Genetic control of development in the parasitic nematode *Haemonchus contortus* by microRNAs

Volume I of I

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Neil Derek Marks

BVetSci BVMS MRes MRCVS

Institute of Biodiversity, Animal Health and Comparative Medicine

School of Life Sciences

College of Medicine, Veterinary and Life Sciences

University of Glasgow

August 2015

Abstract

The parasitic nematode *Haemonchus contortus* has a major impact on the welfare and economic sustainability of small ruminant farming throughout the world. Increasing drug resistance requires the development of novel therapeutic agents. To further this process, we examined the fundamental biology of development in *H. contortus*, specifically, the potential role of microRNAs (miRNAs). miRNAs are short, non-coding RNA molecules that negatively regulate gene expression. In the free-living nematode *Caenorhabditis elegans*, miRNAs regulate a variety of genes including those involved in development. This thesis describes the expression patterns, potential targets and possible functions of miRNAs in *H. contortus* throughout development.

H. contortus miRNAs were first identified by Winter *et al*. (2012). This project details the use of microarrays to quantify their abundance in four life cycle stages (L3, L4, adult males and adult females) and one tissue (adult female gut). 55 miRNAs showed significant temporal or spatial variation in expression, suggesting roles in stage- or tissue-specific gene regulation. Target prediction algorithms identified binding sites in 93 % of annotated genes in *H. contortus*, suggesting extensive miRNA regulation. Gene ontology highlighted enrichment of developmental and metabolic annotation terms in a number of miRNAs, supporting this concept. Further evidence of an important role for miRNAs in development was provided by studies in *C. elegans* using mutant strains and GFP reporters, highlighting intriguing spatial patterns in expression.

The data presented here are the first detailed examination of miRNAs and their targets in *H. contortus*. Significant uptake of labelled RNA was demonstrated by *H. contortus in vitro*, indicating the potential to interfere with miRNA activity through the delivery of miRNA mimics and inhibitors. This is important in defining the function of specific miRNAs and the pathways they regulate, possibly leading to their use as novel therapeutic agents to combat *H. contortus*.

Contents

List of tables

List of figures

Acknowledgements

I would like to express my deepest gratitude to my supervisor Dr. Collette Britton for her continuous support, patience and motivation. Her guidance helped me throughout my research and the writing of this thesis. My secondary supervisor, Prof. Eileen Devaney is also deserving of praise, providing excellent guidance and a positive environment from the moment I started this project. I am incredibly fortunate to have two caring and knowledgeable mentors during my project.

Many people helped me during my research, but particular thanks must go to Dr. Brett Roberts, Dr. Vicky Gillan, Dr. Alan Winter, Ms. Kirsty Maitland and Dr. Gillian Stepek. Without their valuable support in the lab, alongside inexhaustible patience, it would not have been possible to conduct this research. I would like to thank Mr. Henry Gu for always providing support and helpful insights and, more directly by generating the target prediction data. The predictions were made possible by data generously provided by Dr. Axel Martinelli in conjunction with Dr. Roz Laing. An anti-GST antibody was graciously provided by Dr. Ania Owsianka.

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: Neil Derek Marks

Abbreviations

Abbreviation	Word/phrase
aa	Amino acid
ago	Argonaute (gene)
AGO	Argonaute (protein)
alg	Argonaute-like gene (gene)
ALG	Argonaute-like gene (protein)
bp	basepair
C. elegans	*Caenorhabditis elegans*
Daf	Dauer formation
Daf-C	Constitutive dauer formation
Daf-D	Deficient dauer formation
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
gf	Gain-of-function
GO	Gene ontology
GST	Glutathione-S transferase
H. contortus	*Haemonchus contortus*
H. sapiens	*Homo sapiens*
His	Histidine
HITS-CLIP	High-throughput sequencing with cross-linked immunoprecipitation
L1/2/3/4	1^{st}, 2^{nd}, 3^{rd} and 4^{th} larval stage
lf	Loss-of-function
LNA	Locked nucleic acid
lncRNA	Long noncoding RNA
mir-	miRNA gene
miR-	Mature miRNA
miRNA	MicroRNA
mRNA	Messenger RNA
nt	Nucleotide
PCR	Polymerase chain reaction
piRNA	Piwi-interacting RNA
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA
PPP	Prepatent period
qRT-PCR	Quantitative real-time PCR
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNAseq	RNA sequencing
siRNA	Small interfering RNA
tRNA	Transfer RNA
UTR	Untranslated Region
wt	wildtype

1 Introduction to parasitic nematodes of livestock

Haemonchus contortus is a member of the Trichostrongylidae family of nematodes. Parasites in this group are responsible for significant disease, causing major economic and welfare problems throughout the world. In the UK, *Teladorsagia circumcincta* is the most common Trichostrongyloid parasite, affecting 100 % of farms in some surveys (Burgess *et al.*, 2012). The same survey identified *H.contortus* on 50 % of farms. However, *H. contortus* is more pathogenic, responsible for significant losses in production through anaemia and protein loss; large worm burdens can be fatal. Currently, parasitic nematodes are controlled with a variety of anthelmintic drugs. Unfortunately, due to their extensive use, parasites have developed resistance, reducing the treatment efficacy. This is now a major problem in sheep and goats, as well as an increasing problem in cattle (Prichard, 2009). With the increasing consumption of meat throughout the world, particularly in developing nations, there is a greater need for increased efficiency of livestock production. Without an effective control strategy, helminth parasites threaten food security. Genomic and transcriptomic data is becoming available for an increasing number of parasitic nematodes (http://www.sanger.ac.uk/resources/downloads/helminths/) and it is important and timely to apply this knowledge to the design of alternative control approaches. To develop new therapeutic agents, a better understanding of biological processes that can be targeted specifically and efficiently is needed.

1.1 Haemonchus contortus

Haemonchus contortus is an obligate parasitic nematode with a cosmopolitan (worldwide) distribution that primarily infects small ruminants, namely sheep and goats. Adults of the species are dioecious, sexually dimorphic, hematophagous (blood feeding) and responsible for significant pathology and clinical disease.

1.1.1 Taxonomy and host range

Haemonchus contortus is one of 12 species within the *Haemonchus* genus (Table 1-1). All species share a similar life cycle (Fig. 1-1) but are found within different regions and parasitise different members of the order Artiodactyla (even-toed ungulates; Hoberg *et al.*, 2004). Three species (*H. contortus*, *H. placei* and *H. similis*) possess a cosmopolitan distribution and although all three can parasitise a wide range of hosts, *H. contortus* primarily infects small ruminants (sheep and goats), while *H. placei* and *H. similis* primarily infect cattle.

Two species (*H. longistipes* and *H. okapiae*) parasitise a smaller number of hosts and are limited to Africa and southern Asia. *H. longistipes* is a major parasite of camelids while *H. okapiae* infects giraffidae (Lichtenfels *et al.*, 2002). The remaining seven species of *Haemonchus* are only found within sub-Saharan Africa and parasitise various antelope and gazelle (Boomker *et al.*, 1986; Lichtenfels *et al.*, 2001; Sachs *et al.*, 1973).

Table 1-1. Host range and distribution of members of the *Haemonchus* genus.

Species	Distribution	Host(s)[1]	Reference[2]
H. bedfordi	Africa	Giraffidae Antelope	Le Roux, 1929
H. contortus	Cosmopolitan	**Sheep** (*Ovis aries*) **Goat** (*Capra aegagrus hircus*) Other Bovidae Cervidae Camelidae Antilocapridae Giraffidae Alcelaphinae	Rudolphi, 1803
H. dinniki	Africa	Impala (*Aepyceros melampus*) Other antelope	Sachs, Gibbons and Lweno, 1973
H. horaki	Africa	Grey Rhebok (*Pelea capreolus*) Other antelope	Lichtenfels, Pilitt, Gibbons and Boomker, 2001
H. krugeri	Africa	Steenbok (*Raphicerus campestris*) Other Antelope	Ortlepp, 1964
H. lawrencei	Africa	Antelope	Sandground, 1933
H. longistipes	Africa, Southern Eurasia	**Camelidae** Bovidae	Railliet and Henry, 1909
H. mitchelli	Africa	Eland (*Taurotragus oryx*) Other Antelope	Le Roux, 1929
H. okapiae	Africa, Southern Eurasia	Okapi (*Okapia johnstoni*) Other Giraffidae	van de Berghe, 1937
H. placei	Cosmopolitan	**Cattle** (*Bos taurus*) Other bovidae	Place, 1893
H. similis	Cosmopolitan	**Cattle** (Bos taurus) Other bovidae	Travassos, 1911
H. veglia	Africa	Antelope	Le Roux, 1929

[1] Primary host(s) in bold if known.

[2] First description of the species.

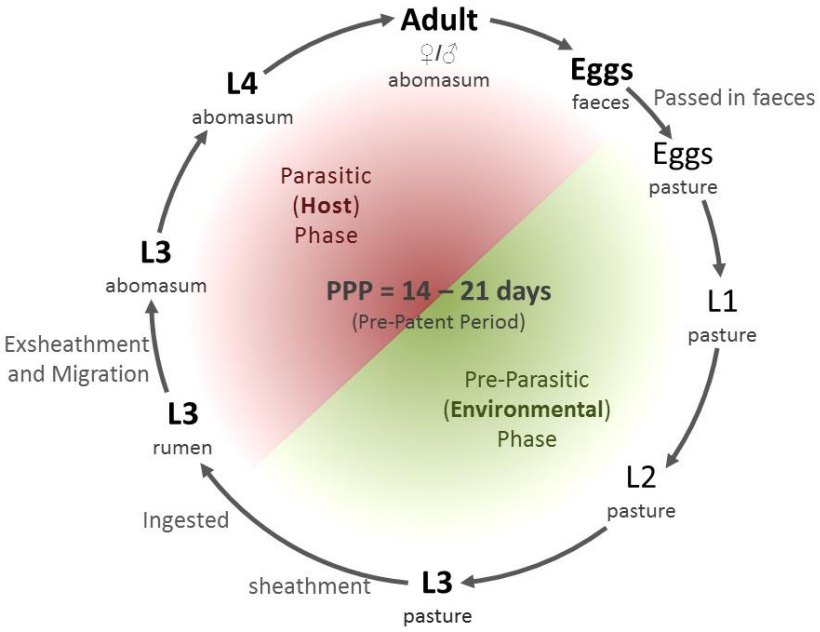


Fig. 1-1: Life cycle of *H. contortus*. Development is split into two phases based on location: environmental and host phases. The environmental (pre-parasitic) phase is variable in duration between five days and several months depending on conditions and host availability. The pre-patent period is the interval between ingestion of infective L3 and release of eggs into the environment.

1.1.2 Appearance

H. contortus is dioecious with adult worms displaying sexual dimorphism. Females are larger than males (~3 cm *versus* ~2 cm) and display characteristic white ovaries that twist around the blood-red intestine. This led to the colloquial name of *H. contortus*: Barber's pole worm (Fig. 1-2, A and B). Males possess a distinctive copulatory bursa at the posterior extremity while females possess a more anterior vulval flap. Morphology is maintained between hosts. Microscopically, cervical papillae and buccal lancets can be identified in both sexes (Fig. 1-2, C and D). Species categorisation is difficult, due to morphological similarities, but can be performed using detailed inspection of cuticular ridges (Lichtenfels *et al.*, 1986)

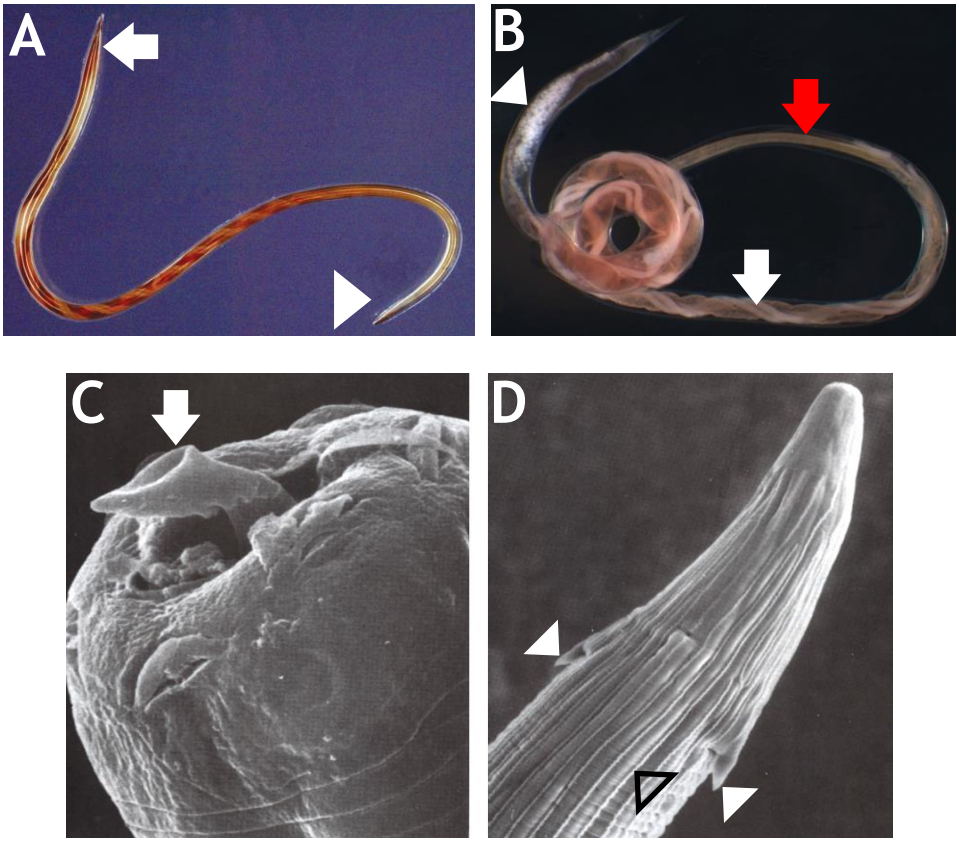


Fig. 1-2: Adult *H. contortus*. **A**: General appearance of female. Arrowhead indicates anterior. Arrow indicates posterior. Taken from British Veterinary Association (Anonymous, 2010). **B**: Detail of intestine (red arrow) and ovaries (white arrow). White arrowhead indicates egg-filled uterus. Taken from Novartis Animal Health (http://www.farmanimalhealth.co.uk/sheep-worms). **C**: Buccal lancet (arrow). **D**: Cervical papillae (arrowheads). Cuticular ridges are also visible (open arrowhead). C and D taken from Veterinary Parasitology 2nd Edition (Urquhart et al., 1996).

1.1.3 Life cycle of *H. contortus*

H. contortus has a direct life cycle (Fig. 1-2) with a single host and two phases: the pre-parasitic (environmental) phase and parasitic (host) phase.

Pre-parasitic phase of the *H. contortus* life cycle

The pre-parasitic phase of *Haemonchus contortus* is typical of other *Trichostrongyloid* nematodes. Adults release eggs into the environment via faeces where they develop into 3rd stage larvae (L3). Eggs hatch into 1st stage larvae (L1), and progress through two moults within the faecal pat, feeding on bacteria, to become L3 in approximately one or two weeks. The majority of L3 will then migrate from the faecal pat, both horizontally (up to 15 cm) and vertically within the local sward (Rose, 1963). Here, the infective L3 arrest development until ingested by a potential host.

H.contortus is primarily considered a tropical parasite, as its development is favoured by warm, humid environments. Under ideal laboratory conditions of high humidity and temperature (33 °C), development from eggs to L3 can take as little as two and a half days (Berberian and Mizelle, 1957), although in the field development generally takes 1 – 2 weeks (Banks *et al.*, 1990). The greatest factor for the successful development of larvae in both temperate and tropical regions is rainfall (Besier and Dunsmore, 1993; Onyali *et al.*, 1990) as it prevents desiccation (Rose, 1963). Temperature is also an important factor for larval development: L3s can develop between 5 °C and 38 °C (Banks *et al.*, 1990; Berberian and Mizelle, 1957). Survivability plummets when temperatures fall outside this range, limiting development in temperate winters (Gibson and Everett, 1976) and tropical dry-seasons (Donald, 1968). In temperate regions, less than 7 % of eggs complete L3 development (Besier and Dunsmore, 1993; Rose, 1963) while in tropical regions 15 % of

eggs complete L3 development (Dinnik and Dinnik, 1958). In all regions, survivability markedly increases beyond the late L2 stage (Berberian and Mizelle, 1957; Todd *et al.*, 1976). Under mild laboratory conditions (10 – 11 °C) larvae can persist for almost two years while still retaining infectivity (Rose, 1963). Once ingested by a suitable host, larvae enter the parasitic phase.

Parasitic phase of the *H. contortus* life cycle

H. contortus parasitise ruminants that feed on herbage, most commonly grasses. Due to their low nutritive value, ruminants graze for extended periods, increasing the risk of infection. Pasture-based farming practices substantially increase this risk. L3 are swallowed by a suitable host, enter the reticulorumen and pass through into the abomasum. Here the protective L2 cuticle that provided environmental resistance is removed (exsheathment). While the exact mechanism of exsheathment is unknown, it likely involves changes in pH as exsheathment can be performed *in vitro* using acids or CO_2 (see chapter 2). Additionally, the abomasum is strongly acidic (pH 3 – 4) while the reticulorumen is weakly acidic (pH 5.5 – 6.5). Larvae are able to arrest development within the host under certain environmental cues. Resumption of development and subsequent release of eggs from adults can generate outbreaks of hyperacute haemonchosis associated with sudden death (Barger and Le Jambre, 1979).

Exsheathed L3 burrow into the abomasal mucosa and moult into L4 within 48 hours before moulting again into adults. Males and female then begin to feed and mate. Females are polyandrous (Redman *et al.*, 2008), releasing 5,000 – 10,000 eggs per day into the gastro-intestinal tract. Eggs are expelled into the environment in faecal material, completing the life cycle. The prepatent period (time between infection and production of eggs) is approximately 18 – 21 days.

1.1.4 Pathogenesis, disease and economic impact of Haemonchosis

L4 and adult *H. contortus* worms are haematophagic (blood feeding) parasites and the pathogenic effects of infection (haemonchosis) are primarily due to this behaviour. Worms bury their heads into the rumen tissue to reach sub-mucosal blood vessels. Haematophagy leads to blood loss, irreversible protein loss and altered volatile fatty acid production within the rumen (Rowe *et al.*, 1988). Tissue damage from the burrowing increases abdominal lymph flow with associated increases in protein content from rumenal tissue damage, but does not produce large influxes of immune cells (Smith *et al.*, 1982).

The severity of clinical disease depends upon the number of adult worms present. Low grade infections (< 3,000 worms) produce a chronic disease resulting in weakness, inappetence and progressive weight loss. Acute infections (3,000 – 30,000 worms) are characterised by anaemia, lethargy, melaena and oedema, while hyperacute infections (> 30,000 worms) will present as sudden death due to severe haemorrhagic gastritis. In naïve lamb populations mortality can reach 50 % (Barger and Southcott, 1978). Resistance to clinical disease increases with previous exposure and age. At two years old, sheep (wethers) display a 36 % mortality rate while older sheep are highly resistant to infection due to a combination of age and prior exposure (Cohen *et al.*, 1972).

During the later stages of pregnancy and early-stages of lactation, adult sheep display an increased susceptibility to infection (Shubber *et al.*, 1981). Pregnancy induces a relaxation in the host's immune response and allows hypobiotic L4 to re-establish infection (Boag and Thomas, 1977). In conjunction with a lowered resistance to ingested L3, eggs are released into the environment ahead of parturition. The resumption of host immunity means that adult sheep are unlikely to succumb to severe clinical disease, but the

periparturient rise in faecal egg output acts as a major source of infection for naïve lambs during early summer (Gibson, 1973).

The major economic impacts of haemonchosis are due to sudden death (acute infections) and poor growth (chronic infection). Specifically, protein loss reduces the growth rate and decreases the market value of the carcass (Abbott *et al.*, 1988). Reduction of milk production can have a major effect on lamb growth as well as affecting efficiency in dairy ewes (Suarez *et al.*, 2009). Wool production is also affected (Barger and Cox, 1984), although fleeces generally account for less than 10 % of a sheep's value in most modern farming systems.

1.1.5 Anthelmintic treatment of *H. contortus* infection

Treatment of haemonchosis is often performed as part of a farm's general worming practice. Broad-spectrum anthelmintics are effective against most gastrointestinal nematodes, including *H. contortus*. There are three main classes of broad-spectrum anthelmintic widely used throughout the world: benzimidazoles (*e.g.* fenbendazole), imidazothiazoles/tetrahydropyramidines (*e.g.* levamisole, morantel) and macrocyclic lactones (*e.g.* ivermectin, moxidectin). Treatment is usually performed at the herd level, but targeted treatment of specific animals may be performed when there are signs of anaemia and a history of haemonchosis. In this situation, narrow-spectrum anthelmintics, such as halogenated salicylanides (e.g. closantel) and substituted phenols (e.g. disophenol) can be used. These drugs are not effective against the same range of parasite species as the broad-spectrum anthelmintics and are therefore not used for a farm's general worming practice.

For many anthelmintics, the mode of action is not fully known, but in general they act selectively on parasite processes absent or non-vital to the host. The widely used benzimidazoles inhibit polymerisation of microtubules in eukaryotes (Lacey, 1990). Selective toxicity between host and parasite is due to differences in affinity to tubulin. Their effectiveness in nematode parasites arises from impairment of reproduction and locomotion. Mutations in parasite tubulin can reduce affinity for these drugs, generating resistance (Lacey and Gill, 1994). Levamisole, pyrantel and morantel (or cholinergic anthelmintics) belong to a class of nicotinic receptor agonists that induce spastic muscle paralysis (reviewed in Martin *et al.*, 2006). Resistance can arise through mutations in the nictonic acetylcholine receptor subunits to which the drugs bind. The macrocyclic lactones (avermectins and milbemycin) induce paralysis of nematode pharyngeal and body wall muscle through interaction with ligand-gated ion-channels. This includes nicotine acetylcholine receptors (Krause *et al.*, 1998) and GABA (γ-aminobutyric acid) receptors (Dye-Holden and Walker, 1990). The importance of the avermectin compounds to global health was recently recognised with a Nobel Prize in physiology or medicine for its discoverers William Campbell and Satoshi Ōmura (in conjunction with Youyou Tu for the discovery of artemisinin).

1.1.6 Anthelmintic resistance in *H. contortus* and other parasites

The extensive use of anthelmintics has resulted in the development of drug resistance (Kaminsky, 2003). Defined as a drug having less than 95 % effectiveness, it is usually measured by a reduction in faecal egg output before and after treatment. Resistance will eventually lead to a complete failure of treatment where an anthelmintic has no effect on an animal's worm burden. Due to shared modes of action, parasites that develop resistance to one anthlemintic are frequently resistant to all other members of that class. Nematodes

resistant to all three classes of broad-spectrum anthelmintic are termed “triple-resistant” and pose a major threat to livestock production.

H. contortus was one of the first parasites reported to display anthelmintic resistance (to phenothiazine; Drudge *et al.*, 1957). Resistance to broad-spectrum anthelmintics is now common around the world (Kaplan, 2004) and there are increasing reports of parasites with triple-resistance (Taylor *et al.*, 2009). The situation is particularly challenging in tropical regions, where the parasite life cycle is continuous throughout the year and numerous treatments result in strong selection pressures. Resistance can be such a problem that farms may have to abandon livestock production, culling the entire herd (Sargison *et al.*, 2005).

Recently, two new anthelmintic classes have been developed, the first for 25 years: monepantel, licensed as Zolvix® by Novartis Animal Health (Kaminsky *et al.*, 2008) and derquantel, marketed as Startect® (in combination with abamectin) by Zoetis (Little *et al.*, 2010). However, resistance to monepantel was predicted to occur rapidly (Dobson *et al.*, 2011) and reports indicate that resistance is already present in Europe (Anonymous, 2014), South America (Mederos *et al.*, 2014) and New Zealand (Scott *et al.*, 2013), albeit in systems with strong selection pressures.

Development of new anthelmintics, as well as anti-infective agents in general, is both expensive and time consuming. Pharmaceutical companies are reluctant to develop a drug that may quickly become obsolete or become limited in use through regulatory constraints. While it may never be possible to prevent resistance, we may be able to mitigate its effects by identifying additional systems fundamental to the parasite. Identifying such systems will make targeted drug development a possibility.

While researchers work on developing novel therapeutics, various groups have been trying to improve the use of the anthelmintics already available. One example is the SCOPS

(sustainable control of parasites, http://www.scops.org.uk) group representing major organisations within the UK. They provide farmers and vets with the latest scientifically-based advice on how to use anthelmintics effectively and minimise the development of anthelmintic. Over the past decade, farmers and veterinarians have begun to realise how they have taken anthelmintics for granted. Responsible use of anthelmintics is vital to reduce selection pressures and limit the development of resistance. Resistance is generally noticed by a farmer only when treatment failure has occurred, by which stage it is too late for preventative measures. For one class of anthelmintics, the benzimidazoles, research suggests that parasites can maintain resistance even when they haven't been exposed to that drug for ten years (Leignel *et al.*, 2010). But recent data suggest resistance may be reversible through good herd management (Leathwick *et al.*, 2015).

Using vaccines as an alternative to anthelmintics

Vaccines against parasites have been a focus of research for many decades (Vercruysse *et al.*, 2004). The earliest to be commercially available was against the bovine lungworm *Dictyocaulus vivparus*, the major causative agent of parasitic bronchitis (husk) in cattle. Originally known as Dictol, now traded under the name Bovilis® Huskvac by MSD Animal health, a single dose consists of 1,000 L3 attenuated by x-rays. It provides protection against clinical disease, but not complete resistance to infection. Sales have declined over the years, even in the face of a rising incidence of husk, due to farmer's preference for broad-spectrum anthelmintic therapies. Aside from Huskvac and Barbervax (see later), the majority of commercially marketed anti-parasite vaccines target veterinary protozoa, such as *Eimeria spp*. (Sharman *et al.*, 2010), *Toxoplasma gondii* (Hiszczyńska-Sawicka *et al.*, 2014), *Neosprora caninum* (Monney and Hemphill, 2014), *Babesia spp*. and *Theileria spp*. (Shkap *et al.*, 2007).

In *H. contortus*, vaccine research has involved irradiated larvae (Sivanathan *et al.*, 1984), larval surface antigens (Piedrafita *et al.*, 2013) and DNA vaccines encoding parasite actin (Yan *et al.*, 2014). However, the majority of work has involved the "hidden antigens" approach using *H. contortus* gut proteins (Knox, 2011). While these are accessible to host antibodies through the parasite's haematophagic feeding, the immune response will not be reinforced through exposure. Early research identified a number of proteins that could induce a protective host immune response, such as cysteine proteases (Boisvenue *et al.*, 1992; Knox and Jones, 1990) and contortin (Munn *et al.*, 1987). The most well studied antigens are two integral membrane glycoproteins: H11 and H-gal-GP (Haemonchus galactose-containing glycoprotein complex).

Native extracts of H11, a 110 kDa integral membrane aminopeptidase, provide excellent protection against experimental infection (Newton and Munn, 1999). However, enzymatically active proteins produced in recombinant insect (Munn, 1977) and *C. elegans* (Roberts *et al.*, 2013) systems are unable to induce protective immunity. A DNA vaccine of H11 also failed to induce significant protection in goats (Zhao *et al.*, 2012). In a similar situation, native extracts of the H-gal-GP complex provide substantial protection while recombinant proteins do not (Smith *et al.*, 1994). Studies have identified four metalloendopeptidases (MEP1-4) within the H-gal-GP complex that generate the majority of antibody-mediated protection, (Newlands *et al.*, 2006; Redmond *et al.*, 1997; Smith *et al.*, 2000).

Recently, efforts have been directed at large-scale isolation of native proteins from the adult *H. contortus* gut in sufficient quantity for commercial production. A vaccine, including both H11 and H-gal-GP, provides up to 95% protection against natural field infections of *H. contortus* (Bassetto *et al.*, 2014b) as well as providing cross-protection against *H. placei* and *H. similis* (Bassetto *et al.*, 2014a). Recently this was released commercially in Australia as Barbervax (http://www.moredun.org.uk/news/moredun-

launches-barbervax-vaccine-barbers-pole-worm). This success shows that vaccination against *H. contortus* is feasible and provides an alternative to drug treatment in regions where *H. contortus* is the dominant parasite and anthelmintics are losing efficacy. However, the vaccine is produced by infecting sheep with sub-clinical doses of *H. contortus* L3 and harvesting adult worms, which limits industrial scale production and raises ethical concerns.

1.1.7 Identifying novel targets for control of *H. contortus*

The development of therapies, such as Barbervax, requires identification of targets that are vital to the parasite, but not to the host. However, progress over the years has been slow and new methods of identifying potential targets are required. An approach that has been successfully applied to *C. elegans* is gene silencing using RNAi (RNA interference).

RNAi involves post-transcriptional regulation of gene expression using small inhibitory RNA (siRNA) molecules (Fire *et al.*, 1998). siRNAs are short (20-25 basepairs), double-stranded RNAs first discovered in plants (Hamilton and Baulcombe, 1999) and later found in mammalian cells (Elbashir *et al.*, 2001). RNAi utilising endogenous siRNA is a major regulatory system in many eukaryotic organisms (Ruvkun *et al.*, 2004), acting to regulate endogenous gene expression as well as protecting against foreign genetic material, such as RNA viruses and transposons (Gitlin *et al.*, 2002; Vance and Vaucheret, 2001). Exogenous siRNA can also be used by researchers to investigate the role of specific genes. Realisation of its importance as both a regulatory system and molecular tool earned Andrew Fire and Craig Mello the 2006 Nobel Prize in Physiology or Medicine.

In *C. elegans*, the use of RNAi has expanded our knowledge of many different areas of nematode biology, including embryogenesis (Du *et al.*, 2015), growth and morphology

(Fraser *et al.*, 2000; Kamath *et al.*, 2003), neuronal development (Nix *et al.*, 2014; Poole *et al.*, 2011), stress response (Volovik *et al.*, 2014), lipid metabolism (Ashrafi *et al.*, 2003) and longevity (Hamilton *et al.*, 2005). RNAi can be a boon to the study of organisms where traditional genetic manipulation is unavailable, such as *Trypanosoma brucei brucei* (Baker *et al.*, 2011). However, the knockdown effect of RNAi is highly variable depending on the delivery method, target of interest and organism involved (Geldhof *et al.*, 2007; Selkirk *et al.*, 2012). This is particularly true for *H. contortus* where numerous attempts have been made at developing RNAi as a molecular tool with limited results (Britton *et al.*, 2012; Geldhof *et al.*, 2006; Samarasinghe *et al.*, 2011).

siRNAs are just one class of RNA molecule involved in post-transcriptional gene regulation. Long, non-coding RNAs (lncRNAs) are single-stranded RNAs greater than 200 nucleotides (nt) in length. They are abundant in many eukaryotic organisms, with almost 15,000 predicted in humans alone (Derrien *et al.*, 2012). lncRNAs regulate gene expression through several mechanisms including generalised effects on transcription (Goodrich and Kugel, 2006), post-transcriptional silencing (Yoon *et al.*, 2013) and epigenetic regulation (Rinn *et al.*, 2007). Piwi-interacting RNAs (piRNAs) are short (26-31 nt) single-stranded RNAs, first identified in *Drosophila melanogaster* (Lin and Spradling, 1997), that maintain germline cells (Siomi *et al.*, 2011). The most well-characterised function of piRNAs is the silencing of mobile genetic elements (*e.g.* transposons) but they can also regulate protein-coding RNAs (Kiuchi *et al.*, 2014). Another category of non-coding RNA, microRNA (miRNA), has gained much attention over the past decade for important roles in post-transcriptional gene regulation. Mature miRNAs are single-stranded and of a similar size to piRNAs and siRNAs (20 – 25 nucleotides), but are involved in a wider range of transcriptional regulation.

The study of RNA-mediated gene regulation has rapidly expanded in the past ten years, with thousands of papers describing their biogenesis, function and usage in target

discovery and as therapeutic agents. This thesis focuses on miRNAs and their potential role in parasitic nematodes, particularly *H. contortus*.

1.2 MicroRNAs

The most well studied role of miRNAs is negative regulation of gene expression at the post-transcriptional level. A group of proteins known as the RNA-induced silencing complex (RISC) uses a mature miRNA to target specific messenger RNAs (mRNAs) through imperfect complementarity, often in the 3′ untranslated region (UTR). When the RISC identifies and binds to a target mRNA, it inhibits translation and promotes degradation, leading to an overall decrease in protein abundance. While poorly understood in comparison, miRNAs are also capable of increasing protein abundance (reviewed in Valinezhad Orang *et al.*, 2014) and regulating gene expression from within the nucleus (reviewed in Roberts, 2014). miRNAs are a highly conserved feature of life, present in the majority of organisms. miRNA regulation can also be highly complex as a single miRNA can affect hundreds of genes across multiple pathways, while a single mRNA can contain multiple mRNA binding sites.

1.2.1 The discovery of miRNAs

The first miRNA, *lin-4*, was discovered almost two decades ago as a genetic mutation affecting correct developmental timing of the non-parasitic, free-living nematode *C. elegans* (Lee *et al.*, 1993). Sequencing of the mutant identified *lin-4* a small, non-coding RNA that down-regulated expression of the protein, LIN-14 by binding to its messenger RNA (mRNA) (Wightman *et al.*, 1993). For several years, this unusual system of post-transcriptional gene regulation was considered limited to nematodes and largely ignored by

the wider community. In 2000, another important miRNA, *let-7,* was also discovered to regulate larval development in *C. elegans* (Reinhart *et al.*, 2000). Very quickly it was found that both *let-7* and its temporal control of development were conserved in a wide range of multicellular organisms (Pasquinelli *et al.*, 2000).

1.2.2 The biogenesis of miRNAs

Since their discovery, much research has been performed on the biogenesis of miRNAs (Fig. 1-3), particularly in *C. elegans* and *D. melanogaster*. Most miRNA genes are transcribed by RNA-Polymerase II (Pol II; Lee *et al.*, 2004), although a small subset can be transcribed by RNA-Polymerase III (Pol III; Borchert *et al.*, 2006). These primary transcripts (pri-miRNAs) are often several kilobases in size and are cleaved to produce a much smaller (60 – 80 bp) stem-loop, or hairpin, structure known as the precursor miRNA (pre-miRNA; Lee *et al.*, 2002). Cleavage of the pre-miRNA occurs in the nucleus and is performed by the RNase III-type protein, Drosha. The pre-miRNA is then exported from the nucleus and further cleaved into the short (18 – 25 nt) mature miRNA duplex by the protein, Dicer (Hutvágner *et al.*, 2001). This duplex is then bound by an Argonaute protein (*e.g.* AGO-1) in association with multiple accessory proteins (*e.g.* AIN-1), forming the RNA-induced silencing complex (RISC; Hammond *et al.*, 2001). One strand (3′ or 5′) of the duplex acts as the guide for the RISC while the remaining strand is usually degraded. The system governing which strand is retained by the RISC involves thermodynamic stability, but is not completely understood (Khvorova *et al.*, 2003). In many instances, both strands can be detected within an organism at the same time, although one strand tends to be more abundant.

Once a mature RISC is complete, miRNAs are able to negatively regulate gene expression through complementary base-pairing to the target mRNA. This usually occurs in the 3′

UTR, although miRNAs may bind to additional sites within an mRNA including introns or the 5′ UTR (Schnall-Levin *et al.*, 2010). While miRNA biogenesis is well understood, the exact nature of the miRNA-mRNA interaction is not completely resolved. However, current evidence suggests that miRNA repression involves a combination of mRNA degradation and prevention of transcription (Jackson and Standart, 2007). Multiple degradative pathways are thought to be involved, including recruitment of endonucleases (Yekta, 2004), deadenylation leading to exonucleolytic cleavage (Orban, 2005) and AU-rich element (ARE) decay (Jing *et al.*, 2005). The specific sequence of a mature miRNA determines which mRNAs are targeted, although different nucleotides are more important than others (reviewed in Grimson *et al.*, 2007). The most notable region is the seed sequence (nucleotides 2 – 7), which helps determine the family to which a miRNA belongs.

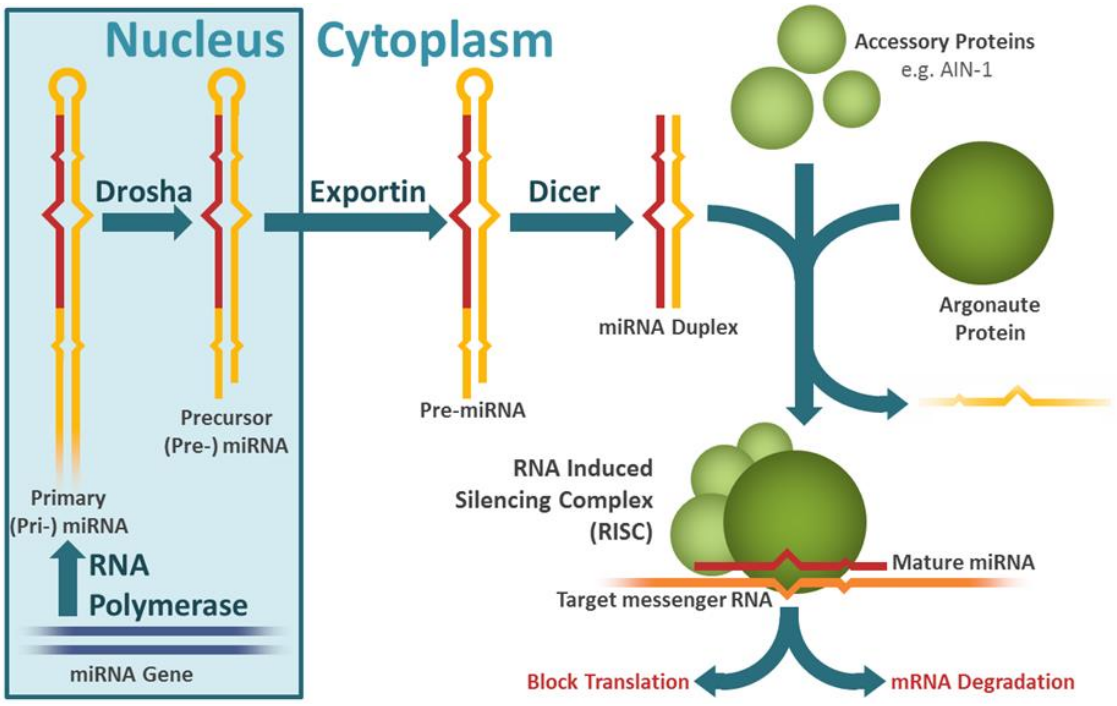


Fig. 1-3: miRNA biogenesis. Primary miRNAs are transcribed from miRNA genes and cleaved into shorter precursor miRNAs before being exported from the nucleus. A second cleavage step by Dicer liberates the double-stranded miRNA duplex which is acquired by a complex of proteins known as the RISC. One strand is removed, while the remaining mature miRNA is used to target specific mRNAs for post-transcriptional regulation. Translation of mRNAs is blocked before degradation is initiated.

1.2.3 The function of miRNAs

As miRNAs are able to target multiple genes, elucidating their exact function has proved challenging. Their effects can also be quite subtle, further complicating matters. While much of the early work on miRNA function has focused on their roles in *C. elegans* and *D. melanogaster*, in recent years there has been a large effort to discover their role in humans.

miRNAs in mammals

Much work has been done to discover the various roles of miRNAs in mammals, principally through human and mouse studies. Studies have linked miRNAs with both physiological and pathological conditions including cardiovascular disease (Small and Olson, 2011), obesity (Lin *et al.*, 2009), DNA damage recognition (Crosby *et al.*, 2009) and kidney development (Trionfini *et al.*, 2014). The largest area of research is oncology with 13,089 articles on the NCBI PubMed database, 2010 - 2014 (search terms: "miRNA" and "cancer"). Research has identified links with miRNAs in tumour formation, metabolism, metastasis and drug resistance (reviewed in Orellana and Kasinski, 2015).

Our knowledge of miRNAs in production animals (principally cows, sheep, pigs and chickens) is much more limited. Similar to miRNA research in general, the initial work has focused on quantifying temporal and spatial miRNA expression profiles. Our understanding of important targets and pathways is lacking. Most frequently, research has focused on economically important performance traits such as fat composition (Jin *et al.*, 2010; Wang *et al.*, 2012), muscle development (Clop *et al.*, 2006; Miretti *et al.*, 2013), fertility (Curry *et al.*, 2011; Torley *et al.*, 2011; Tripurani *et al.*, 2011) and embryonic development (Bannister *et al.*, 2009; Mondou *et al.*, 2012). In cows, there has also been a focus on immunological tissues (Coutinho *et al.*, 2007) including alveolar macrophages

(Vegh *et al.*, 2013). Studies have discovered that bacterial and viral infections cause widespread changes in miRNA expression profiles (Dilda *et al.*, 2012; Glazov *et al.*, 2009; Jin *et al.*, 2014).

miRNAs as therapeutic agents

With the range of potential regulatory systems that miRNAs are involved in, it is no surprise that they are being targeted heavily by research groups and pharmaceutical companies. The first miRNA therapies are already in human clinical trials, primarily covering cancer and infectious diseases (Hydbring and Badalian-Very, 2013).

The most developed therapy currently in clinical trials is an anti-Hsa-miR-122 therapy by Santaris Pharma A/S (Acquired by Roche in August 2014) under the title, Miravirsen. miR-122 is by far the most abundant miRNA in the human liver and targeting this miRNA should help reduce disease caused by chronic hepatitis C infection (Haussecker and Kay, 2010). Endogenously, miR-122 helps regulate hepatic cell identity and lipid metabolism (Chang *et al.*, 2004; Krützfeldt *et al.*, 2005). There are two miR-122 binding sites in a non-coding region of the hepatitis C virus (HCV) genome. The Miravirsen treatment involves administration of a sequence complementary to miR-122 composed of locked-nucleic acids (LNAs). These highly stable, synthetic nucleic acids bind to complementary sequences and, in the case of miR-122, prevent its use by HCV. In this manner, HCV replication is reduced, decreasing the viral load and clinical disease (Elmén *et al.*, 2008).

Regulus Therapeutics (San Diego, CA, USA) is also developing a similar anti-Hsa-miR-122 therapy (RG-101), as well as an anti-Hsa-miR-21 therapy (RG-012) for treatment of Alport disease. This disease is caused by mutations in three type IV collagen genes, resulting in chronic kidney, eye and ear disease. Research identified that in cases of renal

fibrosis, miR-21 is upregulated (Zhong *et al.*, 2011). By decreasing levels of miR-21, Regulus decreased the rate of fibrosis and increased the life-span of Col4a3$^{-/-}$ mice that display similar disease profile to humans with Alport disease (Gomez *et al.*, 2013). RG-101 and RG012 are expected to enter phase I and II clinical trials respectively during 2015.

The first miRNA mimic to enter clinical trials is MRX34, an Hsa-miR-34 replacement therapy developed by miRNA therapeutics Inc. (Austin, TX, USA). Loss of miR-34 has been identified in a wide range of cancers including breast, lung, skin and prostate (Lodygin *et al.*, 2008). MRX34 is currently in phase I clinical trial for the treatment of liver cancer. It is delivered via intravenous injection of liposomes containing the stabilised miRNA. miRNA Therapeutics are also developing therapies involving Hsa-let-7 and Hsa-miR-16.

miRNAs as biomarkers of disease and infection

As well as being used as therapeutic agents, miRNAs are being developed as biomarkers of disease. This is possible because miRNAs have been found in many animal body fluids (Silva *et al.*, 2015; Weber *et al.*, 2010). Linking differential expression (or presence/absence) of miRNAs in these fluids may provide useful tools to medics and veterinarians.

As miRNAs are critical to development in many organisms, it is not surprising that one particularly explored area of biomarker research is developmental disorders, particularly neoplasia (Takasaki, 2015; Zhang *et al.*, 2007a). In addition to detecting neoplasia miRNA biomarkers have also been linked to cancer prognosis (Hur *et al.*, 2015) and the response to treatment (Iqbal *et al.*, 2015). The role of miRNA biomarkers is also being explored in

non-neoplastic diseases such as diabetes mellitus (Chen *et al.*, 2014a), heart failure (Akat *et al.*, 2014) and autoimmune conditions (Carlsen *et al.*, 2013).

While the vast majority of miRNA biomarker studies have focused on human disease, there has been some exploration among veterinary species. This includes both companion animals (*e.g.* heart disease in dogs (Hulanicka et al., 2014; Steudemann et al., 2013)) and production animals (*e.g.* drug residues in cattle (Becker et al., 2011; Melanie et al., 2015)). Parasite-derived miRNAs have also been detected in the sera of mice infected with Schistosoma mansoni (Hoy et al., 2014), dogs infected with Dirofilaria immitis (Tritten et al., 2014) and nodule fluid of cattle infected with Onchocerca ochengi (Quintana et al., 2015). Ideally, parasite-derived miRNAs might provide a genus and species-level diagnostic tool, but much more work is required to evaluate the clinical robustness of such a test.

1.3 miRNAs in *C. elegans*

miRNAs were discovered in *C. elegans* and it remains an important organism for understanding interactions and functions. It is also a useful model for studying *H. contortus* as both belong to the nematode phylogenetic clade V (Blaxter *et al.*, 1998).

1.3.1 Overview of *Caenorhabditis elegans*

C. elegans is free-living, non-parasitic nematode with a temperate, cosmopolitan distribution that has been used in biological research for decades (Brenner, 1974). It is an exemplary model organism, being amenable to culture and manipulation (gross and molecular). It was the first multicellular organism to have its genome sequenced and a wealth of genetic, anatomical and physiological knowledge has been accumulated over the

years. It is also the first organism to have a complete map of its neuronal connections, known as a connectome (White *et al.*, 1986).

The genome of *C. elegans* is approximately 100 million bp, containing 20,470 protein coding genes (Wormbase, release WS227) as well as thousands of non-coding RNAs. There are also an estimated 13,300 RNA genes, the majority of which (12,000) are piwi-interacting RNAs (piRNAs, formerly known as 21U RNAs) (Ruby *et al.*, 2006). The remaining 1,300 RNA genes include transfer RNAs (tRNAs), ribosomal RNAs (rRNAs) and miRNAs (Stricklin *et al.*, 2005). There are currently 434 mature miRNA sequences across 245 genetic loci as recognised by miRBase, the leading datbase for miRNAs (Release 21, June 2014).

Life cycle of *C. elegans*

The life cycle of *C. elegans* (Fig. 1-4) is typical of other free-living nematodes. Eggs are released into the environment where they hatch into L1 and progress through four successive moults before becoming adults. During periods of stress such as over-crowding and food scarcity, L2 can enter a quiescent, environmentally resistant stage known as the dauer. If environmental conditions improve, dauers can re-enter the life cycle as L4 and become reproductive adults.

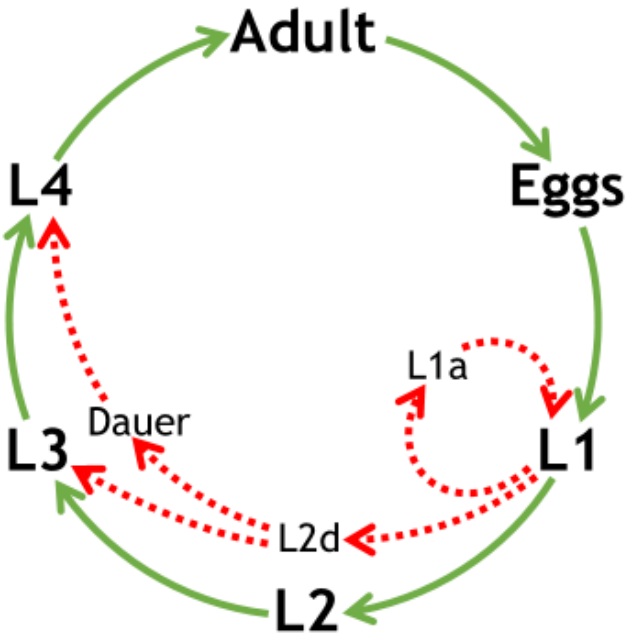


Fig.1-4: life cycle of *Caenorhabditis elegans*. Adult hermaphrodites release eggs that hatch into 1st stage larvae (L1). L1s then progress through four moults to become reproductive adults. If environmental conditions are poor, larvae can arrest development by entering two different quiescent states known as L1 arrest (L1a) and dauer diapause. Dauer development is preceded by a modified 2nd stage (L2) known as the pre-dauer (L2d). Resumption of development can occur as shown when environmental conditions improve.

Developmental arrest in *C. elegans*

Nematodes will often be confronted by periods when environmental conditions are unfavourable. To combat this, numerous species have evolved the ability to arrest development and wait for improved conditions when they can continue a reproductive life-cycle. In *C. elegans*, there are two distinct types of arrested development: the L1 arrest and dauer diapause.

L1 can detect unfavourable conditions and enter these two states. If eggs hatch in the absence of food, larvae will enter L1 arrest. Worms in L1 arrest do not display any morphological differences from standard L1s but possess increased environmental and stress resistance. If food becomes available, arrested L1 will resume development as normal. Any time spent during L1 arrest does not reduce the rate of maturation or life-

span of the worm; it is said to be an ageless state (Johnson *et al.*, 1984). L1 arrest is sometimes referred to as L1 diapause, however, as there is no preparation stage, this is considered an inaccurate term.

If food is present at hatching, but other environmental conditions are poor, worms can enter a diapause state known as the dauer. This is different from L1 arrest, which is simply arrested development. L1 that detect over-crowding, high temperature or limited (but still available) food will enter an alternate form of L2 known as the pre-dauer (L2d) stage (Golden and Riddle, 1984). At the end of L2d, worms can continue normal development (*i.e.* L3, L4 and adult stages) or become dauers.

Dauer larvae possess multiple behavioural, morphological and physiological modifications that promote survivability. They are radially constricted and possess a thickened cuticle with modified alae (Cassada and Russell, 1975). Movement is limited unless physically stimulated. An internal plug blocks the narrowed pharynx that does not pump (Riddle *et al.*, 1981). Fat stores acquired during the extended L2d stage provide energy through glycolysis (Riddle and Albert, 1997). Compared to the typical three week life-span of *C. elegans*, dauers can survive 3 – 6 months before perishing (Klass and Hirsh, 1976). However, the longer the dauer state, the more likely the worm will resume development (Golden and Riddle, 1984). Transcription is generally reduced, although the levels of several stress-resistance genes are increased, such as catalase (Vanfleteren and De Vreese, 1995), heat-shock protein 90 (Dalley and Golomb, 1992) and superoxide dismutase (Larsen, 1993).

One of the biggest triggers for worms to enter dauer diapause is the presence of dauer pheromone (Golden and Riddle, 1982), a diverse cocktail of compounds known as ascarosides (Butcher *et al.*, 2007). Pheromone primarily acts as a signal of population density which is fed into the environmental detection system that determines dauer

formation. However, pheromone also affects a variety of physiological and behavioural systems such as male attraction in hermaphrodites (Simon and Sternberg, 2002), hermaphrodite attraction in males (Izrayelit *et al.*, 2012), solitary or social feeding patterns (Macosko *et al.*, 2009) and olfactory plasticity (Yamada *et al.*, 2010). Environmental and developmental conditions have a drastic impact on the ascaroside composition of pheromone (Kaplan *et al.*, 2011). Production of ascarosides that promote dauer formation increase during unfavourable environmental conditions and decrease when favourable. Young adults produce a different composition of pheromone compared to larval stages. In general, ascaroside production is markedly reduced during the dauer state. Pheromone is highly stable, reducing the need for dauers to maintain the signal and use their limited energy stores. Recent evidence has found that the response to pheromone is not necessarily consistent between members of the same species (Diaz *et al.*, 2014). Distinct populations produce different combinations of ascarosides as well as responding differently to the same ascaroside. The authors suggest that this might be a form of subversive signalling between competing populations of genetically distinct groups. Forcing a different group of worms to enter the dauer state even when there is food provides obvious advantages that favour related genotypes.

Mutations affecting dauer formation (*daf*) can result in constitutive (*daf-c*) or deficient (*daf-d*) phenotypes. This has helped identify four key pathways involved in dauer formation: the guanlyl cyclase pathway, the TGF-β-like pathway, the insulin-like pathway and the steroid hormone pathway. Briefly, guanlyl cyclase receptors on the surface of sensory structures known as amphids relay environmental cues (such as the presence of dauer pheromone) through both the TGF-β-like and insulin-like pathways. Both pathways feed into the steroid hormone pathway, the final common target of which is DAF-12, a nuclear hormone receptor that is involved with development (including the dauer state) and longevity (Antebi *et al.*, 2000). In favourable conditions, activation of DAF-12 promotes

normal progression through the larval stages. Unfavourable environmental conditions prevent the formation of the DAF-12 ligand, inhibiting normal development and promoting dauer formation.

1.3.2 miRNAs and temporal development in *C. elegans*

Temporal development in *C. elegans* (*i.e.* the progression between different life-cycle stages) is well studied and miRNAs play an important role during many stages and decisions. The functions of a few specific miRNAs and miRNA families have been well studied (*e.g.* let-7, miR-35). However, many mutant strains of *C. elegans* lacking miRNAs possess no obvious phenotype under standard laboratory conditions (Brenner *et al.*, 2010).

Embryonic development

Two families of miRNA are essential for embryonic development in *C. elegans*: *mir-35* and *mir-51* (Alvarez-Saavedra and Horvitz, 2010). The *mir-35* family contains eight members (miR-35 – miR-42) while the *mir-51* family contains six (miR-51 – miR-56). Both families contain a high degree of redundancy as the embryonic lethal phenotype was only fully penetrative when all members of a family were absent. There were however variations in phenotypic severity when different combinations were missing.

Early larval development

The first miRNA to be discovered, lin-4 (Lee *et al.*, 1993), was identified as a heterochronic gene essential for temporal development of post-embryonic, early-larval development. Heterochronic genes regulate developmental timings and mutations can

cause cells to enter the next phase earlier (precocious) or later (retarded) than normally expected (Ambros and Horvitz, 1984). One lin-4 target is *lin-14* which encodes a putative transcription factor required for correct developmental timing of cell divisions during L1 and L2. It is known to regulate development of hypodermal, intestinal, neuronal and vulval cell types (Chalfie, 1981). In wildtype (wt) development (Fig. 1-5) levels of LIN-14 protein are high at hatching, while levels of lin-4 miRNA are low. During L1 stage, levels of lin-4 increase causing levels of LIN-14 to decrease through RISC-based degradation of *lin-14* mRNA.

Loss-of-function (*lf*) mutations in *lin-4* cause specific cells to reiterate early larval stages (*i.e.* retarded cell fate), while gain-of-function (*gf*) mutations results in these cells entering a later larval stage before other cells (*i.e.* precocious cell fate). Phenotypically, *lin-4 lf* mutants lack many adult structures, such as the vulva which prevents egg-laying, while *lin-4 gf* mutants are small and poorly formed when compared to the *wt*. As lin-4 negatively regulates lin-14, they share opposing phenotypes, *i.e.* a *lin-4 lf* mutant appears similar to a *lin-14 gf* mutant and vice versa. There are subtle differences as lin-4 regulates other targets, such as *lin-28* and *hbl-1*, two further examples of heterochronic genes (Moss *et al.*, 1997).

While it is clear that miRNAs can prevent a target's mRNA from generating further protein, little is known of how the current protein is removed. One line of evidence is the role of CED-3, a cysteine-aspartate protease (caspase) normally associated with apoptosis (Shaham *et al.*, 1999). Expression of CED-3 is greatest during early larval stages and it can decrease levels of LIN-14 and LIN-28 through proteolytic inactivation (Weaver *et al.*, 2014).

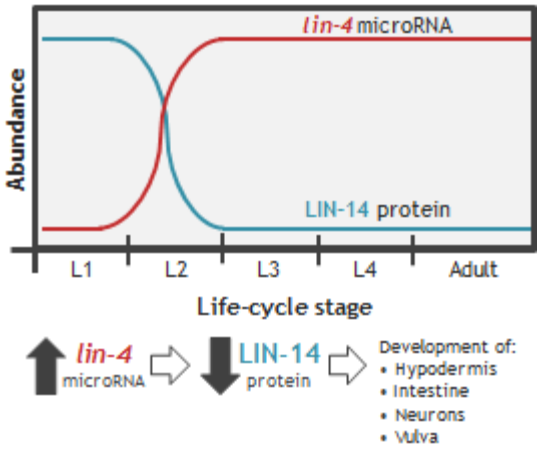


Fig. 1-5: miRNA regulation during early larval stages of *C. elegans*. As levels of the mature miRNA lin-4 increase in late L1, levels of lin-14 mRNA decrease during L2. This results in the developmental progression of certain cells within the hypodermis, intestine, vulva and nervous system. Abundance values are arbitrary representations. Based upon diagram from wormbook (Vella and Slack, 2005).

Late larval development

The second miRNA discovered, let-7, was identified through *C. elegans* mutant suppressor screening (Reinhart *et al.*, 2000). let-7 is expressed during late larval stages in hypodermal seam cells where it regulates developmental timing of the L4/adult moult (Johnson *et al.*, 2003). A null mutation of let-7 results in a lethal (*let*) phenotype due to incorrect development of the cuticle: adults burst through the vulva and die.

In the wildtype worm, levels of let-7 miRNA increase during late L3 stage, peaking during the L4 stage (Fig. 1-6). This reduces abundance of the lin-41 mRNA and subsequently levels of LIN-41 protein decrease. LIN-41 is a ring finger-b box-coiled coil (RBCC) protein that represses LIN-29, an adult-specific zinc-finger transcription factor (Rougvie and Ambros, 1995). When the repression is removed, the LIN-29 protein activates genes that promote alae formation, cuticle synthesis and cessation of moulting (Bettinger *et al.*, 1996). In the absence of let-7, LIN-41 levels remain high, repression of *lin-29* remains and hypodermal development is impeded, resulting in the lethal phenotype. Further research has identified that downregulation of *hbl-1* (Abrahante *et al.*, 2003; Lin *et al.*, 2003a) and

daf-12 (Grosshans *et al.*, 2005) by let-7 play an important role in this larval-to-adult switch.

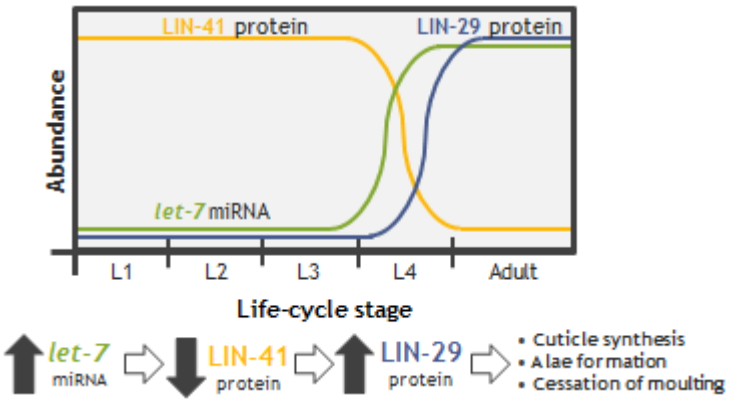


Fig. 1-6: miRNA regulation during the L4/adult moult in *C. elegans*. As levels of the mature miRNA *let-7* increase during late L3, levels of the LIN-41 protein decrease. This removes repression on the LIN-29 protein, promoting cuticle synthesis, alae formation and the cessation of moulting. Abundance values are arbitrary representations. Based upon a diagram from wormbook (Vella and Slack, 2005).

