauvageau M**, **Goff L**, **Rinn JL** and **Pachter L (2013)** Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol*. 31(1): 46–53.

Tritten L, **Burkman E**, **Moorhead A**, **Satti M**, **Geary J**, **Mackenzie C** and **Geary T (2014)** Detection of circulating parasite-derived microRNAs in filarial infections. *PLoS Negl. Trop. Dis.* 8(7): e2971. editor & translator, Public Library of Science.

Tsai W-C, Hsu S-D, Hsu C-S, Lai T-C, Chen S-J, Shen R, Huang Y, Chen H-C, Lee C-H & other authors (2012) MicroRNA-122 plays a critical role in liver homeostasis and hepatocarcinogenesis. *J. Clin. Invest*. 122(8): 2884–97. editor & translator, American Society for Clinical Investigation.

Tsai Y-H, VanDussen KL, Sawey ET, Wade AW, Kasper C, Rakshit S, Bhatt RG, Stoeck A, Maillard I & other authors (2014) ADAM10 regulates Notch function in intestinal stem cells of mice. *Gastroenterology*. 147(4): 822–834.e13. editor & translator, NIH Public Access.

Turchinovich A, **Samatov TR**, **Tonevitsky AG** and **Burwinkel B (2013)** Circulating miRNAs: cell–cell communication function? *Front. Genet*. 4119.

Turchinovich A, **Weiz L**, **Langheinz A** and **Burwinkel B (2011)** Characterization of extracellular circulating microRNA. *Nucleic Acids Res.* 39(16): 7223–7233. editor & translator, Oxford University Press.

Tyagi S, **Vaz C**, **Gupta V**, **Bhatia R**, **Maheshwari S**, **Srinivasan A** and **Bhattacharya A (2008)** CID-miRNA: A web server for prediction of novel miRNA precursors in human genome. *Biochem. Biophys. Res. Commun.* 372(4): 831–834.

Urban JF, **Katona IM** and **Finkelman FD (1991a)** Heligmosomoides polygyrus: CD4+ but not CD8+ T cells regulate the IgE response and protective immunity in mice. *Exp. Parasitol*. 73(4): 500–11.

Urban JF, **Katona IM**, **Paul WE** and **Finkelman FD (1991b)** Interleukin 4 is important in protective immunity to a gastrointestinal nematode infection in mice. *Proc. Natl. Acad. Sci. U. S. A.* 88(13): 5513–7.

Urban JF, **Maliszewski CR**, **Madden KB**, **Katona IM** and **Finkelman FD (1995)** IL-4 treatment can cure established gastrointestinal nematode infections in immunocompetent and immunodeficient mice. *J. Immunol*. 154(9): 4675–84.

Urquhart GM, **J A**, **Duncan JL**, **Dunn AM**, **W JF**, **Armour J**, **Duncam JL**, **Dunn AM** and **W. JF (1996)** *Veterinary parasitology*, 2nd ed. editor & translator, Essex: Cambridge, Mass. : Blackwell Science, 1996.

Valadi H, **Ekström K**, **Bossios A**, **Sjöstrand M**, **Lee JJ** and **Lötvall JO (2007)** Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange

`

between cells. *Nat. Cell Biol.* 9(6): 654–659. editor & translator, Nature Publishing Group.

Vance BA, **Wu W**, **Ribaudo RK**, **Segal DM** and **Kearse KP (1997)** Multiple Dimeric Forms of Human CD69 Result from Differential Addition of N-Glycans to Typical (Asn-X-Ser/Thr) and Atypical (Asn-X-Cys) Glycosylation Motifs. *J. Biol. Chem.* 272(37): 23117–23122.

Vanfleteren JR and **De Vreese A (1995)** The gerontogenes age-1 and daf-2 determine metabolic rate potential in aging Caenorhabditis elegans. *FASEB J.* 9(13): 1355–61.