Mid larval development

While great detail is known about the regulation of early and late larval development in *C. elegans* by the miRNAs lin-4 and let-7 respectively, little is known of regulation during mid larval stages. We do know that three miRNAs are involved in the L2 to L3 developmental transition: miR-48, miR-84 and miR-241 (Abbott *et al.*, 2005). These miRNAs are classified as belonging to the *let-7* family due to sequence conservation of the seed sequence (nucleotides 2 – 8) (Lim *et al.*, 2003). They act redundantly to suppress *hbl-1*: morphological defects in development require all three to be absent. These defects are seen earlier in development than in mutants of let-7 itself, suggesting a similar but non-overlapping role for members of this miRNA family. The L2-to-L3 transition is also regulated by LIN-28 (Ambros and Horvitz, 1984), but miR-48, miR-84 and miR-241 do not affect its expression. Instead, their regulation occurs in parallel and provides another example of miRNAs providing robustness to developmental timing in *C. elegans*.

1.4 miRNAs in parasitic nematodes

The miRNAs of various human and veterinary parasites have also been examined, although knowledge of their functions is still limited (Britton *et al.*, 2014; Liu *et al.*, 2010; Matrajt, 2010). Their potential role in drug resistance is of particular interest in light of expanding anthelmintic resistance (Devaney *et al.*, 2010).

1.4.1 Overview of miRNAs known in parasites

Most of the miRNA research in parasitic nematodes (and parasites in general) is at an early stage. Identifying endogenous miRNAs and profiling their expression has been the major aim of research thus far. An obvious progression is therefore to identify miRNA function, both within the parasite itself and in potential host-parasite interactions. Table 1-2 lists the miRNA complement from a number of human and veterinary parasites. The majority of these figures are likely to be revised in the future as additional life-cycle stages are studied and more complete genomic data become available. Different sequencing techniques and bioinformatic approaches also affect the stringency to which a putative miRNA might be classified as 'real'.

Table 1-2: miRNA complement of parasitic helminths.

Parasite	Host(s)[1]	miRNAs[2]	Reference
Angiostrongylus cantonensis	Rats	631	(Li *et al.*, 2014)
Ascaris lumbricoides	Humans	171	(Shao *et al.*, 2014)
Ascaris suum	Pigs	97	(Wang *et al.*, 2011a)
Baylisascaris schroederi	Giant panda	108	(Zhao *et al.*, 2013)
Brugia malayi	Humans	145	(Poole *et al.*, 2014a)
Brugia pahangi	Dogs, cats	104	(Winter *et al.*, 2012)
Bursaphelenchus xylophilus	Pine trees	57	(Huang *et al.*, 2010)
Dirofilaria immitis	Dogs, Cats	1,076	(Fu *et al.*, 2013)
Echinococcus granulosus	Dogs	38	(Cucher *et al.*, 2011)
Eurytrema pancreatium	Ruminants	27	(Xu *et al.*, 2013a)
Fasciola gigantica	Ruminants	19	(Xu *et al.*, 2012)
Fasciola hepatica	Ruminants	16	(Xu *et al.*, 2012)
Haemonchus contortus	Sheep	192	(Winter *et al.*, 2012)
Schistosoma japonicum	Multiple	78	(Cai *et al.*, 2013)
Schistosoma mansoni	Humans	112	(Marco *et al.*, 2013b)
Taenia saginata	Cows	11	(Ai *et al.*, 2012)
Taenia multiceps	Sheep	1,026	(Wu *et al.*, 2013)
Trichinella spiralis	Humans	24	(Chen *et al.*, 2011)

[1] Primary host(s)

[2] Mature miRNA sequences

1.4.2 miRNAs and parasite-host interactions

Our expanding knowledge of miRNAs in both parasites and their hosts has generated some exciting research in the area of parasite-host interactions. The rat lungworm, *Angiostrongylus cantonensis*, causative agent of human eosinophilic meningoencephalitis, is able to alter local miRNA expression in its host (Yu *et al.*, 2014). One miRNA in particular, miR-146 is known to be involved in astrocyte inflammatory responses (Iyer *et al.*, 2012). Circulating levels of miR-146 also provided a robust biomarker of infection (Chen *et al.*, 2014b). This host stress response due to parasite infection is also seen in soybean plants during soybean cyst nematode infection when 20 miRNAs show significant variation (Li *et al.*, 2012b).

Evidence has accrued over the past few years that parasites actively release miRNAs into the host environment (Hoy and Buck, 2012; H. Gu, unpublished data). This practice is associated with a wide range of nematode (Tritten *et al.*, 2014a, b) and trematode parasites (Bernal *et al.*, 2014). Their exact function remains unclear, but evidence suggests a role in modulating the host immune response to facilitate parasite survival.

Research into *Heligmosomoides polygyrus*, a nematode found in the small intestines of rodents, identified miRNAs associated with Argonaute were released from the parasite's gut within exosomes and allowed for transfer of parasite miRNAs into host cells (Buck *et al.*, 2014). *In vitro* work identified two targets: interleukin-33 receptor (*Il33r*) and dual specificity protein phosphatase 1 (*Dusp1*), genes involved in immune activation and regulation (Hammer *et al*., 2006; Smith and Maizels, 2014).

1.4.3 miRNAs in *Haemonchus contortus*

miRNAs play a crucial role in regulating many developmental processes (Wienholds and Plasterk, 2005), with much of the evidence coming directly from *C. elegans* (Ambros, 2011). *C. elegans* and *H. contortus* are relatively closely related, both belonging to the nematode phylogenetic group clade V (Blaxter *et al.*, 1998). We have therefore speculated that miRNAs are also involved with regulating development in *H. contortus*.

A comprehensive set of miRNAs in *H. contortus* was first identified in our lab by Winter et al (2012). Deep sequencing was performed on two small RNA libraries: L3 and mixed sex adult worms, with the resulting data aligned to the early scaffolds of the *H. contortus* genome, which has now been completed (Laing *et al.*, 2013). From the alignment, two computer algorithms, miRDeep (Friedländer *et al.*, 2008) and MIREAP (Chen *et al.*, 2009), were used to identify 180 unique miRNAs. Homology-based BLAST queries of miRBase (release 15) identified an additional 12 miRNAs, giving a total of 192 miRNAs. Further work with miRBase has expanded this to 195 officially recognised, mature miRNAs in *H. contortus* (miRBase, release 21).

The search for *H. contortus* miRNAs highlighted that the majority are unique and not found in any other species: 135 mature sequences (71 %) are species-specific. While higher than in *C. elegans* (53 %), the *C. elegans* genome has been studied in much greater detail with a wider range of tools. Other members of the *Caenorhabditis* genus (*e.g. C. briggsae, C. remanei*) have also been studied and found to contain miRNAs originally considered unique to *C. elegans*. If other *Haemonchus* species were to be studied, it is likely that some miRNAs would be reclassified as genus-specific. *H. placei* is currently being sequenced by the Sanger Institute under the 50 Helminth Genomes Initiative (http://www.sanger.ac.uk/research/initiatives/globalhealth/research/helminthgenomes/). The other *Haemonchus* species have little economic impact and are unlikely to be a focus

of study in the immediate future. As the sequencing efforts for other genera of parasitic nematodes expands (*e.g. Teladorsagia*, *Ostertagia*, *Nematodirus*, *etc.*), the number of unique miRNAs in *H. contortus* is likely to decrease even further. This will also help us separate miRNAs with a role in general nematode biology from parasite-specific miRNAs and also identify miRNAs conserved within specific nematode families.

There is also a large degree of conservation in the position of miRNAs within the genome. For example, miR-50 was located in a homologous gene in three different nematode species: *H. contortus*, *C. elegans* and the filarial parasite *Brugia malayi*, even though there is little microsynteny between these species (Winter *et al.*, 2012). This may indicate a conservation of function for this miRNA that extends back approximately 350 million years to the last common ancestor of these three nematodes. miR-50 is only found in nematodes, and is likely to be involved in a nematode-specific function. As it is conserved in *C. elegans*, its function is unlikely to be directly involved in parasitism.

Sequence abundance from RNA sequencing (RNAseq) suggested that many miRNAs, including novel and highly-conserved miRNAs, were abundantly expressed and this was confirmed using quantitative PCR (Winter *et al.*, 2012). Significant variation in expression levels across the different life-cycles (L3 and adult stages) under examination were also identified. Such variation may reflect stage-specific regulation of genes, possibly involved in development. Studying these miRNAs will expand our understanding of fundamental *H. contortus* biology and highlight potential targets for therapeutic intervention.

1.5 Aims

miRNAs are essential in regulating genes involved in the growth and development of *C. elegans* and it seems likely that they possess a similar role in *H. contortus*. The overall aims of this thesis were to identify temporally expressed miRNAs and their potential targets to further our understanding of parasite development. The specific aims were to:

- Identify variation in miRNA expression across larval and adult life-cycle stages of *H. contortus*.
- Predict potential targets of temporally and spatially variant *H. contortus* miRNAs using *in silico* prediction tools and biochemical pulldown assays and to identify putative homologues.
- Examine possible functions and target genes of specific developmentally regulated miRNAs of *H. contortus* and investigate functional conservation with *C. elegans*.
- Investigate the impact of miRNA inhibitors and mimics on *H. contortus* L3/L4 development *in vitro*.

2 Materials and Methods

2.1 Reagents, media and oligonucleotides

The appendix (Table 8-1) lists the basic chemicals, buffers and medium used during this project in addition to oligonucleotides used during PCR, qRT-PCR and molecular cloning (Table 8-2). Oligonucleotides used as miRNA mimics are listed in Table 8-3. General methods are described below and detailed methods specific to each chapter are described at the start of the chapter.

2.2 *Caenorhabditis elegans* materials and methods

2.2.1 Culture and maintenance of *C. elegans*

Solid medium growth

The majority of *C. elegans* strains used during this project (Table 6-1) were obtained from the Caenorhabditis Genetics Center (CGC funded by NIH Office of Research Infrastructure Programs (P40 OD010440)). All strains were grown and maintained on 5 cm diameter, triple-vented petri dishes containing NGM (nematode growth medium) agar (Appendix, Table 8-1) incubated at 20°C, unless otherwise stated. Plates were seeded with a thin 'lawn' of *Escherichia coli* OP50 using 200 µL of an overnight culture grown in LB-medium (Appendix, Table 8-1). Worms were transferred to new plates by 'chunking' a section of agar from the old plate or by transferring a small number (~10) of gravid adult hermaphrodites to a fresh plate using a platinum worm-pick.

Liquid medium growth

C. elegans strains were grown in liquid culture when large numbers of worms were required. Nearly starved NGM agar plates of *C. elegans* were washed using 2 mL S-medium (Appendix, Table 8-1) per plate and worms added to S-medium in a conical flask. Due to the extended growth periods required for the growth of dauer larvae, baffled flasks were used to maintain oxygenation. For all stages, 100 – 500 mL S-medium was used in a flask at least four-times the volume (*i.e.* 400 – 2,000 mL) and a concentrated bolus of *E. coli* OP50 added as a food source. This was prepared from an overnight culture grown in superbroth medium (Appendix, table 8-1) at 37 °C, centrifuged at 10,000 x g for 10 minutes and washed twice in PBS (Appendix, table 8-1). The final bacterial pellet was resuspended in 5 % of the original culture volume of S-medium and stored at 4 °C. For each 100 mL of *C. elegans* culture, 10 mL concentrated OP50 was added. Cultures were checked daily for signs of contamination and abundance of food.

To purify worms from a liquid culture, sucrose flotation and/or Ficoll precipitation were performed, adapted from an online protocol (http://diamond.tuebingen.mpg.de/mediawiki/index.php/Dauer_purification_by_Sucrose_Floatation_and_Ficoll_precipitation). When dauers were required, worms were washed with dH_2O. All other stages were washed with M9 buffer (Appendix, table 8-1). The liquid culture was centrifuged at 1,000 x g for five minutes and the worm pellet washed three times in M9/dH_2O. The final pellet of worms was transferred to a 50 mL centrifuge tube and resuspended in M9/dH_2O to a volume of 7 mL. 7 mL of a 60 % sucrose solution was added and mixed by inverting the tube five times. The mixture was centrifuge at 50 x g for one minute followed immediately by a spin of 1,150 x g for three minutes. Debris and dead worms were found in bottom of the tube, below the sucrose. The live worms floated above the sucrose in a yellow band which was gently aspirated using a glass Pasteur pipette and transferred to a fresh 50 mL tube. Worms were centrifuged at 1,650 x g for 5 minutes and washed three times in M9/dH_2O.

For dauer larvae, a further Ficoll precipitation was performed. A 15 % (w/v) solution of Ficoll-400 (GE Healthcare) in 0.1 M NaCl was used in a 2:1 ratio (Ficoll:worms). The final worm pellet generated during sucrose flotation (after washing) was resuspended in 15 mL dH_2O. 30 mL Ficoll solution was added to a fresh 50 mL conical centrifuge tube before worms were very slowly layered on top using a glass Pasteur pipette. The tube was allowed to settle for 10 minutes before being centrifuged at 300 x g for 10 minutes. Debris and dead worms remained on top of the Ficoll while live dauers were found pelleted at the bottom of the tube. The live worm pellet washed three times in dH_2O (centrifuged at 1,000 x g for five minutes). Dauer larvae were allowed to recuperate overnight in 200 mL M9 in a 500 mL conical flask, incubated at 20 °C in an orbital shaker (150 rpm) before being centrifuged at 1,000 x g for five minutes. The resulting worm pellet was stored at – 20 °C.

2.2.2 Storage of *C. elegans*

For short-term storage (2 – 3 weeks), a single, nearly-starved plate of worms was incubated at 15°C instead of 20°C until required. For long-term storage (months – years) a single, nearly-starved, mixed-population plate of worms was washed in 2 mL M9 buffer and 2 mL freezing solution (Appendix 8-1). This solution was then distributed amongst 1.5 mL microcentrifuge tubes and frozen gradually at -80°C.

2.2.3 Generating synchronised populations of *C. elegans*

Age-matched worms were obtained by allowing gravid adults to lay eggs on standard NGM agar plates over the course of 2 – 4 hours before being removed. Eggs were incubated at 20 °C for 2/3 days (depending on the strain involved) until larvae had developed to L4. 50 individual worms were transferred to M9, centrifuged briefly (1,000 x

g, 10 seconds) and excess M9 removed. The resulting worm pellet was stored at -80 °C until required.

In situations requiring large numbers of age-matched worms, synchronisation was achieved through bleaching of adult worms, based upon methods described in Wormbook methods (Stiernagle, 2006). Gravid hermaphrodites were washed from plates using 1 mL sterile distilled water (dH_2O) and transferred to a 15 mL centrifuge tube before adding 3.5 mL sterile dH_2O, 0.25 mL 5 N potassium hydroxide (KOH) and 0.5 mL 10 % sodium hypochlorite (NaOCl) (freshly prepared). Tubes were vortexed for five seconds every two minutes until worms were lysed (approximately 10 minutes). Lysis was confirmed by microscopic examination before centrifuging worms at 1,300 x g for one minute and washing eggs three times in dH_2O. Eggs were left to hatch overnight at 20 °C on unseeded NGM agar plates. Synchronised L1 were centrifuged at 1,150 x g for three minutes, washed twice in M9 and transferred to an OP-50 seeded, 5 cm NGM agar plate. When worms grew to the desired stage (e.g. 37 hours for mid-L4 worms), they were washed from plates in 1 mL M9 before being centrifuged at 1,000 x g for three minutes and washed twice in M9. The final pellet was stored at -80 °C until required.

2.2.4 Microscopic imaging

C. elegans (and *H. contortus*) were examined using either a dissection or high-powered (Zeiss Axioskop 2 Plus) microscope. Worms were mounted on 2 % agarose pads in a 10 µL drop of M9 containing 10 mM levamisole to inhibit movement. For high-powered microscopy, images were taken using a Hamamatsu ORCA-ER camera controller attached to an Apple Macintosh running Improvision Openlab (version 5.5) for image capture. For fluorescent imaging a Netz Power ebq-100 UV source was used.

2.3 *Haemonchus contortus* materials and methods

2.3.1 Obtaining *H. contortus* worms

All *H. contortus* material was obtained from the Moredun Institute (Edinburgh, Scotland, UK) where the life-cycle is maintained through infection of sheep using the (anthelmintic susceptible) MHco3(ISE) strain. L3 were obtained from eggs that developed after being processed from faecal material using the Baermann technique. L4 and adult stages of *H. contortus* were obtained from the abomasum of artificially infected sheep post-mortem, 7 and 21 days post-infection respectively. For L4, the abomasum was incubated at 38°C for 3 – 5 hours in physiological saline (0.85 % NaCl w/v), releasing larvae from the mucosa that were rinsed and snap frozen in liquid nitrogen. Adult worms were sexed using a combination of gross observations and microscopic examination. Worms were snap-frozen and stored in liquid nitrogen.

2.3.2 Culturing *H. contortus* L3 *in vitro*

H. contortus L3 were exsheathed with bleach or CO_2. For bleaching, 1,000 – 3,000 L3 were centrifuged at 1,000 x g for one minute and washed twice in PBS before being resuspended in 180 µL PBS and 5 µL sterilising fluid (Milton, Bournemouth, UK) in a 1.5 mL microcentrifuge tube. Tubes were incubated horizontally at room temperature and vortexed for five seconds every minute. Exsheathment was checked microscopically, usually occurring around 8 – 10 minutes after exposure to bleach. Larvae were then washed three times in PBS followed by twice in Earle's balanced salt solution (EBSS; with calcium, magnesium and phenol red; Life Technologies) containing penicillin (100 IU/mL), streptomycin (100 µg/mL) and amphotericin B (125 µg/mL) (pen/strep/ampho). Larvae were cultured at 37°C in a 5% CO_2 atmosphere for 7 – 10 days.

For CO_2 exsheathment, 10,000 L3 were centrifuged at 1,000 x g for one minute and washed twice in PBS before being resuspended in 5 mL EBSS + pen/strep/ampho in a 50 mL centrifuge tube. Larvae were incubated at 40°C for 15 minutes with 60 rpm rotation then exposed to vigorous CO_2 bubbling using a glass tube capped with foamed glass. The tube was then sealed with laboratory film and incubated at 40 °C for 90 minutes with 100 rpm rotation. Exsheathment was checked microscopically and cultured as described previously.

2.3.3 Extracting *H. contortus* adult-female gut tissue

Live adult female worms were washed several times in Hank's balanced salt solution (HBSS, Life Technologies), then transferred to chilled PBS (4 °C) in a 10 cm petri dish. Under a dissecting microscope, individual worms were manipulated using two 23-gauge needles (Fig. 2-1). An initial incision was performed anterior to the vulva which allowed the worm's organs (principally the gut and the ovaries) to exit the body due to internal pressure. A loop of gut was carefully dragged from the body as far as possible. The 'head' of the worm was removed immediately anterior to the anterior extent of the ovaries. The posterior end of the gut was cut where it entered the body at the vulval incision. Once liberated, the loop of gut was gently washed in PBS to remove any adherent eggs and transferred to TRIzol® (Life Technologies) at room temperature. The gut tissue from 20 adult female worms were added to 1 mL of Trizol and stored overnight at 4 °C. RNA extraction was performed as described in Chapter 2.4.1.

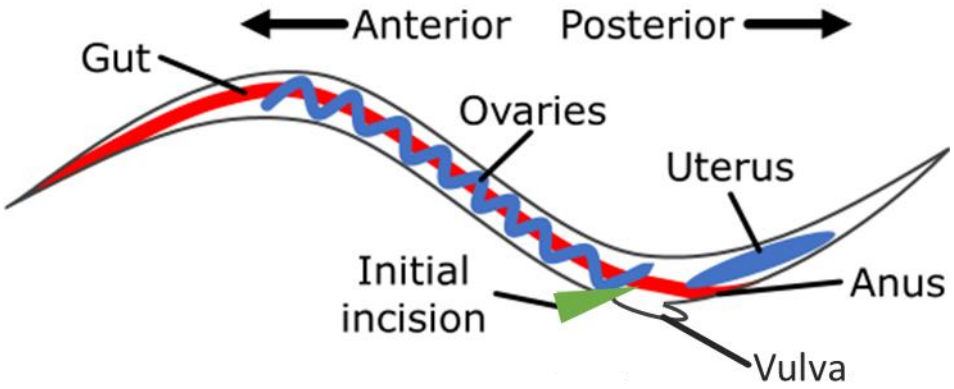


Fig. 2-1: Extraction of gut tissue from adult female *H. contortus*. Structures visible under a dissection microscope are annotated. The initial incision site is marked in green.

2.4 General molecular biology techniques

2.4.1 Total RNA extraction

Worm samples were crushed and ground to a fine powder using a mortar and pestle, pre-chilled with liquid nitrogen. 1 – 2 mL of TRizol® was then added and RNA extracted using a high-stringency version of the manufacturer's instructions. RNA pellets were resuspended in 20 µL DEPC-treated dH_2O and incubated at 50 °C for five minutes. The quantity of RNA in each sample was obtained by measuring the optical density (wavelength 260/280 nm) while quality was visually checked using agarose gel electrophoresis by running 0.5 – 1.0 µg on a 1.0 % gel in TAE buffer. RNA was stored at -80 °C until required.

2.4.2 cDNA synthesis

Contaminating genomic DNA was removed from 1 µg total RNA using a DNase cleanup kit (Invitrogen) following the manufacturer's instructions. Samples were stored at -80 °C until required. RNA was used to generate unbiased or miRNA-biased cDNA using two different synthesis kits. For the study of mRNAs, 1 µg DNase treated total RNA was

converted to cDNA using the AffinityScript multiple temperature cDNA synthesis kit (Agilent Technologies Ltd, Edinburgh, Scotland, UK) following the manufacturer's instructions. For the study of miRNAs, 1 µg DNase treated total RNA was converted to cDNA using a miRNA 1st-strand cDNA synthesis kit (Agilent Technologies), following the manufacturer's instructions. In both situations, an additional reaction was generated for each sample that lacked the reverse transcription enzyme. These RT-negative samples were used as controls during qRT-PCR.

2.4.3 Standard polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) amplification of DNA was performed using Promega GoTAQ® G2 Flexi DNA polymerase kit (M7805) following the manufacturer's instructions. Standard thermocylcer settings involved an initial denaturation at 95 °C for two minutes, followed by 25 – 35 cycles of denaturation (95 °C, 30 seconds), annealing (54 – 60 °C, 30 seconds) and extension (72 °C, 30 – 90 seconds) before a final extension of 72 °C for five minutes and holding period at 4 °C until products were removed. For more stringent amplification, a proof-reading Phusion® high-fidelity DNA polymerase kit (New England Biolabs, MA, USA) was used according to the manufacturer's instructions.

2.4.4 Quantitative polymerase chain reaction (qRT-PCR)

qRT-PCR using standard cDNA

Quantitative real-time PCR (qRT-PCR) amplification of unbiased (total RNA) cDNA was performed in 96-well plates using the Agilent Stratagene Mx3005P spectrofluorometric thermal cycler and Brilliant III Ultra-Fast SYBR® Green QPCR Master Mix (Agilent Technologies), following the manufacturer's instructions and using 5 ng/µL cDNA per

well. For *H. contortus* and *C. elegans*, β-tubulin isotype 1 was used as a normaliser gene (primer sequences can be found in Appendix, Table 8-2), while cDNA from the L3 stage was used as the calibrator. On a single plate, each primer-cDNA combination was repeated in triplicate, while no-template (*i.e.* no cDNA) and no-RT (reverse transcriptase) controls were repeated in duplicate. Using a comparative quantification reaction template, ROX and SYBR fluorescence data was recorded by the thermal cycler, with ROX acting as the reference dye. During the reaction, an initial 10 minutes at 95 °C was followed by 40 (2 step) cycles of 15 seconds at 95 °C and 20 seconds at 60 °C. After amplification, a standard dissociation curve was calculated. Analysis was performed using the MXPro QPCR software (Agilent Technologies).

qRT-PCR using miRNA-biased cDNA

Quantitative real-time PCR (qRT-PCR) amplification of miRNA-biased cDNA was performed in 96-well plates using the Agilent Stratagene Mx3005P spectrofluorometric thermal cycler and miRNA QPCR master mix (Agilent Technologies), following the manufacturer's instructions and using 1.0 ng cDNA per well. Hco-miR-5899-3p and Cel-miR-72-5p were used as normalisers in their respective species (primer sequences can be found in appendix, Table 8-2), while cDNA from the L3 stage was used as the calibrator in both species. On a single plate, each primer-cDNA combination was repeated in triplicate, while no-template (*i.e.* no cDNA) and no-RT (reverse transcriptase) controls were repeated in duplicate. Using a comparative quantification reaction template, ROX and SYBR fluorescence data was recorded by the thermal cycler, with ROX acting as the reference dye. During the reaction, an initial 10 minutes at 95 °C was followed by 40 (3 step) cycles of 10 seconds at 95 °C, 15 seconds at 60 °C and 20 seconds at 72 °C. After amplification,

a standard dissociation curve was calculated. Analysis was performed using the MXPro QPCR software (Agilent Technologies).

2.4.5 Agarose gel electrophoresis

Products of standard PCR amplification were examined using gel electrophoresis on 1 % (for products greater than 200 bp) or 3 % (for products less than 200 bp) agarose gels. Visualisation was carried out using 0.0025 % v/v Safeview nucleic acid stain (NBS Biologicals, Huntingdon, UK) added to agarose melted in TAE (Appendix, Table 8-1). Gels were run at 100 V for 25 – 35 minutes and imaged using the GelDoc Imaging system (Biorad) using variable exposure settings.

2.4.6 Restriction enzyme digestion of DNA

Restriction enzymes were purchased from NEB and digestions carried out according to the manufacturer's instruction. 1 µg DNA and 5 U restriction enzyme were combined in a total volume of 50 µL in a thin-walled PCR tube. Reactions were incubated at 37 °C for one hour before being heat inactivated at 65 °C for 10 minutes. When two restriction enzymes were needed, compatible enzymes were combined in one reaction (double digestion) or performed one after the other (serial digestion).

2.4.7 Molecular cloning of DNA

DNA to be cloned was amplified from a whole organism or cDNA sample by PCR using a proof-reading Phusion® High-fidelity DNA polymerase (NEB). Amplified DNA was isolated from agarose gels using the QIAquick gel extraction kit (Qiagen) following the

manufacturer's instructions. DNA was eluted in 26 µL of dH_2O and stored at -80 °C. To check isolation quality, 3 µL of DNA was combined with 2 µL triple loading dye (Promega) and run on a 1 % agarose gel.

Depending on the amount recovered, 1 – 5 µL of eluted DNA was combined with 1 µL 2 mM dATP and 1 µL Taq DNA polymerase (Promega) and incubated at 72 °C for seven minutes to generate A-overhangs. Once tailed, the product was cloned into the TOPO® cloning system (Invitrogen) following the manufacturer's instructions. Transformed TOP-10 *E. coli* were then inoculated onto two LB plates containing 0.01 % ampicillin, coated with 0.1 mM IPTG and 100 mM X-Gal and incubated at 37 °C overnight.

White colonies were checked for insertion of the DNA fragment by lysing individual white colonies in 100 µL dH_2O and incubating at 95 °C for five minutes followed by PCR amplification using TOPO-vector specific primers (Appendix, Table 8-2). Colonies containing an insert of the correct sizes were restreaked onto LB+Amp agar plates and incubated at 37 °C for six hours before being used to inoculate 10 mL LB medium and incubated at 37 °C overnight in an orbital shaker with 200 rpm rotation. Plasmid DNA was recovered from overnight cultures using a Plasmid Miniprep kit (Qiagen) following the manufacturer's instructions. DNA was eluted in 50 µL dH_2O and stored at -80 °C.

2.5 General protein analysis techniques

2.5.1 Sodium-dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated by 1-dimensional SDS-PAGE using the Mini-PROTEAN® Tetra Cell vertical electrophoresis system and pre-cast 4 – 15 %, 10-well Mini-PROTEAN® TGX™ gels (Biorad). 10 µL Precision Plus Protein™ Dual Color Standard (Biorad) was

used as a marker of molecular weight. Gels were run at 200 V until the dye front almost reached the bottom of the gel and stained (Coomassie Blue or Silver) or used in Western blotting.

2.5.2 Coomassie Blue staining of polyacrylamide gels

Polyacrylamide gels were often visualised using a 0.1 % Coomassie R-250 staining solution (Appendix, Table 8-1), carried out at room temperature on a rocking platform (30 rpm). After electrophoresis, gels were washed in dH_2O and incubated in Coomassie staining solution for 1 – 2 hours. Stain was removed using a methanol-based destain solution (Appendix, Table 8-1) and incubated until the background became clear. Destain solution was replaced when saturated. For storage, gels were incubated in 25 mL gel stabilising solution (Appendix, table A-1) for five minutes, sandwiched between cellulose sheets and dried overnight in a SE 1210 drying frame (Hoefer Inc, MA, USA) according to manufacturer's instructions.

2.5.3 Silver staining of polyacrylamide gels

When a greater sensitivity was required for visualisation of proteins separated using SDS-PAGE, a Silver Stain Plus^{TM} kit (Biorad) was used, following the manufacturer's instructions. For storage, gels were dried in a drying frame as described previously (Chapter 2.5.2).

2.5.4 Western blotting of polyacrylamide gels

Antibody-based detection of proteins involved separation using SDS-PAGE followed by western blotting. Proteins were transferred to a nitrocellulose membrane using the Criterion^{TM} vertical electrophoresis tank (Biorad) and a Tris-glycine blotting buffer (Appendix, table 8-1) run at 100V for one hour using standard techniques. After transfer, the membrane was blocked (1 x 20 minutes), incubated with primary antibody (1 x 1 hour), washed (3 x 20 minutes) and incubated with a horse-radish perodixase (HRP)-conjugated secondary antibody (1 x 1 hour) before a final wash (3 x 10 minutes). All steps were carried out at room temperature. The blocking step, primary and secondary incubations and the primary wash used PBS-T (Appendix, Table 8-1) containing 5 % dried milk powder. The final wash used PBS-T containing no milk. The membrane was then incubated in enhanced chemiluminescent (ECL) substrate (Pierce) following the manufacturer's instructions before being covered in plastic-wrap and exposed to X-OMAT radiographic film (Kodak) for between two seconds to two hours in a dark room. The radiographic film was then developed in a Konica Minolta SRX-101A film processor.

2.5.5 Protein dot blots

Antibody-based detection of proteins was also performed on protein samples without prior SDS-PAGE separation. A grid (1.5 x 1.5 cm) was drawn on dry nitrocellulose blotting paper using pencil. Samples were diluted in dH_2O to generate a range of concentrations (1:1 – 1:100) and 1 µL of a single concentration was slowly pipetted onto the middle of an individual square and air-dried for 5 – 10 minutes before being washed once in PBS-T for five minutes. The membrane was then probed using ECL detection of primary and secondary antibodies as described in Chapter 2.5.4.

2.5.6 Determining protein concentration

Protein concentration was determined using either the Bradford assay or Qubit® fluorometer (Life Technologies). For the Bradford assay 1 – 50 μL of protein was added to 1 mL Bradford reagent (Sigma-Aldrich) in a 2 mL disposable cuvette and incubated at room temperature for five minutes before being analysed using a spectrophotometer (wavelength = 595 nm). Protein concentration was also calculated using the Qubit ® fluorometer following the manufacturer's instructions.

3 miRNAs in *Haemonchus contortus* and *Caenorhabditis elegans*

3.1 Introduction

miRNAs have emerged as important regulators of gene expression in a range of biological and pathological processes. Attempting to identify the functions of miRNAs in any organisms involves a number of steps. First, they must be identified. This is often performed using deep sequencing of small RNA libraries, bioinformatics or both, as was the case for *H. contortus* (Winter *et al.*, 2012). Next, it is important to quantify the level of miRNAs in multiple life-cycle stages or tissue. This chapter describe the use of microarrays to quantify miRNA abundance in *H. contortus* across larval and adult stages. Abundance was also quantified in gut tissue was isolated from adult females. This is the largest life-cycle stage of *H. contortus* and allowed relatively straight-forward isolation of gut material. Isolation of gut tissue from adult males or earlier stages was not possible using available equipment.

Development is a complex process involving many time-sensitive decisions and miRNAs represent one mechanism that regulates and coordinates these decisions (see Chapter 1.3.2). This relies not simply on which genes are being targeted by a miRNA, but when they are being targeted, *i.e.* temporal and spatial regulation. For example, the miRNA let-7 is required for the L4/adult moult in *C. elegans*. If it is not expressed at the correct stage and in the correct tissues, cuticle development is disrupted and the animal dies (Ecsedi *et al.*, 2015). In this way, quantifying temporal and spatial variation in expression is as important for understanding miRNA function as identifying targets.

While much work has gone into identifying miRNAs in parasitic nematodes (see Chapter 1.4), comparatively little is known regarding their expression profiles. As a model organism, developmental expression of miRNAs is well established in *C. elegans*, including larval stages (Kato *et al.*, 2009), the dauer decision (Karp *et al.*, 2011) and adult aging (Ibáñez-Ventoso *et al.*, 2006). Expression profiling has also highlighted the effect of environmental stressors on miRNA levels, such as chronic nicotine exposure (Taki *et al.*, 2014) and spaceflight (Xu *et al.*, 2014). Unfortunately, there are only a handful of datasets available for parasitic nematodes such as the pig roundworm *Ascaris suum* and human filarial worm *Brugia malayi*. For *A. suum*, expression profiles cover male and female adults (Xu *et al.*, 2013b) as well as embryogenesis and early larval stages (Wang *et al.*, 2011a). They found that in adults, approximately two-thirds of unique miRNAs were sex-specific, while the majority of conserved miRNAs were expressed by both sexes. In the zygote, 12 maternally-derived miRNAs complement the 15 expressed by the zygote after fertilisation. miRNA expression in *B. malayi* microfilariae and adult stages has been reported (Poole *et al.*, 2014b), while recent microarray work from our lab has characterised miRNAs in mosquito-derived L3 of *B. pahangi*, L3 isolated from a mammalian host (*i.e.* pre- and post-infection), L4 and adult stages (Winter *et al*., 2015). This identified one particular miRNA (miR-5364) which is significantly upregulated within 24 hours following infection of the mammalian host and may be important for development.

miRNAs can have important roles beyond development, such as effects on the host. This is particularly well studied in plant parasites such as the soybean cyst nematode *Heterodera glycines* (Alkharouf *et al.*, 2006; Ithal *et al.*, 2007), sugar beet cyst nematode *Heterodera schachtii* (Hewezi *et al.*, 2008) and pinewood nematode *Bursaphelenchus xylophilus* (Ding *et al.*, 2015). During infection, these nematodes express a wide-range of miRNAs that co-opt plant genes, most likely aimed at increasing parasite survivability. In animals, parasite-derived miRNAs have been found in host circulation during parasitic

infections with tissue helminths including *Dirofilaria immitis*, *Loa loa*, *Onchocerca ochengi*, *Onchocerca volvulus* and *Schistosoma mansoni* (Hoy *et al.*, 2014a; Tritten *et al.*, 2014a, b). Studies are beginning to illuminate the potential effects of these miRNAs on host biology. For example, the rodent parasite *Heligmosomoides polygyrus* was shown to release miRNAs that can target genes involved in innate immunity (Buck *et al.*, 2014). In *H. contortus*, miRNAs have already been identified and expression data was previously obtained for three stages: L3, adult males and adult female (A. Winter, unpublished data). In this chapter, *H. contortus* miRNA abundance was quantified in the L4 stage and gut tissue and combined with L3/adult data. miRNAs displaying temporal or spatial variation were identified for further study.

3.2 Materials and methods

3.2.2 Microarray analysis of *H. contortus* miRNAs through development and in gut tissue

Total RNA was extracted from L4 and adult female gut tissue as described previously (Chapters 2.3.3 and 2.4.1). 10 µg RNA was combined with 40 µL DEPC-treated dH_2O, 5 µL 3M sodium acetate (NaOAc) and 150 µL 100 % ethanol before being shipped on dry ice to LC Sciences (Houston, TX, USA) for miRNA microarray analysis. The RNA was applied to the same custom arrays used in the previous analysis of *H. contortus* miRNAs, containing 890 sequences: 609 putative *H. contortus* miRNA sequences, 232 *C. elegans* mature miRNA sequences and 49 internal control sequences designed by LC Sciences. Statistical analysis was performed by LC Sciences, including T-tests and analysis of variance (ANOVA) studies across all sample groups. L4 data (from duplicate biological samples) were combined with the previous microarray data provided by Dr. Winter (L3, adult males and adult females), for a total of four *H. contortus* life-cycle stages. miRNA

microarray data from isolated adult female gut tissue (using triplicate biological samples) were compared to the adult female group, generating data on spatial variation in miRNAs in adult females. Raw and normalised data were sent electronically by LC Sciences for further study.

Results for internal control sequences (n = 49) and *C. elegans* miRNAs (n = 232) were removed, except for *Cel-let-7* (3′ and 5′). The remaining *H. contortus* miRNAs were filtered by abundance (microarray signal > 500) before post-hoc analysis (Benjamini and Hochberg false discovery rate) was performed. miRNAs with a strong signal (signal > 500) and statistically significant (corrected p-value < 0.05) change in abundance were termed variant/differentially expressed, while those with a strong signal but statistically insignificant changes in abundance were deemed invariant. The remaining miRNAs with weak signals (< 500) were considered low expressors.

Raw L4 and gut expression data were combined with the L3 and adult data provided by Dr. Winter, normalised and subjected to hierarchical clustering using the matrix visualization and analysis platform GENE-E, version 3.0.240 (http://www.broadinstitute.org/cancer/ software/ GENE-E/index.html). Row distance was calculated using a one-minus Pearson correlator metric and complete linkage (farthest-neighbour approach). All other settings were default. A rooted dendrogram and three-colour heat-map (blue = low, white = 50%, red = high) of the clustering was generated. Groups were labelled according the life-cycle stage which displayed the highest level of expression.

3.2.2 Quantitative real-time PCR (qRT-PCR) confirmation of *H. contortus* miRNA expression

Six *H. contortus* miRNAs (miR-45, miR-60, miR-71, miR-87a, miR-228 and miR-235) were chosen to cover different expression patterns observed in the microarray data. Their abundance was measured across four life-cycle stages (L3, L4, adult male and adult female) using quantitative reverse-transcription PCR (qRT-PCR) as described previously (Chapter 2.4.4). Gut abundance was tested for three of these miRNAs (miR-45, miR-71 and miR87a). L3 samples were used as the calibrator, while the invariant miRNAs miR-50-5p and miR-5899-3p were used as normaliser miRNAs. Three biological replicates were used for all stages, except L4 where only two biological replicates were available. miRNA abundance values were output as relative quantities (dRn) where the calibrator stage (*i.e.* L3) was set to one.

3.2.3 Examining conservation of miRNAs in *H. contortus* and *C. elegans*

The miRNAs of *H. contortus* and *C. elegans* were analysed for conservation status by interrogating the latest version of miRBase (http://www.mirbase.org, release 20, June 2014). miRNA stem-loops were used instead of mature sequences to avoid confusion arising from multiple loci expressing the same sequence. Classification was based on the range of species that possessed a miRNA. If the miRNA was not found in any other species, it was termed 'unique'. In *C. elegans*, miRNAs were termed 'genus-specific' if they were found in any other *Caenorhabditis* species. For both species, if a miRNA was only found in other nematode species, it was labelled 'nematode-specific'. Finally, if a miRNA was found in any other phyla outside of nematodes, it was termed 'widely

conserved'. One miRNA (*lin-4*) was considered 'heminth-specific', but was included in the nematode-specific group for simplicity.

3.2.4 Comparing expression between conserved miRNAs in *H. contortus* and *C. elegans*

For 25 miRNAs, conserved in both *H. contortus* and *C. elegans*, *H. contortus* microarray data were matched against *C. elegans* qRT-PCR abundance data obtained from Kato *et al.* (2009). Comparisons were made across L3, L4 and two adult sexes: *H. contortus* males with *C. elegans* males and *H. contortus* females with *C. elegans* hermaphrodites. Abundance profiles were graphed on independent vertical axes to examine similarities.

3.3 Results

3.3.1 Temporal and spatial expression of miRNAs in *H. contortus*

In *H. contortus* there were 273 mature miRNA sequences expressed by 152 stem-loops across 189 genetic loci (Appendix, Table 8-4). Across L3, L4, adult stages and in gut tissue, microarray analysis identified 55 mature miRNAs with a strong signal (> 500) and statistically significant variation in abundance ($p < 0.05$) (Fig. 3-1). Three miRNAs (miR-252-5p, miR-5885b-5p and miR-5899-3p) displayed a strong signal (> 500), but statistically insignificant variation ($p > 0.05$). The majority of miRNAs tested on the array showed low abundance below the signal threshold (n = 156, 69 %). The data in figure 3-1 represent values normalised across all four stages and gut tissue. The 55 variant miRNAs were expressed by 48 stem-loops across 43 genetic loci, Normalised mean abundance values for the variant miRNAs were highest in gut tissue ($\bar{x}$ = 8,638) and lowest in L3 ($\bar{x}$ =

1,749). For L4, male and female stages, mean abundance values were $\bar{x}$ = 5,907, 4,691 and 6,376 respectively. The strongest abundance value was recorded for miR-5885b-3p in the L4 (signal = 53,743). miR-5885b-3p also displayed the greatest abundance in gut tisse (signal = 44,083) while miR-71-5p was the most abundant miRNA in both L3 and male stages (signal = 13,003 and 27,606 respectively) and miR-5939-3p was the most abundant miRNA in the female stage (signal = 36,550). The L4 data contained 16 miRNAs that showed a significant peak or trough in expression when compared to L3 and adult stages. Examination of the gut data identified 21 miRNAs that were depleted compared to the whole female, while 14 displayed spatial abundance in the gut. Overall microarray signal abundance of miRNAs in L4 and gut RNA samples were generally consistent with L3 and adult stages generated previously by AD Winter

Included in the microarray data was the highly conserved miRNA let-7. Although not identified in the original study by Winter *et al*. (2012), BLAST searches of the latest genome assembly discovered the sequence (Fig. 8-2). Details of the stem-loop sequence and its intragenic position can be found in the Appendix (Fig. 8-3). As the *H. contortus* let-7-5p sequence displayed 100 % identity match to its *C. elegans* counterpart (which was present in the array), the abundance data for *C. elegans* let-7-5p was included as a proxy for *H. contortus*. Its abundance pattern matched that seen in other species (Pasquinelli *et al.*, 2000)

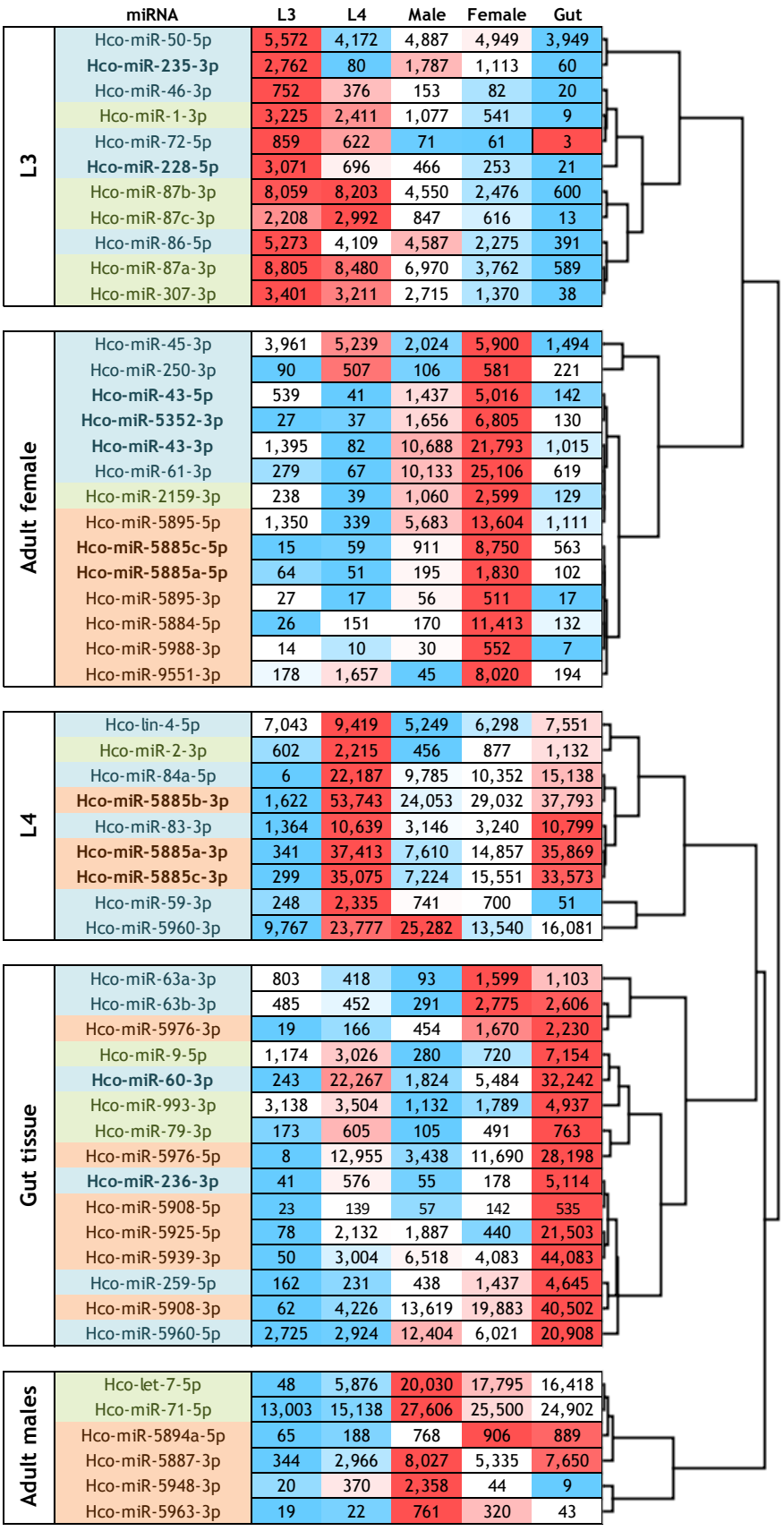
3.3.2 Hierarchical clustering of miRNA expression

Hierarchical clustering of the variant miRNA expression data (Fig. 3-1) identified five major groups (Fig. 3-1) that broadly corresponded to one life-cycle stage or tissue where miRNA expression was greatest. At the first branch node, the L3 and adult female groups were separated from the L4, adult male and gut dgroups. The gut (n = 15) and male (n =

6) groups were largest and smallest respectively and together with the female group (n = 15), accounted for 64 % of differentially expressed miRNAs in *H. contortus*. Eight miRNAs (Hco-lin-4-5p, Hco-miR-60-3p, Hco-miR-79-3p, Hco-miR-83-3p, Hco-miR-84a-5p and Hco-miR-5885a/b/c-3p) in the L4/gut group possessed a double-peak in expression during these two stages.

3.3.3 Comparing abundance of miRNAs from the same stem-loop

From the clustering data (Fig. 3-1) it was apparent that miRNAs derived from different arms of the same stem-loop (pre-miRNA) showed different patterns of abundance. There were seven instances where both arms (3′ and 5′) of a miRNA displayed significant variation in abundance (Table 3-1), two of which belonged to the same stem-loop (mir-5885a and mir-5885c; variation of miR-5885b-5p fell below the significance threshold). Interestingly, these are also the most abundant miRNAs in *H. contortus*. For two miRNAs (mir-43 and mir-5895), abundance of both arms was highest in adult stages. In the remaining five miRNAs (mir-5885a/c, mir-5908, mir-5960 and mir-5976), abundance patterns were markedly different between the two arms. In the most extreme example, all three loci of mir-5885a/c showed a similar pattern of abundance where the 3′ sequences displayed peaks in expression in L4 while the 5′ was only expressed at significant levels in the adult female stage. mir-5885a/c also represented one of three occurrences in *H. contortus* of variant miRNAs where the same mature miRNA sequence was expressed from different stem-loops at distinct genetic loci. The other two were mir-63a/b and mir-87a/b/c. All three of these multi-locus miRNA stem-loops (mir-5885, mir-63 and mir-87) displayed similar patterns of expression between different loci. While expression of the arms of miR-5976 displayed differences, the disparity was not significant enough to warrant separation between major expression clusters.



	miRNA	L3	L4	Male	Female	Gut
L3	Hco-miR-50-5p	5,572	4,172	4,887	4,949	3,949
	Hco-miR-235-3p	2,762	80	1,787	1,113	60
	Hco-miR-46-3p	752	376	153	82	20
	Hco-miR-1-3p	3,225	2,411	1,077	541	9
	Hco-miR-72-5p	859	622	71	61	3
	Hco-miR-228-5p	3,071	696	466	253	21
	Hco-miR-87b-3p	8,059	8,203	4,550	2,476	600
	Hco-miR-87c-3p	2,208	2,992	847	616	13
	Hco-miR-86-5p	5,273	4,109	4,587	2,275	391
	Hco-miR-87a-3p	8,805	8,480	6,970	3,762	589
	Hco-miR-307-3p	3,401	3,211	2,715	1,370	38
Adult female	Hco-miR-45-3p	3,961	5,239	2,024	5,900	1,494
	Hco-miR-250-3p	90	507	106	581	221
	Hco-miR-43-5p	539	41	1,437	5,016	142
	Hco-miR-5352-3p	27	37	1,656	6,805	130
	Hco-miR-43-3p	1,395	82	10,688	21,793	1,015
	Hco-miR-61-3p	279	67	10,133	25,106	619
	Hco-miR-2159-3p	238	39	1,060	2,599	129
	Hco-miR-5895-5p	1,350	339	5,683	13,604	1,111
	Hco-miR-5885c-5p	15	59	911	8,750	563
	Hco-miR-5885a-5p	64	51	195	1,830	102
	Hco-miR-5895-3p	27	17	56	511	17
	Hco-miR-5884-5p	26	151	170	11,413	132
	Hco-miR-5988-3p	14	10	30	552	7
	Hco-miR-9551-3p	178	1,657	45	8,020	194
L4	Hco-lin-4-5p	7,043	9,419	5,249	6,298	7,551
	Hco-miR-2-3p	602	2,215	456	877	1,132
	Hco-miR-84a-5p	6	22,187	9,785	10,352	15,138
	Hco-miR-5885b-3p	1,622	53,743	24,053	29,032	37,793
	Hco-miR-83-3p	1,364	10,639	3,146	3,240	10,799
	Hco-miR-5885a-3p	341	37,413	7,610	14,857	35,869
	Hco-miR-5885c-3p	299	35,075	7,224	15,551	33,573
	Hco-miR-59-3p	248	2,335	741	700	51
	Hco-miR-5960-3p	9,767	23,777	25,282	13,540	16,081
Gut tissue	Hco-miR-63a-3p	803	418	93	1,599	1,103
	Hco-miR-63b-3p	485	452	291	2,775	2,606
	Hco-miR-5976-3p	19	166	454	1,670	2,230
	Hco-miR-9-5p	1,174	3,026	280	720	7,154
	Hco-miR-60-3p	243	22,267	1,824	5,484	32,242
	Hco-miR-993-3p	3,138	3,504	1,132	1,789	4,937
	Hco-miR-79-3p	173	605	105	491	763
	Hco-miR-5976-5p	8	12,955	3,438	11,690	28,198
	Hco-miR-236-3p	41	576	55	178	5,114
	Hco-miR-5908-5p	23	139	57	142	535
	Hco-miR-5925-5p	78	2,132	1,887	440	21,503
	Hco-miR-5939-3p	50	3,004	6,518	4,083	44,083
	Hco-miR-259-5p	162	231	438	1,437	4,645
	Hco-miR-5908-3p	62	4,226	13,619	19,883	40,502
	Hco-miR-5960-5p	2,725	2,924	12,404	6,021	20,908
Adult males	Hco-let-7-5p	48	5,876	20,030	17,795	16,418
	Hco-miR-71-5p	13,003	15,138	27,606	25,500	24,902
	Hco-miR-5894a-5p	65	188	768	906	889
	Hco-miR-5887-3p	344	2,966	8,027	5,335	7,650
	Hco-miR-5948-3p	20	370	2,358	44	9
	Hco-miR-5963-3p	19	22	761	320	43

Fig. 3-1: Differentially expressed miRNAs in *H. contortus* based on RNA microarray abundance. Data displayed as relative heat-map (blue = low, white = median, red = high). miRNAs groups and associated dendrogram based on hierarchical clustering (see Chapter 3.3.3). miRNAs colour-coded based on conservation status (see Chapter 3.2.2): Green = widely-conserved, blue = nematode-only, orange = unique to *H. contortus*.

Table 3-1: *H. contortus* miRNAs with both arms displaying significant variation in expression.

	Microarray abundance data[2]				
miRNA[1]	L3	L4	AM[3]	AF[4]	AF Gut
Hco-miR-43-3p	1,395	82	10,688	21,793	1,015
Hco-miR-43-5p	539	41	1,437	5,016	142
Hco-miR-5885a-3p	341	37,413	7,610	14,857	35,869
Hco-miR-5885a-5p	64	51	195	1,830	102
Hco-miR-5885c-3p	299	35,075	7,224	15,551	33,573
Hco-miR-5885c-5p	15	59	911	8,750	563
Hco-miR-5895-3p	27	17	56	511	17
Hco-miR-5895-5p	1,350	339	5,683	13,604	1,111
Hco-miR-5960-3p	9,767	23,777	25,282	13,540	16,081
Hco-miR-5960-5p	2,725	2,924	12,404	6,021	20,908
Hco-miR-5976-3p	19	166	454	1,670	2,230
Hco-miR-5976-5p	8	12,955	3,438	11,690	28,198
Hco-miR-5908-3p	62	4,226	13,619	19,883	40,502
Hco-miR-5908-5p	39	3,415	10,070	18,103	39,637

[1] Colour indicates conservation. Blue = nematode-specific. Orange = unique to *H. contortus*.

[2] Microarray data displayed in a relative heat-map (blue = low, white – 50 %, red = high).

[3] AM = Adult male; [4] AF = Adult female.

3.3.4 Expression of miRNA clusters in *H. contortus*

The eight miRNA genomic clusters (*i.e.* physical location within the genome) originally identified in *H. contortus* (Winter *et al.*, 2012) were examined in more detail to determine the expression pattern across each cluster (Fig. 3-2). Four clusters were intergenic (*mir-40a/c/d/e*, *mir-5885b/c*, *mir-5902/5903/5919/5932* and *mir-5918c/d*) while three were intragenic (*mir-2/71*, *mir-43/61/5352/5895* and *mir-63a/b*). For the *mir-2/71* and *mir-63a/b* clusters, only one arm of each miRNA possessed a strong signal and significant temporal variation to be deemed variant; the other arms fell below the signal threshold. While abundance and expression profiles were similar for members of the *mir-63a/b* cluster, there were substantial differences between members of the *mir-2/71* cluster. Abundance of miR-2-3p was much lower than miR-71-5p ($\bar{x}$ = 1,056 and 21,230 respectively). In addition, their expression profiles differed, with miR-2-3p being placed in the L4 hierarchical clustering group, while miR-71-5p was placed into the male group. As discussed previously, miRNAs in the *mir-5885b/c* cluster exhibited an interesting expression profile where both stem-loops displayed the same profile, but the 3′ and 5′ arms differed. All miRNAs of the *mir-43/61/5352/5895* cluster with expression above the signal threshold displayed a similar expression profile, peaking during the female stage. For the

major arms (3′, 3′, 3′ and 5′ respectively) abundance varied between a minimum of 1,731 for miR-5352-3p and a maximum of 7,241 for miR-43-3p. Expression of miRNAs in three other clusters (*mir-40a/c/d/e*, *mir-590/5903/5919/5932* and *mir-5918b/c/d*) did not pass the signal threshold (*i.e.* low expressors), while the *mir-250/5983* cluster could not be identified in the latest version of the *H. contortus* genome.

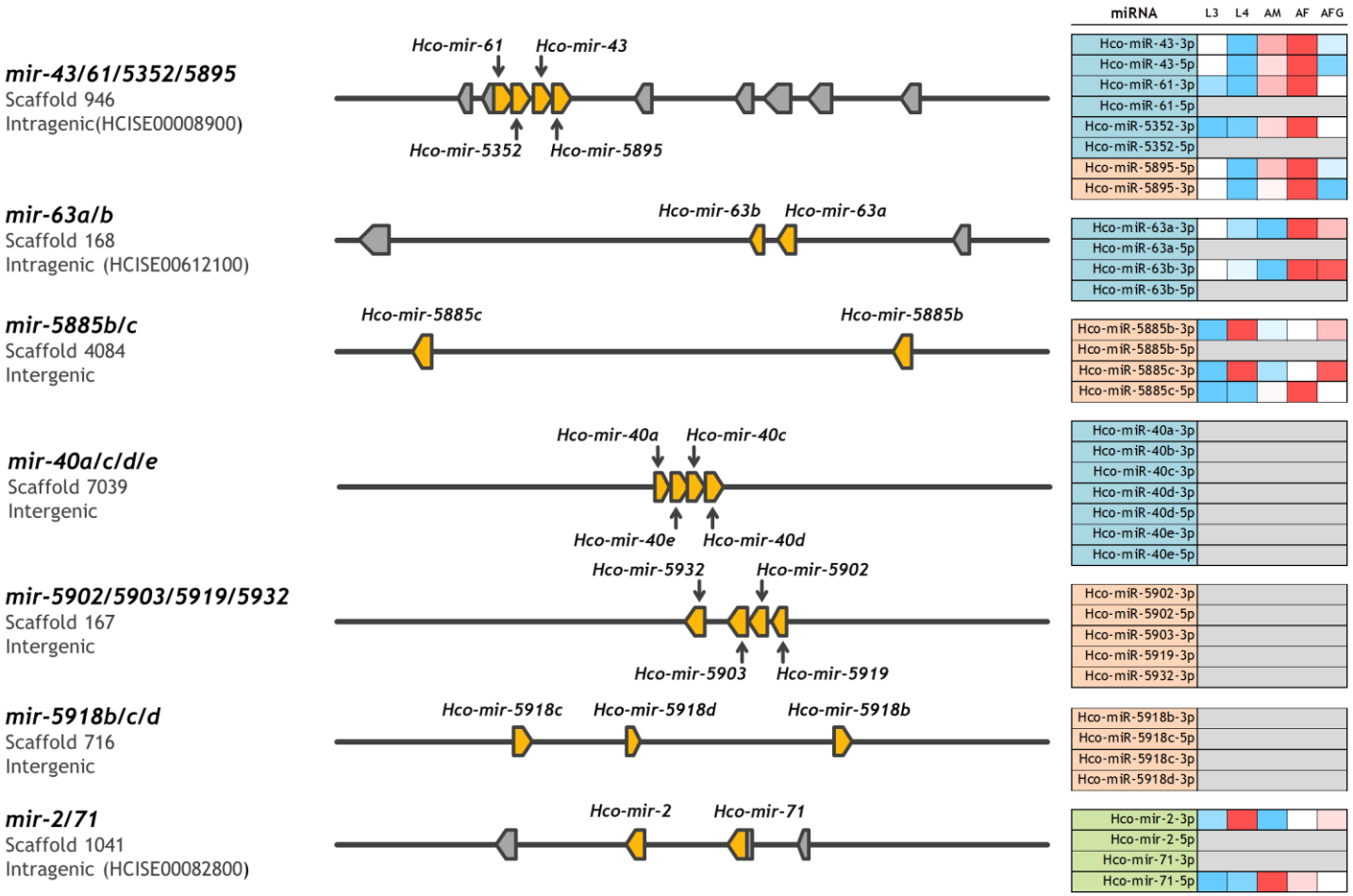


Fig. 3-2: Position and abundance profiles of seven miRNA gene clusters in *H. contortus*. The positions of each miRNA stem-loop (yellow) are marked relative to gene exons (grey) along a 4Kb length of the genome. The abundance of each mature miRNA is listed to the right of each cluster as a relative heat-map (blue = low, white = median, red = high) while miRNAs with low abundance (microarray signal < 500) are marked grey. miRNAs in the table are colour coded based on conservation: green = highly-conserved, blue = nematode-only, orange = unique to *H. contortus*.AM = adult male. AF = adult female. AF Gut = Gut tissue from adult females.

3.3.5 qRT-PCR confirmation of *H. contortus* miRNA microarray data

qRT-PCR was carried out to confirm the microarray abundance data for six *H. contortus* miRNAs (miR-45-3p, miR-60-3p, miR-71-5p, miR-87a-3p, miR-228-5p and miR-235-3p) (Fig. 3-3). In general, the two datasets supported one another. The trends in miRNA abundance (*i.e.* increases, decreases, peaks or troughs) were consistent, although the two techniques did display certain discrepancies. For three miRNAs, miR-87a, miR-228 and miR-235, the expression patterns were comparable between datasets. Both miR-87a and miR-228 displayed a decreasing level of abundance between larval and adult stages, while miR-235 presented a dip in abundance at the L4 stage, that was much lower than L3 or adult stages. For miR-45, abundance was broadly similar between techniques, except for L3 and female stages where microarray expression was relatively higher than qRT-PCR levels. For miR-60, the two datasets indicate a particularly high level of abundance in adult female gut tissue, however, the microarray showed a large peak in abundance during the L4 stage that was not present in the qRT-PCR data. Both techniques reported significant variation between replicates for miR-71 that complicated interpretation, although there was a general increase in abundance between larval and adult stages. As a whole, the qRT-PCR experiments increased our confidence in the microarray, specifically for trends in abundance.

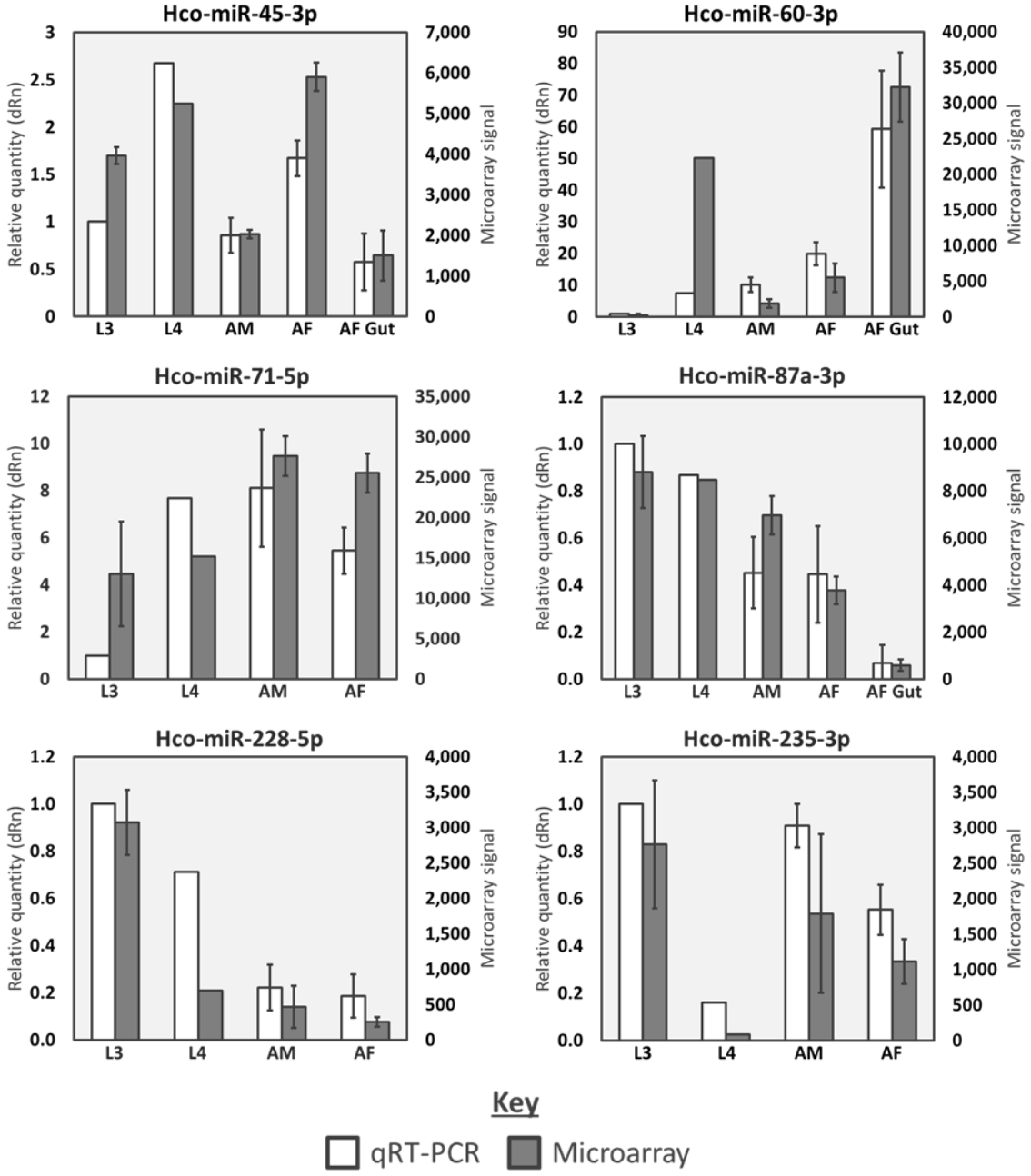


Fig. 3-3: Comparison of abundance of six *H. contortus* miRNAs obtained using qRT-PCR and microarray. qRT-PCR data (relative quantity) displayed in blue, microarray data (signal strength) in yellow. Biological triplicates data were used, except for L4 stage where only duplicate data were available. Error bars represent ± 1 standard deviation. Gut data displayed for miR-45-3p, miR-60-3p and miR-87a-3p only.

3.3.6 Conservation status of *H. contortus* miRNAs

Interrogation of miRBase using *H. contortus* miRNAs (Fig. 3-4) showed that 108 out of 152 stem-loop sequences (71.1 %) were unique to this organism. The remaining stem-loops were identified in other nematodes (n = 27, 17.8 %) or other phyla (n = 17, 11.2 %). This includes the stem-loop lin-4, which was originally classified as helminth-specific, being found in many parasitic and non-parasitic nematodes as well as two species of Platyhelminth: the free-living Planarian *Schmidtea mediterranea* and the (parasitic) Monogenean *Gyrodactylus salaris*. However, as it was the only member of the helminth-specific group, it was merged into the widely-conserved group for simplicity.

When the variant miRNA stem-loops (n = 43) were analysed, comparatively few (n = 13, 30.2 %) were found to be unique in *H. contortus*. Most were identified in other nematodes (n = 20, 46.5 %) or different phyla (n = 10, 23.3 %). The three invariant miRNAs were widely conserved (Hco-miR-252-5p) or unique (Hco-miR-5885b-5p and Hco-miR-5899-3p). In comparison, 93 of the 102 miRNA stem-loops with low abundance (microarray signal < 500), were unique to *H. contortus* and only a handful (n = 9, 8.8 %) were conserved in other organisms. Across the five expression clusters (Fig. 3-1), there was a clear division between early and late developmental stages in the conservation status of miRNAs. Unique miRNAs were not present in the L3 group (n = 11) and only three were found in the L4 group (n= 9), all of which were miR-5885-3p sequences. Almost half (n = 17; 48.6 %) of all miRNAs in the adult groups (male, female and gut; n = 35) were unique to *H. contortus*.

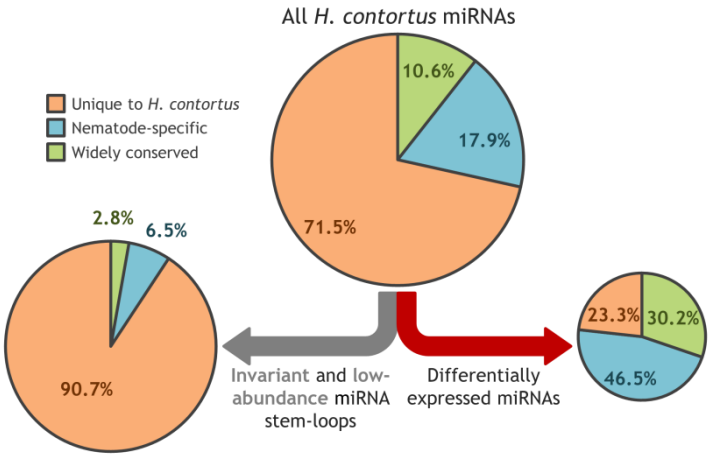


Fig. 3-4: Conservation status of *H. contortus* miRNA stem-loops (top chart). Using microarray abundance data, miRNAs were separated into those with differential expression (right chart) and those with invariant expression or low abundance (left chart). The conservation of each miRNA was also classified: unique to *H. contortus* (orange), specific to nematodes only (blue) or widely conserved (green).

3.3.7 Comparison of conservation between miRNAs in *H. contortus* and *C. elegans*

The conservation status of miRNAs in *H. contortus* was compared against *C. elegans.* Although they live in different ecological niches (*i.e.* parasitic *versus* free-living), *C. elegans* remains the best source of comparison for nematodes, as its miRNAs have been well studied. Interrogation of miRBase (release 21) identified the conservation status of 228 miRNA stem-loops in *C. elegans* expressed by 250 genetic loci (Fig. 3-5). 117 (51.4 %) were unique to *C. elegans* and not found in any other organism. Of the remaining 111 miRNAs, 53 (23.3 %) were genus-specific (*i.e.* limited to *Caenorhabditis spp.*), 45 (19.7 %) were nematode-specific and 13 (5.7 %) were more widely conserved in different phyla. While *C. elegans* possessed 50 % more stem-loops than were recognised in *H. contortus*, the general proportions that belonged to each conservation category (*e.g.* widely conserved, nematode-specific) were similar between these two organisms. Specifically, the largest number of miRNAs belonged to the unique category with fewer being widely

conserved. Both species possessed a similar proportion of nematode-specific miRNAs (19.7 % in *C. elegans* versus 17.9 % in *H. contortus*), although *H. contortus* possessed a higher number of widely-conserved miRNAs (13 in *C. elegans* versus 17 in *H. contortus*). An additional category of genus-specific miRNAs was used for *C. elegans*, reflecting the larger knowledge base available for the genus of this model organism. This included data from *C. brenneri*, *C. briggsae* and *C. remanei* (Shi *et al.*, 2013; De Wit *et al.*, 2009). No such information was available for the *Haemonchus* genus or the wider Trichostrongyloid family.

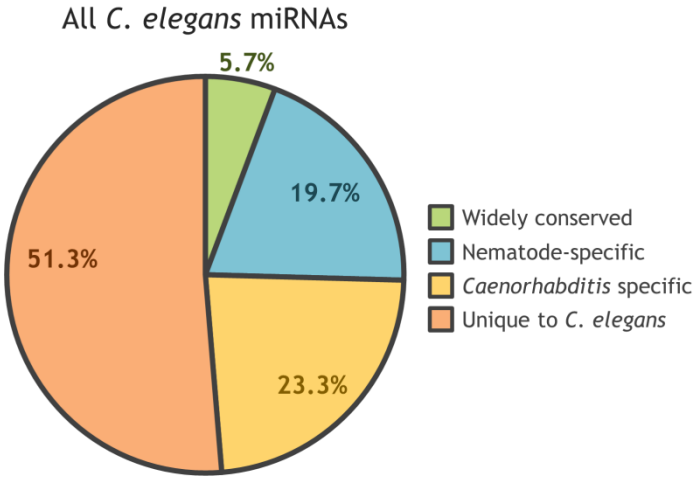


Fig. 3-5: Conservation status of *C. elegans* miRNA stem-loops. The conservation of each miRNA was classified: unique to *C. elegans* (orange), limited to *Caenorhabditis spp.* (yellow), limited to nematodes (blue) or widely conserved (green).

3.3.8 Conservation of miRNA abundance between *H. contortus* and *C. elegans*

Similarly to conservation status, abundance profiles for *H. contortus* miRNAs were compared to those of *C. elegans* where more data is available. The abundance data for 25 miRNAs conserved in both *H. contortus* and *C. elegans* was compared for relative similarities (Appendix, Fig. 8-1). The *H. contortus* data were based on RNA microarrays (see Chapter 3.3.1) while the *C. elegans* data were graphed from qRT-PCR data obtained

from Kato *et al.* (2009). Four miRNAs (lin-4, miR-63, miR-71 and miR-228) showed equivalent expression profiles between the species, while seven miRNAs (miR-1, miR-43, miR-45, miR-234, miR-235, miR-250, miR-252) displayed strong similarity in most, but not all of the stages examined. For miR-43 and miR-45, abundance in *H. contortus* females was low, while in *C. elegans* their abundance was high. miR-1 expression was higher in males than females in *H. contortus*, but not *C. elegans*. The opposite is true for miR-43, miR-45 and miR-234 (lower in males than females in *H. contortus*, but not *C. elegans*). For miR-250 and miR-252, abundance is lower in L3 than L4 in *H. contortus* but not *C. elegans* and vice-versa for miR-235. The remaining fourteen miRNAs (miR-2, miR-46, miR-59, miR-60, miR-61, miR-72, miR-79, miR-83, miR-84, miR-86, miR-87, miR-124, miR-236 and miR-259) did not display noteworthy similarities in abundance between the two species.

3.3.9 Conservation of *mir-5885*

The microarray data (Fig 3.1) showed that miR-5885-3p was the most abundant miRNA in *H. contortus*. It was of interest to determine if this miRNA is related to any known miRNA family in *C. elegans* and if so, compare abundance profiles. Initial analysis showed that the seed sequence (nucleotides 2-7) of Hco-miR-5885 was identical to members of the *C. elegans* Bantam family which contains mir-58, mir-80, mir-81 and mir-82 (Fig. 3-6). Alignments of the stem-loops can be found in the Appendix (Fig. 8-4). However, despite this similarity, temporal abundance profiles for these miRNAs were not conserved between *H. contortus* and *C. elegans* (Fig. 3-7). In *H. contortus*, abundance was highest in L4 and decreased in adults, with males showing 20 – 50 % lower abundance compared to females. In contrast, Cel-miR-80, Cel-miR-81 and Cel-miR-82 were most abundant in males. Cel-miR-58 was the most abundant member of the Bantam family and

overall, one of the most abundant miRNAs among all *C. elegans* miRNAs with 1 – 2.5 million reads. To determine if Hco-mir-5885 is found in other parasitic nematodes, BLAST analysis of the NCBI nr (non-redundant) nucleotide database was carried out. This identified the mature miR-5885 3′ and 5′ arms in seven species (*Angiostrongylus cantonensis*, *Angiostrongylus costaricensis*, *Cylicostephanus goldi*, *Haemonchus placei*, *Heligmosomoides polygyrus*, *Nippostrongylus brasilensis* and *Strongylus vulgaris*). All eight species (including *H. contortus*) belonged to phylogenetic clade V. The sequences displayed 100 % nucleotide identity to Hco-miR-5885a-3p while at the stem-loop level there was 90 – 100 % identity with mir-5885a suggesting that they may have arisen from a common parasitic ancestor. Indeed, *H. placei* contains three *mir-5885* loci as seen in *H. contortus*.

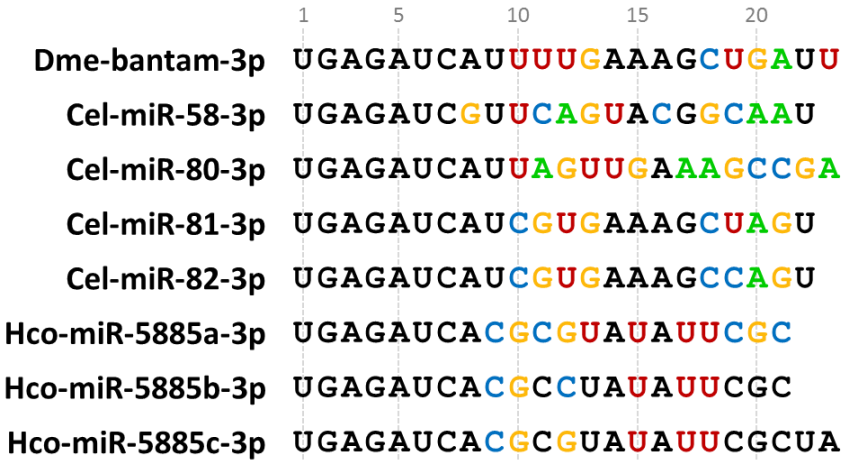


Fig. 3-6: Alignment of Hco-miR-5885a/b/c-3pand *Bantam* family members from *C. elegans* and *Drosophila melanogaster*. All sequences are 3' mature miRNAs. The seed sequence represents nucleotide positions 2 – 7. Nucleotides in the majority at a specific position are coloured black, otherwise they are colour-coded: A = green, C = blue, U = red, G = yellow.

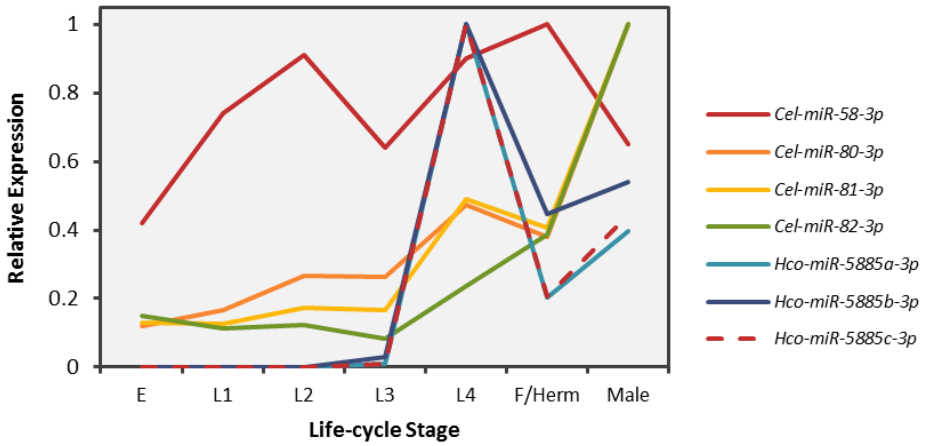


Fig. 3-7: Relative expression of members of the *Bantam* miRNA family in *H. contortus* and *C. elegans*. Values calculated as relative to the largest value for that miRNA. E = embryo. F/Herm = Female (*H. contortus*) or hermaphrodite (*C. elegans*). No embryo, L1 or L2 data available for *H. contortus*. Data for *H. contortus* taken from microarray (see Chapter 3.3.1). Data for *C. elegans* taken from *Kato* et al. (2009).

3.3 Discussion

There are presently 152 miRNA stem-loops identified in *H. contortus* (Appendix, Table 8-4), however, some may remain undiscovered. As bioinformatic identification of miRNAs utilised early drafts of the genome, analysis of a more complete version might identify additional miRNAs, such as the discovery of Hco-let-7. Furthermore, the deep sequencing for miRNA discovery was performed on L3 and adult worms, which would fail to identify miRNAs with no significant expression during these stages, such as those unique to egg and early larval stages. However, the current list of miRNAs most likely covers the vast majority of miRNAs present in *H. contortus*

After identification of the miRNAs in an organism, examination of their expression pattern through development and in particular tissues can help provide information on possible regulatory functions. Initial microarray analysis by Dr. Winter (unpublished data) determined miRNA expression across three major life-cycle stages, namely L3, adult males and adult females of the *H. contortus* MHco3(ISE) strain, which is susceptible to the three major broad-spectrum anthelmintics. Genomic and transcriptomic studies also

utilised this strain (Laing *et al.*, 2013). Both studies included sheathed L3 and L3 exsheathed by exposure to sodium hypochlorite; however, for both miRNA and mRNA expression data, no significant differences were found between the two. In this study, data is only reported for sheathed L3.

Of the 273 mature miRNA sequences available, 223 were investigated using a custom chip including at least one arm from all 152 stem-loops currently identified in *H. contortus*. This meant that 50 sequences were not examined and although their inclusion may have delivered insights, the current data provide ample opportunity to study miRNA variation across late larval and adult stages. The microarray analyses performed during this study expanded on the first dataset by adding miRNA expression data across the L4 stage and gut tissue extracted from adult female worms. To maintain consistency, RNA was extracted from the same strain of parasite using an identical technique and analysed by the same company (LC Sciences). With the addition of L4 data, miRNA expression data is now available for all stages that inhabit or interact with the host, providing a powerful resource for the study of developmental regulation and host-parasite interactions. In addition, the gut data provided a first analysis of the spatial regulation of *H. contortus* miRNAs in an important organ, providing clues for the potential role that miRNAs play in maintaining tissue function. Together, 55 miRNAs were identified as possessing significant temporal or spatial variation in expression, (Fig. 3-1). Three miRNAs (Hco-miR-252-5p, Hco-miR-5885b-5p and Hco-miR-5899-3p) displayed high abundance but insignificant variation ($p > 0.$ and were therefore useful as calibrators during qRT-PCR. The remaining miRNAs (73 %), most of which were unique to *H. contortus*, displayed negligible levels of expression. A similar situation is found in other parasites, such as *Echinococcus granulosus* and *Echinococcus multilocularis* where 73 and 79 % of miRNAs displayed negligible expression (Cucher *et al.*, 2015).

The L4 stage represents an important transitory phase in the development of *H. contortus*. The organism must shift from a non-motile, non-feeding stage (*i.e.* L3) considered to be in developmental arrest, to a parasitic, blood-feeding and reproductive adult (*i.e.* L4). If the L4 stage was simply an extension of the L3 stage, we might expect miRNA expression to be averaged between L3 and adult stages. Indeed, the abundance of most variant miRNAs (n = 35) was either similar to, or trending between, L3 and adult stages. However, 15 miRNAs possessed a significant peak or trough in expression compared to the L3 and adult stages. These miRNAs might regulate genes involved in the transition between environmental and parasitic lifestyles. Such changes include development of the gut (for feeding), motility (to reach the abomasum), adaptation (to host temperature and pH), sensation (to identify the abomasum) and preparation for the adult stages.

The gut is a major organ in *H. contortus* and an important location for host-parasite interactions due to the blood-feeding nature of *H. contortus*. Gut data suggested that miRNAs play an important role in maintaining this tissue. Interestingly, overall miRNA abundance was highest in the gut, followed by the L4 stage. High abundance in both L4 stage and gut tissue suggests these miRNAs may play a role in both development (during the L4 stage) and maintenance (during adult stages) of the gut. A large proportion of the variant miRNAs displayed a significant spatial variation between total adult worms and gut tissue, seen as a depletion (n = 14) or enrichment (n = 15) when compared to the whole female. The differences in expression were marked, suggesting a very tissue-specific regulation of gene expression and also confirming extraction and enrichment of gut tissue by the methods used. Depletion of miRNAs in the gut relative to the adult female worm may be explained by the presence of developing embryos and eggs within the whole female worm. These stages are likely to show much different miRNA expression profiles. Between the adult stages, 11 miRNAs were abundant in the female but displayed low expression in the male, while only five showed the opposite pattern. This suggests a sex-

specific role, possibly covering female reproduction or embryogenesis. Male sexual biology is more straight-forward and less energy intensive than female sexual biology, possibly explaining the disparity as more complicated systems require more regulation.

In the clade III parasite *A. suum* deep sequencing of small RNAs identified 84 miRNAs expressed during embryogenesis and early larval development (Wang *et al.*, 2011a). Their abundance was highly regulated and stage-specific with significant levels present that changed rapidly. Comparison of this data with *H. contortus* highlighted three miRNAs present in both species with significant abundance in *A. suum* zygotes and embryos: miR-43, miR-250 and miR-5352. If functional conservation is maintained, this suggests that the high levels of these miRNAs in *H. contortus* females may be due to high levels in embryos, rather than within actual female tissue.

Extraction of the gut was limited to female worms due to their larger size which aided dissection of the gut from other tissues. More advanced micro-dissection techniques may allow removal of the gut tissue in males, thus providing data on any sex-specific variation of gut miRNAs (and mRNAs). For example, the gut is the site of expression of nutrients such as yolk proteins for the developing oocytes and early embryos in *C. elegans* (Kimble and Sharrock, 1983). Whether these as well as other sex-specific genes are regulated post-transcriptionally may be identified from comparison of male and female gut array data.

Hierarchical clustering reinforced the concept that the variant miRNAs were involved in stage-specific regulation as five groups, each correlating to a stage or tissue, are clearly demarcated (Figs. 3-1). The L3 group is the most distinct, with miRNAs enriched at this stage showing very little expression in the adult stages and negligible gut expression. During the L3 stage, *H. contortus* is considered to be in developmental arrest until ingested by a host. miRNAs may be involved in maintaining this arrest, which would explain the separation seen between the L3 and other groups. A number of miRNAs displayed strong

expression in both the L4 stage and gut tissue. They might regulate the same processes such as gut development and maintenance, or *H. contortus* may use the same miRNAs to regulate different processes in different stages. In *C. elegans*, let-7 is expressed in multiple tissues (*e.g.* nerve cord, intestine, body wall) but phenotypic effects of loss of function are observed only in hypodermal cells. This suggests that let-7 regulates non-essential genes in other tissues, or that the hypodermal targets are poorly expressed in other tissues. Indeed, both scenarios are likely, providing fidelity arising from tissue-specific variation in both targets and regulatory molecules (*e.g.* miRNAs).

At a genomic level, eight miRNA clusters were discovered (Winter *et al.*, 2012), although analysis of the current genome could only identify seven of these (Fig. 3-2). Analysis of the microarray data generated in this study suggest that miRNAs within clusters are usually expressed together, similar to operons in *C. elegans* (Zorio *et al.*, 1994) but further analysis is required to confirm if they share a promoter. An exception to this is the *mir-2/71* cluster which shows contrasting levels of abundance and patterns of temporal expression between its two members.

One particularly interesting miRNA, which has been mentioned several times, is mir-5885. Intriguingly, the expression profiles for its two arms differed significantly from one another. This runs contrary to the accepted notion that while both arms may differ in abundance (due to degradation of one of the arms), the profiles of expression would be consistent. This is not the case for mir-5885a/b/c, where the 3′ arms are abundant during both L4 stage and gut tissue while the 5′ arms are only adult female worms. A similar disparity between arms is also seen in mir-5960 and mir-5976. All three loci (*a*, *b* and *c*) of *mir-5885* share this pattern. Two *mir-5885* loci (*b* and *c*) are part of a cluster (Fig. 3-2), which might explain a shared abundance and expression pattern. Being on separate scaffolds however, *mir-5885a* is, at least, many kilobases distant from the other loci. Due to the distances involved, we can speculate an initial duplication event, producing two loci,

one of which duplicated again, producing the *b* and *c* loci. In addition, microarray data indicate that miR-5885 is the most abundant miRNA in *H. contortus*. Together, multiple loci and high abundance suggest an important regulatory role requiring high levels of expression in L4 and gut tissue. Alignment of the mature miR-5885-3p sequence with other species suggest that it is related to the *Bantam* family of miRNAs due to a shared seed sequence (Fig. 3-5). The widely conserved *Bantam* miRNA, originally discovered in *Drosophila melanogaster*, regulates neuronal growth (Parrish *et al.*, 2009). In *C. elegans* the *Bantam* family (Consisting of *mir-58*, *mir-80*, *mir-81* and *mir-82*) is involved in diverse processes including body-size regulation (Lucas and Lozano, 2011), gut-specific defence against pathogens (Pagano *et al.*, 2013) and dauer formation (Alvarez-Saavedra and Horvitz, 2010). BLAST searches identified sequences identical to miR-5885-3p in seven other nematode species (see Chapter 3.3.7). All are parasitic and belong to phylogenetic clade V, suggesting a shared role unique to these organisms. The homology between the seed-sequences may represent an instance of convergent evolution between clade V parasites and other organisms, or may have arisen through evolutionary divergence, with *mir-5885* acting as a replacement for the other *Bantam* family members, none of which have been identified in clade V parasitic nematodes. In contrast, miR-81 has been found in *A. suum* which belongs to clade III. While Hco-miR-5885-3p and the *bantam* family in *C. elegans* both display strong expression (see Chapter 3.3.8), disparities between their expression profiles suggests evolutionary divergence, although additional evidence is required for confirmation.

This study has focussed on the *H. contortus* ISE isolate, which is susceptible to anthelmintic drugs (R. Laing, personal communication). Expression across the developmental stages could also be compared in different strains of *H. contortus*, specifically strains and/or field isolates that display anthelmintic resistance. This may

allow identification of miRNAs potentially regulating genes associated with resistance (Devaney *et al.*, 2010). Recent microarray analysis in *C. elegans* identified two miRNAs (Cel-miR-85 and Cel-miR-788) significantly altered in worms resistant to ivermectin, compared to N2 wild-type worms (Gillan *et al.*, unpublished data). In addition, one miRNA, *Hco-mir-9551,* was upregulated in Ivermectin resistant *H. contortus*. From the array analysis carried out here on susceptible worms, Hco-miR-9551-3p displayed peaks in expression during L4 and adult females (signal = 1,657 and 8,020 respectively), while other stages showed negligible expression. Efforts are ongoing to identify a potential role for miR-9551 in anthelmintic resistance.

The *H. contortus* transcriptomic studies mentioned earlier (Laing *et al.*, 2013) examined gene expression in multiple life-cycle stages including the L4 as well as gut tissue extracted using the same technique applied here. Among their findings, the authors identified significant variation in gene expression between L3 and L4, with many genes being switched on and off completely. This is not surprising, given the substantial physiological and behavioural changes that occur when the life-cycle resumes. Large differences were also seen between whole females and gut tissue. Together, the transcriptomic data and miRNA microarray data reinforces the view that gene regulation is an important aspect of stage- and tissue-specific development.

qRT-PCR was used to confirm the microarray data for six miRNAs (Chapter 3.2.2). qRT-PCR verification was also performed by Winter *et al.* to confirm some of the deep sequencing data (2012). Overall, qRT-PCR corroborated the expression trends of the chosen miRNAs. While variation was seen in relative values, high or low expression relative to another stage was generally consistent. One notable difference was seen in miR-60-3p, where microarray identified a strong peak in expression during the L4 stage, which was not seen by qRT-PCR. As both results were obtained using the same RNA

samples, the discrepancy may be due to technical differences between qRT-PCR and microarrays.

The majority miRNAs in *H. contortus* were found to be unique to *H. contortus* and only a small proportion were conserved in other species (Winter *et al.*, 2012). A similar situation was detected in the filarial parasite *Brugia pahangi*. . Intriguingly, when the variant miRNAs were analysed individually, the situation was reversed, with most being conserved and only a minority of unique miRNAs. Development is a particularly conserved feature in biology, with even small changes in the genes and their regulation proving lethal, as observed for lin-4 and let-7 in *C. elegans*(Lee *et al.*, 1993; Reinhart *et al.*, 2000). It therefore makes sense that both miRNAs and the genes involved in regulating development are conserved.

C. elegans also possesses similar total proportions of unique and conserved miRNAs to *H. contortus*. This is particularly true if genus information is removed for *Caenorhabditis*, indicating that the majority are 'unique' (71 % in *H. contortus* versus 75 % in *C. elegans*). As genome and miRNA data become available for other *Haemonchus* species and for closely related Trichostrongylid nematodes (eg *Teladorsagia circumcincta*) it is likely that such unique miRNAs will be redefined as genus- or clade-specific. Indeed, recent sequencing of small RNAs from excretory-secretory exosomes of *T. circumcincta* identified some with significant sequence identity to *H. contortus* miRNAs (Tzelos and Gu, unpublished data). In addition, BLAST searches of the NCBI nr database for mir-5885 identified this miRNA in *Haemonchus placei* (see chapter 3.3.4). Further analysis could discover how many of the miRNAs currently identified in *H. contortus* are also present in *H. placei*.

Only three conserved miRNAs (miR-40, miR-124 and miR-277) identified in *H. contortus* did not show significant levels of expression in the *H. contortus* microarray data (signal > 500). Low levels of these miRNAs may be all that is required by *H. contortus* for regulation, or expression may be limited to eggs, L1 or L2 stages or be environmentally regulated (rather than developmentally regulated). It is also possible that these miRNAs are redundant, left over from a shared nematode ancestor and not required in modern *H. contortus*. As mentioned, most miRNAs in *H. contortus* are unique and of these, the vast majority (90 %) also possessed negligible expression. Studies in several species suggest that small RNA pathways as a whole are rapidly evolving, with the exception of certain highly conserved miRNAs (Shi *et al.*, 2013), suggesting particularly important roles for these miRNAs.

Unique miRNAs are more likely to be recently evolved and may not have acquired optimal promoter activity in order to be expressed, or may only be expressed under certain conditions. The majority of miRNAs in an organism are generated *de novo* in the genome (Marco *et al.*, 2013a), rather than through duplication of older miRNAs that are more likely to be conserved. With a large genome and high degree of genetic polymorphism (R. Laing, personal communication), it is not surprising that many miRNAs are unique to *H. contortus*.

Abundance profiles were obtained for miRNAs conserved between *C. elegans* and *H. contortus* (Chapter 3.3.8). While some showed highly conserved patterns of abundance between the species, others were highly varied. If we assume that miRNAs with sequence homology have arisen from a common ancestor rather than through evolutionary convergence, these miRNAs are either targeting different genes, so there is no need for conservation of expression, or are targeting the same genes, in which case gene function and expression are no longer conserved between the species. As miRNAs target multiple genes in various pathways, the situation is likely to be a combination of both.

The results presented in this chapter demonstrated the temporal and spatial nature of miRNA abundance in *H. contortus*. miRNAs displaying variant abundance through development were mostly conserved in other species, particularly the nematode phylum. Comparisons of the abundance profiles for miRNAs common to both *H. contortus* and *C. elegans* found that only a handful of profiles were similar, suggesting that for miRNAs a conserved sequence does not necessarily imply conserved function. The next step is to identify the genes targeted by the variant miRNAs of *H. contortus*, study biological functions and investigate if this implication is true.

4 *In silico* target prediction for miRNAs in *H. contortus* and *C. elegans*

4.1 Introduction

Following discovery of miRNA sequences in an organism and quantifying their abundance, identifying genes targeted by specific miRNAs is the next logical step in understanding their function. However, this has proven a major challenge in miRNA research. In plants, miRNAs must show perfect complementarity to their target gene for effective regulation (reviewed in Voinnet, 2009), allowing accurate prediction of targets (Rhoades *et al.*, 2002). This is not the case for animal miRNAs where the exact mechanism of interaction between miRNAs and mRNAs (messenger RNA) remains poorly understood (Karbiener *et al.*, 2014a). In addition, validation of potential targets has confirmed only a handful of interactions and involves time-consuming and costly procedures (reviewed in Eulalio and Mano, 2015). To narrow down the number of potential targets, a range of computational approaches have been used to predict interactions *in silico*. In general, miRNA binding sites are predicted by aligning a miRNA and the 3′ untranslated region (UTR) of an mRNA. There are several aspects of the interaction that can be examined and most algorithms employ multiple mechanisms for target prediction, weighting the importance of each aspect (*e.g.* seed matches, stability, secondary structure) when calculating the probability of a binding site (Peterson *et al.*, 2014).

The seed sequence covers nucleotides 2 – 7 from the 5′ end of the mature miRNA (Lewis *et al.*, 2003). Seed matches involve identifying perfect Watson-Crick alignment of six or seven nucleotides (6mer, 7mer) between miRNA (covering the seed sequence) and mRNA

sequences (Brennecke *et al.*, 2005). The stability of miRNA-mRNA interactions can also be examined by calculating the Gibbs free energy (G), or more specifically the change in free energy (ΔG); reactions with a lower ΔG are deemed more stable. Identifying the ΔG of a miRNA-mRNA interaction indicates the binding strength, which is used as a proxy for the likelihood of that interaction (Yue *et al.*, 2009). All RNA molecules display complicated secondary structures, particularly after transcription (Mahen *et al.*, 2010). miRNA-mRNA interactions involve two stages, where the miRNA partially binds firstly to a short, exposed region of the mRNA, before the mRNA unfolds and allows complete binding (Long *et al.*, 2007). Predicting mRNA secondary structure helps identify which binding sites are more accessible and therefore more likely to be functional.

In addition to these major techniques, algorithms often employ additional factors to help rank predicted targets, such as the number and position of binding sites within a UTR (Garcia *et al.*, 2011; Grimson *et al.*, 2007; Hon and Zhang, 2007). Ideally sites should be positioned at least 15 nucleotides from a stop codon and away from the middle of long UTRs. While the seed sequence is known to be important, mismatched alignments should not be ignored due to the presence of compensatory sites. These regions (nucleotides 12 – 17) promote stability and offset seed mismatches (Friedman *et al.*, 2009). Additionally, interactions with the seed sequence can include a G:U wobble, where guanine aligns with uracil (rather than a cytosine) without unduly affecting miRNA-mRNA stability (Doench and Sharp, 2004), while adenosine or uracil nucleotides at position one of the mature miRNA (7/8mer-1A or 7/8mer-1U) can improve stability (Betel *et al.*, 2010). Sequence conservation between species can help identify important miRNA-mRNA interactions as well as common promoter regions (Fujiwara and Yada, 2013) and flanking genes (Ohler *et al.*, 2004). This approach was recently used by Winter *et al*. (2015) to narrow down the list of potential target genes of a miRNA conserved in different filarial species.

Many of the algorithms utilise online repositories of miRNA and UTR information, particularly miRBase (Kozomara and Griffiths-Jones, 2011). While miRBase is updated regularly with recently discovered miRNAs, the majority of algorithms are limited to the major research species (*e.g. Homo sapiens, Drosophila melanogaster* and *Caenorhabditis elegans*). This limits their use in species such as in *H. contortus* where we must utilise algorithms that permit custom sequences, such as PITA, RNAhybrid and miRanda.

This chapter focuses on the use of *in silico* target prediction combined with transcriptomic data and gene ontology. Custom 3′ UTR sequences were interrogated using *H. contortus* miRNAs that displayed temporal or spatial variation. To overcome algorithm bias, the results from multiple algorithms were compiled. Transcriptomic data for predicted targets were compared against miRNA expression data to identify inverse correlations. Finally, for selected targets, putative homologues were identified in *C. elegans* and analysed for enrichment of gene ontology annotations to identify molecular functions, biological processes and cellular locations. Results were also compared against predicted targets of miRNAs conserved in *C. elegans*. Together these data identified potential miRNA targets in *H. contortus* and suggest possible biological roles.

4.2 Materials and methods

4.2.1 *In silico* target prediction for variant miRNAs in *H. contortus*

miRNA targets were predicted in *H. contortus* by interrogating a custom 3′ UTR sequence database with 53 mature miRNAs using three algorithms miRanda (John *et al.*, 2004), PITA (Kertesz *et al.*, 2007) and RNAhybrid (Rehmsmeier *et al.*, 2004). Predictions were performed by Henry Gu using the Unbuntu (14.04 LTS) operating system. PITA was used with default settings (no additional settings specified). miRanda settings were set at score =

145, energy = 10, gap open penalty = -9, gap extend penalty = -4 and scaling parameter = 4. RNAhybrid was used with default settings except for allowing up to 20 hits per UTR and providing a p-value threshold of 0.1.

The custom sequence database combined two separate databases, both created by Axel Martinelli of the Wellcome Trust Sanger Institute (Hinxton, England, UK), containing sequences downstream of the stop codon of genes currently annotated in *H. contortus*. The 3′ UTR database contained variable length sequences based on transcriptomic data, while the 3′ DS (downstream) database was generated by taking 1,000 nt downstream of a stop codon. The combined database contained either a UTR or DS sequence for each gene; approximately 95 % of genes had an associated sequence. The choice between UTR and DS sequence was determined by considering the length of the UTR sequence. If this sequence was > 50 nt and < 1,000 nt, it was included. Shorter (< 50 nt) or longer (> 1,000 nt) UTR sequences were replaced with DS sequences.

The 53 miRNAs used for interrogation displayed significant temporal and/or spatial variation as determined through microarray analysis (see Chapter 3.3.2, Fig. 3-1). Results were filtered for each algorithm after interrogation (miRanda: Score > 145 and energy < -10; PITA: seed sequences = 8 and $\Delta\Delta G < -10$; RNAhybrid: p-value < 0.1 and energy < -22). Thresholds were based on work by Marín and Vaníek (2011). For nine miRNAs (Hco-miR-60-3p, Hco-miR-228-5p, Hco-miR-235-3p, Hco-miR-5885a/b/c-3p and Hco-miR-5885a/b/c-5p), additional stringency filters were applied to miRanda (perfect alignment >= 0.7) and RNAhybrid (p-value < 0.05) results; PITA results were not filtered further.

4.2.2 *In silico* target prediction for miRNAs common to *H. contortus* and *C. elegans*

34 miRNAs common to both *H. contortus* and *C. elegans* were identified and used to predict targets in *C. elegans*. Six algorithms were used through their online portal: EIMMo (Wang and El Naqa, 2008), Microcosm (Griffiths-Jones *et al.*, 2006), MirWip (Hammell *et al.*, 2008), PicTar (Lall *et al.*, 2006), PITA (Kertesz et al., 2007) and TargetScan (Jan *et al.*, 2011). Default settings were used throughout. Results were compiled into a list of potential target genes for each miRNA. Genes predicted as a target of a particular miRNA by multiple algorithms were highlighted.

4.2.3 Transcriptomic data for predicted targets in *H. contortus*

Transcriptomic data was acquired for selected *H. contortus* genes predicted as miRNA targets of Hco-miR-60-3p, Hco-miR-228-5p, Hco-miR-235-3p and Hco-miR-5885. Of the six miR-5885 sequences used in target prediction (two arms from three loci), the arms were studied together, *i.e.* miR-5885a/b/c-3p and miR-5885a/b/c-5p. The data, based on RNAseq (RNA sequencing) provided by R. Laing (Laing *et al.*, 2013), was generated by the Wellcome Trust Sanger Institute using Augustus gene prediction (Keller *et al.*, 2011). The data consisted of triplicate biological samples covering six life-cycle stages (eggs, L3, activated-L3, L4, adult-males and adult-females) as well as adult female gut tissue. The data was normalised across all stages by A. Martinelli of the Wellcome Trust Sanger Institute, while the mean and standard deviation expression for each stage were calculated for the genes of interest. Genes with low expression (maximum RNAseq signal < 100) were excluded from further analysis.

Expression values were subjected to hierarchical clustering using the matrix visualization and analysis platform GENE-E, version 3.0.240 (http://www.broadinstitute.org/cancer/

software/ GENE-E/index.html). Row distance was calculated using a one-minus Pearson correlator metric and complete linkage (farthest-neighbour approach). All other settings were default. A rooted dendrogram and three-colour heat-map (blue = low, white = 50%, red = high) of the clustering was generated.

4.2.4 Identifying putative homologues of predicted targets in *H. contortus*

The putative homologues of selected *H. contortus* genes predicted as miRNA targets were identified in *C. elegans*. Protein sequences were acquired using annotation files of the latest version of the *H. contortus* genome provided by R. Laing. Amino acid sequences were then submitted for BLASTp analysis of the NCBI nr (non-redundant) protein sequences database, limited to *C. elegans* sequences. The closest match, query coverage (%) and identity (%) were recorded for each. Genes were considered homologous if both the query coverage and identity were > 40 %. Further confirmation was provided through reciprocal BLAST analysis, where the identified *C. elegans* gene was aligned to the *H. contortus* genome, confirming the best match to the original *H. contortus* gene. Gene functions were identified through Wormbase.

4.2.5 Gene ontology for predicted miRNA targets in *H. contortus* and *C. elegans*

Selected genes, predicted as miRNA targets *in silico* for *H. contortus* and *C. elegans*, were subject to analysis of gene ontology (GO). For *H. contortus* genes, putative *C. elegans* homologues were used while genes predicted in *C. elegans* were used directly. This was necessary as GO requires genes to be functionally annotated and *H.contortus* genes did not

have such information at the time of analysis. For both species, gene names were converted to Wormbase gene IDs using the gprofiler online tool (http://biit.cs.ut.ee/gprofiler/gconvert.cgi). The list of predicted targets for each miRNA was submitted to the AmiGO 2 online tool (http://geneontology.org/page/go-enrichment-analysis) for analysis of GO annotation enrichment. Enrichment was performed using three GO groups: biological processes (BP), molecular functions (MF) and cellular components (CC). Default settings were used for the analysis. Enrichments were ranked based upon the probability (p-value < 0.01) that the sample list contained more annotations than would be expected by chance, given the proportion of genes in the whole (*C. elegans*) genome that are annotated with that GO term.

4.3 Results

4.3.1 The predicted targets of variant miRNAs in *H. contortus*

To predict potential miRNA targets in *H. contortus*, a custom 3′ UTR nucleotide database was interrogated with 53 miRNAs using three algorithms (miRanda, PITA and RNAhybrid). These miRNAs displayed significant temporal and/or spatial variation in expression across L3, L4 and adult stages (male and female) as well as in adult female gut tissue (Chapter 3.3.1). Three variant miRNAs (Hco-let7-5p, Hco-miR-50-5p and Hco-miR-5908-5p) were omitted from the target prediction analysis for various reasons: Let-7 was not identified as a miRNA in *H. contortus* at the time of analysis while miR-50 and miR-5908 were not originally counted as variant miRNAs. In addition, one miRNA (Hco-miR-5885b-5p) that showed insignificant variation was included, to complete the analysis of miR-5885a/b/c. The custom database contained 3′ sequences 50 – 1,000 nucleotides, covering 95 % of the 21,779 genes currently annotated in *H. contortus*.

Across all 53 miRNAs, 20,323 genes were predicted as targets (Fig 4-1), accounting for 93 % of all genes currently annotated in *H. contortus*. miR-5885b-5p was predicted to target the greatest number of genes (15,842), while miR-9-5p was predicted to target the fewest (872); the average number of predicted targets across all miRNAs was 4,428. Between algorithms, RNAhybrid accounted for almost half of all predicted targets (47 %) while the remainder were split between miRanda and PITA (26 % and 27 % respectively). This may be a result of the relatively optimistic prediction parameters used as default settings by RNAhybrid. Four miRNAs (miR-5885b-5p, miR-5885a-5p, miR-5939-3p and miR-5960-5p) were predicted to target over 10,000 genes. Three of these miRNAs were classified as unique miRNAs (*i.e.* only found in *H. contortus*) while miR-5960-5p was classified as nematode-specific, based on analysis of miRBase (see Chapter 3.3.6). With respect to individual binding sites, the DS database provided the majority of hits with 340,985 (68 %) compared to the 157,740 (32 %) provided by the UTR database (Appendix, Fig. 8-5).

Analysis of target prediction based upon miRNA conservation status (widely conserved, nematode-only and unique to *H. contortus*) was performed (Fig 4-2). On average, the unique miRNAs of *H. contortus* were predicted to target the largest number of genes (mean = 5,195 genes), while the widely-conserved and nematode-specific miRNAs were predicted to target fewer genes (3,799 and 4,299 genes respectively). However, the unique group also displayed the greatest variation in number of predicted targets (range 1,788 – 16,242). For each miRNA, the number of predicted targets was examined in association with its mean abundance (across L3, L4, adult stages and gut tissue) as determined through microarray analysis (see Chapter 3.3.1); no significant correlation was identified ($r = -0.18$).

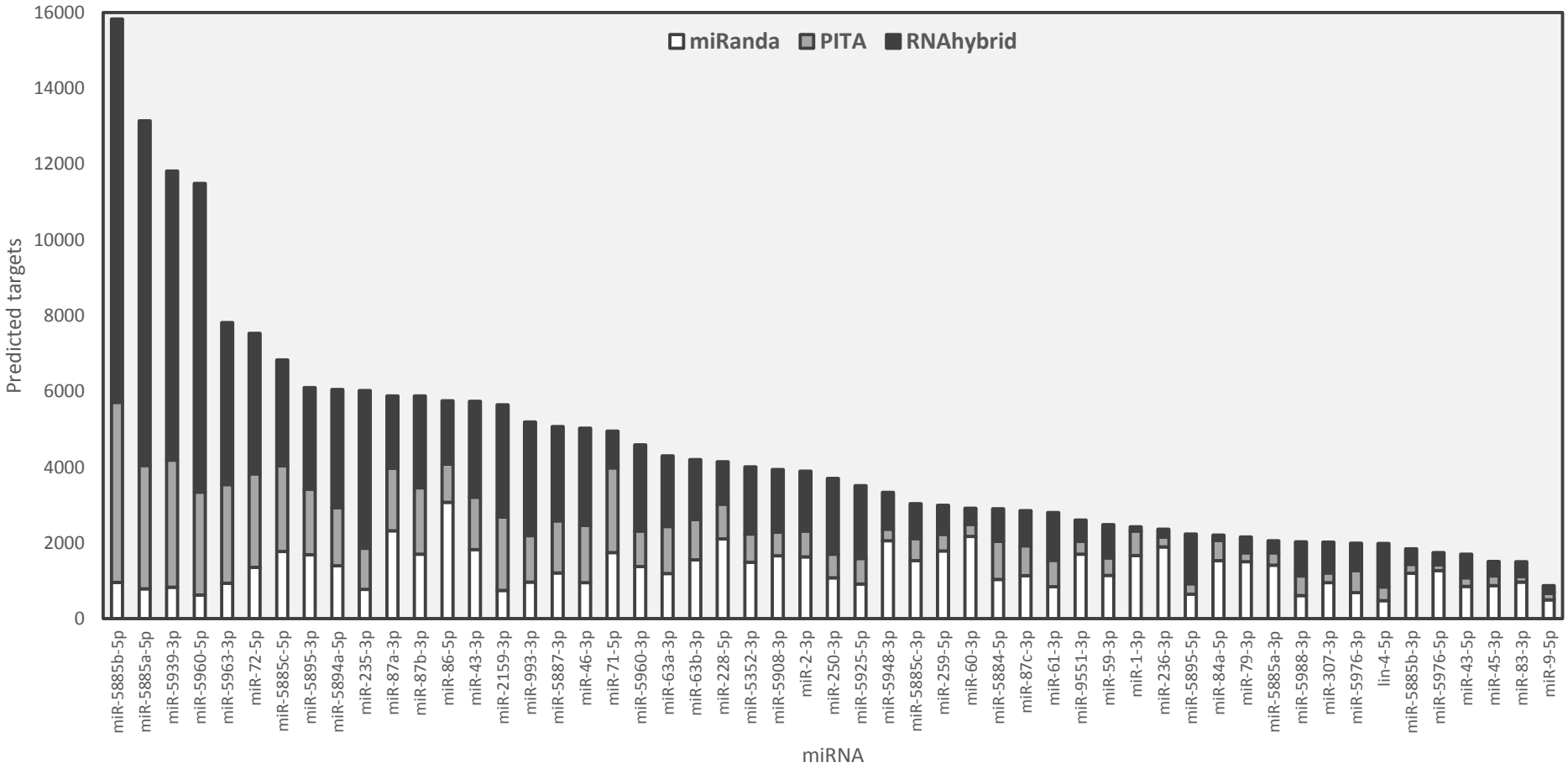


Fig. 4-1: Number of *H. contortus* genes predicted as miRNA targets *in silico*. Colours represent the proportion identified by each algorithm (white = miRanda, grey = PITA and black = RNAhybrid). Genes predicted as a target for a particular miRNA are counted once, regardless of how many binding sites were identified. All 52 miRNAs displayed significant temporal or spatial variation in abundance across L3, L4, adult male and adult female stages as well as adult female gut tissue.

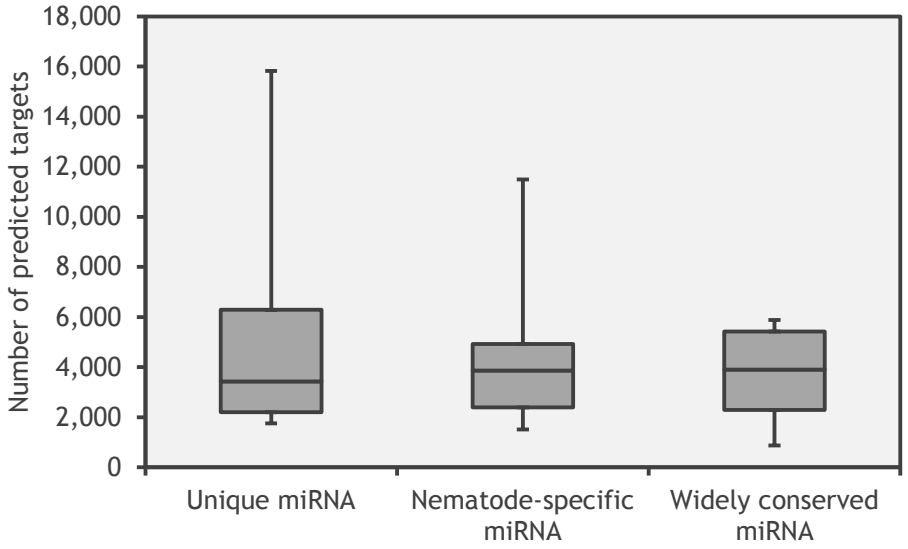


Fig. 4-2. Box-and-whisker plot of number of *H. contortus* genes predicted as targets of variant miRNAs, grouped by conservation status. Conservation groups ('widely conserved', 'only found in nematodes' and 'unique to *H. contortus*') were identified through interrogation of miRBase (see Chapter 3.3.6). The whiskers represent minimum and maximum values.

The target dataset included predictions for multiple loci of the same miRNA (miR-63a/b-3p, miR-87a/b/c-3p, miR-5885a/b/c-3p and miR-5885a/b/c-5p; Table 4-1). In general, the arms of a miRNA stem-loop expressed at different loci (*e.g.* miR-63a-3p and miR-63b-3p) were predicted to target very few common genes (2.7 – 36.2 %), highlighting how large differences in target prediction can arise from small changes in miRNA sequence.

There were seven pairs of miRNAs belonging to a single miRNA duplex, *i.e.* the same locus but different arms: miR-43, miR-5885a/b/c, miR-5895, miR-5960 and miR-5976 (Table 4-2). Significant variation was seen in the number of targets predicted between arms of the same duplex, with average fold difference of 3.9 (range 1.1 – 8.6). Counterintuitively, the arm with the highest mean expression was predicted to target the fewest genes. This was true for all pairs except miR-43. Additionally, the number of targets common to both arms was very low (range 6 – 22 %). Together, these suggest that different arms of a miRNA duplex are involved in different regulatory systems.

Four miRNA clusters were also examined, covering fourteen miRNAs: miR-2/71, miR-43/61/5352/5895, miR-63a/b and miR-5885b/c (see Chapter 3.3.4). The number of predicted targets common to all miRNAs in a cluster differed markedly (Table 4-3). 31.5 % of targets were common to both miR-63a and miR-63b, while only 0.1 % were common to all members of the miR-43/61/5352/5895 cluster. This reflects the similarity in sequences between miR-63a and miR-63b, while also suggesting that although miRNAs in a cluster are co-ordinately expressed, they target different genes.

The miRNA target prediction algorithms also identified individual binding sites within each gene. In total 440,718 miRNA binding sites were predicted with each gene containing an average of 25 binding sites. The highest number of binding sites was seen in three genes (HCISE00489200, HCISE00489600 and HCISE01717200), each with 51 sites with 49, 51 and 51 miRNAs respectively. BLASTp searches identified putative *C. elegans* homologues as IARS-2 (isoleucyl amino-acyl tRNA synthetase, involved in post-embryonic growth), ECH-1.2 (enoyl-CoA hydratase, involved in fatty acid metabolism) and T28D6.6 (a developmentally regulated GTP binding protein). *Cel-iars-2* has been experimentally shown to be enriched in AIN-1 and AIN-2 pulldowns, consistent with miRNA regulation (Wormbase).

Table 4-1: Number of predicted targets for *H. contortus* miRNAs expressed at multiple loci.

miRNA	Targets	Total targets[1]	Common[2]	Percentage[3]
Hco-miR-63a-3p	4,300	6,463	2,034	31.47%
Hco-miR-63b-3p	4,197			
Hco-miR-87a-3p	5,007	8,359	228	2.73%
Hco-miR-87b-3p	6,009			
Hco-miR-87c-3p	2,921			
Hco-miR-5885a-3p	2,060	3,494	1,264	36.18%
Hco-miR-5885b-3p	1,846			
Hco-miR-5885c-3p	3,035			
Hco-miR-5885a-5p	13,140	18,122	4,574	25.24%
Hco-miR-5885b-5p	15,824			
Hco-miR-5885c-5p	6,834			

[1] Total number of targets for all loci expressing a miRNA (*i.e.* removes duplicates).

[2] Common targets between loci of a particular miRNA.

[3] Common targets as a percentage of the total targets.

Table 4-2: Number of predicted target genes for different arms of an *H. contortus* miRNA.

miRNA	Targets	Total Targets[1]	Fold difference[2]	Common[3]	Percentage[4]
Hco-miR-43-3p	5,741	6,887	3.4	559	8.12%
Hco-miR-43-5p	1,705				
Hco-miR-5885a-3p	2,060	13,812	6.4	1,388	10.05%
Hco-miR-5885a-5p	13,140				
Hco-miR-5885b-3p	1,846	16,168	8.6	1,522	9.41%
Hco-miR-5885b-5p	15,824				
Hco-miR-5885c-3p	3,035	8,942	2.3	1,121	12.54%
Hco-miR-5885c-5p	6,834				
Hco-miR-5895-3p	6,101	7,560	2.7	780	10.32%
Hco-miR-5895-5p	2,239				
Hco-miR-5960-3p	4,588	13,188	2.5	2,889	21.91%
Hco-miR-5960-5p	11,489				
Hco-miR-5976-3p	1,998	3,548	1.1	198	5.58%
Hco-miR-5976-5p	1,748				

[1] Total number of targets for both arms of a miRNA (*i.e.* removes duplicates).