Vega-Ramos J, **Alari-Pahissa E**, **Valle J del**, **Carrasco-Marín E**, **Esplugues E**, **Borràs M**, **Martínez-A C** and **Lauzurica P (2010)** CD69 limits early inflammatory diseases associated with immune response to Listeria monocytogenes infection. *Immunol. Cell Biol.* 88(7): 707–715. editor & translator, Nature Publishing Group.

Veldhoen M, **Hocking RJ**, **Atkins CJ**, **Locksley RM** and **Stockinger B (2006)** TGFbeta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. *Immunity.* 24(2): 179–89. editor & translator, Elsevier.

Veldhoen M, **Hirota K**, **Westendorf AM**, **Buer J**, **Dumoutier L**, **Renauld J-C** and **Stockinger B (2008)** The aryl hydrocarbon receptor links TH17-cell-mediated autoimmunity to environmental toxins. *Nature.* 453(7191): 106–9.

Vercruysse J, **Knox DP**, **Schetters TPM** and **Willadsen P (2004)** Veterinary parasitic vaccines: Pitfalls and future directions. *Trends Parasitol.* 20(10): 488–492.

Verheul HM, **Hoekman K**, **Luykx-de Bakker S**, **Eekman CA**, **Folman CC**, **Broxterman HJ** and **Pinedo HM (1997)** Platelet: transporter of vascular endothelial growth factor. *Clin. Cancer Res.* 3(12 Pt 1): 2187–90.

Vermeire JJ, **Cho Y**, **Lolis E**, **Bucala R** and **Cappello M (2008)** Orthologs of macrophage migration inhibitory factor from parasitic nematodes. *Trends Parasitol.* 24(8): 355–63.

Vermeulen A, **Behlen L**, **Reynolds A**, **Wolfson A**, **Marshall WS**, **Karpilow J** and **Khvorova A (2005)** The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA.* 11(5): 674–82.

Vickers KC, **Palmisano BT**, **Shoucri BM**, **Shamburek RD** and **Remaley AT (2011)** MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat. Cell Biol.* 13(4): 423–433. editor & translator, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.

Vidal K, **Grosjean I**, **evillard JP**, **Gespach C** and **Kaiserlian D (1993)** Immortalization of mouse intestinal epithelial cells by the SV40-large T gene. Phenotypic and immune characterization of the MODE-K cell line. *J. Immunol. Methods.* 166(1): 63–73.

Villanueva RA, **Jangra RK**, **Yi M**, **Pyles R**, **Bourne N** and **Lemon SM (2010)** miR-122 does not modulate the elongation phase of hepatitis C virus RNA synthesis in isolated replicase complexes. *Antiviral Res.* 88(1): 119–23.

Vowels JJ and **Thomas JH (1992)** Genetic analysis of chemosensory control of dauer

`

formation in Caenorhabditis elegans. *Genetics*. 130(1): 105–23.

Wadsworth WG and **Riddle DL (1989)** Developmental regulation of energy metabolism in Caenorhabditis elegans. *Dev. Biol*. 132(1): 167–173.

Wahlquist C, Jeong D, Rojas-Muñoz A, Kho C, Lee A, Mitsuyama S, van Mil A, Park WJ, Sluijter JPG & other authors (2014) Inhibition of miR-25 improves cardiac contractility in the failing heart. *Nature*. 508(7497): 531–5. editor & translator, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.

Wang H, Morita M, Yang X, Suzuki T, Yang W, Wang J, Ito K, Wang Q, Zhao C & other authors (2010a) Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. *EMBO J*. 29(15): 2566–76. editor & translator, EMBO Press.

Wang J, **Yu M**, **Yu G**, **Bian J**, **Deng X**, **Wan X** and **Zhu K (2010b)** Serum miR-146a and miR-223 as potential new biomarkers for sepsis. *Biochem. Biophys. Res. Commun*. 394(1): 184–8.

Wang J, **Czech B**, **Crunk A**, **Wallace A**, **Mitreva M**, **Hannon GJ** and **Davis RE (2011a)** Deep small RNA sequencing from the nematode Ascaris reveals conservation, functional diversification, and novel developmental profiles. *Genome Res*. 21(9): 1462–77.