[2] Positive fold difference in the number of predicted targets between each arm of a miRNA

[3] Common targets between arms of a particular miRNA.

[4] Common targets as a percentage of the total targets.

Table 4-3: Number of predicted target genes for *H. contortus* miRNAs expressed within a cluster.

miRNA	Targets	Total targets[1]	Common[2]	Percentage[3]
Hco-miR-2-3p	3,895	6,463	1,045	13.4 %
Hco-miR-71-3p	4,955			
Hco-miR-43-3p	5,741	13,854	10	0.1 %
Hco-miR-43-5p	1,705			
Hco-miR-61-3p	2,805			
Hco-miR-5352-3p	4,109			
Hco-miR-5895-3p	6,101			
Hco-miR-5895-5p	2,239			
Hco-miR-63a-3p	4,389	6,463	2,034	31.5 %
Hco-miR-63b-3p	4,284			
Hco-miR-5885b-3p	1,846	17,402	432	2.5 %
Hco-miR-5885b-5p	15,824			
Hco-miR-5885c-3p	3.035			
Hco-miR-5885c-5p	6,834			

[1] Total number of targets for members of a miRNA clusters (*i.e.* removes duplicates).

[2] Common targets between all members of a miRNA cluster.

[3] Common targets as a percentage of the total targets.

4.3.2 Predicted targets for miR-60, miR-228 and miR-235 in *H. contortus*

Detailed analysis of *in silico* predicted targets was performed for three miRNAs: Hco-miR-60-3p, Hco-miR-228-5p and Hco-miR-235-3p. Microarray analysis (Fig. 3-1) highlighted interesting abundance profiles for these miRNAs: miR-228 peaked at L3 before decreasing through L4 and adult stages, miR-235 showed a trough in abundance during L4 while remaining high in L3 and adult stages, while miR-60 showed peaks in abundance during L4 and in adult gut tissue. Predicted targets were filtered using higher stringency settings generating 1,166, 1,638 and 3,115 genes respectively (Fig. 4-3A). 556 genes were predicted as targets for two miRNAs (*i.e.* miR-60 + miR-228, miR-60 + miR-235 or miR-228 + miR-235), while 28 genes were common targets of all three miRNAs (Fig. 4-3B).

Genes were filtered by the number of algorithms that concurrently identified each gene for a particular miRNA (Fig 4-3C). For miR-60 and miR-235, as the number of concurrent algorithms increased, the number of predicted targets decreased. This was not true for miR-228, where most targets were predicted by two algorithms. A similar situation existed when the number of miRNA binding sites per gene were studied (Fig. 4-3D). For miR-60 and miR-235, the majority of genes were predicted to contain only one binding site, with decreasing numbers containing multiple sites. At most, predicted targets of miR-60 contained a maximum of four binding sites per gene, while miR-235 targets contained a maximum of seven. In contrast, most predicted targets of miR-228 (56 %) were predicted to contain two binding sites, while the maximum number of binding sites was five. Genes with multiple binding sites and identified by all three algorithms provided greater confidence in target predictions.

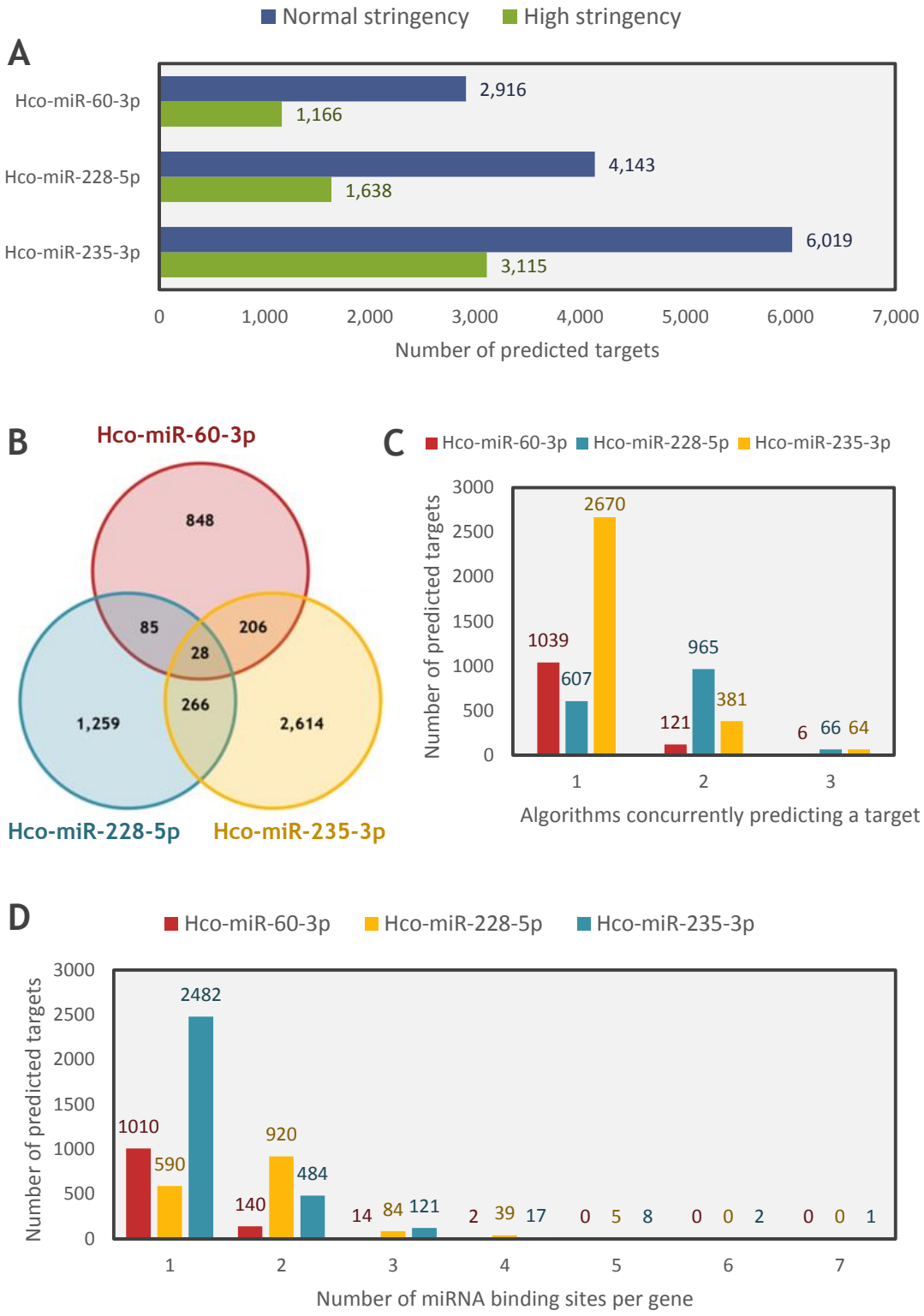


Fig. 4-3: Number of *H. contortus* genes predicted as miRNA targets of Hco-miR-60-3p, Hco-miR-228-5p and Hco-miR-235-3p. **A**: Number of predictions using low and high stringency settings. **B**: Venn diagram of common predicted targets. **C**: Number of predicted targets identified by one, two or three algorithms concurrently. **D**: Number of predicted binding sites per target gene.

4.3.3 Predicted targets for miR-5885 in *H. contortus*

In a similar fashion to targets for miR-60, miR-228 and miR-235, predicted targets for miR-5885 were studied in greater detail. This included six mature miRNA sequences covering three loci (a, b and c), each with two arms (-3p and -5p) that were among the most abundant miRNAs during microarray analysis and displayed interesting expression profiles (Chapter 3.3.3). The predictions for these six miRNAs were studied in two groups: Hco-miR-5885a/b/c-3p and Hco-miR-5885a/b/c-5p. As the -3p arm showed much higher levels of abundance than the -5p arm, it was expected that -3p arm would also be predicted to target more genes. Surprisingly, this was not the case: 3,494 targets were predicted for Hco-miR-5885a/b/c-3p and 18,618 targets for miR-5885a/b/c-5p. Increasing the stringency (as used previously in Chapter 4.3.2) reduced the number of predictions to 2,506 and 15,909 respectively, of which 2,025 genes were common to both arms. This suggests there is a core set of genes regulated by both arms, but that the -5p arm targets an additional set of genes.

4.3.4 Predicting *C. elegans* miRNA targets *in silico*

Prior to the release of the HCISE genome (Laing *et al.*, 2013), large scale *in silico* target prediction was unavailable for *H. contortus*. The alternative was to use algorithms designed for use in the relatively closely related *C. elegans* as a proxy. Targets were predicted for 34 miRNAs common to both *H. contortus* and *C. elegans*. Combining the results from six algorithms, 62,335 binding sites were predicted across 14,190 genes, accounting for 69 % of all genes annotated in *C. elegans*. The majority of genes (77.1 %) were predicted as a target of a particular miRNA by only one algorithm while only 237 (0.6 %) were predicted as targets by all six algorithms (Fig. 4-4A).

Cel-miR-1 generated the greatest number of targets (n = 2,814) while Cel-miR-265 generated the fewest (n = 366) (Fig. 4-4B). Overall, the number of targets predicted for a miRNA in *C. elegans* showed negligible correlation with the number of predictions for the same miRNA in *H. contortus* (data not shown). While different algorithms and stringency settings were used for each species, it suggests that the number of genes targeted by a miRNA is not consistent between species.

In terms of specific algorithms, Microcosm and PITA both generated the most targets (18,457 and 18,044 respectively) while Eimmo generated the fewest (n = 5,812) (Fig. 4-4C). There were six instances where an algorithm failed to identify any targets for a particular miRNA: Cel-miR-45 (Microcosm and PITA), Cel-miR-76 (Eimmo) and Cel-miR-265 (Eimmo and PicTar). TargetScan always predicted a target (minimum = 36). Cel-miR-60, Cel-miR-228 and Cel-miR-235 were predicted to target 1,081, 1,144 and 848 genes respectively. This contrasts with predictions in *H. contortus* where miR-235 was predicted to target the most genes (n = 3,115) and miR-60 the fewest (n = 1,166). This could be a result of poorly annotated *H. contortus* genes or due to functional differences.

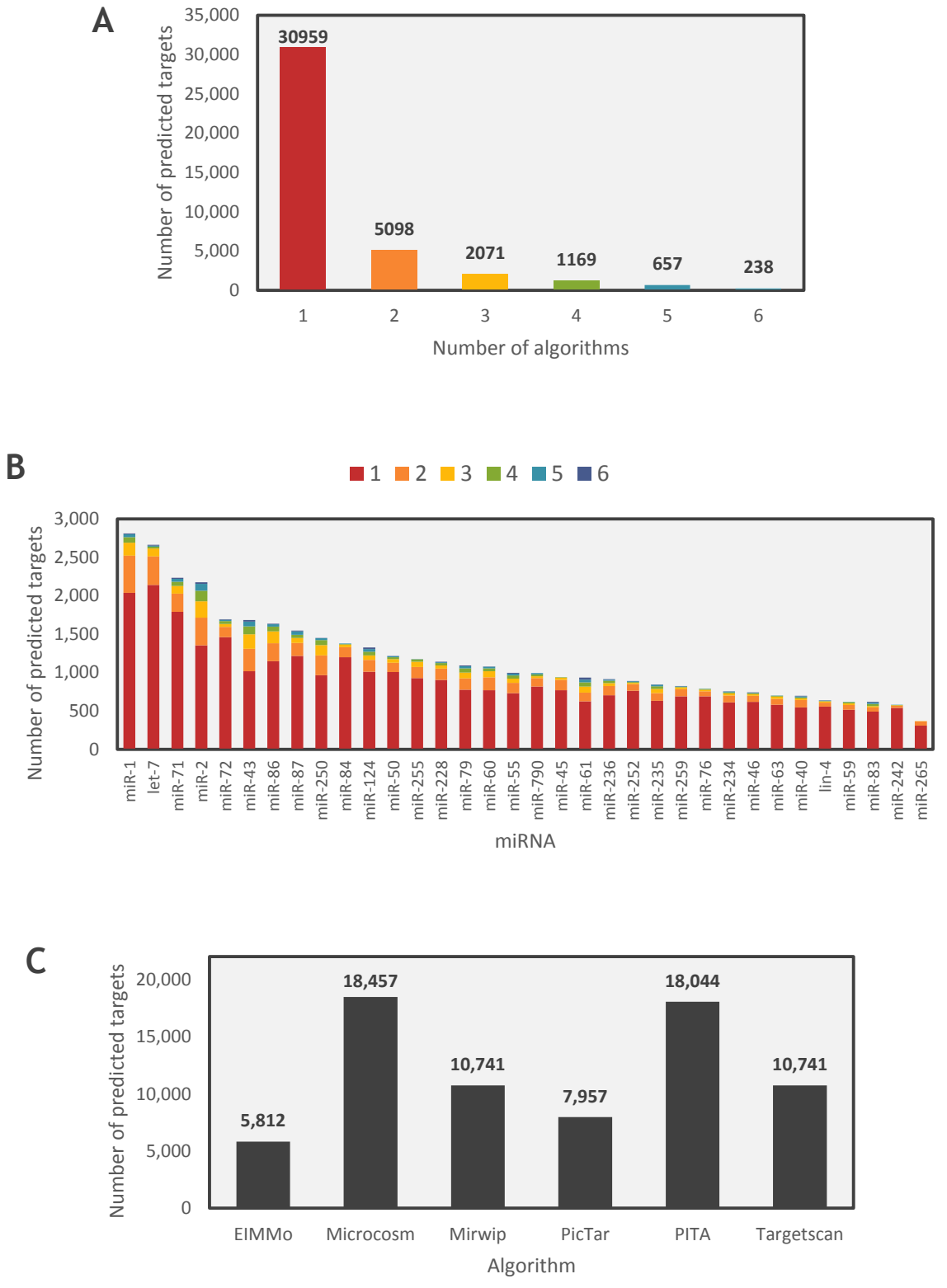


Fig. 4-4: Number of *C. elegans* genes predicted as targets across 34 miRNAs. **A**: Number of targets concurrently predicted by 1 – 6 algorithms. **B**: Number of predicted targets per miRNA. Colours indicate the number of algorithms that concurrently predicted that gene. **C**: Number of predicted targets per algorithm.

4.3.5 qRT-PCR abundance of predicted targets in *C. elegans*

Of the genes predicted as targets of Cel-miR-228, three genes *(hbl-1*, *hpd-1* and *T04A8.7*) were particularly strong candidates, identified by multiple algorithms. If these genes were indeed regulated by miRNAs, their expression should be affected by removal of the miRNA (*i.e.* miR-228) or parts of the RISC (*e.g. alg-1* or *ain-1*). To confirm this, mRNA abundance was examined by qRT-PCR of the mid-L4 stage of three mutant strains lacking *mir-228*, *alg-1* and *ain-1* (MT14446, RF54 and MH2385, respectively; Fig. 4-5A/B/C).

Abundance of *T04A8.7* (Fig. 4-5A) was significantly higher in all three mutant strains compared to wildtype worms, suggesting that *T04A8.7* is regulated by miRNAs. Additionally, abundance was higher in the *alg-1*(-) and *ain-1*(-) mutants than the *mir-228*(-) mutant, suggesting that additional miRNAs regulate T04A8.7. For *hpd-1* (Fig. 4-5B), abundance was significantly higher in the RISC mutants compared to wildtype worms, while expression in *mir-228*(-) mutant was lower. This suggests that *hpd-1* is regulated by miRNAs, but not directly by miR-228. In contrast, hbl-1 abundance(Fig. 4-5C) displayed no significant variation between strains. Tiling array data (Spencer *et al.*, 2011) indicated that abundance of *hbl-1* decreased significantly in L4 and adult stages. As abundance did not increase in the mutant strains, the low abundance of *hbl-1* in the L4 stage was not due to negative miRNA regulation.

Using BLASTp, a homologue of T04A8.7 was found in *H. contortus*, labelled as HCISE HCISE00019000 but hereafter referred to as Hco-T04A8.7. As no mutants were available in *H. contortus*, qRT-PCR was used to examine expression across four different life cycle stages (L3, L4, adult males and adult females; Fig 4-5D). Abundance of *Hco-T04A8.7* was low in L3 and high in L4 and adult stages. This pattern was inverse to the pattern seen for Hco-miR-228 (high in L3, low in L4/adult stages), suggesting that miR-228 regulation of *T04A8.7* is conserved between *C. elegans* and *H. contortus*.

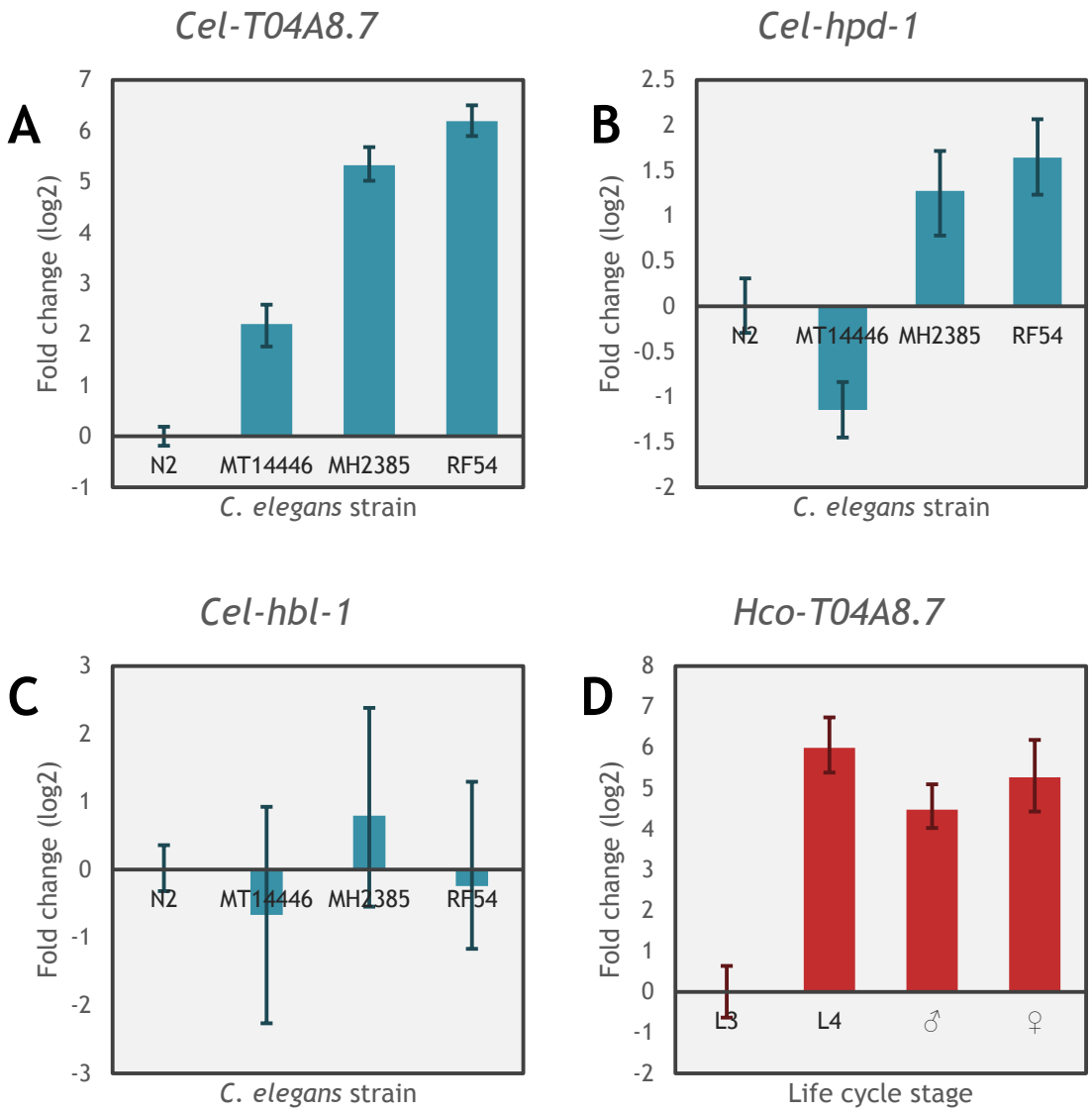


Fig. 4-5: qRT-PCR expression of genes predicted as targets of miR-228. Expression was quantified for three *C. elegans* genes (panels A – C) in wildtype (N2) and mutant strains: *mir-228*(-), *ain-1*(-) and *alg-1*(-) (MT14446, MH2385 and RF54 respectively). Expression was also quantified for the *H. contortus* homologue of *T04A8.7* (panel D) across four life cycle stages: L3, L4, adult males (♂) and adult females (♀). Y-axis represents log-2 fold changes in relative abundance. Vertical bars represent error values calculated from three technical replicates.

4.3.6 Putative homologues of *H. contortus* genes predicted as miRNA targets

To understand the potential function of targets predicted in *H. contortus*, putative homologues were identified in *C. elegans*, the closest relative for which large amounts of functional information was available. Putative homologues were identified for targets of five miRNAs (miR-60-3p, miR-228-5p, miR-235-3p, miR-5885-3p and miR-5885-5p; Figs. 4-7 – 4-11 respectively). The length of proteins of *H. contortus* and *C. elegans* homologues were compared (Fig. 4-6) as a proxy for annotation accuracy. As the *C. elegans* genes were better annotated (being studied over many years by multiple researchers) and assuming a high degree of sequence conservation, well-annotated *H. contortus* genes (and hence proteins) were expected to be of similar length to their putative *C. elegans* homologues. Overall, this proved to be true, with homologues displaying a strong positive correlation in length (r = 0.901). The average variation in length was 115 amino acids, with *H. contortus* proteins tending to be shorter, suggesting incomplete annotation, *e.g.* incorrect start/stop positions or missing exons. In contrast, unspliced *H. contortus* genes are known to be much larger than their *C. elegans* homologues due to an increased number of introns (Laing *et al.*, 2013).

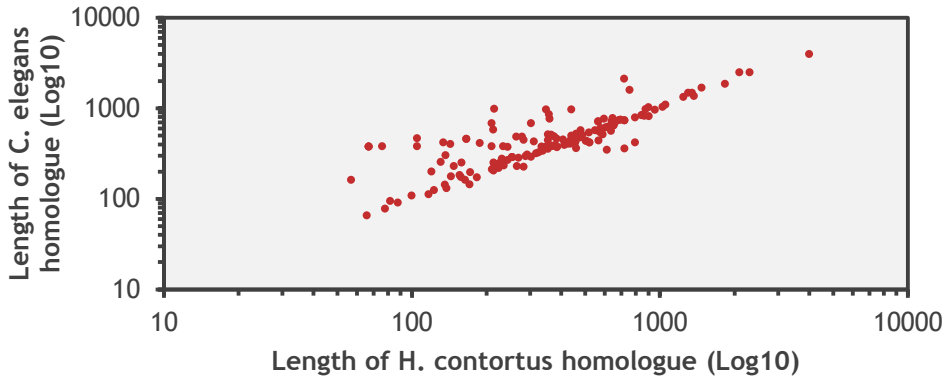


Fig. 4-6: Comparison of protein lengths of *H. contortus* miRNA targets and their putative *C. elegans* homologues. Data covered 160 proteins predicted as the targets of Hco-miR-60-3p, Hco-miR-228-5p, Hco-miR-235-3p, Hco-miR-5885a/b/c-3p and Hco-miR-5885a/b/c-5p. Both *H. contortus* (x-axis) and *C. elegans* (y-axis) values were plotted using a logarithmic (base 10) scale. The correlation coefficient was 0.901.

4.3.7 Abundance data for predicted targets of miRNAs in *H. contortus*

As miRNAs regulate gene expression at the post-transcriptional level, transcriptomic data for predicted targets was studied to identify inverse relationships and support predicted miRNA-mRNA interactions. Selected genes were high stringency targets of five miRNAs (Hco-miR-60-3p, Hco-miR-228-5p, Hco-miR-235-3p, Hco-miR-5885a/b/c-3p and Hco-miR-5885a/b/c-5p) predicted by three algorithms concurrently while also having a minimum level of abundance (normalised RNAseq signal > 100) during at least one life cycle stage. mRNA abundance was compared against miRNA expression obtained using microarrays (Chapter 3.3.1, summarised in Table 4-4). Putative *C. elegans* homologues are indicated when identified.

From 1,166 high-stringency predictions for miR-60, 31 genes fulfilled the selection criteria (3 algorithms, abundance > 100; Fig. 4-7). miR-60 was most abundant during L4 and in gut tissue, so an inverse pattern would be expected of its targets. Five genes in cluster D show such a pattern, namely *HCISE001122000*, *HCISE01174900* (*C. elegans* homologue *C39E9.11*), *HCISE01248700* (*F44B9.8*), *HCISE01456100* (*F44B9.8*) and *HCISE00582700* (*M05D6.2*). In *C. elegans*, *C39E9.11* is orthologous to human *prcc*, known to regulate the cell cycle (Weterman *et al.*, 2001), while *F44B9.8* is orthologous to human *rfc1* which is required for elongation of DNA during DNA replication and transcription (Mossi *et al.*, 1997).

For miR-228, 1,638 high-stringency targets were predicted, of which 99 fulfilled the selection criteria (Fig. 4-8). miR-228 abundance was highest during L3 decreasing in L4 and adult stages.. 22 genes from clusters C and D displayed a significant increase in expression between L3 and L4, however none showed continuing increases in the adult stages. The closest matches were *HCISE00545400* and *HCISE01678400* (*C. elegans sma-*

1). *Cel-sma-1* is a beta-H spectrin, involved in embryonic elongation and morphogenesis of the pharynx, intestine and excretory cell (McKeown *et al.*, 1998).

Of 3,115 high stringency targets predicted for miR-235, 126 were selected for analysis (Fig. 4-9). miR-235 displayed a trough in abundance during L4, with increased abundance in adult stages and negligible levels in gut tissue. Four genes (cluster E) showed an inverse pattern to miR-235 expression: *HCISE01780400* (*C. elegans* homologue *mec-12*), *HCISE02189600* (*atgp-1*), *HCISE01806500* (*mrpl-36*) and *HCISE01840100*. *Cel-mec-12* encodes an alpha-tubulin found within touch receptors (Fukushige *et al.*, 1999), *Cel-atgp-1* encodes a glycoprotein subunit of a cell surface amino acid transporter (Veljkovic *et al.*, 2004) and *Cel-mrpl-36* encodes a structural component of mitochondrial ribosomes (Camon *et al.*, 2003).

56 of the 2,506 high stringency predicted targets of miR-5885a/b/c-3p fulfilled the selection criteria (Fig. 4-10). The abundance of miR-5885a-3p shows significant peaks during L4 and in gut tissue, with lowest abundance was in L3. 27 genes displayed low abundance during L4 and in gut tissue, but only HCISE00173900 (*C. elegans* homologue *pif-1*) and HCISE00098000 (*inx-10*) concurrently showed higher levels in the remaining three stages. *Cel-pif-1* encodes a DNA-dependent ATPase that maintains nuclear and mitochondrial genome integrity(Yoneda *et al.*, 2004) while *Cel-inx-10* encodes a structural component of gap junctions (Starich *et al.*, 2001).

15,909 high stringency targets were predicted for miR-5885a/b/c-5p, of which 1,032 fulfilled the criteria. As this was a high number to examine individually for abundance patterns, further filters were applied: genes were only analysed if they had > 11 binding sites (in addition to the previous criteria). This produced 58 targets (Fig. 4-11). In contrast to its -3p arm, miR-5885-5p only shows a peak of abundance during the adult female stage. The two clearest examples of an inverse pattern were seen in cluster A with

HCISE01323100 (*Hco-rsbp-1*) and *HCISE01154500* (*Hco-fat-2*). *Cel-rsbp-1* encodes a binding protein found in neurons and striated muscle that negatively regulates locomotion (Porter and Koelle, 2010) while *Cel-fat-2* encodes a delta-12 fatty acyl desaturase required for normal growth rates, movement rate, body shape, defaecation rhythm and brood size (Watts and Browse, 2002).

Table 4-4: Microarray abundance data of *H. contortus* miRNAs used for transcriptomic analysis of predicted targets.

miRNA	Expression				
	L3	L4	Male	Female	Gut
Hco-miR-60-3p	243	22,267	1,824	5,484	32,242
Hco-miR-228-5p	3,071	696	466	253	21
Hco-miR-235-3p	2,762	80	1,787	1,113	60
Hco-miR-5885a/b/c-3p[1]	754	42,077	12,962	19,813	35,745
Hco-miR-5885a/b/c-5p[1]	49	43	393	3,738	230

[1] Mean expression across all three loci.

	Gene	Putative homologue	Life cycle stage				
			L3	L4	Male	Female	Gut
A	HCISE00009000	*prx-19*	689	396	368	337	512
	HCISE01746300	*vps-60*	1,791	784	706	948	1,387
	HCISE01950600	-	337	175	93	111	303
B	HCISE01840100	-	318	76,185	7,170	1,237	68,266
	HCISE02181200	*Y61A9LA.11*	80	268	219	176	442
	HCISE00345200	*let-23*	130	687	607	110	1,573
	HCISE00539500	*rab-11.1*	2,242	2,592	1,529	1,620	3,649
	HCISE01346400	*sod-4*	428	666	467	71	2,293
	HCISE01357700	-	43	29	13	7	109
	HCISE01450000	-	359	1,245	605	1,818	6,150
	HCISE01797200	*ell-1*	133	252	128	296	805
	HCISE00738600	*acl-14*	710	283	426	529	1,556
C	HCISE00905300	-	164	142	190	202	230
	HCISE00324300	-	96	82	406	172	180
	HCISE00513500	-	4	22	911	30	45
	HCISE01352100	-	3	0	8,568	7	12
	HCISE00510400	-	0	0	442	0	0
	HCISE02030100	*nipa-1*	138	229	865	227	640
	HCISE01619400	*lnkn-1*	811	956	1,239	829	1,001
D	HCISE00112000	-	1,202	607	758	887	555
	HCISE01174900	*C39E9.11*	1,573	898	608	1,806	981
	HCISE02141100	*rpt-6*	666	635	669	694	303
	HCISE00147200	*hot-3*	177	207	79	261	14
	HCISE01248700	*F44B9.8*	123	91	163	144	49
	HCISE01456100	*F44B9.8*	206	108	207	210	79
	HCISE00582700	*M05D6.2*	440	153	684	598	191
E	HCISE00838400	*rpl-33*	8,779	12,404	3,922	4,217	6,330
	HCISE00021100	*rpl-15*	4,291	4,513	1,569	2,211	1,627
	HCISE00429600	-	2	73	5	13	8
	HCISE00795300	-	36	104	21	42	27
	HCISE00620200	*F47F2.1*	62	46	45	1	0

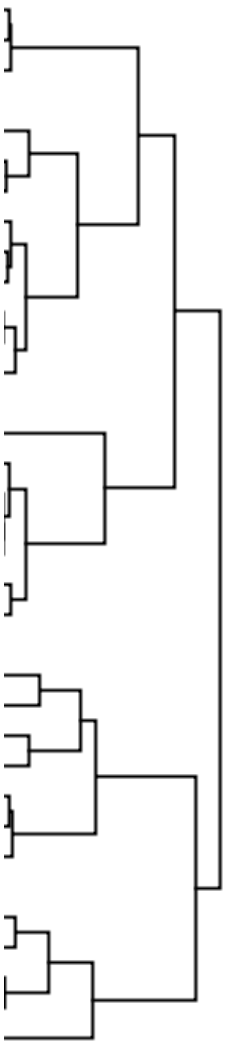
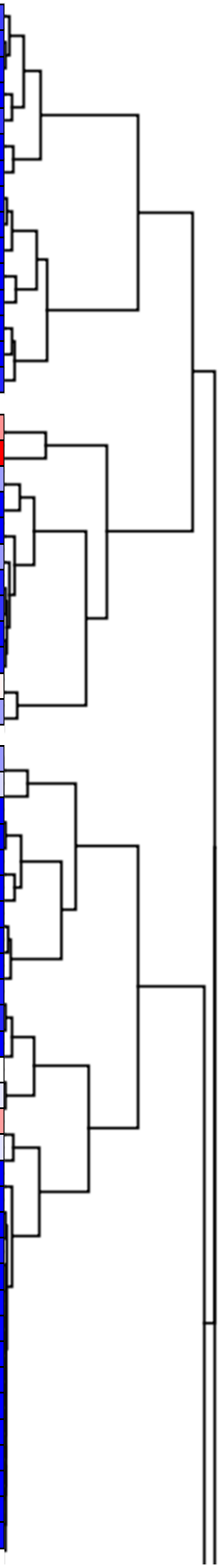
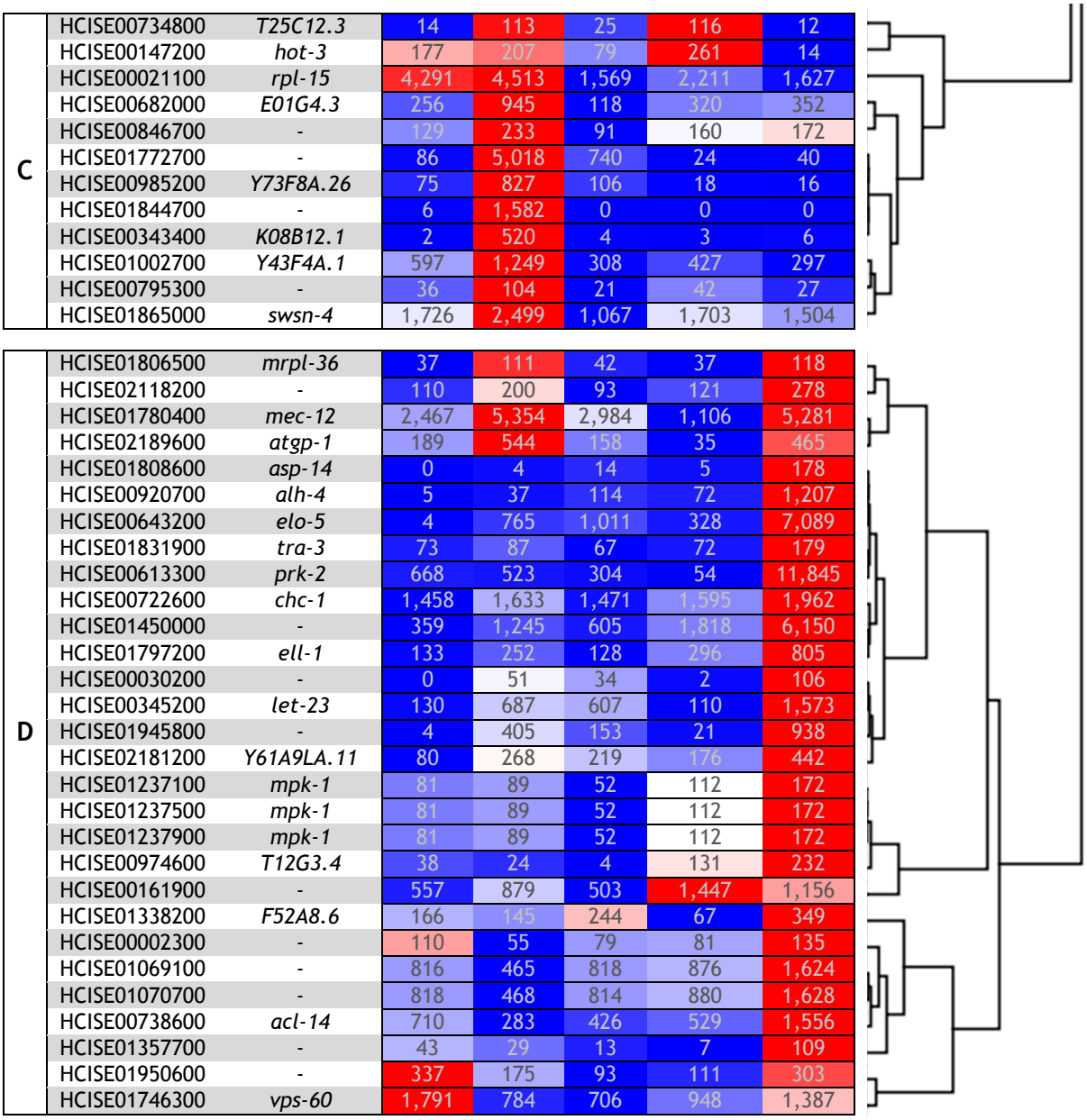


Fig. 4-7: Hierarchical clustering of *H. contortus* genes predicted as targets for Hco-miR-60-3p. Clustering of genes, shown as a dendrogram (right hand side), was based on RNA-seq abundance data highlighted using a relative heatmap (blue = low, white = median, red = high). Putative *C. elegans* homologues are listed when identified. Five genes (marked in red) displayed inverse expression patterns with Hco-miR-60-3p.

	Gene	Putative Homologue	Life cycle stage				
			L3	L4	Male	Female	Gut
A	HCISE01526500	*lev-1*	263	67	21	7	40
	HCISE01451400	-	21,903	11,226	4,967	3,016	4,410
	HCISE01806600	*glrx-5*	367	233	122	105	112
	HCISE00832700	-	285	18	14	61	0
	HCISE01323800	*T08B2.5*	1,892	495	827	483	617
	HCISE00477700	*cal-4*	298	77	16	176	3
	HCISE00849700	-	288	169	182	227	97
	HCISE00472600	*C05D2.10*	185	87	172	271	21
	HCISE01132200	-	84	28	81	162	0
	HCISE01337700	*H35B03.2*	74	66	84	148	36
	HCISE01086900	*Y48G1A.4*	199	228	230	284	117
	HCISE02141100	*rpt-6*	666	635	669	694	303
	HCISE01248700	*F44B9.8*	123	91	163	144	49
	HCISE01456100	*F44B9.8*	206	108	207	210	79
	HCISE00582700	*M05D6.2*	440	153	684	598	191
B	HCISE02181400	-	1,437	518	1,619	1,535	1,307
	HCISE00905300	-	164	142	190	202	230
	HCISE01732500	*F19B6.1*	406	527	1,087	1,485	741
	HCISE00912400	-	599	1,096	1,251	1,922	582
	HCISE01057800	-	62	63	28	300	20
	HCISE00941300	*lmn-1*	342	117	359	1,091	409
	HCISE01789600	*mtr-4*	434	326	532	1,261	384
	HCISE01538600	*rabn-5*	322	284	419	814	340
	HCISE02075200	-	0	2	10	494	13
	HCISE00114100	-	0	1	31	773	49
	HCISE01672000	*elpc-3*	97	45	57	127	88
	HCISE01174900	*C39E9.11*	1,573	898	608	1,806	981
C	HCISE00545400	-	291	490	529	535	364
	HCISE01678400	*sma-1*	4,097	9,811	8,119	7,135	6,526
	HCISE00480600	C38H2.2	110	144	197	159	79
	HCISE01286200	*rpn-2*	1,120	1,319	2,025	1,512	617
	HCISE01001000	-	417	1,117	1,220	639	44
	HCISE00845500	-	323	667	1,010	378	257
	HCISE01615400	*ttr-27*	15	402	357	14	2
	HCISE01534200	-	10	2,133	1,288	103	6
	HCISE00607600	*twk-31*	233	606	431	76	68
	HCISE01749200	-	2,146	785	3,285	1,073	919
	HCISE01666600	-	1,473	580	2,019	662	689
	HCISE00416400	*unc-13*	1,878	762	2,277	572	84
	HCISE00091300	*cnx-1*	3,331	2,742	4,710	1,095	2,903
	HCISE00091700	*cnx-1*	2,134	1,875	3,085	741	1,801
	HCISE02030100	*nipa-1*	138	229	865	227	640
	HCISE01619400	lnkn-1	811	956	1,239	829	1,001
	HCISE00243500	-	86	43	381	141	43
	HCISE00237100	*tbg-1*	126	147	339	133	111
	HCISE00051800	-	37	62	524	0	0
	HCISE01212800	-	0	34	1,975	134	73
	HCISE00927900	-	21	0	1,269	1	2
	HCISE01444500	-	0	0	186	0	2
	HCISE00825700	-	1	4	519	0	4
	HCISE02051500	*fut-1*	1	0	609	0	0
	HCISE01461500	-	10	12	4,599	2	2
	HCISE00985300	-	0	1	1,389	2	0
	HCISE00238300	-	1	0	9,611	7	0
	HCISE02184200	*psr-1*	0	0	830	0	0
	HCISE00510400	-	0	0	442	0	0
	HCISE01352100	-	3	0	8,568	7	12
	HCISE01869100	-	2	13	277	6	7

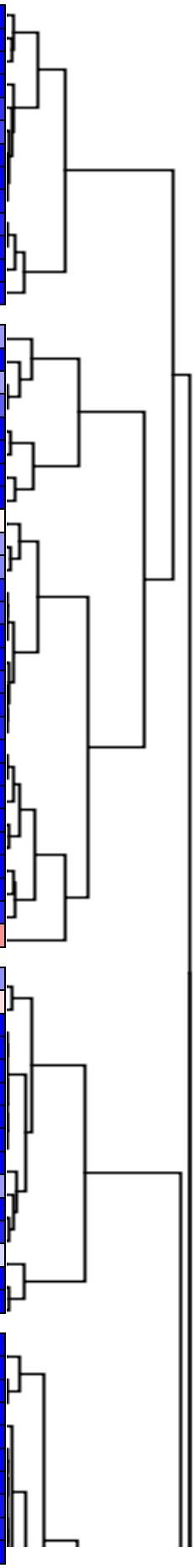


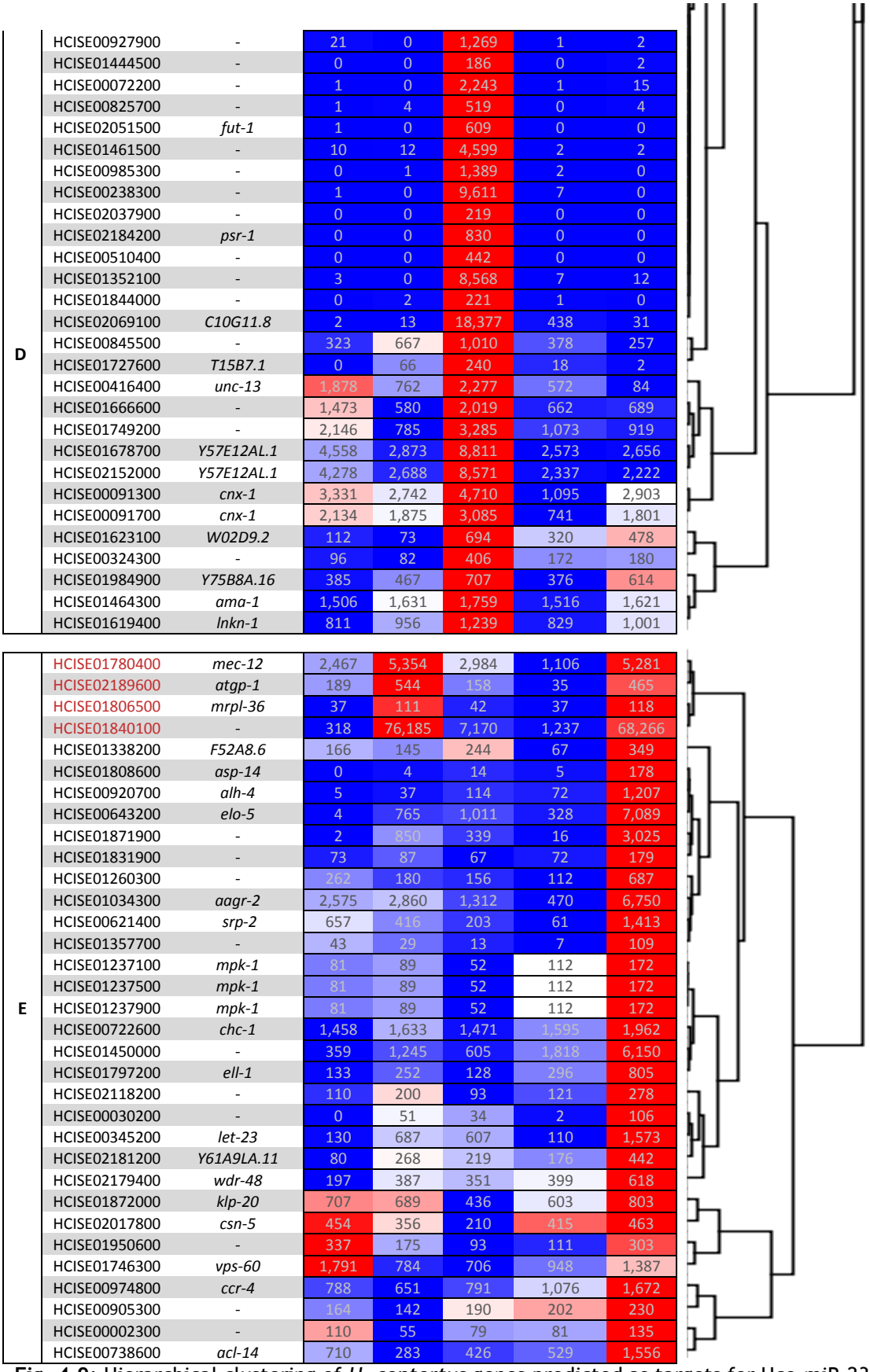


C	HCISE00734800	*T25C12.3*	14	113	25	116	12
	HCISE00147200	*hot-3*	177	207	79	261	14
	HCISE00021100	*rpl-15*	4,291	4,513	1,569	2,211	1,627
	HCISE00682000	*E01G4.3*	256	945	118	320	352
	HCISE00846700	-	129	233	91	160	172
	HCISE01772700	-	86	5,018	740	24	40
	HCISE00985200	*Y73F8A.26*	75	827	106	18	16
	HCISE01844700	-	6	1,582	0	0	0
	HCISE00343400	*K08B12.1*	2	520	4	3	6
	HCISE01002700	*Y43F4A.1*	597	1,249	308	427	297
	HCISE00795300	-	36	104	21	42	27
	HCISE01865000	*swsn-4*	1,726	2,499	1,067	1,703	1,504
D	HCISE01806500	*mrpl-36*	37	111	42	37	118
	HCISE02118200	-	110	200	93	121	278
	HCISE01780400	*mec-12*	2,467	5,354	2,984	1,106	5,281
	HCISE02189600	*atgp-1*	189	544	158	35	465
	HCISE01808600	*asp-14*	0	4	14	5	178
	HCISE00920700	*alh-4*	5	37	114	72	1,207
	HCISE00643200	*elo-5*	4	765	1,011	328	7,089
	HCISE01831900	*tra-3*	73	87	67	72	179
	HCISE00613300	*prk-2*	668	523	304	54	11,845
	HCISE00722600	*chc-1*	1,458	1,633	1,471	1,595	1,962
	HCISE01450000	-	359	1,245	605	1,818	6,150
	HCISE01797200	*ell-1*	133	252	128	296	805
	HCISE00030200	-	0	51	34	2	106
	HCISE00345200	*let-23*	130	687	607	110	1,573
	HCISE01945800	-	4	405	153	21	938
	HCISE02181200	*Y61A9LA.11*	80	268	219	176	442
	HCISE01237100	*mpk-1*	81	89	52	112	172
	HCISE01237500	*mpk-1*	81	89	52	112	172
	HCISE01237900	*mpk-1*	81	89	52	112	172
	HCISE00974600	*T12G3.4*	38	24	4	131	232
	HCISE00161900	-	557	879	503	1,447	1,156
	HCISE01338200	*F52A8.6*	166	145	244	67	349
	HCISE00002300	-	110	55	79	81	135
	HCISE01069100	-	816	465	818	876	1,624
	HCISE01070700	-	818	468	814	880	1,628
	HCISE00738600	*acl-14*	710	283	426	529	1,556
	HCISE01357700	-	43	29	13	7	109
	HCISE01950600	-	337	175	93	111	303
	HCISE01746300	*vps-60*	1,791	784	706	948	1,387

Fig. 4-8: Hierarchical clustering of *H. contortus* genes predicted as targets for Hco-miR-228-5p. Clustering of genes, shown as a dendrogram (right hand side), was based on RNA-seq abundance data highlighted using a relative heatmap (blue = low, white = median, red = high). Putative *C. elegans* homologues are listed when identified. Two genes (marked in red) displayed inverse patterns to Hco-miR-228-5p.

	Gene	Putative homologue	Life cycle stage				
			L3	L4	Male	Female	Gut
A	HCISE00477700	*cal-4*	298	77	16	176	3
	HCISE00849700	-	288	169	182	227	97
	HCISE00112000	-	1,202	607	758	887	555
	HCISE00832700	-	285	18	14	61	0
	HCISE01323800	*T08B2.5*	1,892	495	827	483	617
	HCISE01526500	*lev-1*	263	67	21	7	40
	HCISE00018000	-	2,167	175	77	6	28
	HCISE00636700	-	129	5	2	0	0
	HCISE00912800	*che-2*	118	1	7	0	4
	HCISE01451400	-	21,903	11,226	4,967	3,016	4,410
	HCISE01806600	*glrx-5*	367	233	122	105	112
	HCISE00235400	*F35H12.1*	289	201	138	31	3
	HCISE00021100	*rpl-15*	4,291	4,513	1,569	2,211	1,627
B	HCISE00545400	*C38H2.2*	291	490	529	535	364
	HCISE00912400	-	599	1,096	1,251	1,922	582
	HCISE01732500	*F19B6.1*	406	527	1,087	1,485	741
	HCISE00064800	-	91	398	1,371	2,134	459
	HCISE00734800	*T25C12.3*	14	113	25	116	12
	HCISE00389000	*C45G3.3*	35	135	72	123	35
	HCISE02080400	-	345	398	346	534	89
	HCISE00147200	*hot-3*	177	207	79	261	14
	HCISE01672000	*elpc-3*	97	45	57	127	88
	HCISE00941300	*lmn-1*	342	117	359	1,091	409
	HCISE00516700	*ttc-7*	987	619	694	1,642	885
	HCISE01789600	*mtr-4*	434	326	532	1,261	384
	HCISE01538600	*rabn-5*	322	284	419	814	340
	HCISE01296600	*rnr-2*	250	169	604	1,616	131
	HCISE01057800	-	62	63	28	300	20
	HCISE01081100	-	119	168	131	608	145
	HCISE02075200	-	0	2	10	494	13
	HCISE00114100	-	0	1	31	773	49
	HCISE00472600	*C05D2.10*	185	87	172	271	21
	HCISE01132200	-	84	28	81	162	0
	HCISE01337700	*H35B03.2*	74	66	84	148	36
	HCISE01306900	*T07F8.4*	823	687	1,040	1,326	346
	HCISE01578100	*mrpl-12*	123	101	148	160	49
	HCISE01248700	*F44B9.8*	123	91	163	144	49
	HCISE01456100	*F44B9.8*	206	108	207	210	79
	HCISE00582700	*M05D6.2*	440	153	684	598	191
	HCISE02181400	-	1,437	518	1,619	1,535	1,307
C	HCISE00682000	*E01G4.3*	256	945	118	320	352
	HCISE00846700	-	129	233	91	160	172
	HCISE01772700	-	86	5,018	740	24	40
	HCISE00985200	*Y73F8A.26*	75	827	106	18	16
	HCISE01760800	-	0	251	13	10	0
	HCISE01844700	-	6	1,582	0	0	0
	HCISE00650000	*col-38*	111	26,218	6	4	0
	HCISE00649400	*col-38*	111	26,216	92	4	0
	HCISE01153000	*T01D1.3*	11	148	2	63	0
	HCISE01865000	*swsn-4*	1,726	2,499	1,067	1,703	1,504
	HCISE01002700	*Y43F4A.1*	597	1,249	308	427	297
	HCISE00795300	-	36	104	21	42	27
	HCISE01678400	*sma-1*	4,097	9,811	8,119	7,135	6,526
	HCISE01615400	*ttr-27*	15	402	357	14	2
	HCISE01534200	-	10	2,133	1,288	103	6
D	HCISE01001000	-	417	1,117	1,220	639	44
	HCISE00480600	*C38H2.2*	110	144	197	159	79
	HCISE01286200	*rpn-2*	1,120	1,319	2,025	1,512	617
	HCISE00243500	-	86	43	381	141	43
	HCISE00237100	*tbg-1*	126	147	339	133	111
	HCISE00221800	-	14	59	843	87	0
	HCISE00051800	-	37	62	524	0	0
	HCISE01212800	-	0	34	1,975	134	73
	HCISE00513500	-	4	22	911	30	45
	HCISE01200500	-	0	3	164	0	0





D	HCISE00927900	-	21	0	1,269	1	2
	HCISE01444500	-	0	0	186	0	2
	HCISE00072200	-	1	0	2,243	1	15
	HCISE00825700	-	1	4	519	0	4
	HCISE02051500	*fut-1*	1	0	609	0	0
	HCISE01461500	-	10	12	4,599	2	2
	HCISE00985300	-	0	1	1,389	2	0
	HCISE00238300	-	1	0	9,611	7	0
	HCISE02037900	-	0	0	219	0	0
	HCISE02184200	*psr-1*	0	0	830	0	0
	HCISE00510400	-	0	0	442	0	0
	HCISE01352100	-	3	0	8,568	7	12
	HCISE01844000	-	0	2	221	1	0
	HCISE02069100	*C10G11.8*	2	13	18,377	438	31
	HCISE00845500	-	323	667	1,010	378	257
	HCISE01727600	*T15B7.1*	0	66	240	18	2
	HCISE00416400	*unc-13*	1,878	762	2,277	572	84
	HCISE01666600	-	1,473	580	2,019	662	689
	HCISE01749200	-	2,146	785	3,285	1,073	919
	HCISE01678700	*Y57E12AL.1*	4,558	2,873	8,811	2,573	2,656
	HCISE02152000	*Y57E12AL.1*	4,278	2,688	8,571	2,337	2,222
	HCISE00091300	*cnx-1*	3,331	2,742	4,710	1,095	2,903
	HCISE00091700	*cnx-1*	2,134	1,875	3,085	741	1,801
	HCISE01623100	*W02D9.2*	112	73	694	320	478
	HCISE00324300	-	96	82	406	172	180
	HCISE01984900	*Y75B8A.16*	385	467	707	376	614
	HCISE01464300	*ama-1*	1,506	1,631	1,759	1,516	1,621
	HCISE01619400	*lnkn-1*	811	956	1,239	829	1,001
E	HCISE01780400	*mec-12*	2,467	5,354	2,984	1,106	5,281
	HCISE02189600	*atgp-1*	189	544	158	35	465
	HCISE01806500	*mrpl-36*	37	111	42	37	118
	HCISE01840100	-	318	76,185	7,170	1,237	68,266
	HCISE01338200	*F52A8.6*	166	145	244	67	349
	HCISE01808600	*asp-14*	0	4	14	5	178
	HCISE00920700	*alh-4*	5	37	114	72	1,207
	HCISE00643200	*elo-5*	4	765	1,011	328	7,089
	HCISE01871900	-	2	850	339	16	3,025
	HCISE01831900	-	73	87	67	72	179
	HCISE01260300	-	262	180	156	112	687
	HCISE01034300	*aagr-2*	2,575	2,860	1,312	470	6,750
	HCISE00621400	*srp-2*	657	416	203	61	1,413
	HCISE01357700	-	43	29	13	7	109
	HCISE01237100	*mpk-1*	81	89	52	112	172
	HCISE01237500	*mpk-1*	81	89	52	112	172
	HCISE01237900	*mpk-1*	81	89	52	112	172
	HCISE00722600	*chc-1*	1,458	1,633	1,471	1,595	1,962
	HCISE01450000	-	359	1,245	605	1,818	6,150
	HCISE01797200	*ell-1*	133	252	128	296	805
	HCISE02118200	-	110	200	93	121	278
	HCISE00030200	-	0	51	34	2	106
	HCISE00345200	*let-23*	130	687	607	110	1,573
	HCISE02181200	*Y61A9LA.11*	80	268	219	176	442
	HCISE02179400	*wdr-48*	197	387	351	399	618
	HCISE01872000	*klp-20*	707	689	436	603	803
	HCISE02017800	*csn-5*	454	356	210	415	463
	HCISE01950600	-	337	175	93	111	303
	HCISE01746300	*vps-60*	1,791	784	706	948	1,387
	HCISE00974800	*ccr-4*	788	651	791	1,076	1,672
	HCISE00905300	-	164	142	190	202	230
	HCISE00002300	-	110	55	79	81	135
	HCISE00738600	*acl-14*	710	283	426	529	1,556

Fig. 4-9: Hierarchical clustering of *H. contortus* genes predicted as targets for Hco-miR-235-3p. Clustering of genes, shown as a dendrogram (right hand side), was based on RNA-seq abundance data highlighted using a relative heatmap (blue = low, white = median, red = high). Four genes (marked in red) displayed inverse expression to miR-235.