Wang J, **Mitreva M**, **Berriman M**, **Thorne A**, **Magrini V**, **Koutsovoulos G**, **Kumar S**, **Blaxter ML** and **Davis RE (2012a)** Silencing of germline-expressed genes by DNA elimination in somatic cells. *Dev. Cell*. 23(5): 1072–80.

Wang J, **Lu M**, **Qiu C** and **Cui Q (2010c)** TransmiR: a transcription factor-microRNA regulation database. *Nucleic Acids Res*. 38(Database issue): D119-22.

Wang K, **Zhang S**, **Weber J**, **Baxter D** and **Galas DJ (2010d)** Export of microRNAs and microRNA-protective protein by mammalian cells. *Nucleic Acids Res*. 38(20): 7248–7259. editor & translator, Oxford University Press.

Wang K, **Li H**, **Yuan Y**, **Etheridge A**, **Zhou Y**, **Huang D**, **Wilmes P** and **Galas D (2012b)** The complex exogenous RNA spectra in human plasma: an interface with human gut biota? *PLoS One*. 7(12): e51009. editor & translator, Public Library of Science.

Wang L, **Li Z**, **Shen J**, **Liu Z**, **Liang J**, **Wu X**, **Sun X** and **Wu Z (2015)** Exosome-like vesicles derived by Schistosoma japonicum adult worms mediates M1 type immune- activity of macrophage. *Parasitol. Res*. 114(5): 1865–73.

Wang T, **Van Steendam K**, **Dhaenens M**, **Vlaminck J**, **Deforce D**, **Jex AR**, **Gasser RB** and **Geldhof P (2013)** Proteomic analysis of the excretory-secretory products from larval stages of Ascaris suum reveals high abundance of glycosyl hydrolases. *PLoS Negl. Trop. Dis*. 7(10): e2467. editor & translator, Public Library of Science.

Wang X, Chen W, Hu F, Deng C, Zhou C, Lv X, Fan Y, Men J, Huang Y & other authors (2011b) Clonorchis sinensis enolase: identification and biochemical characterization of a glycolytic enzyme from excretory/secretory products. *Mol. Biochem. Parasitol*. 177(2): 135–42.

`

Warf MB, **Johnson WE** and **Bass BL (2011)** Improved annotation of C. elegans microRNAs by deep sequencing reveals structures associated with processing by Drosha and Dicer. *RNA*. 17(4): 563–77.

Warner JR (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* 24(11): 437–440. editor & translator, Elsevier.

Watanabe Y, **Tomita M** and **Kanai A (2007)** Computational Methods for MicroRNA Target Prediction. *Methods Enzymol.* 42765–86.

Weir MR and **Dzau VJ (1999)** The renin-angiotensin-aldosterone system: a specific target for hypertension management. *Am. J. Hypertens.* 12(12): 205S–213S.

Wightman B, **Ha I** and **Ruvkun G (1993)** Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell.* 75(5): 855–862.

Winter A a. DAD, **Weir W**, **Hunt M**, **Berriman M**, **Gilleard JSJS**, **Devaney E** and **Britton C (2012)** Diversity in parasitic nematode genomes: the microRNAs of Brugia pahangi and Haemonchus contortus are largely novel. *BMC Genomics.* 13(1): 4. editor & translator, BioMed Central Ltd.

Winter AD, Gillan V, Maitland K, Emes RD, Roberts B, McCormack G, Weir W, Protasio A V, Holroyd N & other authors (2015) A novel member of the let-7 microRNA family is associated with developmental transitions in filarial nematode parasites. *BMC Genomics.* 16(1): 331. editor & translator, BioMed Central.