	Gene	Putative homologue	Life cycle stage				
			L3	L4	Male	Female	Gut
A	HCISE01933000	-	835	821	865	621	570
	HCISE00657800	-	0	218	0	0	0
	HCISE01738500	-	70	136	41	8	2
	HCISE00711500	*F40A3.6*	263	728	209	10	7
		-					
B	HCISE01915400	-	133	11	4	13	0
	HCISE01811700	-	955	128	100	50	52
	HCISE01357500	*D1014.2*	122	0	4	1	0
	HCISE01881600	-	639	26	19	2	2
	HCISE00488700	-	553	10	10	2	0
	HCISE01595700	-	460	17	183	0	2
	HCISE00080900	-	248	108	13	2	0
	HCISE01493800	-	257	185	64	30	19
	HCISE00477700	*cal-4*	298	77	16	176	3
C	HCISE01067100	-	327	1,046	542	561	1,170
	HCISE02056500	*bris-1*	322	419	568	495	3,961
	HCISE01641600	*psd-1*	16	71	90	25	595
	HCISE00923100	-	35	25	34	11	121
	HCISE00536000	*bris-1*	243	342	623	516	1,496
	HCISE01871400	*F55B11.1*	56	155	88	11	289
	HCISE00404500	*T02B11.8*	6	147	78	16	340
	HCISE00432500	-	329	528	362	178	1,469
	HCISE01186700	-	6	115	76	55	379
	HCISE00621400	*srp-2*	657	416	203	61	1,413
	HCISE00850100	-	440	186	250	527	804
	HCISE00562600	-	154	64	69	210	371
	HCISE01465800	*trpp-4*	160	76	85	132	163
	HCISE01080300	*mek-1*	251	115	60	123	264
D	HCISE02140000	*mmcm-1*	3	33	283	10	35
	HCISE01522600	-	1	0	23,769	10	24
	HCISE00861700	-	8	6	7,934	4	5
	HCISE01640100	-	1	1	474	60	6
	HCISE00757200	-	3	4	191	17	0
	HCISE00750800	-	31	8	669	58	2
	HCISE00616100	-	0	6	276	66	2
	HCISE01386000	*emc-6*	40	33	120	69	38
	HCISE01211400	*dnj-2*	385	1,628	1,074	764	625
	HCISE01462500	*F56A11.5*	224	597	563	248	424
	HCISE01008100	*M05D6.6*	140	1,134	989	98	448
	HCISE00743700	-	4	801	901	70	0
	HCISE00065000	*R12E2.8*	104	179	197	170	160
E	HCISE00228900	*lig-1*	775	386	572	1,684	245
	HCISE00290500	-	223	262	224	1,111	218
	HCISE00928500	-	2	0	7	165	2
	HCISE00669500	*rnr-1*	175	88	208	1,678	134
	HCISE01312200	*F12F6.1*	172	78	64	581	71
	HCISE00409500	-	115	46	84	377	67
	HCISE01437900	*rnr-1*	599	326	417	2,060	208
	HCISE00425700	*nars-2*	251	240	192	294	120
	HCISE00418000	*F13G3.6*	187	246	173	333	109
	HCISE02141100	rpt-6	666	635	669	694	303
	HCISE01701900	*rpn-2*	618	661	824	695	221
	HCISE00173900	*pif-1*	167	105	170	216	66
	HCISE00098000	*inx-10*	662	314	557	544	132
	HCISE00909700	*F42G9.1*	606	698	1,761	2,008	985
	HCISE00683100	-	210	242	705	1,342	263
	HCISE01731700	*pri-2*	42	85	234	569	79

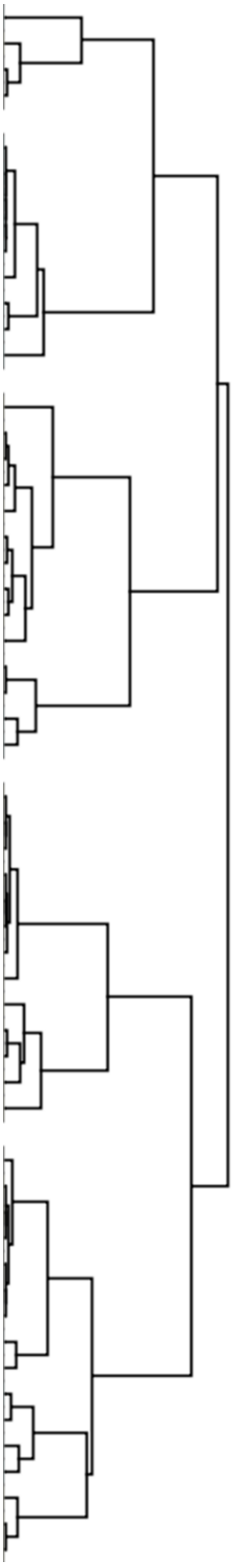


Fig. 4-10: Hierarchical clustering of *H. contortus* genes predicted as targets for Hco-miR-5885-3p. Clustering of genes, shown as a dendrogram (right hand side), was based on RNA-seq abundance data highlighted using a relative heatmap (blue = low, white = median, red = high). Putative *C. elegans* homologues are listed when identified. Two genes (marked in red) displayed inverse expression profiles to miR-5885a/b/c-3p.

	Gene	Putative homologue	Life cycle stage				
			L3	L4	Male	Female	Gut
A	HCISE02052800	HEX-5	56	271	140	21	213
	HCISE00803200	DNJ-24	220	939	1,037	58	1,230
	HCISE01589300	DNJ-24	569	2,687	3,122	128	2,207
	HCISE01960100	F48C1.6	257	343	363	186	328
	HCISE01323100	RSBP-1	470	481	708	148	494
	HCISE01154500	FAT-2	2,610	2,131	2,688	500	2,736
	HCISE00972000	MNK-1	9,838	13,639	4,601	3,772	14,130
	HCISE01034300	AAGR-2	2,575	2,860	1,312	470	6,750
	HCISE00643700	-	182	58,670	19,135	12,877	94,059
	HCISE00361800	-	38	72	47	105	184
	HCISE01260200	K01A11.2	17	36	86	43	167
	HCISE01860000	-	517	341	735	504	1,342
	HCISE02037300	-	0	102	104	30	738
	HCISE00088300	-	148	774	1,217	471	5,290
	HCISE00470600	SPSB-1	549	690	668	959	2,691
B	HCISE01187100	-	161	161	67	186	91
	HCISE00604400	SEM-2	103	64	3	59	40
	HCISE01486300	SRGP-1	654	940	626	830	412
	HCISE00205000	NEP-2	253	695	199	43	2
	HCISE00458300	-	4,501	15,169	3,930	1,897	4,433
	HCISE00873900	Y37D8A.3	57	1,651	161	6	30
	HCISE01604100	-	45	741	39	10	5
	HCISE00195200	-	4	853	2	1	1
	HCISE00193100	F10D11.6	20	1,021	16	82	2
	HCISE02143300	-	0	303	60	5	0
C	HCISE01950600	-	337	175	93	111	303
	HCISE00742100	Y111B2A.12	1,627	587	419	351	1,295
	HCISE00706500	UNC-103	329	222	118	77	28
	HCISE01936600	-	1,493	1,328	620	31	38
	HCISE00377600	F02D8.4	616	451	301	34	9
	HCISE01288600	F45E4.5	172	62	73	15	2
	HCISE00852100	PDFR-1	313	87	149	51	6
	HCISE00585700	DAF-16	3,607	498	526	295	1,191
	HCISE01288700	-	255	32	25	14	1
	HCISE00482000	F23H12.7	115	4	3	0	0
	HCISE00439800	CCB-1	2,610	379	197	34	98
	HCISE01380500	GPA-3	240	25	3	0	13
	HCISE01740300	-	107	51	29	45	43
	HCISE02165800	TUT-1	286	55	115	143	89
	HCISE00387500	-	614	14	119	81	3
D	HCISE00156900	Y67D2.6	429	233	299	793	364
	HCISE01425600	SET-16	1,084	513	223	2,579	529
	HCISE00095400	SET-16	1,038	530	323	2,620	549
	HCISE00794100	Y67D2.6	435	172	214	580	237
	HCISE01416300	-	498	420	403	662	327
	HCISE01234300	CLN-3.2	240	247	352	612	262
	HCISE01954800	ELKS-1	643	568	417	2,289	576
	HCISE01730700	BED-2	58	193	100	896	71
	HCISE01467400	CIT-1.1	628	219	545	636	500
	HCISE00912400	-	599	1,096	1,251	1,922	582
	HCISE02130700	SNR-6	342	415	413	503	391
	HCISE00246200	K01A11.2	24	95	138	146	89
	HCISE01923900	-	25	66	877	897	24
	HCISE00704100	LIP-1	172	67	373	446	128
E	HCISE02042900	SNF-9	18	30	215	20	16
	HCISE01327000	-	6	10	302	0	1
	HCISE01800400	-	912	688	2,894	1,226	492
	HCISE02038100	-	913	729	2,917	1,650	534
	HCISE00657900	ZK757.2	550	103	618	33	4

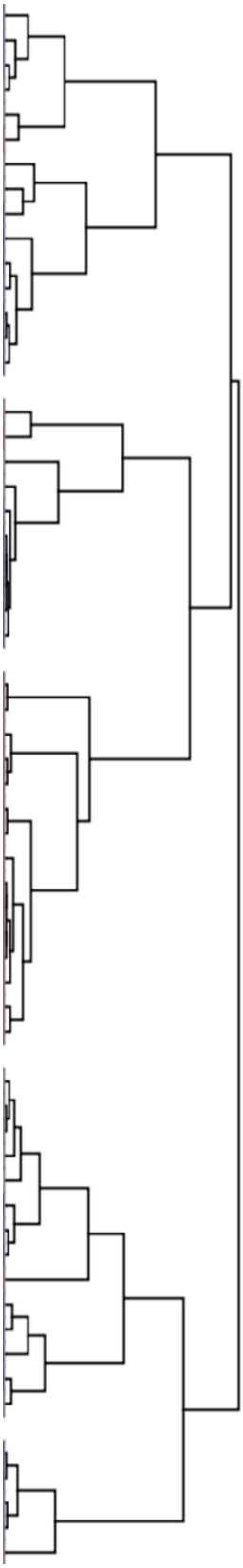


Fig. 4-11: Hierarchical clustering of *H. contortus* genes predicted as targets for Hco-miR-5885-5p. Clustering of genes, shown as a dendrogram (right hand side), was based on RNA-seq abundance data highlighted using a relative heatmap (blue = low, white = median, red = high). Two genes (marked in red) displayed an inverse expression profile to miR-5885a/b/c-5p.

4.3.8 Gene ontology of predicted miRNA targets

To understand the function of predicted *H. contortus* miRNA targets, gene ontology (GO) annotation terms were analysed for enrichment. GO terms were collected for putative *C. elegans* homologues of *H. contortus* genes predicted as targets for five miRNAs (Hco-miR-60, Hco-miR-228, Hco-miR-235, Hco-miR-5885a/b/c-3p and Hco-miR-5885a/b/c-5p). For comparison, the same was done for five miRNAs in *C. elegans* (Cel-let-7, Cel-lin-4, Cel-miR-60, Cel-miR-228 and Cel-miR-235). In total, 160 genes were submitted for *H. contortus* and 177 for *C. elegans*. Three categories of GO annotations were studied: biological processes (BP), molecular functions (MF) and cellular components (CC). In each category, one annotation ("Biological_process", "Molecular_function" and "Cellular_component" respectively) represented the 'root' ontology term. In many cases a gene was only annotated with this term, indicating that no detailed information was available. Fold changes in enrichment between sample and expected values were reported when available. Significantly enriched results (p-value < 0.01) were obtained for predicted targets of the five *C. elegans* miRNAs (Appendix, Tables 8-5 – 8-9; no enrichment of Cel-miR-235 targets) and three *H. contortus* miRNAs (Appendix, Tables 8-9 – 8-13).

The Cel-let-7 dataset, consisting of 46 genes, generated the largest number of enriched terms (n = 63), the majority of which were highly enriched (>5 fold) and highly specific. Many terms referred to nucleotide regulation, nucleotide metabolism, general metabolism and responses to different compounds. No CC terms were enriched. The Cel-lin-4 dataset contained the fewest genes (n = 16), only generating five enriched terms in the BP category; no MF or CC terms were enriched. Aside from the root ("biological_process") and generic ("regulation of biological processes") terms, two signalling terms and one cell communication term were enriched, each with a fold enrichment > 5. In the Cel-miR-60 dataset, with 64 genes, 31 annotations were enriched across all three GO categories. Transport and localisation were the most common BP terms, while the majority of MF

terms involved ribonucleic binding. In the CC category, specific annotations included (apical and plasma) "membrane" and "cell periphery". The Cel-miR-228 dataset, with 49 genes, only generated three annotations, all of which were generic ("cell_component", "cell part" and "cell"), while in the Cel-miR-235 dataset (54 genes), no annotations were enriched.

For the three conserved *H. contortus* miRNAs (miR-60, miR-228 and miR-235), only BP and MF terms were found to be enriched (Appendix, Tables, 8-9 – 8-11, summarised in Table 4-5). The miR-60 dataset (n = 61) generated the largest number of enriched annotations (n = 24). The majority of terms fell into the BP category, with development and metabolism being most common, with an average fold enrichment of 2.5 and 2.4. In the MF category, binding (including "ion binding") was the most common term (average fold enrichment = 2.2). The miR-228 dataset contained 62 genes and generated 18 enriched terms. Again, development and metabolism were the most common terms. The smallest number of enriched annotations was generated by the miR-235 dataset which contained 52 genes. Apart from the root term "Biological_process", only the "anatomical structural development" term was enriched, by 2.5 fold. The two unique miRNAs produced fewer enriched annotation terms (Appendix, Tables 8-12 – 8-13). For miR-5885a/b/c-3p, the "Oxidoreductase activity, acting on CH or CH_2 groups" annotation was enriched. Two of three genes in the entire *C. elegans* genome with this annotation were predicted as targets. The enrichment for targets of miR-5885a/b/c-5p was much broader, covering regulation of development, biological processes, multicellular organisms and protein modification.

Table 4-5: Summary of enriched annotation terms for putative *C. elegans* homologues of *H. contortus* genes predicted as miRNA targets.

	GO Term	miRNA				
		Hco-miR-60	Hco-miR-228	Hco-miR-235	Hco-miR-5885-3p	Hco-miR-5885-5p
BP	Cellular process	●	●			
	Single-organism process	●	●			
	Single-organism developmental process	●	●			
	Single-multicellular organism process	●	●			
	Multicellular organismal process	●	●			
	Multicellular organismal development	●	●			
	Developmental process	●	●			
	Anatomical structure development	●	●	●		
	Embryo development	●				
	Post-embryonic development	●				
	Larval development	●				
	Nematode larval development	●				
	Metabolic process	●	●			
	Primary metabolic process	●	●			
	Macromolecule metabolic process	●	●			
	Organic substance metabolic process	●	●			
	Protein metabolic process	●	●			
	Biological regulation					●
	Regulation of biological process					●
	Regulation of multicellular organismal process					●
	Regulation of developmental process					●
	Regulation of protein modification process					●
MF	Structural molecule activity	●				
	Binding	●				
	Ion binding	●	●			
	Catalytic activity	●	●			
	Oxidoreductase activity, acting on CH or CH2 groups				●	
CC	Cell part					●
	Cell					●

4.4 Discussion

Target prediction is an important step in understanding the biological role of miRNAs. Even with good quality, high resolution abundance data, interpretation of their function is limited. Numerous computational approaches have been developed over the past decade, each utilising a different approach as mentioned previously. No single algorithm has emerged as the method of choice for miRNA target prediction, reaffirming our lack of complete understanding with regards to miRNA-mRNA interactions. It therefore makes

sense to compile the results of multiple algorithms together in an attempt to strengthen their overall predictive power. If a target is predicted by multiple algorithms each using different selection criteria, then it is theoretically more likely to be a true target of miRNA regulation (Coronnello *et al.*, 2012) although this presumption is not universally accepted (Alexiou *et al.*, 2009). The results presented in this chapter combine predictions from three algorithms (miRanda, PITA and RNAhybrid) interrogating a custom database of 3′ UTR sequences. The focus was on miRNAs that showed variant expression and were therefore likely to be involved in developmental regulation.

The ability to perform target prediction in less studied species relies upon custom databases. Although some binding sites have been identified in coding sequence or 5′ UTR, the majority of miRNAs are believed to target the 3′ UTR of the mRNA (Fang and Rajewsky, 2011). The completion of the first draft of the *H. contortus* genome (Laing *et al., 2013*) allowed development of a custom 3′ UTR database (A. Martinelli of the Wellcome Trust Sanger Institute). Approximately 50 % of genes had an associated 3′ UTR sequence, with a significant proportion being either too long (> 1,000 nt) or too short (< 50 nt). In total, only 33 % of annotated genes contained a UTR of acceptable length. In response to this low coverage, a second database was generated using 1,000 nucleotides downstream of the stop codon (DS database). In combination, coverage of genes markedly improved to 95 % of annotated genes.

Interrogation of the combined UTR/DS database was performed by H. Gu by running downloaded versions of the three algorithms for each miRNA and filtered using higher stringency settings than default. In general, these settings were chosen arbitrarily based upon previous work performed in *C. elegans* and the number of targets generated in small scale trials.

For the 53 miRNAs examined, over four hundred thousand (440718) binding sites were predicted across the majority of 3′ UTRs (n= 20,323, 93 %) currently annotated in *H. contortus*. There was significant variation in the number of targets predicted by each algorithm, with almost half (47 %) being identified by RNAhybrid (26 and 27 % for miRanda and PITA respectively). This pattern was not maintained at the individual miRNA level. For example, 80 % of targets for miR-235 were predicted by miRanda. Significant variation was also seen in the total number of targets predicted for each miRNA individually: the difference between the miRNA with the most targets (miR-5885b-5p) and the fewest (miR-45-3p) was 14,642. Strikingly, five miRNAs were predicted to target more than 10,000 genes, almost half of the known genes in *H. contortus*. This disparity was caused by a greater than average number of predictions by RNAhybrid, most likely due to a lower relative stringency used for filtering RNAhybrid predictions compared to miRanda or PITA. The variation was not seen during small scale testing, becoming apparent only after analysis of all 59 miRNAs, highlighting the difficulty of identifying comparable stringency filters between algorithms utilising different prediction criteria.

Previous studies have shown that in eukaryotes, 30 – 60 % of genes within an organism are targeted by miRNAs (Friedman *et al.*, 2009; Yu *et al.*, 2007). The genes of *H. contortus* may indeed be under an unusual degree of miRNA regulation, although the more plausible explanation is that the algorithms were ‘optimistic’ in their predictions and are likely to include large numbers of false positive results (Yue *et al.*, 2009). This complicates interpretation, requiring the data to be filtered using higher stringency settings. However, care must also be taken not to exclude biologically relevant interactions. Achieving a balance between false-positives and false-negatives is difficult when we cannot identify which genes are truly regulated by miRNAs.

With this in mind, some additional filtering was performed on a small scale, specifically for five miRNAs with interesting expression profiles (see Chapter 3). Expression of miR-

228 peaked during L3 before decreasing in L4 and adult stages while miR-235 expression also decreased during L4, but increased again during adult stages. In contrast, miR-60 was abundant in L4 and adult gut tissue. Targets of miR-5885 were studied as it is the most abundant miRNA in *H. contortus*. It also showed a unique expression profile where the more abundant -3p arm peaked in L4 and gut tissue, while the less abundant -5p only peaked in males.

A small number of genes were predicted as common targets of miR-60, miR-228 and miR-235. Interestingly, the proportion of targets shared between any two of these miRNAs was equivalent, even though miR-228 and miR-235 display similar expression profiles. This suggests that patterns in abundance do not correlate to targets and that miRNAs expressed at the same time do not necessarily regulate the same genes. For miR-5885, a much larger proportion of genes were common between its two arms, suggesting a shared regulatory role. Overall, the data suggests genes can be regulated by multiple miRNAs. Complicated regulatory networks have been identified in other organisms, such as *C. elegans* (see Chapter 1.3). This provides robustness, flexibility and fine control to the regulatory network. For example, changes in expression of a miRNA through random mutations might unduly affect the expression (and hence function) of a gene that is regulated by a single miRNA. It is therefore unsurprising that genes are targeted by multiple miRNAs.

As miRNAs usually decrease mRNA abundance, their targets should display inverse expression profiles. RNAseq data was acquired for a number of high-stringency predictions for the five *H. contortus* miRNAs studied in detail. Expression of individual genes varied enormously, with most below 1,000 (normalised signal). Genes tended to display peaks at one particular stage, rather than continuous expression. As the miRNAs studied also displayed peaks, less than a dozen genes were identified that matched or

closely followed the expected inverse expression profile of their corresponding miRNA. This is unlikely to cover the full number of genes targeted by these miRNAs.

There are three methods that may improve usage of transcriptomic data to support miRNA target prediction. The process could be simplified by studying specific transitions or periods, such as L3/L4 or male/female, rather than including all stages. Initial examination of the data in this way identified many more potential targets. There are also a number of advanced statistical models developed specifically to identify potential miRNA-mRNA interactions that might prove useful in future studies (Calura *et al.*, 2014; Guo *et al.*, 2014; Naifang *et al.*, 2013). Another approach would be to reverse the order of operations, by using target prediction algorithms to support transcriptomic data. Rather than look at transcriptomic data from a subset of genes (predicted by algorithms), hierarchical clustering could be performed on all genes in *H. contortus*. Then clusters of genes that displayed inverse expression profiles would be identified before determining if any of the genes in these clusters were predicted as targets. In theory, these clusters should contain a significantly higher number of predicted targets than groups that do not mirror the expression profile of a miRNA.

Correlations between transcriptomic analysis and miRNA target prediction have been used effectively in a number of situations (Li *et al.*, 2009a; Ruike *et al.*, 2008). However, miRNAs often have a subtle effect (Wang *et al.*, 2011b), with potentially only 30 % of the expression of a gene being determined by miRNAs (Selbach *et al.*, 2008). miRNAs modulate transcript abundance rather than acting as an on-off switch. Such subtle changes may be missed in large scale RNAseq data. More detailed analysis of expression of specific miRNA targets using qRT-PCR will help confirm inverse patterns identified here and this has been carried out on a small scale for miR-228 targets (see Chapter 6). While transcriptomic analysis is somewhat simplistic and cannot confirm predictions it does provide a valuable piece of supporting evidence in the identification of miRNA targets.

A comparative genomics approach, searching for conservation of binding sites in homologues in closely related species, can also add weight to target prediction, an approach used recently by Winter *et al.* (2015). This made used of annotated genome data from several filarial species to identify targets for a Clade III specific miRNA, miR-5364. As more annotated data becomes available for parasitic nematodes related to *H. contortus* (*e.g. Teladorsagia circumcincta*, *Necator americanus*), comparative approaches will be applied to provide further evidence for specific miRNA-mRNA interactions.

Although genome sequence is available, large scale annotation of *H. contortus* genes is still in its infancy. In contrast, *C. elegans* is a well-studied, model organism, providing decades of collective information on gene function. Putative *C. elegans* homologues of *H. contortus* miRNA targets were therefore identified to gain an idea of function. For the three conserved miRNAs (miR-60, miR-228 and miR-235), *C. elegans* homologues were identified for over half of selected *H. contortus* targets. As a unique miRNA, it was assumed that fewer homologues would be identified for targets of miR-5885 as it was more likely to regulate *H. contortus*-specific genes. Surprisingly, this was incorrect, as similar proportions were identified and may reflect its similarity to the conserved bantam miRNA family (see Chapter 3).

To understand if any patterns existed in the genes targeted by miRNAs, the putative homologues identified for miR-60, miR-228, miR-235 and miR-5885 were subjected to gene ontology analysis. Performed separately for targets of each miRNA, annotations of putative homologues were compared against a background of all annotated genes in *C. elegans*. When an annotation was found to be more common than would be expected by random chance ($p < 0.01$), it was considered enriched. This allowed for patterns in

functions, processes or location to be identified. For all miRNAs studied, annotations were more common than expected (negative enrichment for unclassified genes). These genes have a greater number of annotations, which may reflect conserved functions that have been identified across species. If so, this supports the idea that miRNAs target developmentally important genes, which tend to be conserved between organisms.

The greatest number of enriched terms was seen for targets of miR-60 and miR-228. All of the enriched annotations for targets of miR-228 were also enriched for targets of miR-60, suggesting a common regulatory function, particularly in regards to metabolism. In addition, miR-60 targets were also enriched for developmental annotations including post-embryonic and larval development, which further strengthens the argument that miRNAs are involved in regulating this fundamental process. The single term enriched in miR-235 targets (anatomical structure development) was also enriched in targets of miR-60 and miR-228 indicating a common role. For miR-5885-3p, the enriched term was highly specific (oxidoreductase activity, acting on CH or CH_2 groups) which arose from two genes (out of three in the genome) with this annotation. Targets of miR-5885-5p were enriched for regulatory terms (*e.g.* regulation of development, regulation of protein modification), suggesting that targets higher-level genes and not genes at the end of a pathway. The results presented here highlight the complexity of miRNA regulation: miRNAs can target a wide or narrow range of systems (*e.g.* miR-60 versus miR-235); miRNAs with different expression profiles can target similar systems (*e.g.* miR-60 and miR-228) and miRNAs expressed from the same stem-loop (*i.e.* miR-5885-3p and miR-5885-5p) do not necessarily regulate the same systems.

In addition to *H. contortus* miRNA target prediction, a computational approach was also used in *C. elegans* for miRNAs common to both species. As a model organism, many

more algorithms were available for use and gene ontology was straightforward. However, while many miRNAs are conserved between the two species, expression profiles were not (shown in Chapter 3), suggesting that functions are not conserved either. Comparison of predicted targets of miR-60, miR-228 and miR-235 supports this idea, as only one gene (*let-23*) was found to be predicted as a target of the same miRNA (miR-60) in both species. *Cel-let-23* encodes a EGF tyrosine kinase receptor involved in development of the posterior epidermis, vulva and male spicules while linked to quiescent behaviour during lethargus (Moghal and Sternberg, 2003; Skorobogata *et al.*, 2014; Yoon *et al.*, 2000). The low number of common targets may reflect the low proportion of predicted targets for which putative homologues were identified, or that targets are not conserved between these two species. The former issue could be addressed through further identification of putative homologues, while the latter situation would only become clear after large scale *in vitro* validation of targets was performed in *H. contortus*. Interestingly, a prominent gene ontology category for predicted targets of Cel-miR-60 was transport and localisation (membrane and cell periphery). miR-60 is abundant in the gut (Wormbase) and this suggests a possible role in regulating genes associated with transport or secretion within the gut. Initial analysis of *C. elegans mir-60*(-) mutants shows no obvious phenotype (see Chapter 6) but further work will aim to examine any changes in localisation of gut proteins using available antibodies. This is also valid for *H. contortus* miR-60 analysis, particularly with the importance of gut proteins as vaccine antigens (Chapter 1.1.6).

The targets of one particular *C. elegans* miRNA (miR-228) were studied in greater detail by quantifying gene expression in different mutant strains.

Genes were selected based on the strength of *in silico* target predictions using multiple algorithms. Of the three genes studied in *C. elegans*, *T04A8.7*, was predicted as a target or miR-228 by four of six algorithms and encodes a lysosomal enzyme involved in glucose metabolism. The remaining genes (*hbl-1* and *hpd-1*) were predicted by three algorithms

and identified as important for development: *hbl-1* determines normal lifespan and negatively regulates dauer formation (Lin *et al.*, 2003b), while *hpd-1* regulates post-embryonic developmental timing (Lee, 2003). HITS-CLIP performed by Zisoulis *et al.* (2010) using *alg-1*(-) mutants found that all three genes were regulated by miRNAs. In addition, *hbl-1* is a confirmed target of let-7 (Roush and Slack, 2009). The qRT-PCR data presented here confirmed miRNA regulation of *T04A8.7* and *hpd-1*. More specifically, T04A8.7 expression was affected by a loss of miR-228, strongly suggesting regulation. Surprisingly, the qRT-PCR data suggested that *hbl-1* was not regulated by miRNAs. As both life cycle stage (L4) and mutant strain (RF54) were the same, experimental differences between qRT-PCR and HITS-CLIP may explain this disparity. The expression of T04A8.7 was also examined in *H. contortus* using variation between stages, rather than between mutant strains. However, even though the abundance profile of Hco-T04A8.7 displayed an inverse pattern to abundance of Hco-miR-228, no miRNA binding sites were identified in the 3′ UTR of Hco-T04A8.7. It may be that miR-228 regulates Hco-T04A8.7 through binding sites in other mRNA regions (*e.g.* 5′ UTR). Or, miR-228 may not regulate T04A8.7 in *H. contortus* and the inverse patterns of expression seen here were coincidental. This would indicate divergence in the regulation of miR-228 between *H. contortus* and *C. elegans*.

The results presented in this chapter represent the first computational study of miRNA targets in *H. contortus*. While use of different algorithms, stringency settings and transcriptomic data can help narrow down the list of potential target genes, these remain unconfirmed predictions. Experimental validation of miRNA-mRNA interactions is a logical next step in confirming true targets and will be discussed in the next chapter.

5 Generating an anti-Argonaute antibody for target validation

5.1 Introduction

Validating targets is one of the most challenging aspects of miRNA research. As described in Chapter 4, computational approaches have predicted numerous miRNA-mRNA interactions in the past decade. However, without validation, these interactions remain speculative (reviewed in Karbiener *et al.*, 2014). The most common methods for target validation involve gain/loss-of-function experiments. Cell lines can be transfected with miRNA mimics (Lim *et al.*, 2005), inhibitors (Nicolas *et al.*, 2008) or sponges (Kluiver *et al.*, 2012), artificially increasing, decreasing or silencing miRNA activity. If a particular mRNA is under miRNA regulation, abundance (quantified using Northern blots, qRT-PCR or high-throughput sequencing) should change following transfection with an inhibitor or miRNA sponge. miRNA loss-of-function experiments are often performed in conjunction with a reporter system where a 3′ UTR containing a miRNA binding site is cloned into a vector downstream of a reporter gene, such as luciferase (Nicolas, 2011) or fluorescent protein (Chen *et al.*, 2014c). The vector is then co-transfected into cells with a miRNA mimic or inhibitor and abundance of the reporter gene quantified. To confirm specific interactions, the UTR sequence can be mutated at the predicted miRNA binding site. This is currently considered the 'gold standard' approach to confirming individual miRNA-mRNA interactions.

However, these techniques are inefficient when studying multiple targets and are also restricted to cells or organisms amenable to transfection, limiting their direct use in *H. contortus* or other parasitic helminths. One method that can be used in live animals is

HITS-CLIP (high-throughput sequencing with cross-linked immunoprecipitation). Originally developed to probe RNA interactions with NOVA proteins in the brain of mice (Ule *et al.*, 2003), it has been successfully used to identify miRNA-mRNA interactions in mammalian cells (reviewed in Clark *et al.*, 2014) and *C. elegans* (Stefani *et al.*, 2015; Than *et al.*, 2013a; Zisoulis *et al.*, 2010b, 2011). This method involves cross-linking RNA and protein using UV irradiation. An antibody raised against a component of the RISC can then be used to pull down and isolate the complex before removing the proteins and sequencing the remaining RNA. The resulting RNA library contains all miRNAs and mRNA fragments that were being processed in the live animal at the time of irradiation. An advanced version, PAR-CLIP (photactivatable ribonucleoside enhanced CLIP), uses photoactivatable nucleotides that are incorporated into RNA during transcription (Hafner *et al.*, 2010), but is limited to cell cultures.

When studying miRNA interactions, the antibody should be specific to the miRNA-mediated RISC (miRISC) and exclude the siRNA-mediated RISC (Tang, 2005). As the major component of the RISC, Argonaute is often the target of choice, but other proteins such as AIN (Argonaute-interacting protein) have been used as well (Zhang *et al.*, 2007b). While antibodies to *H. sapiens* and *C. elegans* Argonaute proteins are available, the aim of the work presented here was to generate an antibody to *H. contortus* Argonaute for use in HITS-CLIP.

This chapter describes the identification of an Argonaute gene in *H. contortus* using bioinformatic and experimental techniques. Two antibodies, specific to *H. contortus* Argonaute, were generated by immunising rabbits with either a short synthetic peptide or larger recombinant protein, expressed in *E. coli*. Both peptide and polyclonal antibodies were tested with various extracts of *H. contortus*.

5.2 Materials and methods

5.2.1 Identification of an Argonaute gene in *H. contortus*

An Argonaute-like gene (*alg*) was identified in *H. contortus* by searching the draft genome with *C. elegans* Argonaute proteins ALG-1 and ALG-2. Draft versions of the *H. contortus* genome were downloaded from the Wellcome Trust Sanger website (http://www.sanger.ac.uk/resources/downloads/helminths/haemonchus-contortus.html), while the *C. elegans* sequences were acquired from WormBase. *H. contortus* nucleotide scaffolds were interrogated locally using tBLASTn and aligned in Geneious (version 5, Biomatters limited).

Confirmation of the 5′ region was performed using 5′ RACE (Rapid amplification of cDNA ends; Frohman, 1993). In brief, RNA extracted from a mix of male and female *H. contortus* adult worms was converted to cDNA by Dr. B. Roberts using a reverse transcriptase and oligo-dT adapter kit (Life Technologies). An argonaute-specific reverse primer (Appendix, Table 8-2) was used in conjunction with a universal forward primer that annealed to the poly(A) tail of the mRNA. Amplification was performed using a proof-reading polymerase (Phusion®, NEB) and the PCR product cloned and sequenced as described previously (see Chapters 2.4.3 and 2.4.7). The nucleotide and protein sequences were re-confirmed following the release of the completed *H. contortus* genome (Laing *et al.*, 2013). Nucleotide polymorphisms were detected throughout the gene, but all changes were synonymous (*i.e.* no change to the protein sequence).

Homologues of Hco-ALG-1 were identified in a range of species, principally parasitic nematodes, using BLASTp searches of the NCBI (nr) database and aligned in Geneious using ClustalW2. Putative domains were identified using the Pfam (Finn *et al.*, 2014) online tool (http://pfam.xfam.org/) while a phylogenetic tree was generated in MEGA (version 6.06, MEGA software) using the Neighbour-joining method, poisson model and

bootstrap clustering (1,000 replicates). Nucleotide sequences were obtained for nematode-specific Arognaute genes and used to calculate dN/dS ratios in MEGA with the Nei-Gojobori model and bootstrap procedure (500 replicates).

5.2.2 Cloning of *H. contortus* Argonaute fragments

Three different sized cDNA fragments of an *H. contortus alg* gene were amplified and cloned into vectors with fusion tags and transformed into *E. coli* strains optimised for protein expression. Fragments were amplified using a proof-reading polymerase from female *H. contortus* cDNA with primers that inserted restriction enzyme sites at the 5′ and 3′ ends of the amplicon (Appendix, Table 8-2). The short fragment (462 bp) was cloned using EcoRI (5′) and XhoI (3′) sites while the medium (1,479 bp) and long fragments (1,827 bp) were cloned with NcoI (5′) and BglII (3′) sites. Amplified fragments were purified using agarose gel extraction (Chapter 2.4.7) and cut using a double restriction-enzyme digest (Chapter 2.4.6). Two different enzymes were used (3′ and 5′) to clone the inserts in the correct direction for expression. For the short fragment, the pGEX-5X-1 vector (GE Healthcare) was used which added an N-terminal glutathione S-transferase (GST) tag, while the medium and long fragments were inserted into the pQE-60 vector (Qiagen), adding a C-terminal hexahistidine (His) tag. Vectors were linearized using the same restriction enzymes used to cut the fragments and desphosphorylated using CIP (Calf intestinal alkaline phosphatase, NEB) following the manufacturer's instructions. Vector and insert were ligated using T4 DNA ligase (NEB) following the manufacturer's instructions before being transformed into chemically competent strains of *E. coli* using heat-shock following standard protocols: pQE-60 with M15 [pREP4] cells (Qiagen), pGEX-5X-1 with BL21 cells (Stratagene). Successfully transformed cells were selected using antibiotics (ampicillin for pGEX, amplicillin + kanamycin for pqE-60) and the

presence of the insert confirmed using PCR amplification using primers covering both insert and vector. Further confirmation was obtained by submitting purified plasmid (see Chapter 2.4.7) for sequencing (MWG Eurofins).

5.2.3 Induction of recombinant protein expression in *E. coli*

The glycerol stock of a particular recombinant *E. coli* strain was used to inoculate 10 mL LB media containing 100 µg/mL ampicillin (and 25 µg/mL kanamycin for the pQE vector). Cultures grown overnight (approximately 16 hours) at 37 °C in an orbital shaker (200 rpm) were used to inoculate 400 mL LB media plus ampicillin (± kanamycin) in a 2 L conical flask and incubated at 37 °C (with shaking) until the optical density (OD600) reached 0.5 before adding 1.0 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside, Sigma-Aldrich). The culture was then incubated at 27 °C for six hours. In addition, a range of conditions (described in Chapter 5.3) were used in an attempt to optimise the expression of soluble recombinant protein. After growth, cultures were centrifuged at 10,000 x g for 10 minutes, and bacterial pellets washed three times in PBS. Pellets were then stored at -20 °C until required for cell lysis and protein purification.

5.2.4 Cell lysis of transformed *E. coli*

Frozen pellets of *E. coli* were thawed on ice for 10 minutes before being resuspended in lysis buffer (see appendix, Table 8-1) and incubated on ice for 20 minutes. For each fusion tag (GST or His) a different buffer was used but both contained an EDTA-free protease inhibitor (Roche) and 1 mg/mL lysozyme (Sigma-Aldrich). Suspensions were sonicated on ice for six cycles of 10 seconds on and 10 seconds off before being centrifuged at 16,000 x g for 10 minutes at 4 °C. Material was used immediately for IMAC purification

(immobilised metal affinity chromatography) or kept at 4 °C for up to three days. For the codon-optimised, His-tagged recombinant protein, insoluble material was resuspended in GST lysis buffer containing 8 M urea and incubated at room temperature for 20 minutes on a magnetic stirrer. The suspension was then centrifuged (16,000 x g for 10 minutes) before separating the supernatant and pellet.

5.2.5 Purification of recombinant proteins

Recombinant protein was purified using IMAC by incubating cell lysates with an agarose-based resin, removing unbound material and eluting bound protein. For GST-tagged proteins, the resin was cross-linked to glutathione (Pierce), while His-tagged lysates were incubated with resin incorporating nickel (Ni-NTA) ions (His-Pur, Thermo Scientific). 4 mL of glutathione or Ni-NTA resin slurry (2 mL resin volume) was centrifuged at 700 x g for 30 seconds and washed in two resin volumes of GST or His equilibration buffer respectively (Appendix, Table 8-1). Resin and cell lysate were incubated in a 15 mL falcon tube at 4 °C overnight on a rocking platform.

After binding, the resin/lysate mixture was poured into a polypropylene column (Qiagen) and allowed to settle for five minutes. The column was opened and the flow through collected and reapplied to the column. The resin was washed twice with five resin volumes of GST or His wash buffer. Bound protein was eluted from the resin by applying 8 x 1 mL of GST or His elution buffer (Appendix, Table 8-1). Each elution fraction was collected individually.

5.2.6 Factor Xa cleavage of GST-tag

The GST tag was cleaved from recombinant proteins with using Factor Xa protease (New England Biosciences). The purest IMAC elution fractions were visualised using SDS-pAGE stained with Coomassie blue and quantified using a Qubit® fluorometer (Life Technologies) as described previously (Chapter 2.5.6). 2 mM $CaCl_2$ and 1.0 µg Factor Xa per 50 µg protein were added to the combined elution fractions before being incubated at 23 °C for six hours.

5.2.7 Protein concentration and dialysis

Protein in eluted fractions was concentrated using Pierce 7 mL, 9K MWCO (molecular weight cut-off) concentrators (Life Technologies). Eluted fractions were combined and centrifuged at 3,000 x g until the desired volume was reached. After concentration, recombinant protein samples generated under denaturing conditions (*i.e.* 8 M urea) were dialysed to remove the majority of urea. 2.5 mL of sample was added to a 3 mL, 7K MWCO (molecular weight cut-off) slide-a-lyzer™ dialysis cassette (Life Technologies) following the manufacturer's instructions. The cassette was placed in 1 L PBS containing 6 M urea and stirred at 4 °C. Buffer was replaced three times with PBS containing decreasing levels of urea (4 M, 2 M and 1.5 M) urea at 2, 4 and 6 hours. The final incubation, in PBS + 1.5 M urea, was allowed to proceed overnight. The dialysed protein sample was then removed from the cassette and kept at 4 °C for up to seven days or at -80 °C for long-term storage.

5.3 Results

5.3.1 Structure of the *H. contortus* Argonaute-like gene (*alg-1*)

BLASTp analysis of the nascent *H. contortus* genome using *C. elegans* ALG-1 sequence identified a single Argonaute-like gene (HCISE00973200), henceforth referred to as *Hco-alg-1* (Fig. 5-1A). Analysis of later versions of the *H. contortus* genome identified a second Argonaute gene (HCISE00645900, *Hco-alg-2*). The alg-2 gene is likely to be incomplete as it was detected at the end of a scaffold (#1960). Approximately 14 % of the 3′ end is therefore not known. Based on predicted exons, *alg-2* is > 99 % similar to *alg-1*, while introns also display a high degree of similarity 50 – 100 %). This suggests that a gene duplication event occurred relatively recently, resulting in two *argonaute* genes in *H. contortus*. Unfortunately as *alg-2* was discovered at the end of the project, no further information was gained. All of the following work refers specifically to *Hco-alg-1*.

While the majority of the *Hco-alg-1* sequence aligned well to Cel-ALG-1 (95 % coverage, 83 % identity), the 5′ showed minimal similarity. Sequence (Appendix, chapter 8.2.3) was therefore confirmed using 5′ RACE and searches of the completed genome after release. The unspliced *Hco-alg-1* gene, excluding 5′ and 3′ UTR sequences, was 5,569 nucleotides long, while the spliced length was 3,003 nt. The translated protein contained 1,001 amino acids, with a mass of 111 kDa (estimated using Geneious, version 5). There were 21 exons with an average length of 143 nt (range 39 – 249 nt) and 20 introns with average length of 124 nt (range 33 – 427 nt). Although the Cel-ALG-1 protein was similar in length (1,002 aa), the gene was more compact (unspliced = 3,417 nt) with six exons (average length = 502, range 39 – 1,463) and five introns (average length = 82 nt, range 52 – 142). This pattern of larger unspliced genes with more introns is common when *H. contortus* genes are compared to *C. elegans* (R. Laing, personal communication).

Homologous ALG sequences were identified in a range of species, primarily covering nematodes (20 sequences) as well as non-nematodes (four sequences). The nematode species included two from clade I (*Trichinella spiralis* and *Trichuris trichuria*), seven from clade III (*Ascaris suum*, *Brugia malayi*, *Dirofilaria immitis*, *Loa loa*, *Onchocerca volvulus*, *Toxocara canis* and *Wucheria bancrofti*) and eleven from clade V (*Ancylostoma ceylanicum*, *Caenorhabditis brenneri*, *Caenorhabditis briggsae*, *C. elegans*, *Caenorhabditis remanei*, *Dictyocaulus viviparous*, *H. contortus*, *Heligmosomoides polygyrus*, *Necator americanus*, *Nippostrongylus brasiliensis* and *Oesophagostomum dentatum*). The four non-nematode species included *Danio rerio*, *Drosophila melanogaster*, *Homo sapiens* and *Mus musculus*.

The nematode proteins were labelled as ALG-1 or ALG-2, except for *T. spiralis* which was labelled as Eukaryotic initiation factor 2C (EIF2C) and seven unnamed hypothetical proteins. The non-nematode proteins were simply referred to as Argonaute. Although three proteins were officially labelled as ALG-2, they displayed greater homology to ALG-1. From current genome data in NCBI only one Argonaute-like gene can be identified in each parasitic nematode, with the exception of *A. suum* which contains three (*alg-1*, *alg-4* and *alg-6*) and H. contortus (*alg-1* and *alg-2*). In contrast, two miRNA-specific *alg* genes are present in *C. elegans* (*alg-1* and *alg-2*).

While the majority of the Argonaute proteins were found to be highly conserved between species, alignments highlighted marked differences in both length and sequence of the N-terminal (Fig. 5-1B). A four residue motif, PRRP (position 155 – 158 in Hco-ALG-1), was used as the boundary between the variable (N-terminal) and conserved (C-terminal) regions. This motif was present in all Argonaute proteins except in *D. rerio*, *N. americanus* and *O. dentatum*. In *D. rerio*, there was a substitution producing a PQRP motif, while in the latter two species there was no sequence covering this motif.

Considering the high degree of conservation in other species, it is likely that this represents poor assembly or annotation of Argonaute in these two species.

The large conserved region (covering 847 aa in *H. contortus*) contained three putative domains: DUF1785 (domain of unknown function, 65 aa), PAZ (Piwi, Argonaute and Zwille, 121 aa) and PIWI (P-element induced wimpy testis, 417 aa). The length and position of these domains remained consistent between species. In contrast, the length and sequence of the variable region differed markedly between species, ranging from 19 aa in *D. viviparous* to 160 aa in *C. briggsae*. Nematode Argonaute proteins tended to have a longer variable region than vertebrate homologues. Interestingly, the variable region displayed strong conservation within, but not between, clades (Fig. 5-2).

ClustalW alignment of the full Argonaute protein across 24 species was used to generate a phylogenetic tree (Fig. 5-1C). As expected from the high level of conservation, there was little diversity between the species as indicated by the low level of substitutions per site (scale bar in Fig. 5-1C). The nematode proteins were clustered into groups that matched their phylogenetic clade. Sequences from clades III and V were more similar to one another than to the clade I sequences.

To understand the level of selection pressure for Argonaute, non-synonymous/synonymous substitution ratios (dN/dS or Ka/Ks) were calculated from the nucleotide sequences of nine nematode species (Table 5-1). Across the entire gene, purifying selection pressures (dN/dS < 1) were maintaining the protein sequence. However, when separated, the variable region was found to be under positive selection pressures (dN/dS > 1).

Table 5-1: dN/dS ratio of Argonaute genes (and its regions) from nine nematode species.

Region	dN	dS	dN/dS
Entire gene	166.47	466.20	0.36
Variable region	20.18	8.86	2.28
Conserved region	146.29	457.21	0.32

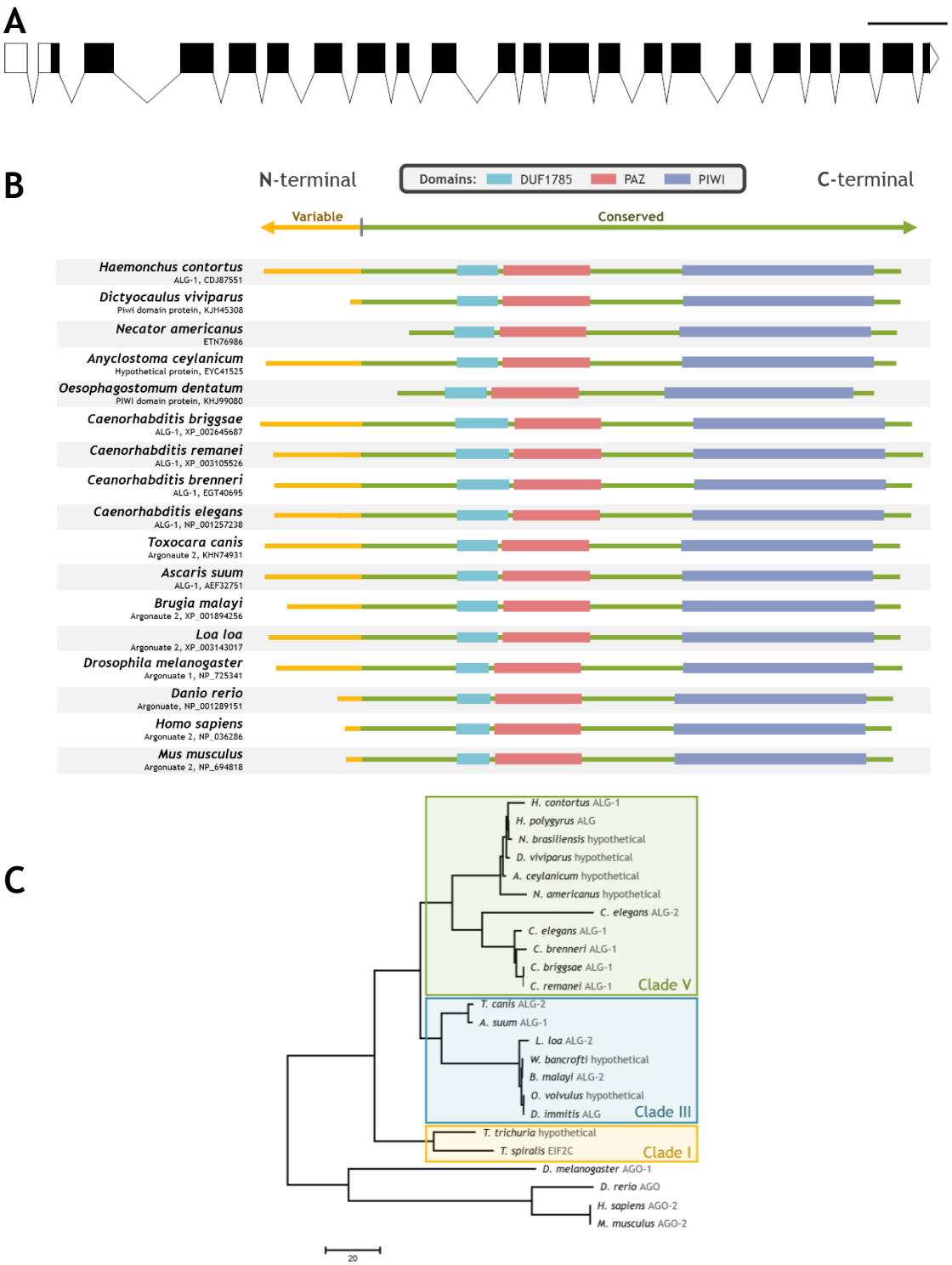


Fig. 5-1: Domain structure and phylogeny of Argonaute. **A**: Intron/exon structure of *Hco-alg-1*. UTRs are marked in white, exons in black, introns are represented by bent lines. Horizontal bar represents 500 nucleotides. Generated using Exon-Intron Graphic Maker (http://wormweb.org/exonintron). **B**: Alignment of Argonaute proteins from 17 species. Based upon conserved and variable regions, highlighted in green and yellow respectively. Putative domains, identified using pfam, are marked. Official names and NCBI accession numbers are listed for each protein. **C**: Phylogenetic tree of Argonaute proteins. Nematode phylogenetic clades (Blaxter *et al.*, 1998) are indicated (yellow = 1, blue = III, green = V). Generated in MEGA (version 6) using the Neighbour-joining method. Scale bar represents number of amino acid substitutions per site.

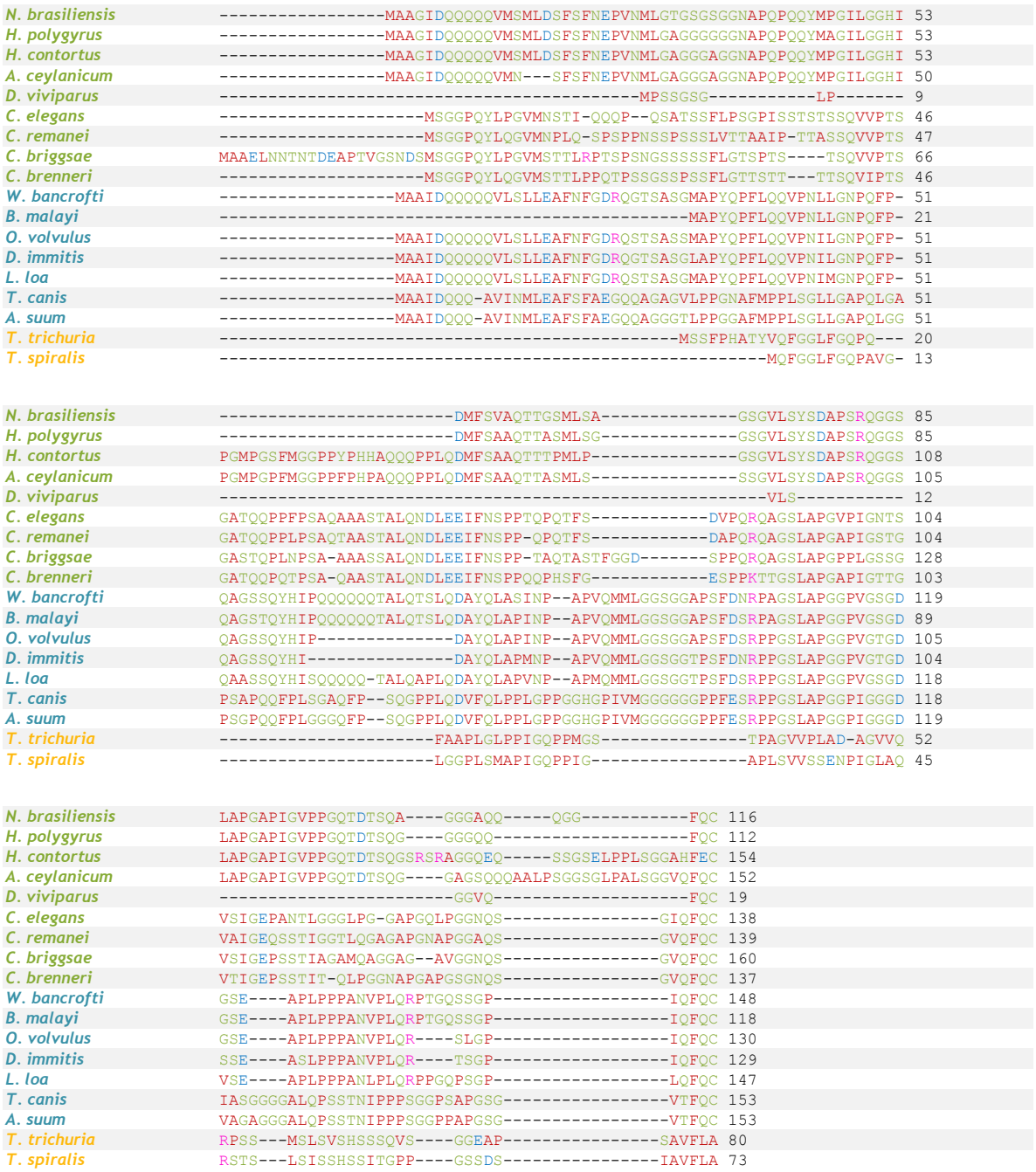


Fig. 5-2: ClustalW2 alignment of N-terminal variable region of 18 nematode Argonaute proteins. Species are colour coded according to phylogenetic clade (yellow = 1, blue = III, green = V). Amino acids are colour coding using the ClustalW colour scheme, based upon physiochemical properties (red = small/hydrophobic, blue = acidic, magenta = basic, green = Hydroxyl/ sulfhydryl/amine/G, grey = others).

5.3.2 Generating a peptide antibody against *H. contortus* ALG-1

The first example of HITS-CLIP in nematodes used an anti-Argonaute antibody (Zisoulis *et al.* 2011) raised in rabbits using a peptide sequence taken from the N-terminal variable sequence of Cel-ALG-1 (Fig. 5-3). Although this region was not conserved, we tested the reactivity of this antibody (Pierce #PA1031, Thermo Scientific, Rockford, IL, USA) against lysates generated from adult *H. contortus*. SDS-PAGE and western blot analysis failed to detect any protein when probed with the *C. elegans* ALG-1 peptide antibody (Fig. 5-4A). This antibody also failed to detect Argonaute in non-denatured *H. contortus* adult lysates. An anti-Human Argonaute antibody (Diagenode, Belgium) raised against an unspecified 400 aa region of Hsa-AGO2 was also tested but failed to cross-react with Hco-ALG-1 (data not shown).

In response, an *H. contortus* anti-ALG peptide antibody was generated following the same principles used for the *C. elegans* antibody. A peptide covering 22 aa of the variable region (Fig. 5-3) was synthesised by GenScript (Piscataway, NJ, USA) with an acetylated N-terminal to improve conformation and a C-terminal cysteine for use in keyhole limpet haemocyanin (KLH) conjugation. Two rabbits were immunised four times over the course of eight weeks before being exsanguinated. The serum was combined before antibodies were affinity purified, lyophilised and shipped. When tested on SDS-denatured lysates of *H. contortus* L3 and adult worms, the antibody detected two bands at 55 and 14 kDa (Fig. 5-4B); the expected size of Hco-ALG-1 was 110 kDa. Appearance of the 55 kDa band was inconsistent, present on approximately half of the blots, while the 14 kDa was often only faintly visible. Use of increased protein/antibody concentrations, different life cycle stages and modification of the lysis procedure (*e.g.* non-denaturing techniques, addition of protease inhibitors) did not improve detection. In contrast, an anti-actin antibody consistently showed reactivity of a band of the appropriate size (~ 50 kDa, Fig. 5-4C).

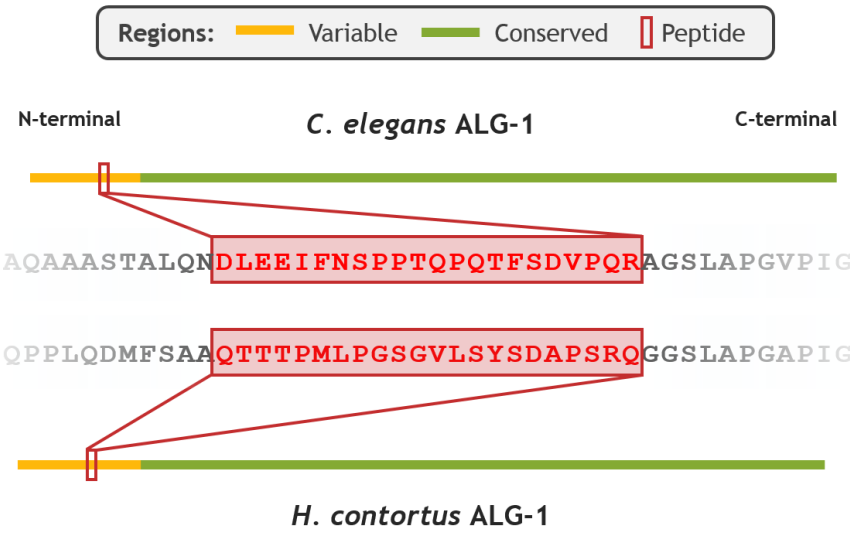


Fig. 5-3: Amino acid sequences used to generate anti-Argonaute peptide antibodies in *C. elegans* and *H. contortus*.

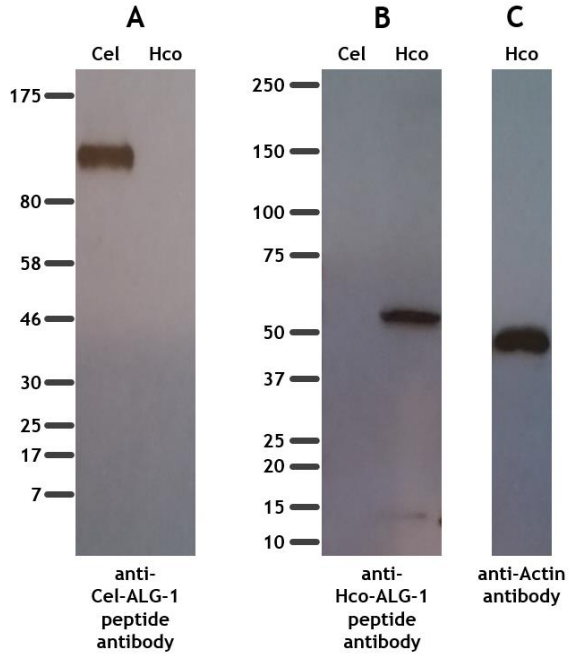


Fig. 5-4: Western blots of *C. elegans* and *H. contortus* lysates probed with anti-Argonaute peptide antibodies. **A**: Blot probed with an anti-Cel-ALG-1 peptide antibody (1:5,000). **B**: blot probed with anti-Hco-ALG-1 peptide antibody (1:1,000). **C**: Blot probed with anti-Actin antibody (1:1,000). Secondary anti-IgG antibodies were HRP-conjugated and used at 1:10,000 dilution. Blots were exposed to radiographic film for 30 seconds.

5.3.3 Expression of a GST-tagged recombinant Argonaute protein fragment

As the anti-Argonaute peptide antibody failed to recognise a protein of the predicted size for Hco-ALG-1 (110 kDa), larger protein fragments were expressed as recombinant proteins with the intention of generating a polyclonal antibody. The first fragment to be expressed covered 154 aa (16 kDa) of Hco-ALG-1 and was fused at its N-terminal to glutathione S-transferase (GST) with a Factor Xa recognition site and expressed in *E. coli* BL21. In total, the fusion protein was 365 aa, with an estimated mass of 42 kDa (Fig. 5-5).

After induction (Fig. 5-6A) cells were lysed and the soluble fraction was purified using IMAC. Multiple bands were observed on SDS polyacrylamide gels of the eluted samples stained with Coomassie blue (Fig. 5-6B). The most abundant band was at the expected size of the recombinant protein, approximately 42 kDa. At least four bands were visible in a cluster at 25 – 37 kDa. While not fully understood, it is known that GST tags can degrade during SDS-PAGE producing a ladder effect. Two faint bands were visible at ~ 75 and 14 kDa, however, both were visible in uninduced cell cultures, suggesting that they were contaminating proteins that could bind to GST.

Purified protein samples were then incubated with Factor Xa (Fig. 5-6C). The majority of the fusion protein and GST fragments disappeared, replaced by a very strong band at 26 kDa and a faint band at 16 kDa. A gel slice of the 16 kDa band was sent to Glasgow Polyomics for mass spectrometry, confirming this as the ALG-1 protein fragment. Western blots probed with an anti-GST antibody indicated that the bands at 26 – 42 kDa contained GST (Fig. 5-6D). In addition, after treatment with Factor Xa, a small amount of fusion protein was detected, indicating that cleavage was not 100 % efficient.

On Coomassie stained gels, the abundance of liberated protein (16 kDa) was much lower than expected, given the intensity of the GST-tagged protein. Cleaved material was

purified again with IMAC, but GST was still present in the flow through and some target protein was lost during this process, making it difficult to produce pure, isolated target protein. This was a potential problem, as GST is highly immunogenic and an unknown proportion of antibodies would have been produced to the GST instead of the target protein. The low yields, inefficient cleavage and unsatisfactory purification meant that the GST-tagged protein was not used for antibody generation.

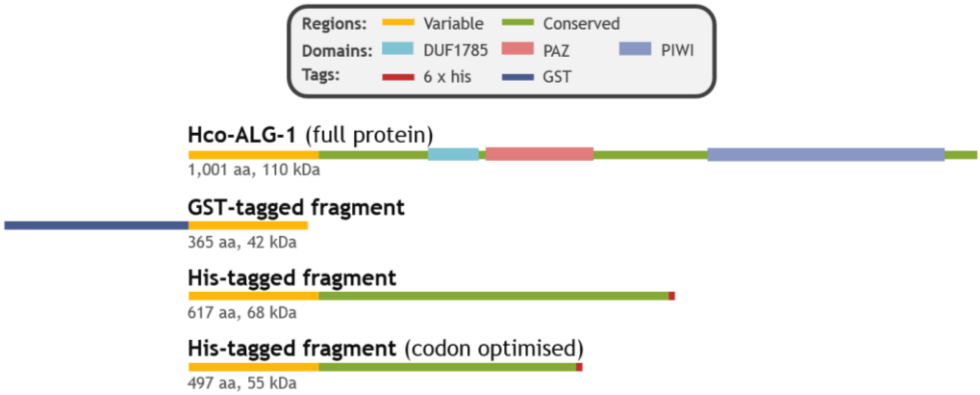


Fig. 5-5: Fragments of *H. contortus* ALG-1 expressed as recombinant proteins. Three different protein fragments were expressed in *E. coli* with either an N-terminal GST-tag or C-terminal His-tag. Molecular masses (kDa) were estimated using Geneious (version 5, Biomatters).

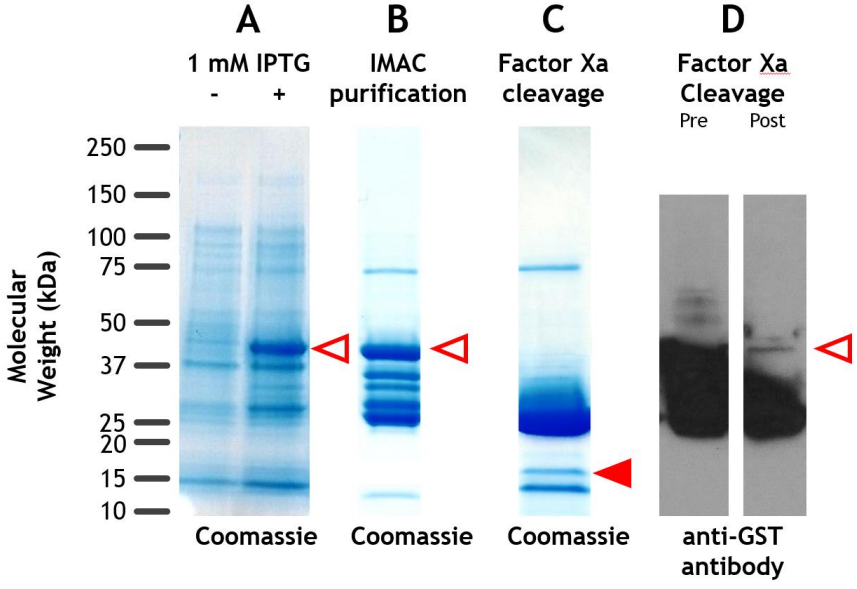


Fig. 5-6: Recombinant *H. contortus* ALG-1 protein with GST tag expressed in *E. coli*. Proteins were subjected to SDS-PAGE and Coomassie staining (A – C) or western blotting (D). **A**: Induction of bacterial cultures ± 1 mM IPTG with incubation at 27 °C for six hours. **B**: IMAC purification using glutathione-agarose resin. **C**: Cleavage of the GST-tag using Factor Xa. **D**: Western blot of IMAC purified material pre- and post-factor Xa cleavage, probed with an anti-GST antibody (1:1,000). Open arrowheads indicate GST-tagged protein. Closed arrowhead indicates GST-free target protein.

5.3.4 Expression of a His-tagged recombinant Argonaute protein fragment

As removal of the GST-tag was proving difficult, an alternative His-tag was used, incorporating six histidine residues in a row. The smaller tag (six aa versus 211 aa for GST) was fused to a larger portion of *Hco-alg-1*. It was hoped that this would increase the number of immunogenic epitopes and, based on sequence conservation, would also allow the antibody to be used in other nematode species, increasing its value. In total, the recombinant His-tagged protein, expressed in *E. coli* was 615 aa, with an expected molecular mass of ~ 68 kDa (see Fig. 5-5).

The His-tagged recombinant protein was induced (Fig. 5-7A) and lysed as described previously. Unfortunately, the vast majority of protein was insoluble (Fig. 5-7B). In addition, yields were lower than expected. Various modifications to induction were performed, including the use of enriched culture media (*e.g.* TB, superbroth), altering the induction temperature and duration (16, 20 or 37 °C for 48, 24 or 4 hours respectively) and varying the IPTG concentration (0.1, 0.5 or 1.0 mM IPTG). The lysis buffer was also modified using additives such as imidazole (20 mM), NaCl (100 – 500 mM), Triton-X-100 (1 %) or glycine-glycine (0.5 – 1.0 M). However, expression levels remained low, while solubility was only marginally improved by the addition of 1 M glycine-glycine (Fig. 5-7C).

The cell lysates were subjected to IMAC using Ni-NTA resin, but purity of the eluted fractions was low (Fig. 5-7D) with multiple bands visible between 10 and 100 kDa. Approximately 5 – 10 % of protein in the eluted fractions contained the recombinant protein at 68 kDa as confirmed by western blots using an anti-His(C-term)-HRP antibody (Life Technologies). Interestingly, when blots were probed with the anti-Hco-ALG-1 peptide antibody (see Chapter 5.3.2), the His-tagged recombinant protein was recognised

alongside additional bands of approximately 30, 40 and 60 kDa (Figs. 5-7E and F). A 25 kDa band was also identified by both antibodies. This may indicate protein degradation or cross-reactivity with other proteins in the sample. Detection of these additional proteins or fragments was unaffected by the addition of protease inhibitors (data not shown), suggesting they were not due to proteolytic degradation. Cobalt resin was also tested as a replacement for Ni-NTA during IMAC, but did not improve purity of the eluted fractions (data not shown).

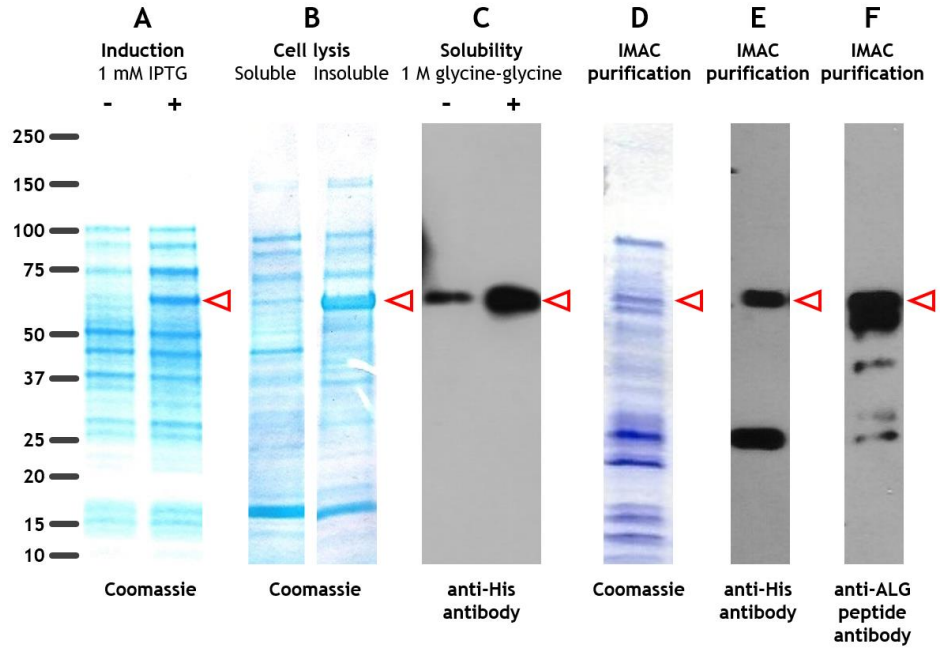


Fig. 5-7: Expression of recombinant His-tagged *H. contortus* ALG-1 protein fragment in *E. coli*. Cell lysates and purified proteins were subjected to SDS-PAGE and Coomassie blue staining (A, B and D) or Western blotting (C, E and F). Open arrowhead indicates position of recombinant protein (~68 kDa). **A**: Induction of bacterial cultures ± 1 mM IPTG incubated at 27 °C for six hours. **B**: Cell lysis showing the majority of recombinant protein was in the insoluble fraction (2nd lane). **C**: Improved solubility of recombinant protein following induction in the presence of 1 M glycine-glycine. Blot probed with anti-His (C-terminal) antibody (1: 10,000). **D**: IMAC purification using Ni-NTA-agarose resin. **E**: IMAC purified material probed with anti-His antibody (1:10,000). **F**: IMAC purified material probed with anti-Hco-ALG-1 peptide antibody (1:1,000). Secondary anti-IgG antibodies were HRP-conjugated and used at 1:10,000 dilution. Blots were exposed to radiographic film for 30 seconds.

5.3.5 Codon optimisation of *Hco-alg-1* for expression in *E. coli*

The *H. contortus alg-1* sequence was codon-optimised for *E. coli*, with the aim of improving expression of the recombinant protein, increasing solubility and aiding purification. Organisms display a preference in their codon usage and synonymous nucleotide changes were made to the *Hco-alg-1* sequence to match the preferred codons used by the expression system, in this case, *E. coli*.

A 1,479 nt fragment from the 5′ end of *Hco-alg-1* was uploaded to the GeneArt® online service (Life Technologies). Except for the 3′ and 5′ restriction enzyme sites, 342 nt (23 %) were changed, improving the codon quality significantly (Appendix, Fig. 8-6). The optimised sequence was then synthesised and inserted into the pMK-RQ-Bb vector. Upon receipt, the fragment, encoding a 499 aa protein with an estimated mass of 55 kDa (Fig. 5-5), was cloned into pQE-60 vector and transformed into *E. coli* M15 cells.

Upon induction (1 mM IPTG, incubated at 27 °C for six hours), the codon-optimised protein displayed slightly higher expression than the non-optimised fragment (Fig. 5-8A). Unfortunately, solubility was not improved, but the total amount of protein present in the insoluble fraction was much greater than previously obtained (Fig. 5-8B). The insoluble material was resuspended in cell lysis buffer containing 8 M urea to solubilise the protein (Fig. 5-8C) for purification using IMAC (Ni-NTA resin). While multiple bands were visible, a much larger percentage of the total protein in the eluted fractions (~ 90 %) consisted of the recombinant protein (Fig. 5-8D) compared to the non-optimised His-tagged protein (Chapter 5.3.4). Western blots using an anti-His antibody (Fig. 5-8E) and the anti-Hco-ALG-1 peptide antibody (Fig. 5-8F) identified a band of 55 kDa. Some smaller (40 – 50 kDa) and larger (100 – 140 kDa) bands were detected by the peptide antibody suggesting recognition of other proteins. Modifications to the IMAC protocol, including the use of Cobalt resin, additives to the buffers (*e.g.* triton-X-100, glycerol) and a

low pH elution buffer (instead of high imidazole), were tested to improve elution purity, but made no significant difference (data not shown).

Suitably purified, the recombinant protein was concentrated to approximately 1 mg/mL before dialysis. The limit of urea required by the company used to generate the antibody (Dundee Cell Products, Scotland) was 1.5 M urea and the recombinant protein was dialysed to this level in PBS. Unfortunately, below 2 M urea, a significant proportion of the protein became insoluble.

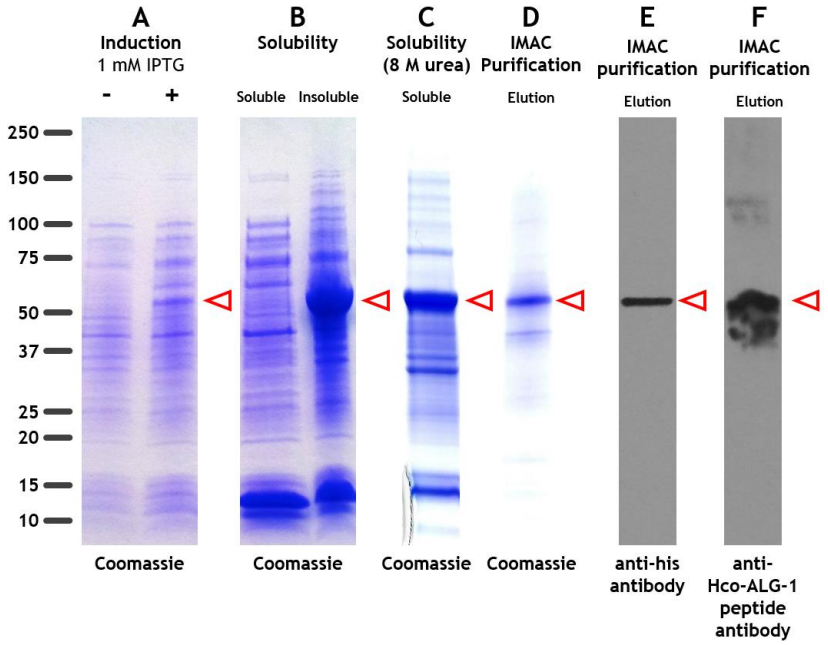


Fig. 5-8: Codon-optimised recombinant *H. contortus* ALG-1 protein with His-tag expressed in *E. coli*. Proteins were subjected to SDS-PAGE and Coomassie staining (A - D) or western blotting (E and F). **A**: IPTG induction of bacterial cultures ± 1 mM IPTG incubated at 27 °C for six hours. **B**: Cell lysis with the majority of recombinant protein in the insoluble fraction (2nd lane). **C**: Cell lysis showing solubility of the recombinant protein in 8 M urea. **D**: IMAC purification using Ni-NTA-agarose resin. **E**: IMAC purified material probed with anti-His antibody (1:10,000). **F**: IMAC purified material probed with anti-Hco-ALG-1 peptide antibody (1:1,000). Open arrowheads indicate position of recombinant protein (~55 kDa). Secondary anti-IgG antibodies conjugated to HRP and used at 1:10,000 dilution. Blots exposed to radiographic film for 30 seconds.

5.3.6 Testing anti-Hco-ALG-1 polyclonal antibody

Approximately 4 mg of recombinant, codon-optimised, His-tagged Hco-ALG-1 protein fragment was sent to Dundee Cell Products as a mixture of insoluble and soluble protein in PBS containing 1.5 M urea. Two rabbits were immunised four times (~ 0.2 mg per immunisation). These rabbits were previously selected by checking their sera for a lack of reactivity against the recombinant protein (data not shown). Pre-immunisation and final-bleed sera were supplied along with purified IgG antibody. These were used to probe Western blots of *H. contortus* whole worm lysates (mixed adult stages) and recombinant Hco-ALG-1 protein (Fig. 5-9).

On western blots of the worm lysates, multiple bands were detected between 60 and 250 kDa (Figs. 5-9A, B and C). Surprisingly, there was no visible differences in the number of bands detected by each of the sera (pre-immunisation, final-bleed and purified IgG), although there was some minor variation in the relative intensity of each band. No differences were apparent between probes from the two rabbits (data not shown). Worryingly however, was the poor detection of the recombinant protein. Only at high concentrations (1:100), was the purified antibody able to detect a band at 55 kDa. At such high levels, there was a large amount of background noise on the blots. The antibody also detected a faint band at 11 kDa, but this was not detected by the anti-His (Fig. 5-9E) or anti-Hco-ALG-1 peptide antibodies (Fig. 5-9F). Dot blots using purified IgG on the recombinant protein were carried out by Dundee Cell Products and in our lab, both producing titres of 1:100 – 1:200. Due to the low reactivity and poor specificity towards Argonaute, the antibody was considered unsuitable for use in HITS-CLIP. Further work to improve the solubility of the recombinant protein is on-going.

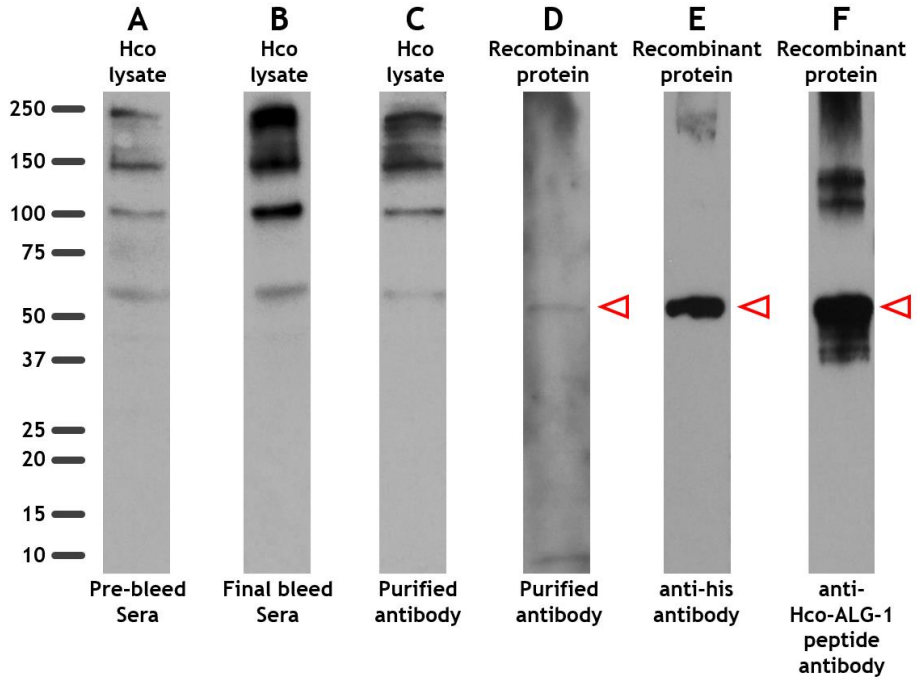


Fig. 5-9: Western blots testing antibody generated in rabbits immunised with recombinant Hco-ALG-1 protein. **A**: *H. contortus* lysate probed with pre-immunisation sera (1:250). **B**: *H. contortus* lysate probed with final-bleed sera (1:250). **C**: *H. contortus* lysate probed with purified IgG antibody (1:250). **D**: Recombinant Hco-ALG-1 protein probed with purified antibody (1:100). **E**: Recombinant Hco-ALG-1 protein probed with anti-His(C-terminal) antibody (1:10,000). **F**: Recombinant Hco-ALG-1 protein probed with anti-Hco-ALG-1 peptide antibody (1:1,000). The open arrowhead indicates position of the recombinant protein (~55 kDa). Secondary anti-IgG antibodies were HRP-conjugated and used at 1:10,000 dilution. Blots were exposed to radiographic film for 15 seconds, except for D, which was exposed for 60 seconds.

5.4 Discussion

Validating mRNA targets is important for understanding the biological function of miRNAs without relying solely on bioinformatic target predictions. A number of experimental techniques have been used successfully to confirm miRNA-mRNA interactions (reviewed in Ritchie *et al.*, 2013), such as the oncogenic targets of Hsa-miR-124 in human renal cell carcinoma (Butz *et al.*, 2015) or the role of Hsa-miR-96 in the development of pulmonary hypertension (Wallace *et al.*, 2015). However, these techniques are restricted to organisms and cell lines that can be genetically manipulated or cultured for extended periods of time and are therefore unsuitable for studying miRNA-

mRNA interactions in *H. contortus*. In contrast, HITS-CLIP can be used in organisms without the need for genetic manipulation, only requiring UV irradiation. Isolation of the cross-linked RISC and its RNA passengers requires an antibody specific to one of the protein components, often Argonaute (Jain and Parker, 2013). In the current work, commercially available antibodies for *C. elegans* ALG-1 and *H. sapiens* AGO-1 were tested for cross-reactivity with *H. contortus* ALG-1. Unfortunately, these failed to recognise any protein by western blotting and therefore an *H. contortus* antibody was required.

The first step to generate an antibody was identifying Argonaute in *H. contortus*. The number of Argonaute genes is highly variable between species. In *Saccharomyces pombe* (fission yeast) there is only one Argonaute gene, while *H. sapiens* contain eight and in *C. elegans* there are 27. Multiple Argonaute genes in an organism tend to be functionally specialised, with two in *C. elegans* (*alg-1* and *alg-2*) and four in *H. sapiens* (*ago1 – ago4*) (reviewed in Buck and Blaxter, 2013 and Hutvagner and Simard, 2008). Searches of the nascent versions of the *H. contortus* identified a single argonaute gene (*Hco-alg-1*), while later versions contained two highly similar argonaute genes. Sequence of the second argonaute gene (*Hco-alg-2*) was incomplete, as it was at the end of a scaffold. Approximately 14 % of the highly conserved 3′ UTR was missing. However, due to the high degree of similarity between the remaining sequences of *Hco-alg-2* and *Hco-alg-1* (> 99 % between coding regions), it is unlikely that their functions differ significantly. These two genes may display different expression patterns, which could lead to different functions. Confirmation of (potential) variation requires further examination. From the data so far, it is quite possible that *alg-1* and *alg-2* represent a relatively recent gene duplication event.

Unsurprisingly, alignment of predicted Argonaute proteins from multiple organisms identified a high degree of conservation. The position and size of three putative domains (DUF1785, PAZ and PIWI) were particularly well conserved. While functional information is available for PAZ and PIWI, no particular role has been identified for DUF1785. The PAZ domain is named after the Piwi, Argonaute and Zwille proteins in which it is commonly found (Cerutti *et al.*, 2000). One of the subdomains in PAZ is an oligonucleotide binding fold (Yan *et al.*, 2003) and is involved in recognising and initiating binding of small RNAs (Frank *et al.*, 2010; Lingel *et al.*, 2004). The PIWI domain was named after the PIWI (P-element induced wimpy-testis) proteins of *D. melanogaster* (Cox *et al.*, 1998). It is found in many proteins that bind nucleic acids, particularly those involved in RNA cleavage (reviewed in Luteijn and Ketting, 2013). Multiple studies have identified the PIWI domain as the catalytic core of Argonaute proteins (Bouasker and Simard, 2012; Tolia and Joshua-Tor, 2007). Although not considered a specific domain, the N-terminal of Hsa-AGO-2 is believed to initiate unwinding of RNA duplexes during assembly of the RISC (Kwak and Tomari, 2012). This N-terminal sequence aligns with the start of the 'conserved region' as defined earlier (Chapter 5.3.1) and likely plays a similar role in other species, including *H. contortus*.

Sequence alignment shows that the majority of the Argonaute protein is conserved between species, with only the N-terminal region (covering 15 % of Hco-ALG-1) being highly variable. Using x-ray crystallography, the 3D structure of Argonaute proteins have been determined from yeast (Nakanishi *et al.*, 2012) and *Homo sapiens* (Schirle and MacRae, 2012). However, these provide little insight into the conformation or possible function of the N-terminal, as their variable region is relatively short (< 30 aa). Natural selection would have likely removed this region if it had no purpose and it is therefore interesting that in nematodes, this region is maintained and highly conserved between related species (*e.g.* filarial nematodes), but variable between clades. No functional domains could be

identified in the variable region and its role may be clade-specific, possibly involving signalling, processing, localisation or interactions with other proteins. Its importance could be examined most readily in *C. elegans* by testing the ability of constructs lacking this region to rescue *alg* mutant phenotypes.

After identification of Hco-ALG-1 and following the method used by Zisoulis *et al.* (2011), an antibody was raised against a short peptide sequence from the N-terminus. While the antibody detected two bands on western blots (14 and 55 kDa), these were much smaller than the expected size of *H. contortus* ALG-1 (~ 110 kDa). If we assume that these bands represent Argonaute, why were they smaller than expected? Even though broad-spectrum protease inhibitors were present, these bands may be degradation fragments of ALG-1. However, degradation generally creates multiple bands of gradually decreasing size (*i.e.* a ladder effect) rather than two distinct bands of very different sizes. In addition, the actin controls did not display any degradation. If instead we assume that these bands were not Argonaute, two questions are raised: what protein do these bands represent and why was ALG-1 not detected? The two bands may simply be cross-reacting with the peptide antibody. The *H. contortus* genome was checked for sequences that matched the peptide, but none were found outside of ALG-1. However, we know that the genome is incomplete (Chapter 4.3.3), so proteins that display cross-reactivity may be missing from the sequence data.

As to why ALG-1 was not detected, it is possible, that the abundance of ALG-1 fell below the detection threshold of western blots. Analysis of the transcriptome showed that the amount of *Hco-alg-1* mRNA was low in most life cycle stages but relatively high in adult females and particularly high in eggs (Appendix, Fig. 8-7). There were no differences detected by the peptide antibody between lysates generated by different stages (L3, adult

males and adult females) although eggs and early larval stages were not available for testing.

Another possible explanation is that the antibody was unable to detect Argonaute due to post-translational modification. Bacteria are unlikely to modify fragments of Hco-ALG-1 in the same manner as the full protein in *H. contortus*. For example, it may be that the region containing the peptide sequence is cleaved from the conserved region or the peptide sequence might be hidden by modifications. Hsa-AGO-2 has been found to be modified with phosphate groups, hydroxyl groups and ubiquitin (Johnston and Hutvagner, 2011). Such modifications can be removed using enzymes and chemical treatments. If modified groups were blocking recognition by the antibody, their removal would allow Hco-ALG-1 to be identified on western blots and possibly used for HITS-CLIP.

The most curious aspect of the peptide antibody was its ability to detect recombinant Hco-ALG-1 fragments expressed by *E. coli*. The peptide antibody was set aside to develop a polyclonal antibody raised against larger regions of the Hco-ALG-1 protein. When these fragments were probed with the peptide antibody, it was found to be both specific and sensitive. Why this antibody was able to detect recombinant ALG-1 fragments, but not native Argonaute remains unclear and is the focus of additional study. Recognition may reflect the difference in abundance between recombinant and native proteins (high and low concentrations respectively).

Three recombinant protein fragments were generated from Hco-ALG-1. As there were originally concerns over the possible existence of multiple Argonautes in *H. contortus*, the first fragment covered only the N-terminal variable region. The identification of a second argonaute gene in *H.* contortus that contains an identical variable region, renders this approach moot. While expression levels and purity were satisfactory, removal of the GST

tag was a major stumbling block. Various refinements to induction, lysis, purification and cleavage were unable to reduce the amount of GST present. It was unclear how effective antibodies raised from this mixture of GST and ALG-1 fragment would have been in identifying *H. contortus* Argonaute. It was therefore decided to use a less immunogenic His-tag that did not require removal.

As the His-tag was much smaller than the GST tag, a larger argonaute fragment was used , covering 60 % of Hco-ALG-1. Unfortunately, expression and solubility of the initial His-tagged construct was poor, impeding purification of the recombinant fragment.

Following recommendations from Dr. Stepek, a fragment of *alg-1* was codon optimised for expression in *E. coli*. Expression of this codon-optimised His-tagged protein was indeed higher than the non-optimised protein, but poor solubility remained a problem. GST-tags markedly improve the solubility of recombinant proteins (Niiranen *et al.*, 2007), which likely explains why the short ALG-1 fragment was soluble in that expression system. In contrast, the vast majority (> 95 %) of the His-tagged fragments was insoluble. The pH of the buffers used during cell lysis was kept at or below pH 7.4, which should be far enough away from the pI of Hco-ALG-1 (8.5 – 9.5) to promote solubility. Changes to ionic strength of the buffer, through the addition of NaCl (up to 0.5 M), did not improve solubility.

The recombinant protein was solubilised using 8 M urea and the protein refolded by dialysis in decreasing concentrations of urea. As the concentration of urea fell below 2 M, much of the protein became insoluble. Protein refolding during dialysis can be problematic (reviewed inYamaguchi and Miyazaki, 2014), leading to misfolded proteins that do not match the original protein. Antibodies generated from these epitopes are less likely to bind to the target, decreasing their usefulness. However, insoluble protein can

generate a stronger immune response, by acting as a long-lasting nidus when injected (Dundee Cell Products, personal communication).

The protein, in both soluble and insoluble forms, was used to generate a polyclonal antibody. However, when tested on lysates of *H. contortus*, the purified antibody and final-bleed sera showed no significant difference in reactivity compared to the pre-immunisation sera. In addition, when tested on the recombinant protein, only two bands were faintly detected, suggesting that the recombinant protein failed to illicit a strong immune response. Even if the antibody showed strong reactivity to the recombinant protein, it may not have detected the native protein, due to refolding errors as an insoluble protein. Our lab is currently exploring the use of co-expressed molecular chaperones (*e.g.* HSP70, Clontech) to increase solubility of the recombinant protein. Early results are encouraging, with higher levels of the recombinant protein found in the soluble fraction after cell lysis. Expression will be scaled up and, if IMAC purity is acceptable, a new antibody will be generated, opening up the possibility of HITS-CLIP in *H. contortus*. To complement this approach, we also examined miRNA-mRNA interactions and possible functions using *C. elegans* as a model organism as discussed in Chapter 6.

6 Investigating the biological impact of miRNAs

6.1 Introduction

In addition to confirming miRNA-mRNA interactions, understanding the role of miRNAs in a wider biological context is equally, if not more, significant. Using a range of techniques, the potential function of miRNAs in *H. contortus* and *C. elegans* was examined. While these organisms occupy different ecological niches (parasitic *versus* free-living), both belong to same phylogenetic clade (Blaxter *et al.*, 1998). The relative ease of genetic manipulation available in *C. elegans* provided opportunities that are currently unavailable in *H. contortus* while still highlighting the importance of miRNAs in nematode development.

Microarray analysis of *H. contortus* gut tissue (Chapter 3.3.1) showed that many miRNAs display spatial variation in expression (*i.e.* between the gut and other tissues). While informative, this analysis was limited to the gut tissue of adult females, the easiest stage to dissect due to its larger size. To further explore spatial variation, we exploited the availability of *C. elegans* reporter strains which direct expression of a marker gene, such as GFP, under the control of a miRNA promoter. This allowed us to examine spatial variation in greater resolution throughout the entire organism.

We also used these reporter strains to examine the effects of stress on miRNA expression. Numerous experiments have found that miRNAs play an important role in mediating an organism's response to stress (reviewed in Leung and Sharp, 2010). For example, both miR-7 in *D. melanogaster* (Li *et al.*, 2009b) and miR-43/84 in *C. elegans* (Burke *et al.*, 2015) protect development from fluctuations in temperature. Reporter strains linked to miRNAs that are conserved in both *C. elegans* and *H. contortus* were exposed to common

stressors (*e.g.* temperature, osmotic and radiation). As stress is a threat that all organisms must cope with, it is plausible that conserved miRNAs involved in regulating stress responses in *C. elegans* may also play a similar role in *H. contortus*.

C. elegans can also respond to certain stressors, such as high population density and low abundance of food, by arresting development and entering an alternate life-cycle stage known as the dauer (see Chapter 1.3.1). Dauers are long-lived, quiescent and resistant to environmental stress. If conditions improve, they can resume development and become reproductive adults. It is already known that miRNAs are involved in many aspects of the dauer state. The sensory neurons that detect unfavourable conditions (Bargmann and Horvitz, 1991) require multiple miRNAs (miR-81, miR-82, miR-124 and miR-234) to prevent an inappropriate response (Than *et al.*, 2013a). Once a signal has been detected, a complex feed-back loop between the miRNA let-7 and *daf-12* (a nuclear hormone receptor) mRNA regulates the decision to enter the dauer state (Hammell *et al.*, 2009). Cel-let-7 is also involved in the resumption of development if conditions improve, by modulating the expression of the transcription factor, *hbl-1* (Karp and Ambros, 2012). We further explored the role of miRNAs in dauer development by examining mutant strains that lacked specific miRNAs and components of the RISC complex.

The morphological and behavioural similarities shared between the infective larvae of parasitic nematodes, such as *H. contortus* L3, and *C. elegans* dauers are well known: they are environmentally resistant and arrest development waiting for ‘conditions to improve’, *i.e.* when ingested by a host. It has even been suggested that the dauer state may be a pre-adaptation for parasitism (reviewed in Crook, 2014). Therefore, exploring the possible roles of conserved miRNAs in *C. elegans* dauer formation and maintenance may provide insights into the developmental arrest of parasitic nematodes.

Finally, the effect of miRNAs on development in *H. contortus* was studied directly using miRNA mimics and inhibitors (see Chapter 1.2.3). Mimics are double-stranded RNA molecules designed to function in an identical manner to endogenous miRNAs, while inhibitors are anti-sense oligonucleotides that bind to, and obstruct the function of, endogenous miRNAs. Both systems are based on the sequence of a mature miRNA, allowing the study of the effects of specific miRNA over-expression, using a mimic, or depletion, in the presence of an inhibitor

By using this range of functional methods, the biological role of specific miRNAs in regulating development in *H. contortus* and the wider nematode phylum was investigated,

6.2 Materials and methods

6.2.1 Environmental stress on *C. elegans* miRNA::GFP reporter strains

C. elegans GFP reporter strains for *mir-60*, *mir-228* and *mir-235* (VT1733, VT1485 and VT1488 respectively) were acquired from the CGC (https://www.cbs.umn.edu/research/resources/cgc) and grown as described previously (Chapter 2.2.2). Mixed populations of worms were washed off plates using M9 buffer and centrifuged at 1,000 x g for one minute. Excess M9 was removed and the worm pellet transferred to seeded NGM plates. Four different stresses were examined: low temperature (4°C), high temperature (30°C), UV-B radiation (100 mJ/cm^2) or osmotic 'shock' (NGM agar with 300 mM NaCl). Worms were incubated under these conditions for 24 hours except for the UV test which involved 30 seconds of irradiation, after which worms were incubated under standard conditions (Chapter 2.2.1) for 24 hours. At two time points (4 and 24 hrs), worms were examined using a UV microscope as described previously (Chapter 2.2.4).

6.2.2 Generating a crude *C. elegans* dauer pheromone extract

In order to generate synchronised populations of dauers, a crude dauer pheromone was extracted from liquid cultures of wild type *C. elegans* (N2 Bristol strain) following protocols by Golden and Riddle (1982) and Andersen (http://www.princeton.edu/~eca/resources/Protocols/Dauer-Pheromone-Prep.pdf). Liquid cultures containing 400 mL S-medium and N2 worms from two 10 cm plates were incubated at 20 °C in a 2 L baffled flask. Cultures were fed pellets of OP50 at days one and seven and incubated until the dauer population was maximised (12 – 14 days). At this point (approximately 60 – 80 % dauers), cultures were centrifuged at 10,000 x g for 20 minutes and the resulting pellet of worms and bacteria discarded. Remaining worms were removed using a Buchner funnel and Whatman filter paper (Grade 1) before bacteria were removed using 0.22 μM PES vacuum filter units (Millipore). The filtrate was transferred to 15 cm Petri dishes without lids and placed in a fan-assisted incubator at 100 °C until the volume was reduced by 50 % (approximately 6 – 8 hours). The temperature was then reduced to 60 °C and the filtrate allowed to dry completely. The resulting solids were crushed using a mortar and pestle into a fine powder and thoroughly mixed with 100 % ethanol. Solids were allowed to settle over the course of five minutes before transferring the transferring the liquid to a fresh 15 cm Petri dish and dried at 60 °C. This process was repeated three times. Solid material was then resuspended in dH_2O and stored in 1 mL aliquots at -20 °C.

6.2.3 Quantifying dauer formation in *C. elegans*

Dauer formation was characterised in *C. elegans* mutant strains by allowing adults to lay eggs on NGM agar plates containing crude dauer pheromone extract. Plates were generated by placing pheromone on the bottom of 5 cm Petri dishes immediately before the addition of 4 mL peptone-free NGM agar cooled to 50 °C. Plates were allowed to cool for

20 minutes before adding 10 µL heat-treated *E. coli* OP50 (HT-OP50). Heat-treatment inhibited bacterial growth and involved heating cultures at 90 °C for 45 minutes (vortexed every five minutes) and freezing overnight at -20°C. After addition of HT-OP50, plates were allowed to dry for 45 minutes before 10 – 20 gravid *C. elegans* adult worms were allowed to egg-lay for 4 – 6 hours. Adults were then removed before incubating plates at 27 °C for 72 hours. The total number of worms on each plate was counted before plates were flooded with 1 % SDS and incubated at 20 °C for 20 minutes. Worms that survived SDS-exposure and possessed a dauer-like appearance (*i.e.* long and thin) were counted as dauers.

6.2.4 Examining recovery from the dauer state in *C. elegans*

Synchronised populations of *C. elegans* dauers, generated as described previously, were transferred to peptone-free NGM agar plates containing 0 – 1 % yeast extract. After 24 hours, the total number of worms on each plate was counted before plates were flooded with 1 % SDS and incubated at 20 °C for 20 minutes. Surviving worms were counted as dauers.

6.2.5 Identifying dye filling (*dyf*) phenotype in *C. elegans*

C. elegans wild type and mutant strains were stained using DiI (1,1'-Dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate, Sigma-Aldrich) and FITC (fluorescein 5(6)-Isothiocyanate, Sigma-Aldrich) to assess the competency of amphid/phasmid dye filling. DiI staining involved transferring a plate of mixed-stage worms to 1 mL M9 containing 10 µg DiI and incubated in 1.5 mL eppendorf tubes for 4 hours at 20 °C with 20 rpm rotation. Tubes were covered in aluminium foil to prevent DiI bleaching. Worms were then washed

three times in M9 (one minute at 1,000 x g). For FITC staining, the surface of an 8 cm seeded NGM plate was covered in 200 µL M9 + 50 µL 2 mg/mL FITC in dimethylformamide. The FITC was allowed to dry for two hours, after which mixed stage worms were transferred and incubated on plates at 20 °C overnight. Worms were then transferred to a FITC-free plate for 45 minutes. For both staining techniques, worms were visualised microscopically as described previously (Chapter 2.2.4) using either a FITC or Texas Red (DiI) UV filter with a 5 ms exposure time. Dye filling was characterised across the amphid (ADF, ASH, ASI, ASJ, ASK and ADL) and phasmid (PHA and PHB) neurons.

6.2.6 Testing *C. elegans* dauer pheromone on *H. contortus* L3

H. contortus L3 were exsheathed using CO_2 as described previously (Chapter 2.3.2). 1,000 exsheathed L3 were maintained in 250 µL EBSS containing 0 – 6 % dauer pheromone previously filtered using 0.22 µm PES syringe filters. Worms were incubated at 37°C in a 5 % CO_2 environment for four days. 20 µL aliquots were examined microscopically and the number of sheathed L3, exsheathed L3 and L4 counted.

6.2.7 Uptake of fluorescently labelled RNAs by *H. contortus* L3 and L4

Approximately 1,000 *H. contortus* L3 were exsheathed using bleach as described previously (Chapter 2.3.2) and incubated in 100 µL EBSS (with penicillin, streptomycin and amphotericin) in 96-well plates. Wells were spiked with 1 µL 50 µM fluorescently labelled RNA (500 – 1,000 nM final concentration). Two different RNAs were used: Silencer® Cy™3 labelled Negative control No. 1 siRNA (AM 4621, Ambion) and Cy™3 dye-labelled Anti-miR Negative control #1 (AM17011, Ambion). Plates were incubated at

37 °C with a 5 % CO_2 atmosphere. 20 µL aliquots were removed daily, washed in M9 and observed under UV microscope using the Texas Red filter, as described previously.

6.2.8 Incubating *H. contortus* L3 with miRNA mimics and inhibitors

The effects of two miRNA inhibitors (Hco-miR-60-3p and Hco-miR-5885a-3p) and one miRNA mimic (Hco-miR-235-3p) were examined on the L3/L4 transition of *H. contortus* cultured *in vitro*. The miRIDIAN miR-235 mimic was acquired by submitting the mature miRNA sequence to Dharmacon. The inhibitors (Appendix, table 8-3) were designed following protocols described by Horwich and Zamore (2008). In brief, 3′ and 5′ flanking sequences consisting of five nucleotides each were added to the reverse complement sequence of the mature miRNA. Sequences were examined using mFold (http://mfold.rna.albany.edu/?q=mfold) for secondary structures that might affect binding before being synthesised (MWG Eurofins) as a 2'O-methyl oligonucleotide to enhance binding (Majlessi *et al.*, 1998). The first and last three bases were modified using phosphorthioate bonds to improve stability (Wan *et al.*, 2014). *H. contortus* L3 were exsheathed using CO_2 and incubated as described previously (Chapter 2.3.2). The EBSS medium was spiked with 1 µM mimic or 1 µM inhibitor. Daily observations and larvae counts were performed microscopically. Exsheathment, incubation and observations were carried out by Dr. C. Britton.

6.3 Results

6.3.1 Localising miRNA expression in *C. elegans*

In addition to temporal variation, miRNAs can display spatial variation between cells and tissues and so regulate target genes expressed there. In *C. elegans*, miRNA promoter-GFP reporter strains provide one means to examine both spatial and temporal expression. We focused on three miRNAs identified in Chapter 3 that showed significant changes in abundance between *H. contortus* L3 and L4: miR-60, miR-228 and miR-235. Hco-miR-228 and Hco-miR-235 are most abundant during L3, while Hco-miR-60 is most abundant during L4. The expression of these miRNAs was examined in *C. elegans* using strains VT1485 (*Cel-mir-228*::GFP), VT14988 (*Cel-mir-235*::GFP) and VT1733 (*Cel-mir-60*::GFP).

For the *mir-60* reporter (Fig. 6-1), GFP was expressed in gut cells (particularly in the first two anterior cells) at all stages from the late embryo onwards, suggesting a role in regulating genes involved in gut maintenance. This is consistent with the significant enrichment of miR-60 in gut tissue extracted from *H. contortus* (Chapter 3.3.1). Strikingly, GFP was also expressed in the nuclei of gut cells, suggesting that Cel-miR-60 is involved in non-canonical gene regulation (see Chapter 6.4). GFP under control of the *mir-228* promoter (Fig. 6-2) was widely expressed in nervous tissue, including the amphid/phasmid sensillia, labial/cephalic neurons and pre-anal/lumbar ganglia, as well as non-nervous tissue such as the vulva and seam cells. Expression was consistent in neurons of all stages from L1 to adult animals, with no obvious expression in eggs. This pattern of expression suggests that Cel-miR-228 may regulate genes involved in maintenance of hypodermal and neuronal tissue. The *mir-235* reporter strain (Fig. 6-3) displayed very low GFP expression across all stages. It was faintly detectable in two locations: vulval epithelium (during L4 stage) and rectal gland (L4 and adult stages).

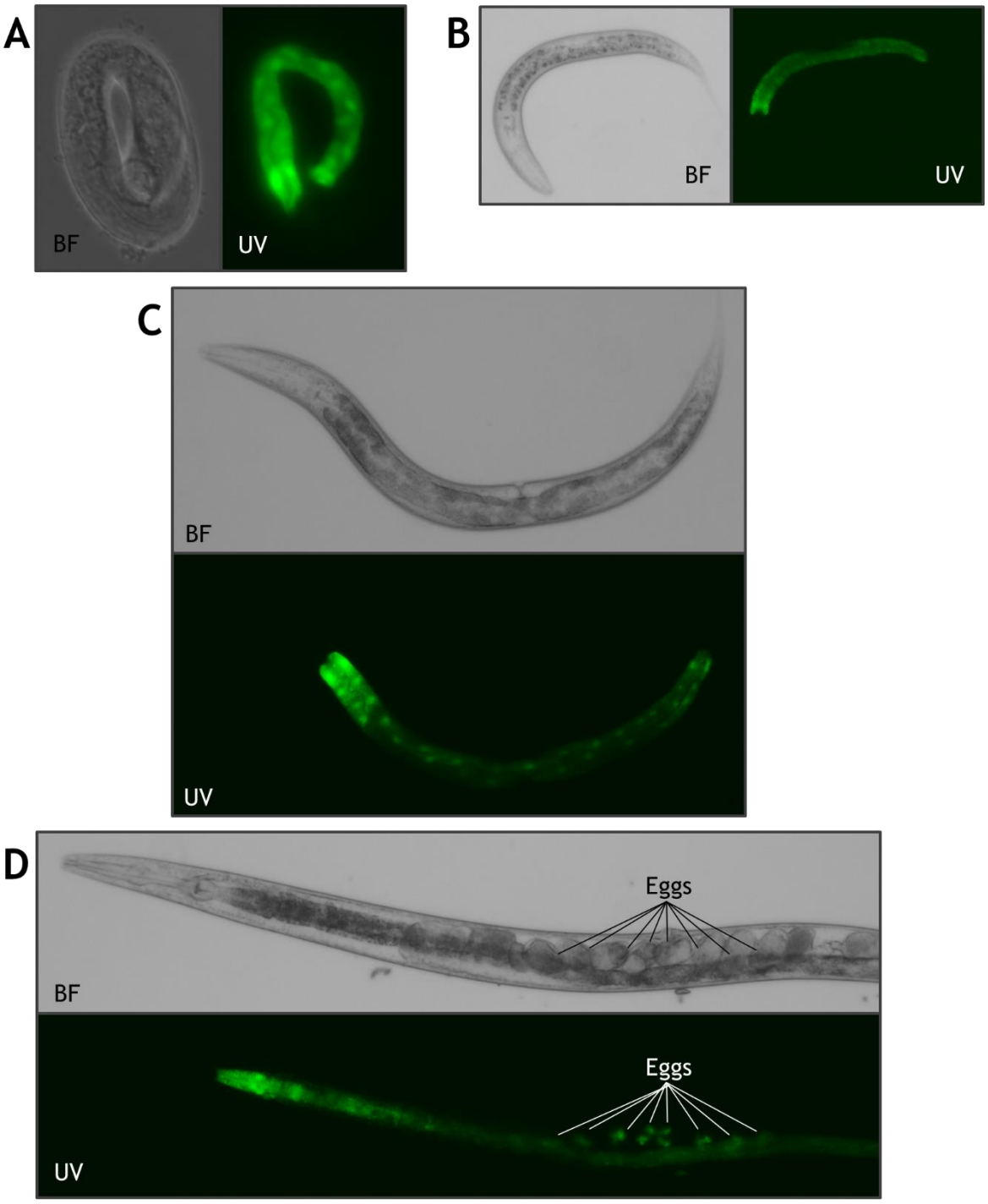


Fig. 6-1. Expression of *C. elegans mir-60*::GFP reporter in strain VT1733. Panels consist of a brightfield (BF) or GFP (UV) image. Magnification at 10x (B –D) or 40x (A). **A**: Egg. **B**: L3. **C**: L4. **D**: Adult hermaphrodite (anterior). GFP was found within the gut cells of all stages, with high levels present in gut nuclei.

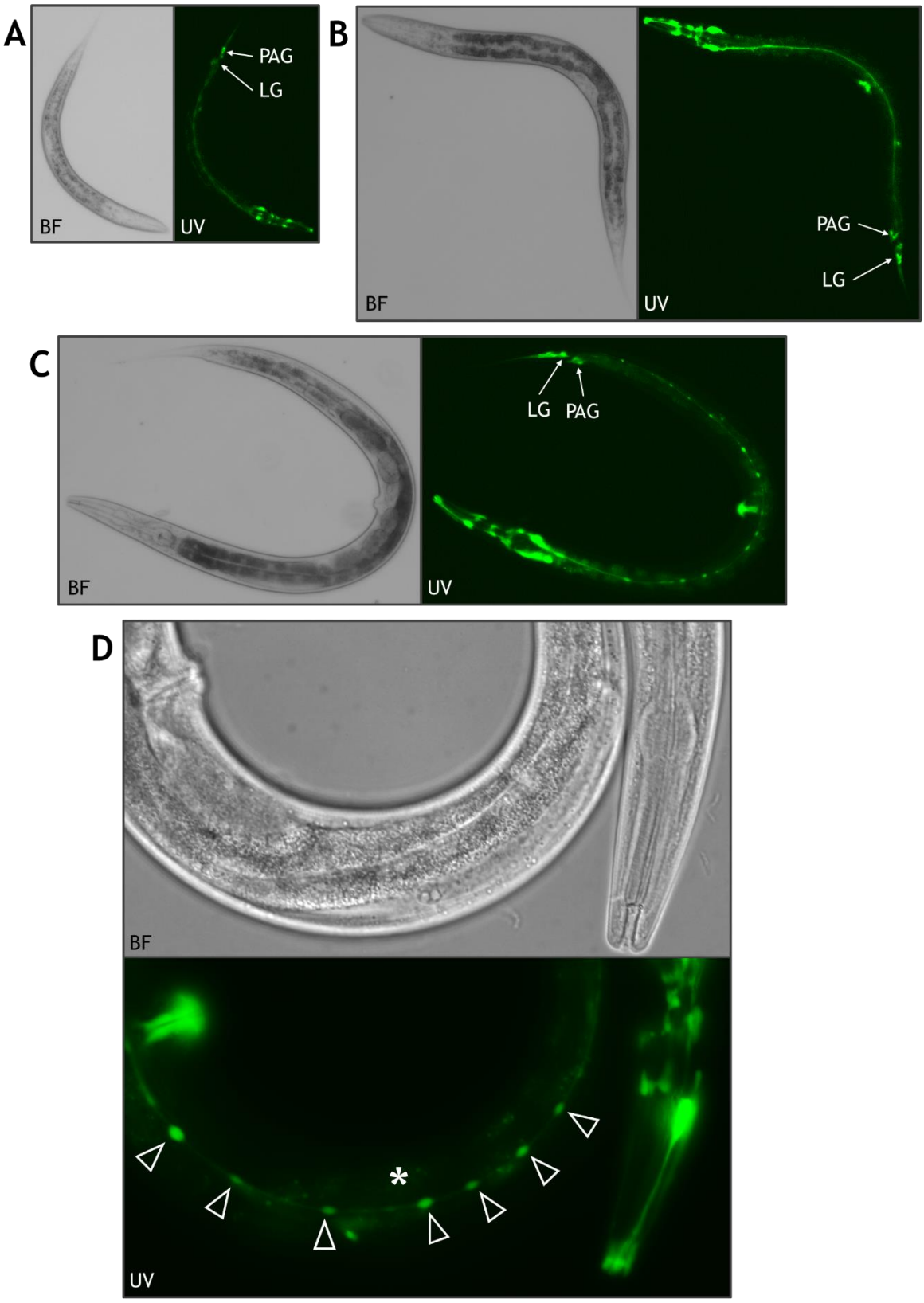


Fig. 6-2. Expression of *C. elegans mir-228*::GFP reporter in strain VT1485. Panels consist of a brightfield (BF) or GFP (UV) image. Magnification at 10x (A – C) or 40x (D). **A**: L3. **B**: L4. **C**: Adult hermaphrodite. **D**: Young adult. The mottled pattern in panel D (asterisk) was due to gut-granule auto-fluorescence. Lumbar and pre-anal ganglion (LG/PAG) indicated by arrows in panels A – C. Open arrowheads in panel D indicate seam cells. Gut granule auto-fluorescence can be seen in (asterisk).

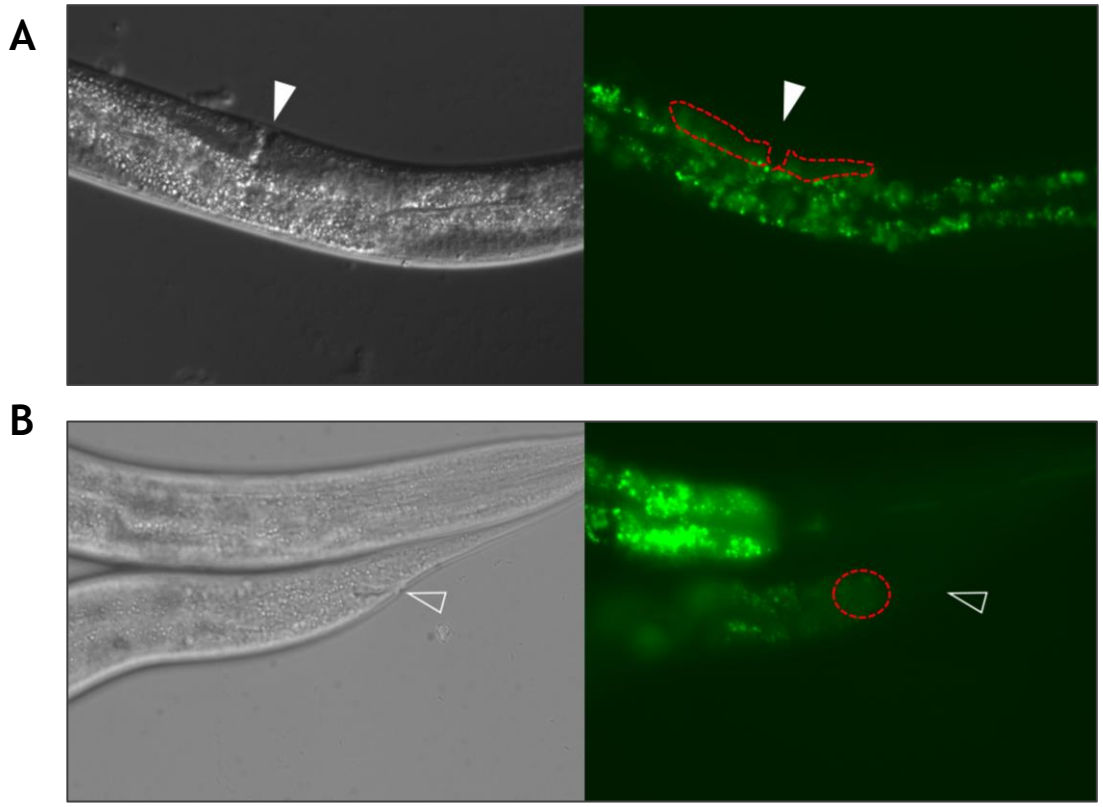


Fig. 6-3: Expression of *C. elegans mir-235*::GFP reporter in strain VT1488. 40x magnification. Most of the fluorescent signal arose from gut granule auto-fluorescence. Only areas demarcated by a red dotted line represent GFP expression; the remaining 'green' signal indicates non-GFP auto-fluorescence which appears different microscopically (yellow-green). Left panel is bright-field and right panel is UV fluorescence. **A**: Mid-body and vulva. Arrowhead indicates vulva. B: Anterior and posterior. Open arrowhead indicates anus.

6.3.2 The effect of environmental stress on miRNA expression

miRNAs can be regulated by both developmental and environmental factors (reviewed in Hudder and Novak, 2008). Altering miRNA expression during unfavourable conditions allows an organism to quickly respond to, and resist, environmental stresses. To determine if Cel-miR-60, Cel-miR-228 and Cel-miR-235 were environmentally regulated, the three GFP reporter strains used previously (Chapter 6.3.1) were subjected to four environmental stresses: low temperature (4 °C), high temperature (30 °C), UV radiation (UV-B, 100 mJ/cm^2) and osmotic stress (300 mM NaCl). Fluorescence was observed at 4 and 24 hours post-stress but showed no apparent difference in location or abundance when compared to worms grown under normal conditions (data not shown).

6.3.3 Dauer formation in *C. elegans*

As discussed previously, the dauer larva is an alternate L3 stage in *C. elegans* that many consider analogous to the infective stage of parasitic nematodes. To study the possible involvement of miRNAs in dauers, we examined dauer formation in mutant strains that lacked specific miRNAs or components of the RISC. Accurate observation of dauer formation among synchronised populations of different strains was most readily achievable using dauer pheromone. Pheromone is a complex mixture of ascarosides that act as a signal of organism density (Butcher *et al.*, 2007). As the composition changes depending upon environmental conditions and developmental stages (Kaplan *et al.*, 2011), it was necessary to characterise pheromone activity before comparing dauer formation.

Characterising the efficacy of dauer pheromone

Dauer pheromone was extracted from exhausted liquid cultures of wild type *C. elegans*. The response of worms to pheromone was examined and optimised under three primary conditions: batch variation, temperature and food signal.

Variation between batches was examined using increasing concentrations of pheromone (Fig. 6-4A). Temperature (27 °C) and food signal (10 µL of HT-OP50) were kept constant. All batches were able to stimulate at least 80 % dauer formation with 2.5 % v/v pheromone. The stability of ascarosides is demonstrated by batch G, which had been stored at -20 °C for 12 years but was still effective in promoting dauer formation.

Temperature is an important regulator of dauer development. Based on previous studies by Golden and Riddle (1982, 1984), we focused on the range of 25 – 27 °C (Fig. 6-4B). Below this temperature, dauer formation was inefficient, requiring restrictively high concentrations of pheromone, while at higher temperatures dauers formed even in the

absence of pheromone (data not shown). Food signal was kept constant (20 µL), while the concentration of pheromone ranged from 0 – 3.75 %. There was a large jump in dauer formation between 25 and 26 °C (15 and 72 % respectively), while most dauers were generated at 27 °C (92 %). As observed in the batch studies previously, dauer formation decreased above 2.5 % pheromone, but only at 27 °C, suggesting that a combination of high temperature and pheromone concentration are antagonistic towards dauer formation.