Wiskocil R, **Weiss a**, **Imboden J**, **Kamin-Lewis R** and **Stobo J (1985)** Activation of a human T cell line: a two-stimulus requirement in the pretranslational events involved in the coordinate expression of interleukin 2 and gamma-interferon genes. *J. Immunol.* 134(3): 1599–1603.

de Wit E, **Linsen SE V**, **Cuppen E** and **Berezikov E (2009)** Repertoire and evolution of miRNA genes in four divergent nematode species. *Genome Res.* 19(11): 2064–74.

Witkos TM, **Koscianska E** and **Krzyzosiak WJ (2011)** Practical Aspects of microRNA Target Prediction. *Curr. Mol. Med.* 11(2): 93–109. editor & translator, Bentham Science Publishers.

Woods K, **Thomson JM** and **Hammond SM (2007)** Direct regulation of an oncogenic micro-RNA cluster by E2F transcription factors. *J. Biol. Chem.* 282(4): 2130–4.

Wu CW, Ng SC, Dong Y, Tian L, Ng SSM, Leung WW, Law WT, Yau TO, Chan FKL & other authors (2014) Identification of microRNA-135b in stool as a potential noninvasive biomarker for colorectal cancer and adenoma. *Clin. Cancer Res.* 20(11): 2994–3002.

Wu J, Meng Z, Jiang M, Zhang E, Trippler M, Broering R, Bucchi A, Krux F, Dittmer U & other authors (2010) Toll-like receptor-induced innate immune responses in non-parenchymal liver cells are cell type-specific. *Immunology.* 129(3): 363–74. editor & translator, Wiley-Blackwell.

`

Xu J and **Wong C (2008)** A computational screen for mouse signaling pathways targeted by microRNA clusters. *RNA*. 14(7): 1276–83.

Xu J, **Zhu X**, **Wu L**, **Yang R**, **Yang Z**, **Wang Q** and **Wu F (2012)** MicroRNA-122 suppresses cell proliferation and induces cell apoptosis in hepatocellular carcinoma by directly targeting Wnt/ß-catenin pathway. *Liver Int*. 32(5): 752–60.

Yáñez-Mó M, Siljander PR-M, Andreu Z, Bedina Zavec A, Borràs FE, Buzas EI, Buzas K, Casal E, Cappello F & other authors (2015a) Biological properties of extracellular vesicles and their physiological functions. *J. Extracell. Vesicles*. 4(0): .

Yáñez-Mó M, Siljander PR-M, Andreu Z, Zavec AB, Borràs FE, Buzas EI, Buzas K, Casal E, Cappello F & other authors (2015b) Biological properties of extracellular vesicles and their physiological functions. *J. Extracell. vesicles*. 427066. editor & translator, Co-Action Publishing.

Yang P, **Baker KA** and **Hagg T (2006)** The ADAMs family: Coordinators of nervous system development, plasticity and repair. *Prog. Neurobiol*. 79(2): 73–94.

Yates DM, **Portillo V** and **Wolstenholme AJ (2003)** The avermectin receptors of Haemonchus contortus and Caenorhabditis elegans. *Int. J. Parasitol*. 33(11): 1183–1193.

Yau TO, Wu CW, Tang C-M, Chen Y, Fang J, Dong Y, Liang Q, Ng SSM, Chan FKL & other authors (2015, November 26) microRNA-20a in human faeces as a non-invasive biomarker for colorectal cancer. *Oncotarget*. editor & translator, Impact Journals.

Yi D, **Xu L**, **Yan R** and **Li X (2010)** Haemonchus contortus: cloning and characterization of serpin. *Exp. Parasitol*. 125(4): 363–70.

Yoshioka Y, **Konishi Y**, **Kosaka N**, **Katsuda T**, **Kato T** and **Ochiya T (2013)** Comparative marker analysis of extracellular vesicles in different human cancer types. *J. Extracell. vesicles*. 2. editor & translator, Co-Action Publishing.

Yuan YRY, **Pei Y**, **Ma JJB**, **Kuryavyi V**, **Zhadina M**, **Meister G**, **Chen HY**, **Dauter Z**, **Tuschl T** and **Patel DJ (2005)** Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol. Cell*. 19(3): 405–419.