The inhibitory effects of food on dauer formation were examined using increasing concentrations of HT-OP50 (Fig. 6-4C). Temperature (27 °C) and pheromone (2.5 %) were kept constant. With no food available (0 µL HT-OP50), few adults remained on the plate to lay eggs and many of the resulting larvae entered L1 arrest. 10 µL of HT-OP50 was sufficient to keep adults on the plate and prevent newly hatched worms from entering L1 arrest. As expected, increasing the abundance of food lead to decreasing levels of dauer formation. At 50 – 80 µL HT-OP50, the decrease in dauer formation levelled off, suggesting that the 'anti-dauer' food signal could no longer overcome the 'pro-dauer' pheromone signal.

From these studies, an optimised method was developed to maximise dauer formation: a concentration of pheromone (1 – 2.5 %) was calculated per batch to generate 85 – 90 % dauers in wild type *C. elegans* when incubated at 27 °C in the presence of 10 µL HT-OP50.

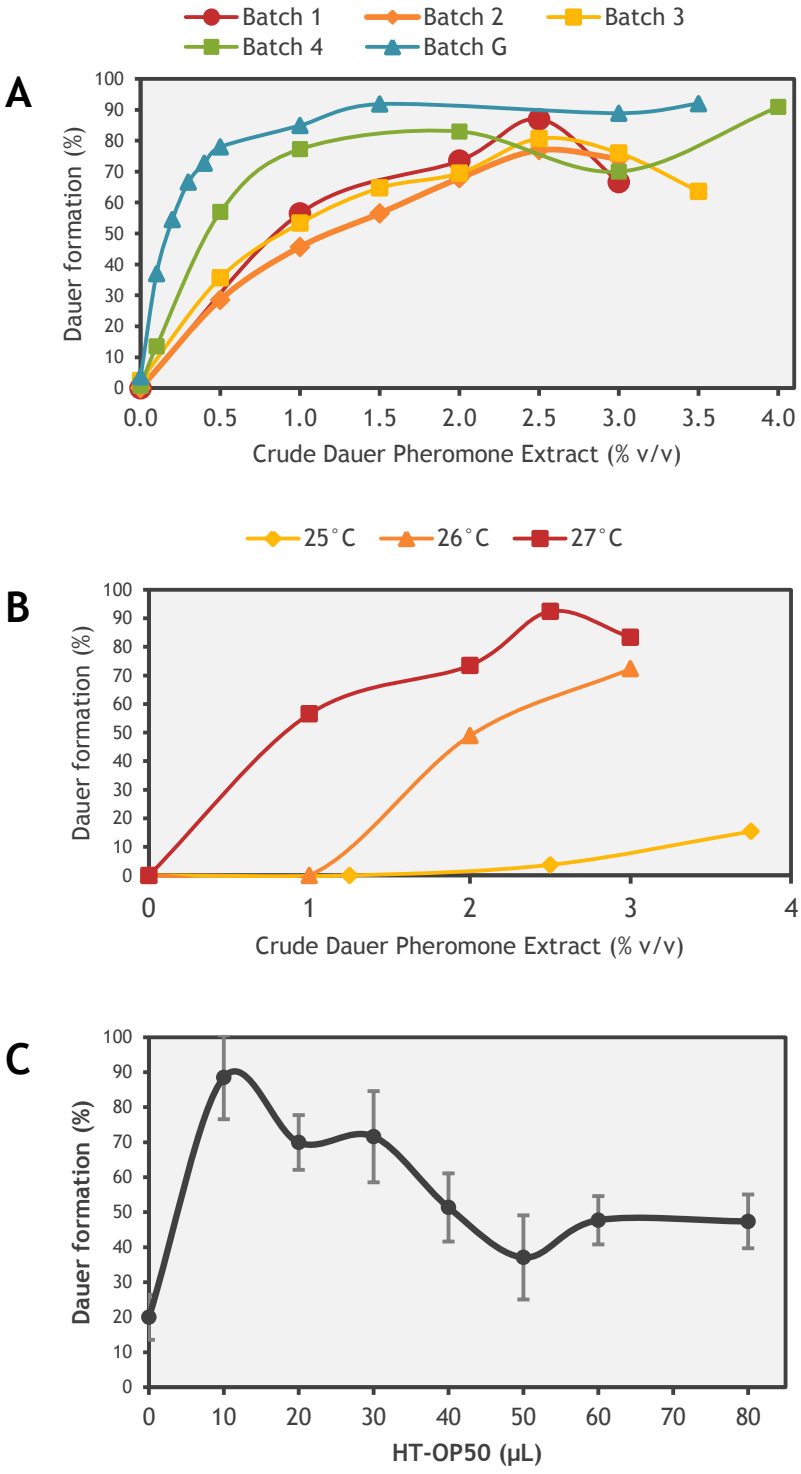


Fig. 6-4: Characterising the efficacy of *C. elegans* dauer pheromone on wild type N2. **A**: Variation in dauer formation between pheromone batches. **B**: Effect of temperature on dauer formation. **C**: Effect of food on dauer formation. Heat-treated *E. coli* OP50 (HT-OP50) used as a food source. Panel C was performed in triplicate with vertical bars representing ± 1 standard deviation.

Dauer formation in *C. elegans* miRNA mutant strains

Using this optimised method, dauer formation was examined in *C. elegans* mutants (Fig. 6-5). Ten mutant strains were examined (detailed in Table 6-1), two of which (RF54 and MH2385) lacked critical components of the RISC (*alg-1* and *ain-1* respectively), while the remaining eight strains lacked at least one miRNA. The double mutant strain CLB061 was a cross of MT14446 (*mir-228*-/-) and MT17997 (*mir-235*-/-) generated by Dr. C. Britton.

The two RISC mutants displayed significantly decreased dauer production with no more than 20 % of larvae becoming dauers. As Cel-ALG-1 is miRNA-specific, the reduction suggests an essential role for miRNAs in the regulation of dauer formation. Furthermore, mutant strains lacking different miRNAs displayed reductions in dauer formation suggesting that different miRNAs are involved in regulating dauer formation to a greater or lesser degree. The three miRNA mutant strains (MT14446, MT16471 and MT17997) that lacked a single miRNA (*mir-228*, *mir-60* and *mir-235* respectively) displayed slightly lower levels of dauer formation when compared to the wild type, although the difference was statistically insignificant (ANOVA, corrected $p < 0.05$). All five mutant strains lacking more than one miRNA displayed significantly lower levels of dauer formation. The effect was greater for strains MT17341 and MT17676 (lacking six and four miRNAs respectively), displaying < 10 % dauer formation. Common to both strains was the lack of *mir-61*, *mir-247* and *mir-250*. The similar level of reduction they showed in dauer formation suggests that these three miRNAs are mostly responsible.

To determine whether the response to pheromone was linear in mutant strains and if reductions in dauer formation could be overcome with high concentrations of pheromone, dauer formation was examined in four strains (CBL061, MT14446, MT17997 and RF54) using a range of pheromone concentrations (Fig. 6-6). In general, increasing the concentration of pheromone led to a higher percentage of dauer formation, with the largest

increase seen between 0 and 1 % pheromone. Increasing the concentration further did not rescue the dauer formation defect in the mutant strains tested. This suggests that the reduction in dauer formation observed is not a result of reduced sensitivity to pheromone but a defect in activation of the dauer response.

Table 6-1: *C. elegans* mutant strains examined in dauer formation assays.

Strain	Genotype	Description
N2 (Bristol)	WT$^+$	Wild type
MH2385	*ain-1*(ku322) X	RISC mutant
RF54	*alg-1*(gk214) X	RISC mutant
MT14446	*mir-228*(n4382) IV	miR-228 mutant
MT15981	*mir-87*(n4104) V; mir-233(n4761) X	miR-87/233 mutant
MT16471	*mir-60*(n4947) II	miR-60 mutant
MT17428	*mir-72*(n4130) II; nDf47 X	miR-72/73/74 mutant
MT17431	nDf49 II; nDf59 V; *mir-247*(n4505) X	miR-42/43/44/61/247/250 mutant
MT17676	*mir-45*(n4280) II; nDf59 V; *mir-247*(n4505) X	miR-45/61/247/250 mutant
MT17997	*mir-235*(n4504) I	miR-235 mutant
CBL061	*mir-228*(n4382) IV; *mir-235*(n4504) I	miR-228/235 mutant

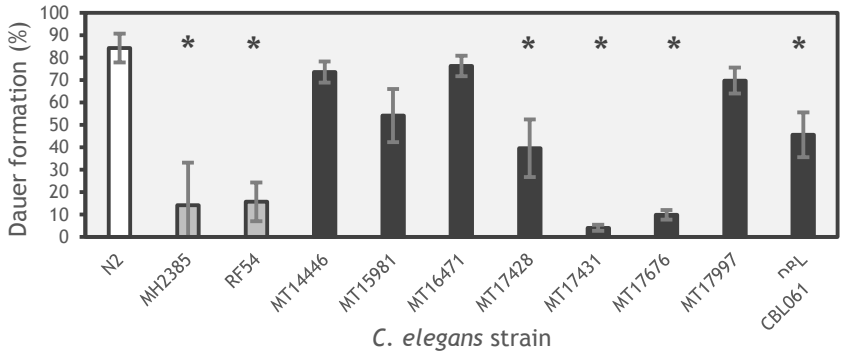


Fig. 6-5: Dauer formation in *C. elegans* miRNA and RISC mutants. Generated using 2 % pheromone (batches two and four). Strains are coded by type of mutation: white = wild type, grey = RISC mutant, black = miRNA mutant. Vertical lines represent ± 1 standard deviation. Asterisks highlight strains that displayed significantly lower dauer formation compared to the wild type (ANOVA, corrected $p < 0.05$).

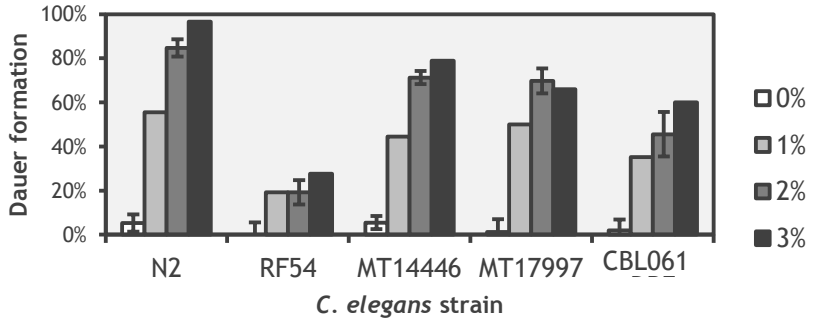


Fig. 6-6: Effect of increasing pheromone on dauer formation in *C. elegans* mutants. Eggs were laid on plates containing 0 – 3 % dauer pheromone (batch 3). N2 = Wild type. RF54 = *alg-1*(-). MT14446 = *mir-228*(-). MT17997 = *mir-235*(-). CBL061 = *mir-228*(-); *mir-235*(-). Experiments using 0 and 2 % pheromone were performed in triplicate, while 1 and 3 % pheromone experiments were performed in duplicate. Vertical lines represent ± 1 standard deviation from the mean.

6.3.4 Dye filling phenotypes of miRNA mutants

Some dauer-defective strains of *C. elegans* have altered amphid structures that restrict dye-filling and inhibit chemosensation (Hudder and Novak, 2008). To exclude pheromone detection as an explanation of the reduced dauer formation seen in miRNA mutants, dye-filling of sensory neurons was examined (Fig. 6-7). Dye filling of amphid and phasmid neurons was assessed in wild type *C. elegans* (Fig. 6-7) and five mutant strains (CBL061, RF54, MH2385, MT14446 and MT17997). The mutant strains displayed the same pattern and level of dye filling as wild type worms, indicating that the reduction in dauer formation observed in these strains (particularly the *alg-1*(-) and *mir-228/235*(-) mutant strains) was not due to an inaccessibility of the sensory neurons. The same strains were also stained using FITC, but no differences were detected between wild type and mutant worms (data not shown).

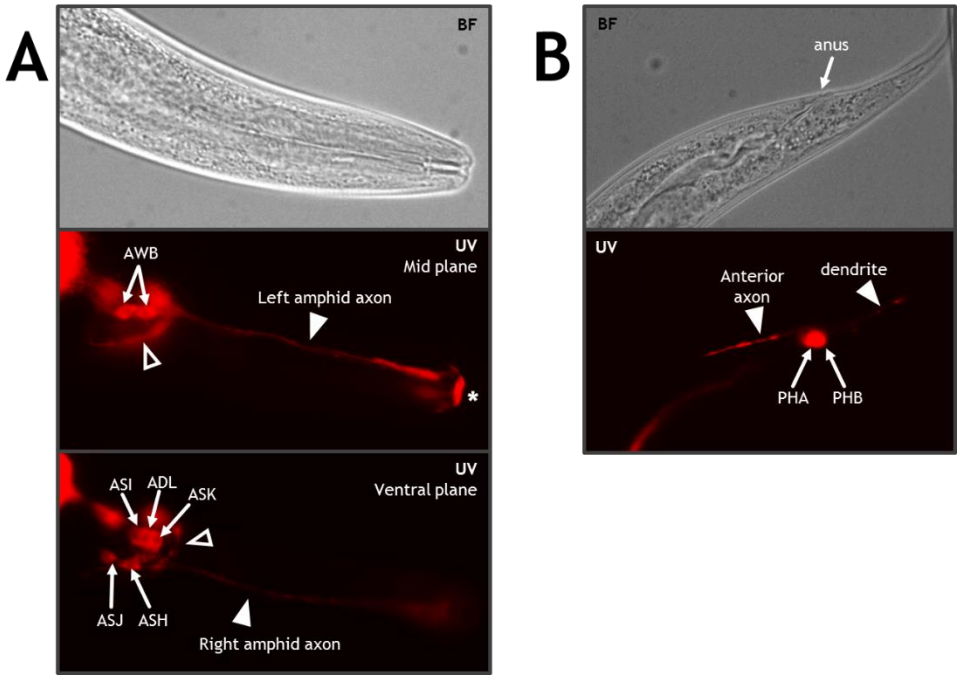


Fig. 6-7: Neurons of *C. elegans* N2 stained with DiI. Young adult stage observed at 40 x magnification under brightfield (BF) or Red (UV) filters. Images exposed for 5 (BF) or 50 (UV) ms. **A**: Dorsoventral-oblique view of amphids. **B**: Lateral-oblique view of phasmids. Arrows indicate neuron cell bodies. Arrowheads indicate axons/dendrites. Open arrowhead indicates nerve ring. Asterisk indicates amphid sensillium.

6.3.5 Effect of *C. elegans* dauer pheromone on *H. contortus* L3

As the pheromone extract was a strong promoter of dauer formation in *C. elegans*, we questioned whether it might also have an effect on *H. contortus*. We investigated whether pheromone was able to maintain the L3 state and inhibit developmental progression to L4. CO_2 exsheathed *H. contortus* L3 were exposed to filtered dauer pheromone (0 – 6 % v/v) and incubated for 72 hours before the number of sheathed L3, exsheathed L3 and L4 were counted. Even with high concentrations of pheromone, there was no significant difference in the number of worms that developed to L4 (Fig. 6-8).

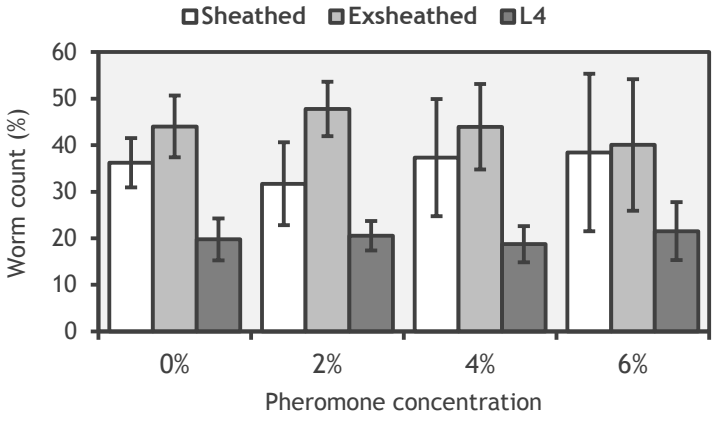


Fig. 6-8: *In vitro* culture of *H. contortus* L3 exposed to *C. elegans* dauer pheromone. L3 were exsheathed using CO_2 and cultured as previously described (Chapter 2.3.2) except for the inclusion of 0 – 6 % v/v *C. elegans* crude dauer pheromone extract. The number of worms in each stage (sheathed L3, exsheathed L3 and L4) was observed after 72 hours. Experiment performed in triplicate with vertical lines representing ± 1 standard deviation from the mean.

6.3.6 Recovery of dauers in *C. elegans* miRNA mutant strains

For the dauer state to be useful, worms must be able to resume a reproductive life-cycle. To determine if miRNAs had any influence on recovery, dauers of wild type and miRNA mutant strains (CBL061, MT14446 and MT17997) were exposed to increasing concentrations of yeast extract, used as a food signal. Yeast extract was used in place of

bacterial cultures to improve precision and repeatability. 24 hours after being removed from pheromone plates, dauer and post-dauer worms were counted (Fig. 6-9).

With no food signal, less than 2 % of worms exited the dauer state, suggesting that there was some instability in maintaining dauers in the absence of pheromone. As the abundance of food increased, more worms resumed development in the N2 and mutant strains. While a small number of worms resumed development at 0.01 % yeast extract, the biggest change was seen at 0.025 % yeast extract, with the majority of worms exiting the dauer state. At this concentration, there was also a noticeable difference in recovery between the different strains. All three mutant strains displayed a higher level of recovery (64 – 70 %) than the wild type worms (51 %). At higher concentrations of yeast extract (0.05 – 0.1 %), there were no significant differences in recovery between the different strains. These data suggest that lack of miR-228 and miR-235 subtly increased activation required to resume development. Due to limitations in the availability of pheromone, this study was performed once, but could be expanded further.

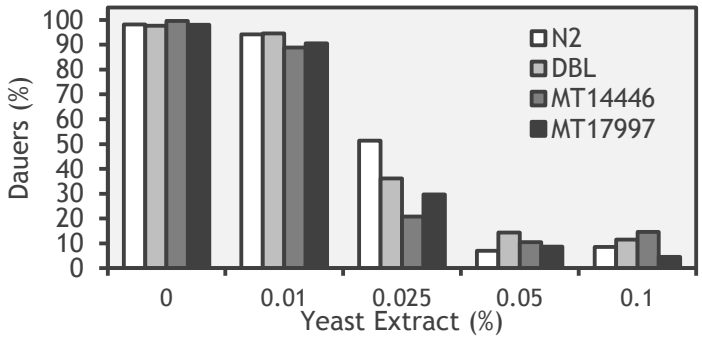


Fig. 6-9: Recovery of *C. elegans* wild type and miRNA mutant strains from the dauer state. Worms were removed from pheromone plates and exposed to increasing concentrations of yeast extract. Recovery was examined in wild type worms (N2) and three mutant strains: MT14446 (miR-228(-)), MT17997 (miR-235(-)) and DBL (miR-228/235(-)).

6.3.8 The effect of miRNA mimics and inhibitors on *H. contortus*

Uptake of fluorescent small RNAs by *H. contortus* larvae

While the use of deletion mutants to examine miRNA function is productive in *C. elegans*, this approach is not yet possible in parasitic nematodes. However, examination can be performed in parasites using miRNA mimics and inhibitors. Prior to testing in *H. contortus*, it was important to confirm that exogenous RNA could enter the organism. *H. contortus* L3 were exsheathed using bleach and incubated for four days in the presence of short, fluorescently labelled, double-stranded RNA. Using UV microscopy, labelled RNA was visible within the gut tissue of L4 worms (Fig. 6-10). Approximately 95 % of L4 worms showed uptake of fluorescent RNA (data not shown). In contrast, fluorescence was not seen in any L3, suggesting that RNA cannot penetrate the cuticle or sensory organs and must be ingested. Additionally, fluorescence was not visible in any other L3 tissues (*e.g.* pharynx or body wall), suggesting that RNA was only taken up the gut and remained in this tissue.

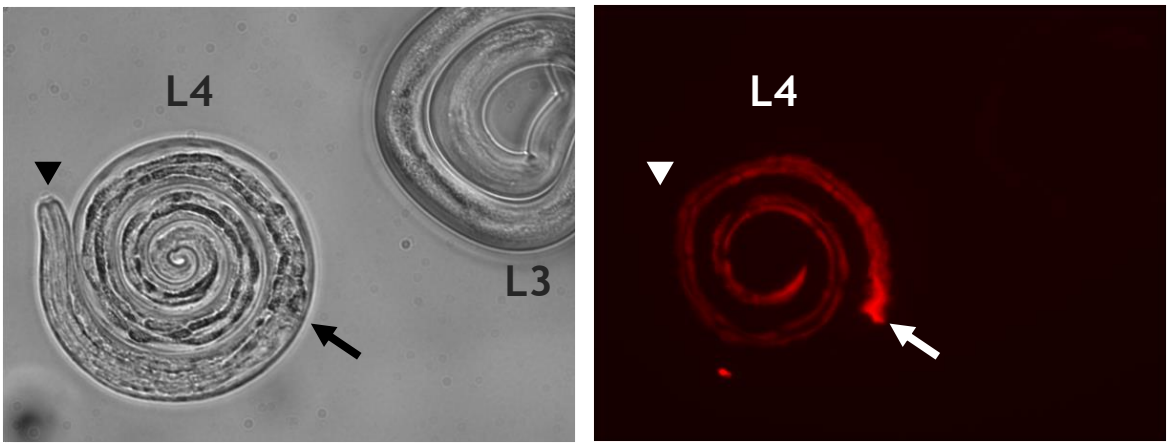


Fig. 6-10: Uptake of fluorescently labelled short RNA by *H. contortus* L4. L3 were exsheathed using bleach and incubated for four days with 0.5 µM Cy3 labelled short RNA. 10 x magnification with a 10 ms exposure time. The first panel is a bright field image, while the second panel was viewed using a Texas red fluorescent filter. Fluorescent signal is detectable in the gut tissue of L4 but not L3. The arrow indicates the pharyngeal/gut junction, while the arrowhead indicates the mouth.

Incubating *H. contortus* L3 with miRNA mimics and inhibitors

Having demonstrated uptake of labelled RNA by *H. contortus* L4, we examined the effects of mimics and inhibitors in this stage when cultured *in vitro*. We focused on three miRNAs that showed significant changes in expression between L3 and L4: Hco-miR-60-3p, Hco-miR-235-3p and Hco-miR-5885a-3p. The abundance of miR-60 and miR-5885 increases dramatically between L3 and L4 and we therefore tested miRNA inhibitors to artificially reduce natural levels. In contrast, miR-235 shows a marked decrease in expression from L3 to L4, so miRNA mimics were used to increase normal levels. L3 were exsheathed using CO_2 and incubated with each mimic or inhibitor individually and in combination. Negative controls consisted of a scrambled mimic or inhibitor using sequences not found in the *H. contortus* genome, which should therefore have no effects on miRNA abundance. Larvae were observed over the course of five days, but showed no difference in size, movement or transition to L4 when compared to negative controls (data not shown). Studies are ongoing to examine effects in L4 co-cultured with Caco2 gut epithelial cells that enhance L4 development and survival *in vitro* (Dr. C. Britton, unpublished data).

6.4 Discussion

To begin exploring the biological roles of miRNAs in nematodes, a range of techniques available in the model organism *C. elegans* were utilised. These included GFP reporter strains to observe spatial expression of miRNAs and any changes induced during stress response. In addition, a role for miRNAs in dauer formation and recovery was examined using *C. elegans* mutants. Finally the function of selected miRNAs in *H. contortus* development was examined using mimics and inhibitors.

C. elegans miRNA::GFP reporter strains were used to observe spatial variation in miRNA expression. Three miRNAs (miR-60, miR-228 and miR-235) were speculated to be involved in regulating development due to their temporal expression pattern (see Chapter 4.3.1). For Cel-miR-60, GFP expression suggests it is involved in regulating gut-specific genes. Expression did not vary between life cycle stages and was detected in embryos which had already undergone morphogenesis and gut formation. This pattern suggests a role in gut maintenance and function rather than gut development. qRT-PCR data for Cel-miR-60 (Kato *et al.*, 2009; Appendix, Fig. 8-1) showed a peak in expression in male worms (which were not examined using GFP reporters), perhaps indicating an additional role in sex-specific gene regulation. There was also greater expression in posterior gut cells than in anterior cells. Such patterns can be seen when promoters are incomplete (Ruvinsky and Ruvkun, 2003) or expression is linked to gut cell endoreduplication (reviewed in Hedgecock and White, 1985). Intriguingly, GFP was seemingly present in the nuclei of gut cells. However, without nuclear co-staining, its exact position could not be confirmed. If miR-60 was indeed present in the nuclei of gut cells, this suggests that miR-60 is involved in non-canonical regulation. While post-transcriptional regulation of mRNAs in the cytoplasm (canonical regulation) is the most well studied aspect of miRNA biology, numerous miRNAs are known to regulate gene expression from within the nucleus (reviewed in Roberts, 2014). Non-canonical miRNA functions are not fully understood, but cover a wide range of interactions including regulation of nuclear mRNAs using intronic binding sites (Chi *et al.*, 2009), regulation of lncRNAs (Morris *et al.*, 2008), transcriptional gene silencing (Kim *et al.*, 2008) and transcriptional gene activation (Morris *et al.*, 2008). On the basis of the data presented here, it is not possible to determine the function of Cel-miR-60, but it is likely to be involved in the regulation of gut-specific processes at both a nuclear and cytoplasmic level.

GFP expression of the *mir-228* reporter was found to be abundant in nervous tissue, including the amphid/phasmid, cephalic/labial neurons and pre-anal/lumbar ganglia. GFP was also detected in the seam cells and in the cells that form the vulval lumen. Fluorescence was consistent between life cycles stages although qRT-PCR data (Kato *et al.*, 2009; Appendix Fig. 8-1) showed that expression decreased during development. A detailed morphological analysis by Pierce *et al.* (2008) of a different *mir-228*::GFP reporter identified fluorescence arising from support cells (*i.e.* sheath/socket cells) rather than ciliated sensory neurons. Their study also identified Cel-miR-228 as an orthologue of the miR-183 family, which is widely conserved in vertebrates and found in sensory neurons (Kloosterman *et al.*, 2006).

In contrast to the *mir-60* and *mir-228* reporters, expression of the *mir-235*::GFP reporter was very low in *C. elegans* larvae and adult hermaphrodites. Fluorescence was faintly detectable in two regions: the vulval epithelium and rectal gland. Low level expression is also consistent with qRT-PCR data (Kato et al., 2009; Appendix, Fig. 8-1). A greater understanding of the role of miR-235 was provided by Kasuga *et al.* (2013) who found that expression of a miR-235::GFP reporter was limited to neuron support cells during L1 diapause. Furthermore, by examination of a miR-235 null mutant, Kasuga *et al.* showed that miR-235, through targeting *nhr-91*, is required for blast cell quescience during L1 diapause (Kasuga *et al.* 2013). While the L1 diapause was not examined in this study, GFP expression in dauers showed no difference in localisation or intensity when compared to non-quiescent stages. From microarray data, the expression profile of miR-235 in *H. contortus* is somewhat different to that of Cel-miR-235. Both organisms display high expression in adult males compared to other stages, however in *H. contortus* expression also peaks during L3. As there is no equivalent to L1 arrest in *H. contortus*, Hco-miR-235 might instead be involved in arresting development during L3.

The three miRNA::GFP reporter strains were also exposed to environmental stresses, consisting of high and low temperatures, high salinity and UV radiation, but displayed no apparent differences in GFP expression to reporter strains grown in the absence of stress. Under the limited conditions tested, it appears that miR-60, miR-228 and miR-235 are not environmentally regulated. However, it is still possible they are affected by alternative stresses, as observed in the response of miR-235 to a lack of food during L1 arrest. In addition, Kudlow *et al.* (2012) showed an increase in some gut miRNAs in response to infection with the bacterial pathogen *Pseudomonas aeruginosa*. While miR-60 was not reportedly increased in that study, it may respond to other infectious agents or chemical insults.

While examining spatial expression can give an indication of possible miRNA function, we wished to investigate a possible role of miRNAs in development, specifically dauer diapause. Understanding how miRNAs may regulate dauer development is relevant to understanding developmental arrest in parasitic nematodes. After characterising the response of wild type worms to pheromone, mutant strains were examined for their ability to generate dauers. The ability of mutants lacking components of the RISC (*i.e. alg-1* or *ain-1*) form dauers was poor. As these proteins, particularly ALG-1, are more closely involved with miRNA-based gene regulation (rather than other non-coding RNAs), these data suggest that miRNAs are very important to formation of dauers in *C. elegans*. The miRISC displays a certain degree of redundancy (*e.g. alg-2* can sometimes replace *alg-1*), which may explain why a small number of worms were able to become dauers. Interestingly, the *alg-1* and *ain-1* mutant strains showed a similar reduction in dauer formation, indicating that both proteins are required. The loss of individual miRNAs had only a negligible effect on dauer formation, reinforcing the idea of redundancy within the miRNA regulatory framework. It is possible that the reduction in dauer formation was not

due to a direct interaction in the dauer pathway, but a consequence of reduced fitness. The concept of over-lapping function became more significant when mutants lacking multiple miRNAs were examined. Two strains displayed very low levels of dauer formation equivalent to that seen in the RISC mutants. The strains lacked multiple miRNAs, but common to both was the absence of *mir-61*, *mir-247* and *mir-250*. miR-61 is known to regulate LIN-12/NOTCH signalling, which is important for cell-cell interactions during development and particularly important for vulval development (Yoo and Greenwald, 2005). Our microarray data also showed that Hco-miR-61 is upregulated in adult female worms, suggesting its function may be conserved in parasitic nematodes. Hco-miR-250 also displays a peak in expression in adult female worms. miR-247, which was not identified in *H. contortus*, is closely related to miR-61 but does not share the same phenotypic effects on vulval development when over- or under-expressed (Miska *et al.*, 2007). Northern blots show that miR-247 is only expressed in *C. elegans* L3 and dauers (Zou *et al.*, 2013). Exploring the role of mutants lacking individual miRNAs from this family might yield further information on their role in dauer formation.

The roles of *C. elegans* miR-228 and miR-235 in dauer formation were studied singly and in combination. DiI and FITC staining of all three strains and the *alg-1*(-) mutant showed that they possessed a normal dye filling phenotype, indicating that the dauer formation defect was not a result of pheromone being unable to physically enter the sensory neurons. The defect in dauer formation was therefore likely due to downstream signalling events in the dauer decision pathway.

The effect of *C. elegans* dauer pheromone was also studied directly in *H. contortus*. Exsheathed L3 were incubated with pheromone and the number of L3 and developing L4 examined. Even in the presence of high concentrations of pheromone, no inhibition of development was observed. While of interest to examine, this is consistent with the finding of strong divergence in the components of the dauer pathway between different

species (Elling *et al.*, 2007). In addition, even between worm populations of the same species there are differences in both the composition of dauer pheromone and the response of each population to different ascaroside compounds (Diaz *et al.*, 2014). It is speculated that this "dishonest signalling" promotes intra-species competition when resources are limited and supporting worms of a similar genotype.

As the *C. elegans* miRNA mutants displayed a striking reduction in dauer formation, it was also of interest to examine resumption of development after dauer diapause. A subtle difference was detected between wild type and mutant strains. With the same concentration of food, more mutant worms resumed development than did wild type worms, suggesting an increased activation in response to food or increased sensitivity to food. The double *cel-mir-228/235*(-) mutant did not display a stronger phenotype than the single miRNA mutants. This effect could be explored in greater detail in additional experiments by using a greater range of food signal concentrations (providing higher resolution data between 0.01 and 0.025 %) and additional strains. The *alg-1*(-) mutant strain was not examined in the current study due to the low number of dauers formed.

To see if miRNAs were involved in developmental arrest in *H. contortus*, the effects of miRNA mimics and inhibitors was examined. Two of the miRNAs studied (Hco-miR-60 and Hco-miR-5885) displayed strong abundance in gut tissue, while Hco-miR-235 displayed an L4-specific nadir in abundance. Unfortunately, even in combination, the addition of mimics and inhibitors to these miRNAs had no observable effect on development in L3 cultured *in vitro*. However, the lack of an effect may have been due to imperfect culture conditions. Although L3 cultured *in vitro* displayed significant morphological changes (particularly in gut tissue) after being exsheathed, they did not

develop into adults and instead died within 7 – 10 days. In addition, qRT-PCR comparison of miRNA abundance in freshly exsheathed L3 and L3/L4 cultured for four days showed an increase in expression of miR-60 and miR-5885 (data not shown), matching data from L4 grown *in vivo*. Although only performed once, it suggests that culture conditions can initiate at least part of the L3/L4 transition, but additional signals are required to complete L4 development. However, the absence of an effect may also be due to the specific miRNAs examined. The miRNA mimics/inhibitors may be causing a subtle effect not observable microscopically or redundancy in the regulatory network may blunt their effect as is observed for *C. elegans* miRNA mutants (Miska *et al.*, 2007).

The combination of functional studies presented here highlighted a number of important roles for miRNAs in both *C. elegans* and *H. contortus*. By utilising additional techniques, our knowledge of the biological impact of miRNAs in nematode development will increase further.

7 Final discussion

H. contortus poses major economic and welfare concerns in small ruminant production throughout the world. Understanding the fundamental processes that control growth and development of *H. contortus* may provide targets for new drugs that are required to combat anthelmintic resistance. miRNAs are now recognised as an integral component of developmental regulation in many organisms (reviewed in Ivey and Srivastava, 2015) and the data presented here are the first indications that miRNAs may also regulate development in *H. contortus*.

Winter *et al.* (2012) first identified miRNAs in *H. contortus* and subsequent microarray analysis quantified their abundance through development (Chapter 3). 55 miRNAs showed significant temporal variation, of which 35 also displayed significant spatial variation in the gut tissue of adult female worms. These abundance patterns suggests a role for *H. contortus* miRNAs in stage and tissue-specific regulation of genes, which may include genes involved in growth or development that require tightly regulated expression. *In silico* identification of potential miRNA targets found that 93 % of the 21,799 genes currently annotated in *H. contortus* contained at least one predicted miRNA binding site (Chapter 4). This is likely to be higher than the number of genes actively targeted by miRNAs; for comparison, 30 – 60 % of humans mRNAs are predicted to be regulated by miRNAs (Friedman *et al.*, 2009; Yu *et al.*, 2007). Our understanding of the mechanisms that govern miRNA-mRNA interactions is currently incomplete (reviewed in Pasquinelli, 2012 and Valinezhad Orang *et al.*, 2014), which limits the accuracy of target prediction algorithms (Peterson *et al.*, 2014; Witkos *et al.*, 2011). Predictions must therefore be viewed as a useful screening tool for miRNA-mRNA interactions when combined with additional sources of evidence.

One such source includes transcriptomic data (Chapter 4), which can be applied to identify inverse patterns of abundance between miRNAs and mRNAs. While this approach is complicated by a number of factors (*e.g.* regulation by multiple miRNAs, positive miRNA regulation and regulation via transcription factors or epigenetic modification), it can identify those genes most likely to be under strong negative miRNA regulation. However, definitive confirmation of a miRNA-mRNA interaction requires validation, which is a rapidly expanding field of miRNA research. HITS-CLIP is one of the few techniques that can be applied to non-model species, such as *H. contortus*, without the need for genetic modification or long-term cultures. Expression constructs for fragments of *H. contortus* ALG-1 were generated during this study (Chapter 5) and although the antibodies generated were unsuitable for immunoprecipitation, efforts are ongoing to develop a new antibody. HITS-CLIP has the potential to identify the majority of miRNA-mRNA interactions in an organism and therefore validate target prediction data.

While large-scale analysis of miRNAs, target predictions and gene expression can provide a wealth of information, there are limitations to the depth of knowledge that can be acquired. In response, a small number of miRNAs were examined in greater detail: Hco-miR-60-3p, Hco-miR-228-5p, Hco-miR-235-3p and Hco-miR-5885. All four displayed significant variation between L3 and L4 stages, which is an important transition between free-living and parasitic stages of *H. contortus* and an ideal point of intervention for potential therapeutic agents. If L3 can be prevented from developing into L4, pathology (which arises from the feeding of L4 and adult stages) would be minimised and the life cycle interrupted. Control of earlier pre-parasitic stages (*i.e.* eggs, L1 and L2) would be impractical, requiring treatment of the environment. An important aspect of drug design is to minimise adverse effects on the host. Searches of the sheep genome (Jiang *et al.*, 2014)

failed to identify these four miRNAs, therefore, activity of these nematode-specific miRNAs is unlikely to have a detrimental effect on host gene expression.

In addition to potential therapeutic agents, miRNAs might also play a role as biomarkers for infection. Specifically, the blood of a host could be examined for the presence of nematode, or preferably, parasite-specific miRNAs, providing an intriguing diagnostic tool, particularly if miRNA abundance is linked to parasite burden. This link was not detected between parasite-derived miRNAs in the sera of dogs infected with *Dirofilaria immitis* (Hoy *et al.*, 2014b), but may exist in other situations. *H. contortus* specific miRNAs have been identified in the abomasal mucosa and lymph nodes, but not sera, of infected sheep (H. Gu, unpublished data). The exact role of these miRNAs in the host is currently being investigated.

Currently, miR-60 has only been identified in *H. contortus* and several species of *Caenorhabditis*. In *C. elegans*, miR-60 is exclusively expressed in gut tissue at consistent levels throughout development. It is also thought to be involved in regulating life span through germline-based repression (Lucanic *et al.*, 2013). The gut-specific nature of miR-60 seems to be conserved in *H. contortus*, as microarray analysis showed high levels of abundance in the gut tissue of adult female worms. There was also a significant peak in abundance in L4, with low abundance during L3. This increase may reflect the changes in gut activity between non-feeding (L3) and feeding (L4) stages, suggesting that Hco-miR-60 could be required to initiate growth and resume development of gut tissue after inactivity. This is supported by gene ontology of predicted Hco-miR-60 targets that found enrichment of gene annotations covering ‘development’ and ‘metabolism’. In addition, abundance of Cel-miR-60 is also significantly lower in the (non-feeding) dauer than in any other larval stage (Karp *et al.*, 2011), supporting a role for miR-60 in active gut tissue.

Interestingly, gene ontology of predicted Cel-miR-60 targets identified enrichment of more specific annotation terms, such as ‘localisation’, ‘transport’ and ‘carbohydrate binding’, regulation of which is likely to be important to gut function. In addition it is suggested that Cel-miR-60 may modulate pathogen responses within the gut (Kudlow *et al.*, 2012).

While miR-228 has only been identified in nematodes (*e.g. Ascaris suum, Caenorhabditis spp*. and *H. contortus*), it is believed to belong to the widely conserved *mir-183* family (Pierce *et al.*, 2008). In vertebrates, *mir-183* family members are expressed in sensory hair cells, while in invertebrates (*e.g. C. elegans* and *D. melanogaster*) they are expressed in neuroglia (supporting cells), suggesting an evolutionary divergence. Based on the conserved location of the miRNA family in *C. elegans* and *D. melanogaster*, it seems likely that the role of miR-228 in nervous tissues is conserved in *H. contortus* and other nematodes. Hco-miR-228 is most abundant during L3, matching expression of Cel-miR-228, which peaks during L1 – L3 (Kato *et al.*, 2009). The requirement for miR-228 might decrease over time as the development of nervous tissue is completed during L3 (Sulston *et al.*, 1983). Predicted targets of Cel-miR-228 showed negligible enrichment of GO terms, while Hco-miR-228 targets were enriched for ‘metabolism’ and ‘ion binding’ annotations. This may reflect the role of glial cells in maintaining homeostasis and providing nutrients to the sensory neurons (reviewed in Oikonomou and Shaham, 2011).

The nematode specific miR-235 belongs to the widely conserved *mir-25* family, which includes the oncogenic Hsa-miR-17/92 cluster (reviewed in Mogilyansky and Rigoutsos, 2013). In *C. elegans*, miR-235 has an important role in regulating blast-cell quiescence during L1 arrest (Kasuga *et al.*, 2013). When food availability is low, levels of miR-235 increase and, acting through the insulin/insulin-like growth factor signalling pathway, development of neuroblasts and mesoblasts is arrested. Except for L1 arrest, abundance was limited to adult males, indicating an additional sex-specific role. Abundance of miR-235 was also low in the *C. elegans* dauer, suggesting that another mechanism is

responsible for arrested development during dauer diapause. Additionally, while a *mir-235* mutant strain was unable to enter L1 diapause, its ability to form dauers was not significantly reduced compared to wild type worms (Chapter 6). In *H. contortus*, expression of miR-235 was markedly different to *C. elegans*, with high levels detected in L3 and adult stages. If the quiescence-maintaining function of miR-235 is conserved between these species, Hco-miR-235 may have been co-opted to regulate L3 arrest.

miR-5885 was first identified as a unique miRNA in *H. contortus*, however, work presented here identified highly similar stem-loop sequences in a further eight parasitic nematodes (Chapter 3). Hco-miR-5885 is expressed by three loci (a, b and c) with the 3′ arms being the most abundant miRNAs in *H. contortus*. In a similar manner to Hco-miR-60, expression peaks during L4 and in gut tissue, suggesting a role in gut development or maintenance. Gene ontology provided little insight into the function of miR-5885-3p, but when aligned to *C. elegans* miRNAs, its seed sequence was found to be identical to those of miR-58/80/81/82, indicating that it belongs to the same group of miRNAs, known as the *bantam* family. The original bantam miRNA was identified in *D. melanogaster* as a regulator of cell proliferation and apoptosis (Brennecke *et al.*, 2003). In *C. elegans*, this family regulates locomotion, body size, egg laying and dauer formation (Alvarez-Saavedra and Horvitz, 2010), although there are some spatial and functional differences between individual members (Than *et al.*, 2013b). As miR-58/80/81/82 have not been identified in *H. contortus* nor, based on current genomic data, in other clade V parasitic nematodes, miR-5885-3p might have evolved as the version of a bantam miRNA used by parasitic nematodes. Interestingly, the 5′ arm of miR-5885 shows no similarity to the 5′ arms of miR-58/80/81/82 (or any other confirmed miRNA), which might be interpreted as stronger evidence of evolutionary convergence instead of replacement and divergence.

The 5′ arm of miR-5885 is also unusual as its expression profile is different from the 3′ arm. This phenomenon, known as arm switching (or arm selection) was once thought only

to exist between organisms at an evolutionary level (de Wit *et al.*, 2009), but is now known to occur dynamically within an organism (Griffiths-Jones *et al.*, 2011). While the mechanisms underlying switching are unclear, it is triggered in a wide range of situations and changes can occur between tissues, developmental stages and different sexes (Marco *et al.*, 2013b; Ninova *et al.*, 2014). Interestingly, abnormal situations including neoplasia (Li *et al.*, 2012a), abiotic stress (Hu *et al.*, 2014) and bacterial infection (Siddle *et al.*, 2015) can also induce arm switching. Only a small percentage of miRNAs appear to undergo arm switching; in humans, for example, seven miRNAs are reported to display tissue-specific switching (Kuchenbauer *et al.*, 2011). The only other example of developmental arm switching in *H. contortus* is seen with Hco-miR-5960. While in many cases, the two arms of a duplex regulate different genes (Marco *et al.*, 2012), a large number of genes predicted as targets of Hco-miR-5885 were common to both arms. With high levels of expression and large number of common targets, it is possible that miR-5885 regulates a wide range of targets and pathways, with arm switching providing flexibility by altering the same genes in different developmental stages and tissues.

While the data presented here highlights the spatiotemporal variation of miRNAs in *H. contortus*, additional experiments are required to fully understand and appreciate their biological impact. One obvious approach would be to expand miRNA expression profiles by studying additional stages (*i.e.* eggs, L1 and L2) or increasing temporal resolution of the current data. For example, the L4 abundance data was based on worms taken from hosts at seven days post-infection. If abundance was examined over shorter periods (*e.g.* two, four and six days post-infection), novel patterns as worms develop within the host might be identified. Expanding spatial abundance data to additional tissues would be difficult using the extraction techniques used to isolate gut tissue. However, *in situ* hybridisation presents an alternative approach where organisms are incubated with labelled locked nucleic acid

probes, complementary to mature miRNA sequences, before being stained and visualised (Lagendijk *et al.*, 2012). For miRNAs conserved in *Ascaris*, miRNA sequencing of extracted tissues including neurons should be feasible due to its large size (Rosa *et al.*, 2014).

Experiments using *C. elegans* highlighted the importance of miRNAs in this organism and provided insights for *H. contortus*. Dauer formation was clearly affected by the absence of miRNAs and expanding on this approach should indicate which specific miRNAs (individually or as a group) are required for dauer formation. Preliminary results from the dauer recovery assay were not statistically significant, but additional experiments should clarify a role for miRNAs. While not examined during this study, it is likely that miRNAs are also involved in maintaining the dauer state. Such involvement might be observable in miRNA mutant strains as reduced survivability (*i.e.* shorter dauer lifespan) or stability (*i.e.* inability to remain as dauers).

While HITS-CLIP remains a potentially powerful strategy if an anti-Argonaute antibody can be generated, additional techniques are available to validate targets albeit on a smaller scale. An alternative approach was recently used to validate targets of miR-71 in the filarial parasite *Brugia malayi* (Liu *et al.*, 2015). Parasites were transiently transfected with a plasmid containing a luciferase reporter linked to a promoter and 3′ UTR containing predicted miRNA binding sites. When incubated with miR-71, activity (as measured using a dual luciferase assay) decreased, indicating that the particular 3′ UTR sequence was being negatively regulated by miR-71. With appropriate transfection tools, this technique may be applied in *H. contortus* to confirm miRNA-mRNA interactions

Even with target validation, identifying the biological impact of a miRNA requires further experimental approaches. Initial studies using miRNA inhibitors and a mimic were unable to detect any obvious effects on *H. contortus* larval development. This may be due to *in*

vitro culture conditions not matching *in vivo* conditions or the redundancy of miRNA regulation. Improving culture conditions and testing inhibition of additional miRNAs with significant L3/L4 variation, singly and in combination, may yet yield insights into the roles of miRNAs. This process is particularly useful in identifying miRNAs that may be used as the basis of novel therapeutic agents. In addition, identifying specific genes and pathways affected by altered miRNA levels would provide a novel approach to identify genes important for parasite development. Specific miRNAs or, alternatively, key genes regulated by these miRNAs could be used as valid therapeutic targets. As targeting one miRNA may potentially target dozens of genes in a pathway, altering miRNA abundance and activity may prove to be a more effective approach.

The data presented in this study represent the first detailed examination of miRNAs and their potential target genes in *H. contortus*. A significant number of miRNAs display temporal variation in abundance, suggesting a role in development, while spatial variation highlights the tissue-specific nature of miRNAs in *H. contortus*. A number of parasite-specific miRNAs showed significant changes in abundance between the L3 and L4 stages. Targeting of these potentially important miRNAs with inhibitors will shed further light on their biological functions and may lead to the development of novel therapeutics for the control of parasitic nematodes.

8 Appendix

8.1 Buffers and reagents

Table 8-1: Composition of buffers and chemicals used during this study.

Buffer/media	Composition
Coomassie Blue stain	0.25 g Coomassie Brilliant Blue R-250 + 100 mL methanol + 25 mL glacial acetic acid + 125 mL dH_2O. Stir for 2 – 3 hours and filter before use. Do not autoclave. Can be reused up to 5 times if filtered after every use.
Coomassie destain solution	200 mL methanol + 50 mL glacial acetic acid + 250 mL dH_2O. Do not autoclave.
Elution buffer for GST-tagged proteins	150 mM NaCl + 50 mM Tris + 10M reduced glutathione. Adjust pH to 8.0.
Elution buffer for His-tagged proteins	300 mM NaCl + 20 mM NaH_2PO_4 + 250 mM imidazole. Adjust pH to 4.5. For denaturing conditions add 8 M urea.
Equilibration/Wash buffer for GST-tagged proteins	150 mM NaCl + 50 mM Tris. Adjust pH to 8.0.
Equilibration buffer for His-tagged proteins	300 mM NaCl + 20 mM NaH_2PO_4 + 10 mM imidazole. Adjust pH to 7.5. For denaturing conditions add 8 M urea.
Lysis buffer for GST-tagged proteins	Phosphate buffered saline (PBS) + 1 mg/mL lysozyme + 10 % Triton-X-100. Adjust pH to 7.5.
Lysis buffer for His-tagged proteins	300 mM NaCl + 20 mM NaH_2PO_4 + 10 mM Tris + 1 % (v/v) Triton-X-100. Adjust pH to 7.5. For denaturing conditions add 8 M urea.
Wash buffer for His-tagged proteins	300 mM NaCl + 20 mM NaH_2PO_4 + 25 mM imidazole. Adjust pH to 7.5. For denaturing conditions add 8 M urea.
Freezing solution	100 mM NaCl + 50 mM KH_2PO_4 + 30% glycerol (v/v) + 400 mL dH_2O. Autoclave. Add 3 mM $MgSO_4$. Adjust volume to 500 mL with dH_2O.
LB medium	10 g bacteriological tryptone + 5 g yeast extract + 5 g NaCl in 800 mL dH_2O. Adjust volume to 1 L using dH_2O. Autoclave.
M9 buffer	6 g Na_2HPO4 + 3 g KH_2PO4 + 5 g NaCl + 0.25 g $MgSO4{\cdot}7H_2O$ + 800 mL dH_2O. Adjust volume to 1 L using dH_2O. Autoclave.
NGM (nematode growth medium) agar	17 g bacteriological agar + 2.5 g bacteriological peptone + 3 g NaCl in 800 mL dH_2O. Adjust volume to 972 mL using dH_2O. Autoclave. Cool to 55 °C before adding 1 mL cholesterol (5 mg/mL in 100 % ethanol) + 1 mL 1M $MgSO_4$ + 1 mL 1M $CaCl_2$ + 25 mL 1M potassium phosphate buffer (pH 6.0)
NGM agar (without peptone)	As above, but excluding peptone.
PBS	1 x sachet of 0.01 M Phosphate buffered saline powder (Sigma-Aldrich) + 1 L dH_2O. Autoclave.
PBS-T (0.1 %)	1 L PBS + 1 mL Tween 20
S Basal medium	5.9 g NaCl + 1.0 g K_2HPO_4 + 6.0 g KH_2PO_4 + 1 mL cholesterol (5 mg/mL in 100% ethanol) in 800 mL dH_2O. Adjust volume to 1 L using dH_2O. Autoclave.
S medium	977 mL S-Basal + 10 mL 1 M potassium citrate + 10 mL trace metals solution + 3 mL 1 M $CaCl_2$ + 3 mL 1M $MgSO_4$ + 1L S-Basal. Do not autoclave.
SDS (1 %)	10 g SDS (sodium dodecyl sulphate) in 800 mL dH_2O. Use fume hood. Use magnetic stirrer and heat to 65°C until dissolved. Adjust volume to 1 L using dH_2O.
Superbroth medium	32 g tryptone + 20 g yeast extract + 5 g NaCl + 5 mL 1 M NaOH + 800 mL dH_2O. Adjust volume to 1 L with dH_2O. Autoclave.
TAE buffer (50x)	242 g tris base + 57.1 mL glacial acetic acid + 100 mL 0.5M EDTA in 800 mL dH_2O. Adjust volume to 1 L using dH_2O. Do not autoclave. Before use, dilute 1 in 50.
Tris-glycine buffer	70 mL 10x TG Transfer Buffer (Biorad) + 140 mL methanol + 490 mL dH2O. Create on day of use.

8.2 Nucleotide and protein sequences

8.2.1 Oligonucleotides

Table 8-2: Oligonucleotides used as primers during PCR and qRT-PCR.

Sequence (5' - 3')	Notes
TGACTAGAGACACATTCAGC	Forward primer for Hco-miR-45-3p
TATTATGCACATTTTCTGGTTC	Forward primer for Hco-miR-60-3p.
TGAAAGACATGGGTAGTGAGAC	Forward primer for Hco-miR-71-3p.
GTGAGCAAAGTCTCAGGTGTGG	Forward primer for Hco-miR-87-3p.
AAATGGCACTGCATGAATTCACGG	Forward primer for Hco-miR-228-5p.
TATTGCACTCGCCCCGGCCTGA	Forward primer for Hco-miR-235-3p
TGAGATCACGCGTATATTCGC	Forward primer for Hco-miR-5885a-3p
GGGTGTACGTGGTGGTCTGGTA	Forward primer for Hco-miR-5885a-5p
TACCCGTAATGTACATAGCTTGAG	Forward primer for Hco-miR-5899-3p. Used as a normaliser in qRT-PCR.
GACGAGCTGCCTCAGT	Universal reverse primer for miRNAs in qRT-PCR
CCACATAGGGAGCAAGCCATTTCAGTG	Forward primer for *Cel-hbl-1*.
ATTGTAGAGGGAGGCGATTGAGGTTGG	Reverse primer for *Cel-hbl-1*.
ACTGCTCAACACGCTATTCGACAAGAC	Forward primer for *Cel-hpd-1*.
CGCGAGACGTTCCTTGAGGTTATCG	Reverse primer for *Cel-hpd-1*.
GGTCAACACACGCCCACCAAAAATAG	Forward primer for *Cel-T04A8.7*.
AGCGACTGGAAACAGCGAAAAAGTTCG	Reverse primer for *Cel-T04A8.7*.
GGTGACGATTTTGCTGGTGG	Forward primer for *HCISE00019000* (*Hco-T04A8.7*).
TCAAATCCTTGCCCACCCTC	Reverse primer for *HCISE00019000* (*Hco-T04A8.7*).
ATGGCCGCTGGGATTGATCAGCAA	Forward primer for *Hco-alg-1*.
CATTCGAAATGTGCGCCACC	Reverse primer for *Hco-alg-1*.
GTAAAACGACGGCCAG	Forward primer M13 (-20) for TOPO vector.
CAGGAAACAGCTATGAC	Reverse primer M13 for TOPO vector.
ATGGCCGCTGGGATTGATCAGCAACAAC	Forward primer for *Hco-alg-1* for GST-tagged recombinant protein. Contains EcoRI.
CCCCTCGAGTTAGCATTCGAAATGTGCGCCACC	Reverse primer for *Hco-alg-1* for GST-tagged recombinant protein. Contains XhoI.
CCCCCATGGCCGCTGGGATTGATCAGCAAC	Forward primer for *Hco-alg-1* for Long His-tagged recombinant protein. Contains NcoI.
CCCAGATCTAAACTCGGCTTTTCTTACAAG	Reverse primer for *Hco-alg-1* for long His-tagged recombinant protein. Contains BglII.
CGTTGTTCCATCACCCAAGGTATC	Forward primer for Hco-ß-tubulin. Used as a normaliser in mRNA qRT-PCR.
CTGTGTAAGCTCAGCAACTGTCGAA	Reverse primer for Hco-ß-tubulin. Used as a normaliser in mRNA qRT-PCR.

Table 8-3: Oligonucleotides used as miRNA inhibitors.

Sequence (5' - 3')[1]	Notes
uaaCCUUGAACCAGAAAAUGUGCAUAAUACCuag	Inhibitor of Hco-miR-60-3p
cuaUAGCGAAUAUACGCGUGAUCUCAACcuu	Inhibitor of Hco-miR-5885a/c-3p

[1] All bases are 2'O-Methyl RNA. Upper case letters indicate phosphodiester bases. Lower case letters indicate phosphorthioate bases

8.2.2 miRNAs

Table 8-4: Sequences of the 273 mature miRNAs in *H. contortus*.