Zamanian M, Fraser LM, Agbedanu PN, Harischandra H, Moorhead AR, Day TTA, Bartholomay LLC, Kimber MJM, Bartholomay LLC & other authors (2015) Release of Small RNA-containing Exosome-like Vesicles from the Human Filarial Parasite Brugia malayi. *PLoS Negl. Trop. Dis*. 9(9): e0004069 (A. Hoerauf, Ed.). editor & translator, Public Library of Science.

Zamore PD, **Tuschl T**, **Sharp PA** and **Bartel DP (2000)** RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*. 101(1): 25–33. editor & translator, Elsevier.

Zang X, **Yazdanbakhsh M**, **Jiang H**, **Kanost MR** and **Maizels RM (1999)** A novel serpin expressed by blood-borne microfilariae of the parasitic nematode Brugia malayi

`

inhibits human neutrophil serine proteinases. *Blood*. 94(4): 1418–28.

Zaph C, Troy AE, Taylor BC, Berman-Booty LD, Guild KJ, Du Y, Yost EA, Gruber AD, May MJ & other authors (2007) Epithelial-cell-intrinsic IKK-beta expression regulates intestinal immune homeostasis. *Nature*. 446(7135): 552–6. editor & translator, Nature Publishing Group.

Zeng C, **Wang R**, **Li D**, **Lin X-J**, **Wei Q-K**, **Yuan Y**, **Wang Q**, **Chen W** and **Zhuang S-M (2010)** A novel GSK-3 beta-C/EBP alpha-miR-122-insulin-like growth factor 1 receptor regulatory circuitry in human hepatocellular carcinoma. *Hepatology*. 52(5): 1702–12.

Zeng Y and **Cullen BR (2005)** Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *J. Biol. Chem.* 280(30): 27595–27603.

Zhang G, **Li Y**, **Zheng S**, **Liu M**, **Li X** and **Tang H (2010)** Suppression of hepatitis B virus replication by microRNA-199a-3p and microRNA-210. *Antiviral Res.* 88(2): 169–75.

Zhang J, **Li S**, **Li L**, **Li M**, **Guo C**, **Yao J** and **Mi S (2015)** Exosome and Exosomal MicroRNA: Trafficking, Sorting, and Function. *Genomics. Proteomics Bioinformatics*. 13(1): 17–24.

Zhang L, Hou D, Chen X, Li D, Zhu L, Zhang Y, Li J, Bian Z, Liang X & other authors (2012) Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. *Cell Res.* 22(1): 107–26. editor & translator, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences.

Zhao J, Luo R, Xu X, Zou Y, Zhang Q, Pan W, Makaula P, Sadalaki J, Muula A & other authors (2015) High-throughput sequencing of RNAs isolated by cross-linking immunoprecipitation (HITS-CLIP) reveals Argonaute-associated microRNAs and targets in Schistosoma japonicum. *Parasit. Vectors*. 8(1): 589. editor & translator, BioMed Central.

Zhu L, Zhao J, Wang J, Hu C, Peng J, Luo R, Zhou C, Liu J, Lin J & other authors (2016a) MicroRNAs Are Involved in the Regulation of Ovary Development in the Pathogenic Blood Fluke Schistosoma japonicum. *PLoS Pathog*. 12(2): e1005423.

Zhu L, **Liu J**, **Dao J**, **Lu K**, **Li H**, **Gu H**, **Liu J**, **Feng X** and **Cheng G (2016b)** Molecular characterization of S. japonicum exosome-like vesicles reveals their regulatory roles in parasite-host interactions. *Sci. Rep.* 625885. editor & translator, Nature Publishing Group.

Zisoulis DG, **Lovci MT**, **Wilbert ML**, **Hutt KR**, **Liang TY**, **Pasquinelli AE** and **Yeo GW (2010)** Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. *Nat. Struct. Mol. Biol.* 17(2): 173–179. editor & translator, Nature Publishing Group.

`