Mature miRNA	Stem-loop[1]	Array[2]	miRBase[3]	Mature miRNA Sequence
Hco-let-7-3p	*Hco-let-7*	No	N/A	CUAUUACAACUUGCUAGCUUGCC
Hco-let-7-5p		Yes	N/A	UGAGGUAGUAGGUUGUAUAGUU
Hco-lin-4-3p	*Hco-lin-4*	Yes	N/A	GCGAAUUGUGACCUCAGGAUAC
Hco-lin-4-5p		Yes	MIMAT0023345	UCCCUGAGACCUCAAUUGCGA
Hco-mir-1-3p	*Hco-miR-1*	Yes	MIMAT0023401	UGGAAUGUAAAGAAGUAUGUAG
Hco-mir-1-5p		Yes	N/A	GUAUGUAUGGUAUGUAUG
Hco-mir-2-3p	*Hco-miR-2*	Yes	MIMAT0023415	UAUCACAGCCAGCUUUGAUGUGC
Hco-mir-2-5p		Yes	N/A	ACGUCAAAGUGGUGGGUGAUUUG
Hco-mir-7-5p	*Hco-miR-7*	Yes	MIMAT0023387	UGGAAGACAGGAGAUUGCGUCGU
Hco-mir-9-5p	*Hco-miR-9*	Yes	MIMAT0023348	UCUUUGGUUAUCUAGCUGUAUGA
Hco-mir-9-3p		Yes	N/A	AUAAGGCUGGACGACUAAAGGAA
Hco-mir-40a-3p	*Hco-miR-40*	Yes	MIMAT0023444	UCACCGGGUGUCUUGCAGUGAG
Hco-mir-40b-3p		Yes	MIMAT0023323	UCACCGGGUGUCUUGUAGCGGA
Hco-mir-40c-3p		Yes	MIMAT0023429	UCACCGGGAGUUCUGUGGUGA
Hco-mir-40d-3p		Yes	MIMAT0023430	UCACCGGGUAUCUUGCAGCGGG
Hco-mir-40d-5p		Yes	N/A	CGUUGUAGGACCCCCUGUGAGG
Hco-mir-40e-3p		Yes	N/A	UCACCGGGAUUUCUGCAGUGA
Hco-mir-40e-5p		Yes	MIMAT0023428	ACUGCAGGACCGCUCAGGUGAAG
Hco-mir-43-3p	*Hco-miR-43*	Yes	MIMAT0023336	UAUCACAGUGUCAUUGGGUCGCA
Hco-mir-43-5p		Yes	N/A	CAACCCGAUGUAGCUGUUAUACU
Hco-mir-45-3p	*Hco-miR-45*	Yes	MIMAT0023339	UGACUAGAGACACAUUCAGCU
Hco-mir-45-5p		Yes	N/A	CUGGAUGUUUCUCUGAGUCAUA
Hco-mir-46-3p	*Hco-miR-46*	Yes	MIMAT0023327	UGUCAUGGAGUCGCUCUCUUCA
Hco-mir-46-5p		Yes	N/A	AAGAGAGCCGUCCGUUGACAGU
Hco-mir-50-3p	*Hco-miR-50*	Yes	N/A	CCAGAACUAUUAGACAUAUCGAA
Hco-mir-50-5p		Yes	MIMAT0023329	UGAUAUGUCUGGUAUUCUUGGG
Hco-mir-55-3p	*Hco-miR-55*	Yes	MIMAT0023441	UACCCGUAUAAGCCGUCUGCCGGC
Hco-mir-55-5p		Yes	N/A	GCGGCAGGGGGUAUGCGAGUGUCC
Hco-mir-59-3p	*Hco-miR-59*	Yes	MIMAT0023356	UCGAAUCGUCACUCUUGAUGCUC
Hco-mir-59-5p		Yes	N/A	UCGUCAAGAGUACGAUCCGGAA
Hco-mir-60-3p	*Hco-miR-60*	Yes	MIMAT0023468	UAUUAUGCACAUUUUCUGGUUCAA
Hco-mir-60-5p		Yes	N/A	GAGCUGGAAACUGUCAUAAAAUC
Hco-mir-61-3p	*Hco-miR-61*	Yes	MIMAT0023335	UGACUAGACUGUUACUCGGCGU
Hco-mir-61-5p		Yes	N/A	CCCGGGUGACUGGCUGGUCCUG
Hco-mir-63a-3p	*Hco-miR-63*	Yes	MIMAT0023318	AAUGACACUGUUGCGAACUGGGAU
Hco-mir-63a-5p		Yes	N/A	ACCUGUUCGACGGGAGUCAUCGU
Hco-mir-63b-3p		Yes	MIMAT0023440	CUAUGACACUGUCGCGAAUUGGGG
Hco-mir-63b-5p		Yes	N/A	CUCCGAUUCGCCGGUAGUCAUCGU
Hco-mir-71-3p	*Hco-miR-71*	Yes	N/A	AUCACUAUUCUGUUUUUCGCC
Hco-mir-71-5p		Yes	MIMAT0023358	UGAAAGACAUGGGUAGUGAGAC
Hco-mir-72-3p	*Hco-miR-72*	Yes	N/A	AGGCAAGAUGUUGAGGCGGACUC
Hco-mir-72-5p		Yes	MIMAT0023498	AGGCAAGAUGUUGGCAUAGCUGA
Hco-mir-76-3p	*Hco-miR-76*	Yes	MIMAT0023496	UUCGUUGUUUAUGAAGCCUGG
Hco-mir-79-3p	*Hco-miR-79*	Yes	MIMAT0023326	AUAAAGCUAGGUUACCAAAGCUA
Hco-mir-79-5p		Yes	N/A	UCUUUGGUGAUUUGGCUUAUAUGA
Hco-mir-83-3p	*Hco-miR-83*	Yes	MIMAT0023355	UAGCACCAUGUAAAUUCAGCA
Hco-mir-83-5p		Yes	N/A	CUGAAUUUAUUUGCGAGCUGGA
Hco-mir-84a-3p	*Hco-miR-84*	Yes	N/A	AUAACAUUUAUCCUACUUCGUC
Hco-mir-84a-5p		Yes	MIMAT0023368	UGAGGUAGUUUUAAAUGUUAUGA
Hco-mir-84b-5p		No	MIMAT0023507	UGAGGUAGUUAAGCACAGGUA
Hco-mir-86-3p	*Hco-miR-86*	Yes	N/A	GCUGGGCCGAGGUUCGCUUAGG
Hco-mir-86-5p		Yes	MIMAT0023350	UAAGUGAAUGCUUUGCCACAGUCU
Hco-mir-87a-3p	*Hco-miR-87*	Yes	MIMAT0023347	GUGAGCAAAGUUUCAGGUGUGC
Hco-mir-87a-5p		Yes	N/A	CCGCCUGAACCUUUCGUCUCAACCU
Hco-mir-87b-3p		Yes	MIMAT0023349	GUGAGCAAAGUCUCAGGUGUGG
Hco-mir-87b-5p		Yes	N/A	CCGCCUGACACUUGACUCAAACCU
Hco-mir-87c-3p		Yes	MIMAT0023504	GGGAGCAAAGUUUCAGGUUC
Hco-mir-124-3p	*Hco-miR-124*	Yes	MIMAT0023338	UAAGGCACGCGGUGAAUGCCAA
Hco-mir-124-5p		Yes	N/A	CGCUUUCAUCGGUGACUUUAGU
Hco-mir-133-3p	*Hco-miR-133*	Yes	MIMAT0023317	AUUGGUCCCCUCCAACCAGCUAA
Hco-mir-190-5p	*Hco-miR-190*	No	MIMAT0023502	GGAUAUGUAUGGGUGAGCUCGUU
Hco-mir-228-5p	*Hco-miR-228*	Yes	MIMAT0023389	AAAUGGCACUGCAUGAAUUCACGG
Hco-mir-234-3p	*Hco-miR-234*	Yes	MIMAT0023375	AUUAUUGCUCGAGAAUACCCGU
Hco-mir-235-3p	*Hco-miR-235*	Yes	MIMAT0023388	UAUUGCACUCGCCCCGGCCUGA
Hco-mir-236-3p	*Hco-miR-236*	Yes	MIMAT0023379	CUAAUACUGUCAGGUAAUGAUGC
Hco-mir-236-5p		Yes	N/A	GCGUCUUACCUGGGCAAUAUUAGA
Hco-mir-242-5p	*Hco-miR-242*	Yes	MIMAT0023369	CUUGCGUAGGCCUUCGUGUAACGA

Mature miRNA	Stem-loop[1]	Array[2]	miRBase[3]	Mature miRNA Sequence
Hco-mir-250-3p	*Hco-miR-250*	Yes	MIMAT0023330	UCGAAUCACAGCCAACUGCUGCU
Hco-mir-250-5p		Yes	N/A	CGGCAGUUGCCUCGUGAUCCGUGAUC
Hco-mir-252-5p	*Hco-miR-252*	Yes	MIMAT0023372	CUAAGUAGUGGUGCCGCAGGUAAC
Hco-mir-255-3p	*Hco-miR-255*	Yes	MIMAT0023411	AAACUGAAAAGAUUUGUUACAG
Hco-mir-259-3p		Yes	N/A	CAACGAAUUCGACUGAGAUUGGG
Hco-mir-259-5p	*Hco-miR-259*	Yes	MIMAT0023333	UAAUCUCAUCCGAAUCUGUUGCA
Hco-mir-265-3p	*Hco-miR-265*	Yes	MIMAT0023508	TGAGGGAGGAAGGGAATAAT
Hco-mir-277-3p	*Hco-miR-277*	Yes	MIMAT0023377	AAAAUGCACCUACCUGGUAUGCA
Hco-mir-277-5p		Yes	N/A	CAUACCAGUGGCAGCAUUUCCA
Hco-mir-307-3p	*Hco-miR-307*	Yes	MIMAT0023407	UCACAACCUCCUUGAGUAAGUCGA
Hco-mir-307-5p		Yes	N/A	UGCUUGCUCAGUCGGUUGUUCGAUG
Hco-mir-790-5p	*Hco-miR-790*	Yes	MIMAT0023374	CUUGGCACUCGAGAACGACUGCG
Hco-mir-993-3p	*Hco-miR-993*	Yes	MIMAT0023413	UAAGCUCGUCGCUACAGGCAGAA
Hco-mir-993-5p		Yes	N/A	CUACCUGAUAGUACGAGAUUGUU
Hco-mir-2159-3p	*Hco-miR-2159*	Yes	MIMAT0023509	UAAGGGGCCAAAUAAAGGGGGGG
Hco-mir-5352-3p	*Hco-miR-5352*	Yes	MIMAT0023381	UUGCACAUGAUGUACGACCUGCA
Hco-mir-5352-5p		Yes	N/A	CGGGCGUGCUCGUUGUGCAGUG
Hco-mir-5884-5p	*Hco-miR-5884*	Yes	MIMAT0023316	UAGGGUACUGACAUUGAAUGAGU
Hco-mir-5884-3p		Yes	N/A	CCAUUUAAUUUCGUACCCUGGA
Hco-mir-5885a-3p	*Hco-miR-5885*	Yes	MIMAT0023319	UGAGAUCACGCGUAUAUUCGC
Hco-mir-5885a-5p		Yes	N/A	GGGUGUACGUGGUGGUCUGGUA
Hco-mir-5885b-3p		Yes	MIMAT0023417	UGAGAUCACGCCUAUAUUCGC
Hco-mir-5885b-5p		Yes	N/A	GGGUAUGGGUGGUGGUCUGGUA
Hco-mir-5885c-3p		Yes	MIMAT0023484	UGAGAUCACGCGUAUAUUCGCUA
Hco-mir-5885c-5p		Yes	N/A	CAGAGUAUACGUGGUGUUCUGGUA
Hco-mir-5886a-3p	*Hco-miR-5886*	Yes	MIMAT0023320	GAACAAUCUUUGACUUUGGUU
Hco-mir-5886b-3p		Yes	MIMAT0023462	GAACACUCUUUGACUUUGGUUUGA
Hco-mir-5886b-5p		Yes	N/A	AAACUGAGUCAUUGAUGUGUUCUU
Hco-mir-5887-3p	*Hco-miR-5887*	Yes	MIMAT0023321	UUACCUGAUGGUGUAGCAAAGGGA
Hco-mir-5887-5p		Yes	N/A	CCUUUGCUACGCCAUCGGGUGAGC
Hco-mir-5888-5p	*Hco-miR-5888*	Yes	MIMAT0023322	AAGCGGAUCUAGGGAUGCUUCUG
Hco-mir-5888-3p		Yes	N/A	GAAGCAUCCCUAGGUCCUCUACA
Hco-mir-5889-3p	*Hco-miR-5889*	Yes	MIMAT0023324	UUUCUGUAAGGAAUGGAGAUGU
Hco-mir-5889-5p		Yes	N/A	GAAGCAUCCCUAGGUCCUCUACA
Hco-mir-5890a-3p	*Hco-miR-5890*	No	MIMAT0023325	UACCCCUUUUCAUUUUUAUGC
Hco-mir-5890b-3p		Yes	MIMAT0023371	UACCCCUUUUCAUUUUUAUGC
Hco-mir-5890b-5p		Yes	MIMAT0023370	CAUAGGACUGAAUUGUGGGAG
Hco-mir-5890c-5p		Yes	MIMAT0023378	CAUAUAACUGAACUGUGGGAG
Hco-mir-5890c-3p		Yes	N/A	CUACCCCUCUUCAUUUUUAUGC
Hco-mir-5890d-3p		Yes	MIMAT0023393	CUACCCCUUUCCAUUCUUAUGC
Hco-mir-5890d-5p		Yes	N/A	ACAGGACUGAACUGUGGGAGGA
Hco-mir-5890e-5p		Yes	MIMAT0023402	AUAGGACUGAACUGUGGGAGA
Hco-mir-5890f-3p		Yes	MIMAT0023410	AUAGGACUGAACUGUGGGAGGA
Hco-mir-5890f-5p		Yes	MIMAT0023409	CUACCCCUUUCCAUUCUUAUGC
Hco-mir-5891-5p	*Hco-miR-5891*	No	MIMAT0023328	AUAGCGAUAUCCUCUGAAUGGU
Hco-mir-5891-3p		Yes	N/A	CAUUUAGAGGAUAUCGCUGUCC
Hco-mir-5892a-5p	*Hco-miR-5892*	Yes	MIMAT0023331	AAUUAACAUUUGGAUCAUCUGAA
Hco-mir-5892b-5p		Yes	MIMAT0023486	AAUUAACAUUUGGAUCAUCUGAGC
Hco-mir-5893-3p	*Hco-miR-5893*	No	MIMAT0023332	UACAAAUUUAUGAACUCCGAAG
Hco-mir-5894a-5p	*Hco-miR-5894*	Yes	MIMAT0023334	CCUAUAUCUGAUGGUGUAGCGGGA
Hco-mir-5894a-3p		Yes	N/A	UCGCUACACCAUCAGUUAUGGACU
Hco-mir-5894b-3p		Yes	MIMAT0023400	GCCUAUAUCUGAUGGUGUAGC
Hco-mir-5895-5p	*Hco-miR-5895*	Yes	MIMAT0023337	UACCCGUAGCAUCUGUAUGUCU
Hco-mir-5895-3p		Yes	N/A	AUAUGCAUUUGCUGUGGGUGAC
Hco-mir-5896-5p	*Hco-miR-5896*	Yes	MIMAT0023340	AUGGGCAAUGUGUCCUGUUCAGA
Hco-mir-5896-3p		Yes	N/A	UGGCAGGCACGUUGUUCAAUA
Hco-mir-5897a-3p	*Hco-miR-5897*	Yes	MIMAT0023341	UUUUGUAUAACUCUUCUAGAAAU
Hco-mir-5897b-3p		No	MIMAT0023364	UUUUGUAUGCCUUAUCUGGAAU
Hco-mir-5897c-3p		Yes	MIMAT0023380	UUUUGUAUGUAUUAUCUGGAAUU
Hco-mir-5898-3p	*Hco-miR-5898*	Yes	MIMAT0023343	CAGUUUGGAACAAUCGAAGCC
Hco-mir-5898-5p		Yes	MIMAT0023342	UUUUGAUUGUUACUAAACUUUU
Hco-mir-5899-3p	*Hco-miR-5899*	Yes	MIMAT0023344	UACCCGUAAUGUACAUAGCUUGAG
Hco-mir-5899-5p		Yes	N/A	CAGGCUAUGCAUUCCGUGUGCA
Hco-mir-5900-3p	*Hco-miR-5900*	Yes	MIMAT0023346	AAUUCUAUCCAAGUGACUCUGC
Hco-mir-5901-3p	*Hco-miR-5901*	Yes	MIMAT0023351	UUAUGGUCAGAAGGAGUGAACUA
Hco-mir-5901-5p	*Hco-miR-5901*	Yes	N/A	AUUCACUCCUUCUGACCAUAGCG
Hco-mir-5902-3p	*Hco-miR-5902*	Yes	MIMAT0023353	AAGAGGAGCGUUGUCACAGUGA
Hco-mir-5902-5p		Yes	MIMAT0023352	UCUGGGAUAACACCUCUCUUCA
Hco-mir-5903-3p	*Hco-miR-5903*	Yes	MIMAT0023354	UACCGGGAUAUGGAACGUCUGAA
Hco-mir-5904a-3p	*Hco-miR-5904*	Yes	MIMAT0023357	AGUGUGAUAGUGGUUUUGCAA
Hco-mir-5904a-5p		Yes	N/A	GUAAACCCACUAUUUUCACUUCG
Hco-mir-5904b-3p		Yes	MIMAT0023373	GUAAACCCUCUACUUCACUUCGAA
Hco-mir-5904b-5p		Yes	N/A	CGAGUGUCGUGGUGGCUUUUACAA

Mature miRNA	Stem-loop[1]	Array[2]	miRBase[3]	Mature miRNA Sequence
Hco-mir-5904c-3p	*Hco-miR-5904*	Yes	MIMAT0023418	CAAACCCACUAUUUUCACUUCGA
Hco-mir-5905-5p	*Hco-miR-5905*	No	MIMAT0023359	UCUGAAAGGCUGAAGUUUUUGGC
Hco-mir-5906-5p	*Hco-miR-5906*	Yes	MIMAT0023360	UGGCAAUGAGCUAGCUAAUUGUU
Hco-mir-5907-3p	*Hco-miR-5907*	Yes	MIMAT0023361	UGAACCCUUUGUCGUGACAUCGA
Hco-mir-5907-5p		Yes	N/A	GAGUCCAACAGAGGUAGUCACU
Hco-mir-5908-3p	*Hco-miR-5908*	Yes	MIMAT0023363	UUCAUCAGGCCUCAGUCAUCUACAU
Hco-mir-5908-5p		Yes	MIMAT0023362	GAUGAUUGUGCACUGGUGAGG
*Hco-mir-5908-5p**		Yes	N/A	UCAUCAGGCCUCAGUCAUCUA
Hco-mir-5909-3p	*Hco-miR-5909*	Yes	MIMAT0023365	UCUGCAACUCUUCAUUCUUGUC
Hco-mir-5909-5p		Yes	N/A	GAAGAAUGGAGAGUUGCAGGAG
Hco-mir-5910-5p	*Hco-miR-5910*	Yes	MIMAT0023366	CCUCCCCAAUUCGAUCGAGGCGA
Hco-mir-5911a-5p	*Hco-miR-5911*	No	MIMAT0023367	UAAGCGCUUCAUCCUUAUAGAC
Hco-mir-5911a-3p		Yes	N/A	UGUGAGGCUGACUUGCUUGGGAG
Hco-mir-5911b-5p		No	MIMAT0023386	UAAGGUUGACUUGCUUAGAAGCU
Hco-mir-5911c-5p		No	MIMAT0023397	UGAGGCUGACUUGCUUAGGAGCU
Hco-mir-5912a-3p	*Hco-miR-5912*	No	MIMAT0023376	UAUUCGAAUCAUCUGCUCCCAGCC
Hco-mir-5912b-3p		No	MIMAT0023423	AUAUUCGAGUCCCUGCUCCCAGC
Hco-mir-5913-3p	*Hco-miR-5913*	Yes	MIMAT0023382	AUAAAGGUCUAGUAGUGAUGCA
Hco-mir-5914-3p	*Hco-miR-5914*	Yes	MIMAT0023383	UAGUGACUAGCAUAGGCAGAGG
Hco-mir-5915-5p	*Hco-miR-5915*	Yes	MIMAT0023384	UUAAGUAGUGACGCCAUUGUAA
Hco-mir-5916-3p	*Hco-miR-5916*	Yes	MIMAT0023385	GCAGACCAUCUGAUCAGAGAGAU
Hco-mir-5917-3p	*Hco-miR-5917*	No	MIMAT0023390	GAAAACCGGGUAGACCUGGAGCU
Hco-mir-5918a-3p	*Hco-miR-5918*	Yes	MIMAT0023391	UAGUCUCCUCGAUCGUGUUACCCA
Hco-mir-5918a-5p		Yes	N/A	GGUAGCACGAUAGGAAGACAAGC
Hco-mir-5918b-3p		Yes	MIMAT0023421	UCGCUCUACAUCGUGUUACCCA
Hco-mir-5918c-5p		Yes	MIMAT0023422	UAAUACGAUUAUGAGGCAAGCU
Hco-mir-5918c-3p		Yes	N/A	CUAGCCUAUACAUCGUGUUACC
Hco-mir-5918d-3p		Yes	MIMAT0023442	CUCGGCUUCUCAUCGUGUUACCCA
Hco-mir-5919-3p	*Hco-miR-5919*	Yes	MIMAT0023392	GUUUGGUUGCCGUUUUGAGGU
Hco-mir-5920-5p	*Hco-miR-5920*	Yes	MIMAT0023394	GUUUUGUCAUUCGUAUUGUGAAGC
Hco-mir-5921-3p	*Hco-miR-5921*	Yes	MIMAT0023396	CCUAUUAGUUUUCCUUGCCGAA
Hco-mir-5921-5p		No	MIMAT0023395	CUGCAAUGAAAUCUAGUAGACCUA
Hco-mir-5922-5p	*Hco-miR-5922*	No	MIMAT0023398	CGGGUGACUUGAGUACAUGAACU
Hco-mir-5923-5p	*Hco-miR-5923*	Yes	MIMAT0023399	UCUGCCAAAACGCUCUCUUUUG
Hco-mir-5924a-3p	*Hco-miR-5924*	Yes	MIMAT0023404	AACUGCUUCCAGCUGUGCUGACAU
Hco-mir-5924a-5p		No	MIMAT0023403	CAACACAGCCAGAUGCAGUAG
Hco-mir-5924b-3p		Yes	MIMAT0023419	AACUGCUUCCAGCUGUGCUGACAU
Hco-mir-5924b-5p		Yes	N/A	GUCAACACAGCCAGAAGCAGUAGU
Hco-mir-5925-3p	*Hco-miR-5925*	Yes	MIMAT0023405	AUAUGGAGUUCCUGGGGGGAUGA
Hco-mir-5925-5p		Yes	N/A	TTCTTCCCGGGACCTCCATATCA
Hco-mir-5926-3p	*Hco-miR-5926*	No	MIMAT0023406	UGGCACUCCGUGGAACGAAACGGA
Hco-mir-5927-5p	*Hco-miR-5927*	Yes	MIMAT0023408	CCCACUCAAUAAACAAUUCGU
Hco-mir-5927-3p		Yes	N/A	GAAUUGUUUAUUAAGUGGGCG
Hco-mir-5928a-5p	*Hco-miR-5928*	No	MIMAT0023412	UGAAGUAGAGUAGAUUUGUAGGAUC
Hco-mir-5928b-5p		No	MIMAT0023438	UGAAGUAAGGUAGAUUAGUAAGGU
Hco-mir-5928c-5p		No	MIMAT0023445	UGAAGUAAGGUAGACUAGUAAGGU
Hco-mir-5928d-5p		No	MIMAT0023456	UGAAGUAGAAUAGUUGGGUAUGGU
Hco-mir-5928e-5p		Yes	MIMAT0023481	UUCAAGUAAAGUAGGUUAGUAGGAUC
Hco-mir-5929-5p	*Hco-miR-5929*	Yes	MIMAT0023414	UGAAUCGCUUUGUUCACUUUCAUGG
Hco-mir-5930-5p	*Hco-miR-5930*	Yes	MIMAT0023416	CAGGGGCGACGUAGAACACAUGCA
Hco-mir-5930-3p		Yes	N/A	CAUGAGUUGCUAGUCGCCCGACAC
Hco-mir-5931-5p	*Hco-miR-5931*	No	MIMAT0023420	UGCAGGUCUUGAUUGGAAGGCU
Hco-mir-5932-3p	*Hco-miR-5932*	Yes	MIMAT0023424	AUUUUGGUCGGUGGUUAGAGUCA
Hco-mir-5933-5p	*Hco-miR-5933*	Yes	MIMAT0023425	GAAUGGUGAUUGUUGCGUCCGCA
Hco-mir-5933-3p		Yes	N/A	UGGAUAUCGGUGACUGUUCUG
Hco-mir-5934-3p	*Hco-miR-5934*	Yes	MIMAT0023426	UUAGUCUGAUCAGAAGCGUGGG
Hco-mir-5935-3p	*Hco-miR-5935*	No	MIMAT0023427	AGGGAUCGUAAUGUGGCAUCGA
Hco-mir-5936-3p	*Hco-miR-5936*	No	MIMAT0023431	AUACUGUGAAAGCUCCUUUGUC
Hco-mir-5937-3p	*Hco-miR-5937*	Yes	MIMAT0023432	AAUCGUGGACUAUGGAAGCAGC
Hco-mir-5938-5p	*Hco-miR-5938*	No	MIMAT0023433	CGCCGACUUUACCAAGCUUGGCA
Hco-mir-5938-3p		Yes	N/A	UCGGGUAUGGAUGUCGGCAUA
Hco-mir-5939-3p	*Hco-miR-5939*	Yes	MIMAT0023434	UGCCGUGGUCGCCAUGUGUUGGAAUC
Hco-mir-5940-3p	*Hco-miR-5940*	Yes	MIMAT0023435	GGAGGGAGUGUGGGCUCGGC
Hco-mir-5941-3p	*Hco-miR-5941*	No	MIMAT0023436	UAACGAUGUGCUGACCUCUGC
Hco-mir-5942-3p	*Hco-miR-5942*	Yes	MIMAT0023437	GUGGACGCAGAUUGAAUCGACC
Hco-mir-5943-5p	*Hco-miR-5943*	No	MIMAT0023439	CGUCGGAGCGUCCCGCUGAGGC
Hco-mir-5943-3p		Yes	N/A	UGCUCAGCAGGGCGUCCCGGCGG
Hco-mir-5944-3p	*Hco-miR-5944*	Yes	MIMAT0023443	UGAAAGGUCCUAUGUGUGGGGCAC
Hco-mir-5945-3p	*Hco-miR-5945*	Yes	MIMAT0023447	GAGGACUCCAUCGCCUGGACUGG
Hco-mir-5945-5p		Yes	MIMAT0023446	AGUUUUGGCGGGAAGUAUUAG
Hco-mir-5946a-5p	*Hco-miR-5946*	Yes	MIMAT0023448	AGGAUUUACACCUGUUGUCGUGAG
Hco-mir-5946b-5p		No	MIMAT0023470	UAGAUUUACACCGGUUAUCGUGA
Hco-mir-5946c-5p		Yes	MIMAT0023471	UGAGCUUUACACCUGUUGUCAUGA

Mature miRNA	Stem-loop[1]	Array[2]	miRBase[3]	Mature miRNA Sequence
Hco-mir-5947-3p	*Hco-miR-5947*	No	MIMAT0023449	UCAGCUUCCUUGUAGAACGGA
Hco-mir-5948-3p	*Hco-miR-5948*	Yes	MIMAT0023450	GAUCAAUAUACCCAAGGCUCGGAA
Hco-mir-5949-3p	*Hco-miR-5949*	Yes	MIMAT0023451	CGUGGGCUUCACGACACUGGG
Hco-mir-5950-3p	*Hco-miR-5950*	No	MIMAT0023452	UGGAUUCAGUUGUUGUAGUCGGC
Hco-mir-5951-3p	*Hco-miR-5951*	Yes	MIMAT0023453	GGUUCCGUCAGGACGACGAAGGAA
Hco-mir-5952-3p	*Hco-miR-5952*	No	MIMAT0023454	UGAGGGACUGGUCGCGGCAGAAUA
Hco-mir-5953-3p	*Hco-miR-5953*	No	MIMAT0023455	AUUCGUGCGAUUCUGGAUGCG
Hco-mir-5954-3p	*Hco-miR-5954*	No	MIMAT0023457	GUCAUUUUUUCAGGACAGAGU
Hco-mir-5955-5p	*Hco-miR-5955*	No	MIMAT0023458	UGACAAGACUCCAAGGACUGGUCA
Hco-mir-5956-3p	*Hco-miR-5956*	No	MIMAT0023459	UGUUGCGGCCUGUACUUCUACCAA
Hco-mir-5957-3p	*Hco-miR-5957*	Yes	MIMAT0023460	GGUCAGGGCGAUACUUCAAGCAGG
Hco-mir-5958-3p	*Hco-miR-5958*	Yes	MIMAT0023461	UAUAGCCGGGUACUUUACCAGCAC
Hco-mir-5959-3p	*Hco-miR-5959*	Yes	MIMAT0023463	UGCUUUUCUAACGCUCAAGGGCUA
Hco-mir-5960-5p	*Hco-miR-5960*	Yes	MIMAT0023464	GUGGUCCGGAGUCGGAGGGUUAU
Hco-mir-5960-3p		Yes	N/A	UUCCCUUCGGCACAAACUCGACCGACA
Hco-mir-5961-5p	*Hco-miR-5961*	Yes	MIMAT0023465	UCCACGAACUACGCAUUUCUGA
Hco-mir-5962-3p	*Hco-miR-5962*	Yes	MIMAT0023466	AGUGGGAUAGGCGCUGCACUCAGA
Hco-mir-5963-3p	*Hco-miR-5963*	Yes	MIMAT0023467	CUAUGGCAGAGGGGCGACCAUC
Hco-mir-5964-3p	*Hco-miR-5964*	No	MIMAT0023469	CUUUCAACUCUCGACUGUCCGAA
Hco-mir-5965-3p	*Hco-miR-5965*	No	MIMAT0023472	GACAAUUCUGACAUUUCGGA
Hco-mir-5966-3p	*Hco-miR-5966*	Yes	MIMAT0023473	UCUCUACCUAGCAAUCCUCACGUA
Hco-mir-5967-3p	*Hco-miR-5967*	No	MIMAT0023474	UUUCGUGAUACCGUUAUCCGGC
Hco-mir-5968-3p	*Hco-miR-5968*	Yes	MIMAT0023475	UCAUACUACAGUAGUUCUUGACGA
Hco-mir-5969-3p	*Hco-miR-5969*	No	MIMAT0023476	GGCUUCAUAUUAUGUGUC
Hco-mir-5970-5p	*Hco-miR-5970*	No	MIMAT0023477	CAUUGAUCGAUAUGACUCA
Hco-mir-5970-3p		Yes	N/A	AGUAAGGCGGUCGGUCGGUUGA
Hco-mir-5971-3p	*Hco-miR-5971*	Yes	MIMAT0023478	GUAGAUGAUUUCAAAAAUCACGCGUCCUGACCCG
Hco-mir-5971-5p		Yes	N/A	GGGGUGAUGAAUAGUGAAUGGGAAUCUCUAUGA
Hco-mir-5972-3p	*Hco-miR-5972*	Yes	MIMAT0023479	GCUUCGUCUGAGGCUUAGGCUUAGGC
Hco-mir-5973-3p	*Hco-miR-5973*	No	MIMAT0023480	CUUGCUGGAAUGUCCAGGG
Hco-mir-5973-5p		Yes	N/A	CCGGCCAUUCUUCUGCCGAGAU
Hco-mir-5974-3p	*Hco-miR-5974*	Yes	MIMAT0023482	AGAAUCAGUUUCGAUUGCAGAA
Hco-mir-5975-3p	*Hco-miR-5975*	Yes	MIMAT0023483	UUGUAGUAAUCGGCUGACCCUCCA
Hco-mir-5975-5p		Yes	N/A	GAGAGUCGGCCGAUUACUACAACC
Hco-mir-5976-5p	*Hco-miR-5976*	Yes	MIMAT0023485	UGGUCAAAUAUCACCAUCUGC
Hco-mir-5976-3p		Yes	N/A	AGAUACGUGAUGUUCGACCGAC
Hco-mir-5977-3p	*Hco-miR-5977*	Yes	MIMAT0023487	UAUGUUUAGUUAUAGAACCACGU
Hco-mir-5977-5p		Yes	N/A	GGACUUCUGUGAAAGUGGAGUGUUGCU
Hco-mir-5978-5p	*Hco-miR-5978*	No	MIMAT0023488	UUCGAGACUAGAGUCACUCCGCA
Hco-mir-5979-5p	*Hco-miR-5979*	No	MIMAT0023489	UGGAUGAUCGCGAAAGUGCCAUGA
Hco-mir-5979-3p		Yes	N/A	AUGGCACUUUUGCGAUCAUUCAUG
Hco-mir-5980-5p	*Hco-miR-5980*	No	MIMAT0023490	AGGCACCUACAGAACUAU
Hco-mir-5981-5p	*Hco-miR-5981*	No	MIMAT0023491	UGUUUACAAAAUUAGGACCCUGC
Hco-mir-5981-3p		Yes	N/A	AGGGUCCUAAUUUUGUGAACAGC
Hco-mir-5982-5p	*Hco-miR-5982*	Yes	MIMAT0023492	AGCUCCUCAGGAGUUGGUAGGUA
Hco-mir-5982-3p		Yes	N/A	CUCACCGACUCCUAAGAAGCUAG
Hco-mir-5983-5p	*Hco-miR-5983*	Yes	MIMAT0023493	UUUGGCACAUGUUAGGCCGGCA
Hco-mir-5983-3p		Yes	N/A	CGAGUCUAUCGUGUCUCAAAGC
Hco-mir-5984a-3p	*Hco-miR-5984*	Yes	MIMAT0023494	GUGUACUCUUUAGACGGUUCUA
Hco-mir-5984a-5p		Yes	N/A	AAACCGUCUAAAGAGUACGCUU
Hco-mir-5984b-5p		No	MIMAT0023495	GCGUACUCUUUAGACGGUUU
Hco-mir-5985-1-5p	*Hco-miR-5985*	No	MIMAT0023497	AAACACUGACUUGUAACAUCGGC
Hco-mir-5985-2-5p		No		AAACACUGACUUGUAACAUCGGC
Hco-mir-5986-5p	*Hco-miR-5986*	No	MIMAT0023499	UGGAAUGUUGUGAAUAUUCUC
Hco-mir-5987-3p	*Hco-miR-5987*	Yes	MIMAT0023501	AACCCGAUUGUGUUACAGCCUU
Hco-mir-5988-3p	*Hco-miR-5988*	Yes	MIMAT0023503	UCCCUGAGAAUUCCAUCGACGAC
Hco-mir-5989-5p	*Hco-miR-5989*	Yes	MIMAT0023505	GUAGUCACCUGGGCGAACAGACA
Hco-mir-5990-5p	*Hco-miR-5990*	Yes	MIMAT0023506	AAUCAUCAAUUUGAUGUAUUCUG
Hco-mir-5991-3p	*Hco-miR-5991*	Yes	MIMAT0023500	UGAGGUAGUAGGUGGUGUUCGA
Hco-mir-9551-3p	*Hco-miR-9551*	Yes	MIMAT0035652	TATCACAGCATTTTACTGAGCC

[1] Stem-loops are colour-coded by conservation status. Orange = unique to *H. contortus*. Blue = nematode-specific. Green = widely conserved.

[2] Indicates if microarray expression data is available.

[3] miRBase mature miRNA identification (version 20).

8.2.3 Argonaute

Nucleotide sequence of *Hco-alg-1* (HCISE0000973200; spliced):

ATGGCCGCTGGGATTGATCAGCAACAACAACAGGTTATGAGTATGTTAGATAGTTTCTCATTCAACGAGCCCGTTAACATGCTCGGTGCTG
GTGGCGGAGCTGGTGGCAATGCACCACAACCACAGCAATACATGCCGGGTATATTGGGTGGCCATATTCCTGGAATGCCAGGCTCCTTCAT
GGGCGGTCCTCCTTATCCTCATCATGCGCAGCAACAACCTCCGTTACAGGATATGTTTTCTGCTGCTCAGACAACCACGCCAATGTTACCT
GGCTCGGGTGTGCTATCTTATTCTGATGCTCCTTCACGTCAGGGAGGAAGCTTGGCCCCTGGGGCACCTATTGGTGTACCTCCGGGTCAAA
CGGATACATCTCAGGGCTCTCGCTCTCGCGCTGGAGGACAAGAACAATCCAGCGGAAGCGAACTGCCGCCGCTTTCCGGTGGCGCACATTT
CGAATGCCCACGTCGACCAAACCACGGTCTCGAAGGACGTTCCATTCTGCTGCGAGCAAACCATTTCGCAGTGCGCATGCCAGGTGGTACC
ATACAGCATTATCACGTTGATGTATCACCAGACAAATGTCCACGAAGAGTAAATCGGGAGATTATTTGCTGCATGATCCGCTCTTTTGGCA
AGTACTTCAGTACAAGTAGACCGGTCTATGATGGGAAGAGGAATATGTACACCCGGGAACCGTTGCCGATTGGACGAGAAAAGATGGAATT
TGAGGTGACTCTTCCTGGTGACTCGGCAGTGGAGCGTCAGTTTACTGTGACTGTGAAATGGGCAGGCCAAGTTTCGTTGTCAACATTAGAG
GATGCAATGGAGGGGCGCATACGTCAAGTACCTTTCGAAGCCGTACAGGCTATGGATGTCATTTTAAGACATCTTCCTAGTCTTAAGTACA
CACCTGTGGGGAGATCATTCTTCTCACCACCGGCTCACCCTCATCAGCCACCTCAACAACATGGACAGTATCACATGGAAAGTAAATTGGG
TGGTGGACGTGAAGTTTGGTTTGGATTCCATCAATCTGTTCGACCGTCGCAGTGGAAAATGATGCTTAACATTGATGTGTCTGCAACGGCA
TTCTATCGATCCATGCCCGTCATAGAGTTCGTCGCAGAAGTTTTGGAGCTTCCCGTGCAGGCTCTTACTGAACGGCGTGCTTTGTCCGATG
CGCAGCGTGTGAAGTTCACGAAGGAGATTCGCGGTCTGAAAATTGAAATCACTCATTGTGGGCAAATGCGCCGTAAGTATCGCGTTTGCAA
TGTGACAAGAAGACCTGCACAAACGCAAACGTTCCCACTACAATTGGAGACAGGTCAGACCATTGAATGCACAGTGGCGAAGTACTTCTAT
GACAAATATCGCATTCAGCTGAAGTATCCACATCTTCCATGCCTTCAGGTTGGACAAGAACAGAAGCATACTTATTTGCCACCAGAAGTAT
GTCACGTAGTGCCAGGGCAACGATGCATCAAGAAATTAACTGATACACAAACGTCGACCATGATCAAGGCAACGGCTCGCTCTGCTCCTGA
AAGAGAGCGTGAAATTGCAAGTCTTGTAAGAAAAGCCGAGTTTTCGGCGGATCCATTCGCCCATGAGTTTGGTATTGCTATCAACTCGGCT
ATGACCGAAGTGAAAGGCCGTGTCCTGTCTGCGCCAAAGTTGCAGTATGGCGGTCGTAACAAAGCGACAGCTCTGCCTAATCAGGGAGTAT
GGGATATGCGTGGAAAGCAGTTCCACACTGGAATTGATGTGAAGGTGTGGGCAATTGCCTGTTTTGCGCAACAACAACACGTCAAGGAGAA
CGATCTGCGTAATTTCACAGCTCAGCTACAACGTATATCCAACGATGCGGGTATGCCAATTGTTGGTCAGCCTTGCTTCTGCAAGTATGCC
ATGGGAGTCGATCAAGTTGAGCCGATGTTCAAATACTTGAAGCAAACTTTCTCGGGAATTCAACTTGTAGTTGTTATTCTGCCCGGCAAAA
CACCCGTCTATGCTGAAGTAAAGCGTGTAGGCGATACAGTGCTTGGAATCGCAACGCAGTGCGTCCAAGCTAAAAACGTTATCAAAACTAC
GCCACAGACTCTCTCGAATTTGTGCCTGAAGATGAACGTGAAGTTGGGTGGTGTGAACAGCATTCTTCTTCCGGCCGTTCGCCCGCGGATT
TTCAACGAACCTGTGATATTCTTTGGCTGTGATATCACACATCCTCCAGCCGGAGATTCACGTAAGCCGTCTATCGCGGCCGTTGTAGCCA
GCATGGATGCACATCCGAGCAGATACGCCGCCACGGTTCGAGTGCAACAGCACCGTCAAGAAATCATTTCGGACTTGACGTACATGGTTCG
AGAGCTGCTTGTACAATTCTATCGTAACACACGTTTCAAACCCAGGCGCATCGTTGTCTACAGAGACGGCGTTTCGGAAGGACAATTTTAT
TACGTTCTGCAGTACGAGTTGCGTGCAATGCGTGAGGCCTGTATGATGCTCGAAAGCGGCTATCAGCCGGGAATGACATTCATCGCGGTTC
AAAAACGACATCACACACGTCTATTTGCTGTAGATAAGAAGGACCAAGTTGGCAAAGCTTTCAATATTCCACCGGGTACAACTGTTGATGT
CGGCATTACTCATCCTACGGAGTTTGATTTTTACCTCTGCTCACATGCGGGTATTCAAGGAACATCTCGTCCATCGCACTATCATGTGCTA
TGGGATGATAACCAGCTTACTGCTGACGAGCTTCAGCAGCTTACGTATCAGATGTGCCACACATATGTGCGATGTACACGTTCGGTTTCGA
TACCAGCACCCGCCTATTATGCCCATCTGGTAGCTTTCCGAGCTCGGTATCATCTTGTAGACAGGGAACACGACTCAGGAGAAGGCAGTCA
GCCATCGGGAACCTCAGAAGACACTACACTTTCCAACATGGCCAGAGCCGTACAGGTGCATCCTGATGCAAATAACGTGATGTATTTTGCG
TAA

Protein sequence of Hco-ALG-1 (HCISE0000973200):

MAAGIDQQQQQVMSMLDSFSFNEPVNMLGAGGGAGGNAPQPQQYMPGILGGHIPGMPGSFMGGPPYPHHAQQQPPLQDMFSAAQTTTPMLP
GSGVLSYSDAPSRQGGSLAPGAPIGVPPGQTDTSQGSRSRAGGQEQSSGSELPPLSGGAHFECPRRPNHGLEGRSILLRANHFAVRMPGGT
IQHYHVDVSPDKCPRRVNREIICCMIRSFGKYFSTSRPVYDGKRNMYTREPLPIGREKMEFEVTLPGDSAVERQFTVTVKWAGQVSLSTLE
DAMEGRIRQVPFEAVQAMDVILRHLPSLKYTPVGRSFFSPPAHPHQPPQQHGQYHMESKLGGGREVWFGFHQSVRPSQWKMMLNIDVSATA
FYRSMPVIEFVAEVLELPVQALTERRALSDAQRVKFTKEIRGLKIEITHCGQMRRKYRVCNVTRRPAQTQTFPLQLETGQTIECTVAKYFY
DKYRIQLKYPHLPCLQVGQEQKHTYLPPEVCHVVPGQRCIKKLTDTQTSTMIKATARSAPERERE IASLVRKAEFSADPFAHEFGIAINSA
MTEVKGRVLSAPKLQYGGRNKATALPNQGVWDMRGKQFHTGIDVKVWAIACFAQQQHVKENDLRNFTAQLQRISNDAGMPIVGQPCFCKYA
MGVDQVEPMFKYLKQTFSGIQLVVVILPGKTPVYAEVKRVGDTVLGIATQCVQAKNVIKTTPQTLSNLCLKMNVKLGGVNSILLPAVRPRI
FNEPVIFFGCDITHPPAGDSRKPSIAAVVASMDAHPSRYAATVRVQQHRQEIISDLTYMVRELLVQFYRNTRFKPRRIVVYRDGVSEGQFY
YVLQYELRAMREACMMLESGYQPGMTFIAVQKRHHTRLFAVDKKDQVGKAFNIPPGTTVDVGITHPTEFDFYLCSHAGIQGTSRPSHYHVL
WDDNQLTADELQQLTYQMCHTYVRCTRSVSIPAPAYYAHLVAFRARYHLVDREHDSGEGSQPSGTSEDTTLSNMARAVQVHPDANNVMYFA

Nucleotide sequence of *Hco-alg-1* fragment used for GST-tagged recombinant protein expression:

ATGGCCGCTGGGATTGATCAGCAACAACAACAGGTTATGAGTATGTTAGATAGTTTCTCATTCAACGAGCCCGTTAACATGCTCGGTGCTG
GTGGCGGAGCTGGTGGCAATGCACCACAACCACAGCAATACATGCCGGGTATATTGGGTGGCCATATTCCTGGAATGCCAGGCTCCTTCAT
GGGCGGTCCTCCTTATCCTCATCATGCGCAGCAACAACCTCCGTTACAGGATATGTTTTCTGCTGCTCAGACAACCACGCCAATGTTACCT
GGCTCGGGTGTGCTATCTTATTCTGATGCTCCTTCACGTCAGGGAGGAAGCTTGGCCCCTGGGGCACCTATTGGTGTACCTCCGGGTCAAA
CGGATACATCTCAGGGCTCTCGCTCTCGCGCTGGAGGACAAGAACAATCCAGCGGAAGCGAACTGCCGCCGCTTTCCGGTGGCGCACATTT
CGAA

Nucleotde sequence of Hco-alg-1 fragment used for long His-tagged recombinant protein expression:

ATGGCCGCTGGGATTGATCAGCAACAACAACAGGTTATGAGTATGTTAGATAGTTTCTCATTCAACGAGCCCGTTAACATGCTCGGTGCTG
GTGGCGGAGCTGGTGGCAATGCACCACAACCACAGCAATACATGCCGGGTATATTGGGTGGCCATATTCCTGGAATGCCAGGCTCCTTCAT
GGGCGGTCCTCCTTATCCTCATCATGCGCAGCAACAACCTCCGTTACAGGATATGTTTTCTGCTGCTCAGACAACCACGCCAATGTTACCT
GGCTCGGGTGTGCTATCTTATTCTGATGCTCCTTCACGTCAGGGAGGAAGCTTGGCCCCTGGGGCACCTATTGGTGTACCTCCGGGTCAAA
CGGATACATCTCAGGGCTCTCGCTCTCGCGCTGGAGGACAAGAACAATCCAGCGGAAGCGAACTGCCGCCGCTTTCCGGTGGCGCACATTT
CGAATGCCCACGTCGACCAAACCACGGTCTCGAAGGACGTTCCATTCTGCTGCGAGCAAACCATTTCGCAGTGCGCATGCCAGGTGGTACC
ATACAGCATTATCACGTTGATGTATCACCAGACAAATGTCCACGAAGAGTAAATCGGGAGATTATTTGCTGCATGATCCGCTCTTTTGGCA
AGTACTTCAGTACAAGTAGACCGGTCTATGATGGGAAGAGGAATATGTACACCCGGGAACCGTTGCCGATTGGACGAGAAAAGATGGAATT
TGAGGTGACTCTTCCTGGTGACTCGGCAGTGGAGCGTCAGTTTACTGTGACTGTGAAATGGGCAGGCCAAGTTTCGTTGTCAACATTAGAG
GATGCAATGGAGGGGCGCATACGTCAAGTACCTTTCGAAGCCGTACAGGCTATGGATGTCATTTTAAGACATCTTCCTAGTCTTAAGTACA
CACCTGTGGGGAGATCATTCTTCTCACCACCGGCTCACCCTCATCAGCCACCTCAACAACATGGACAGTATCACATGGAAAGTAAATTGGG
TGGTGGACGTGAAGTTTGGTTTGGATTCCATCAATCTGTTCGACCGTCGCAGTGGAAAATGATGCTTAACATTGATGTGTCTGCAACGGCA
TTCTATCGATCCATGCCCGTCATAGAGTTCGTCGCAGAAGTTTTGGAGCTTCCCGTGCAGGCTCTTACTGAACGGCGTGCTTTGTCCGATG
CGCAGCGTGTGAAGTTCACGAAGGAGATTCGCGGTCTGAAAATTGAAATCACTCATTGTGGGCAAATGCGCCGTAAGTATCGCGTTTGCAA
TGTGACAAGAAGACCTGCACAAACGCAAACGTTCCCACTACAATTGGAGACAGGTCAGACCATTGAATGCACAGTGGCGAAGTACTTCTAT
GACAAATATCGCATTCAGCTGAAGTATCCACATCTTCCATGCCTTCAGGTTGGACAAGAACAGAAGCATACTTATTTGCCACCAGAAGTAT
GTCACGTAGTGCCAGGGCAACGATGCATCAAGAAATTAACTGATACACAAACGTCGACCATGATCAAGGCAACGGCTCGCTCTGCTCCTGA
AAGAGAGCGTGAAATTGCAAGTCTTGTAAGAAAAGCCGAGTTTTCGGCGGATCCATTCGCCCATGAGTTTGGTATTGCTATCAACTCGGCT
ATGACCGAAGTGAAAGGCCGTGTCCTGTCTGCGCCAAAGTTGCAGTATGGCGGTCGTAACAAAGCGACAGCTCTGCCTAATCAGGGAGTAT
GGGATATGCGTGGAAAGCAGTTCCACACTGGAATTGATGTGAAGGTGTGGGCAATTGCCTGTTTTGCGCAACAACAACACGTCAAGGAGAA
CGATCTG

Nucleotide sequence of *Hco-alg-1* fragment optimised for expression in E. coli. Includes codon for six histidine residues (His-tag):

ATGGCCGCAGGTATTGATCAGCAGCAGCAACAGGTTATGAGCATGCTGGATAGCTTTAGCTTTAATGAACCGGTTAATATGCTGGGTGCCG
GTGGTGGTGCGGGTGGTAATGCACCGCAGCCGCAGCAGTATATGCCTGGTATTCTGGGTGGTCATATTCCGGGTATGCCAGGTAGCTTTAT
GGGTGGTCCGCCTTATCCGCATCATGCACAACAGCAGCCTCCGCTGCAGGATATGTTTAGCGCAGCACAGACCACCACCCCGATGCTGCCT
GGTAGCGGTGTTCTGAGCTATAGTGATGCACCGAGTCGTCAGGGTGGTAGCCTGGCACCGGGTGCACCGATTGGTGTTCCGCCTGGTCAGA
CCGATACCAGCCAGGGTAGCCGTAGCCGTGCAGGCGGTCAAGAACAGAGCAGCGGTAGCGAACTGCCACCGCTGTCAGGTGGTGCACATTT
TGAATGTCCGCGTCGTCCGAATCATGGTCTGGAAGGTCGTAGCATTCTGCTGCGTGCAAATCATTTTGCAGTTCGCATGCCTGGTGGCACC
ATTCAGCATTATCATGTTGATGTTAGTCCGGATAAATGCCCTCGTCGTGTTAATCGTGAAATTATCTGCTGTATGATTCGCAGCTTCGGCA
AATATTTCAGCACCAGCCGTCCGGTTTATGATGGTAAACGTAATATGTATACCCGTGAACCGCTGCCGATTGGTCGTGAAAAAATGGAATT
TGAAGTTACCCTGCCTGGTGATAGCGCAGTTGAACGTCAGTTTACCGTTACCGTTAAATGGGCAGGTCAGGTTAGCCTGAGCACCCTGGAA
GATGCAATGGAAGGCCGTATTCGTCAGGTTCCGTTTGAAGCAGTTCAGGCAATGGATGTTATTCTGCGTCATCTGCCGAGCCTGAAATATA
CACCGGTTGGTCGTAGCTTTTTTAGCCCTCCGGCACATCCTCATCAGCCACCGCAGCAGCATGGTCAGTATCATATGGAAAGCAAACTGGG
TGGTGGCCGTGAAGTTTGGTTTGGTTTTCATCAGAGCGTTCGTCCGAGCCAGTGGAAAATGATGCTGAATATTGATGTGAGCGCAACCGCA
TTTTATCGTAGCATGCCGGTTATTGAATTTGTTGCAGAAGTTCTGGAACTGCCGGTTCAGGCACTGACCGAACGTCGTGCACTGTCAGATG
CACAGCGTGTGAAATTTACCAAAGAAATTCGCGGTCTGAAAATCGAAATTACCCATTGTGGTCAGATGCGTCGTAAATATCGTGTTTGTAA
TGTTACCCGTCGTCCGGCACAGACCCAGACCTTTCCGCTGCAACTGGAAACCGGTCAGACCATTGAATGTACCGTTGCCAAATACTTCTAT
GATAAATATCGCATCCAGCTGAAATACCCGCATCTGCCGTGTCTGCAGGTTGGTCAAGAGCAGAAACATACCTATCTGCCTCCGGAAGTTT
GTCATGTTGTTCCGGGTAGATCTCATCATCATCACCACCAC

8.4 Gene ontology

The following tables detail enriched gene ontology terms for predicted miRNA targets in *C. elegans* and *H. contortus*. Enriched annotations (p-value < 0.01) were identified in three categories: biological processes (BP), molecular functions (MF) and cellular component (CC). Background = number of annotations among all *C. elegans* genes. Sample = number of annotations in gene list submitted. Expected = number of genes expected to possess an annotation by chance in an equally sized sample list.

Table 8-5: Enriched gene ontology terms for 46 predicted miRNA targets of Cel-let-7.

	GO Term	Frequency			Fold enrichment	p-value
		Background	Sample	Expected		
	Unclassified	10,443	12	23	0.51	0.00E+00
	Regulation of transcription, DNA-templated	806	14	2	> 5	2.10E-06
	Regulation of nucleic acid-templated transcription	806	14	2	> 5	2.10E-06
	Regulation of RNA biosynthetic process	812	14	2	> 5	2.31E-06
	Regulation of RNA metabolic process	823	14	2	> 5	2.75E-06
	Regulation of nitrogen compound metabolic process	846	14	2	> 5	3.90E-06
	Regulation of nucleobase-containing compound metabolic process	846	14	2	> 5	3.90E-06
	Regulation of cellular macromolecule biosynthetic process	877	14	2	> 5	6.17E-06
	Regulation of macromolecule biosynthetic process	883	14	2	> 5	6.72E-06
	Regulation of cellular biosynthetic process	887	14	2	> 5	7.12E-06
	Regulation of biosynthetic process	888	14	2	> 5	7.22E-06
	Cellular response to organic substance	401	10	1	> 5	2.43E-05
	Cellular response to chemical stimulus	412	10	1	> 5	3.13E-05
	Response to organic substance	415	10	1	> 5	3.34E-05
	Regulation of gene expression	1,032	14	2	> 5	4.77E-05
BP	Regulation of primary metabolic process	1,062	14	2	> 5	6.80E-05
	Regulation of cellular metabolic process	1,075	14	2	> 5	7.91E-05
	Transcription, DNA-templated	600	11	1	> 5	9.63E-05
	Nucleic acid-templated transcription	605	11	1	> 5	1.05E-04
	RNA biosynthetic process	610	11	1	> 5	1.14E-04
	Regulation of metabolic process	1,324	15	3	> 5	1.45E-04
	Response to chemical	639	11	1	> 5	1.81E-04
	Regulation of macromolecule metabolic process	1,199	14	3	> 5	3.02E-04
	Regulation of developmental process	690	11	2	> 5	3.88E-04
	Cellular response to organic cyclic compound	295	8	1	> 5	4.09E-04
	Response to organic cyclic compound	295	8	1	> 5	4.09E-04
	Response to hormone	296	8	1	> 5	4.20E-04
	Cellular response to hormone stimulus	296	8	1	> 5	4.20E-04
	Cellular response to endogenous stimulus	312	8	1	> 5	6.23E-04
	Response to endogenous stimulus	313	8	1	> 5	6.38E-04
	RNA metabolic process	909	12	2	> 5	7.53E-04
	Nucleobase-containing compound biosynthetic process	751	11	2	> 5	8.94E-04
	Aromatic compound biosynthetic process	773	11	2	> 5	1.19E-03
	Heterocycle biosynthetic process	784	11	2	> 5	1.36E-03
	Cellular nitrogen compound biosynthetic process	793	11	2	> 5	1.52E-03

	Organic cyclic compound biosynthetic process	803	11	2	> 5	1.72E-03
	Gene expression	997	12	2	> 5	1.99E-03
	Cellular macromolecule biosynthetic process	1,013	12	2	> 5	2.34E-03
	Anatomical structure development	3,855	23	9	2.67	2.41E-03
	Macromolecule biosynthetic process	1,020	12	2	> 5	2.52E-03
	Biological regulation	3,584	22	8	2.74	2.95E-03
	Single-organism developmental process	4,579	25	10	2.44	3.19E-03
	Developmental process	4,593	25	10	2.43	3.38E-03
	Reproductive process	1,496	14	3	4.18	4.25E-03
	Cellular response to lipid	285	7	1	> 5	4.65E-03
	Cellular response to steroid hormone stimulus	285	7	1	> 5	4.65E-03
	Steroid hormone mediated signalling pathway	285	7	1	> 5	4.65E-03
	Response to lipid	285	7	1	> 5	4.65E-03
	Response to steroid hormone	285	7	1	> 5	4.65E-03
	Regulation of biological process	3,074	20	7	2.91	4.70E-03
	Hormone-mediated signalling pathway	287	7	1	> 5	4.86E-03
	Nucleic acid metabolic process	1,112	12	2	4.82	6.14E-03
	Single-organism process	7,476	32	17	1.91	7.01E-03
	Multi-organism process	934	11	2	> 5	7.33E-03
MF	Unclassified	12,036	15	27	0.56	0.00E+00
	DNA binding	872	16	2	> 5	2.65E-08
	Sequence-specific DNA binding transcription factor activity	490	13	1	> 5	3.21E-08
	Nucleic acid binding transcription factor activity	490	13	1	> 5	3.21E-08
	Sequence-specific DNA binding	477	12	1	> 5	3.74E-07
	Nucleic acid binding	1,492	17	3	> 5	7.96E-06
	Heterocyclic compound binding	2,816	21	6	3.33	1.30E-04
	Organic cyclic compound binding	2,820	21	6	3.33	1.33E-04
	Steroid hormone receptor activity	285	7	1	> 5	2.69E-03
	Binding	5,100	26	11	2.28	3.69E-03

Table 8-6: Enriched gene ontology terms for 16 predicted miRNA targets of Cel-lin-4.

	GO term	Frequency			Fold enrichment	p-value
		Background	Sample	Expected		
BP	Unclassified	10,443	2	8	-0.26	0.00E+00
	Biological_regulation	3,584	12	3	4.59	2.96E-04
	Regulation of biological process	3,074	11	2	4.9	8.75E-04
	Single organism signaling	1,338	8	1	>5	1.94E-03
	Signaling	1,338	8	1	>5	1.94E-03
	Cell communication	1,373	8	1	>5	2.36E-03

Table 8-7: Enriched gene ontology terms for 64 predicted miRNA targets of Cel-miR-60.

	Term	Frequency			Fold Enrichment	p-value
		Background	Sample	Expected		
BP	Unclassified	10,443	11	32	-0.34	0.00E+00
	Single-organism cellular process	4,617	39	14	2.76	3.06E-08
	Cellular process	5,864	41	18	2.28	2.72E-06
	Localization	2,792	28	9	3.27	3.53E-06
	Single-organism process	7,476	46	23	2.01	5.20E-06
	Establishment of localization	2,152	24	7	3.64	1.09E-05
	Transport	2,078	23	6	3.61	2.97E-05
	Biological_process	10,105	52	31	1.68	5.92E-05
	Single-organism localization	1,331	18	4	4.41	8.64E-05
	Regulation of biological process	3,074	26	9	2.76	5.82E-04
	Regulation of cellular process	2,472	23	8	3.03	7.13E-04
	Single-organism transport	1,256	16	4	4.15	1.23E-03
	Biological regulation	3,584	27	11	2.46	3.17E-03
MF	Unclassified	12,036	17	37	-0.46	0.00E+00
	Molecular_function	8,512	46	26	1.76	3.08E-04
	Purine ribonucleoside triphosphate binding	1,116	14	3	4.09	4.72E-03
	Purine nucleoside binding	1,116	14	3	4.09	4.72E-03
	Purine ribonucleoside binding	1,116	14	3	4.09	4.72E-03
	Ribonucleoside binding	1,120	14	3	4.08	4.92E-03
	Nucleoside binding	1,121	14	3	4.07	4.97E-03
	Purine ribonucleotide binding	1,123	14	3	4.07	5.07E-03
	Ribonucleotide binding	1,139	14	3	4.01	5.96E-03

	Purine nucleotide binding	1,140	14	4	4.01	6.02E-03
	Carbohydrate derivative binding	1,180	14	4	3.87	8.89E-03
CC	Unclassified	13,903	21	43	-0.49	0.00E+00
	Plasma membrane	587	14	2	> 5	1.02E-06
	Cell periphery	658	14	2	> 5	4.30E-06
	Cellular_component	6,645	42	20	2.06	1.06E-05
	Cell part	4,261	31	13	2.37	2.03E-04
	Cell	4,272	31	13	2.37	2.15E-04
	Membrane	3,075	25	9	2.65	6.97E-04
	Plasma membrane part	286	8	1	> 5	1.17E-03
	Apical plasma membrane	45	4	0	> 5	5.24E-03
	Cytoplasm	2,125	19	7	2.92	5.37E-03

Table 8-8: Enriched gene ontology terms for 49 predicted miRNA targets of Cel-miR-228-5p.

	GO term	Frequency			Fold enrichment	p-value
		Background	Sample	Expected		
BP	Unclassified	10,443	10	25	-0.39	0.00E+00
	Single-organism cellular process	4,617	26	11.2	2.31	7.07E-03
CC	Unclassified	13,903	18	34	-0.53	0.00E+00
	Cellular_component	6,645	32	16	1.98	1.85E-03

Table 8-9: Enriched gene ontology terms for 61 predicted miRNA targets of Hco-miR-60-3p.

	GO term	Frequency			Fold Enrichment	p-value
		Background	Sample	Expected		
BP	Unclassified	10,443	16	46	-0.35	0.00E+00
	Biological_process	10,105	75	45	1.68	6.95E-08
	Single-organism process	7,476	62	33	1.87	1.11E-06
	Single-organism developmental process	4,579	46	20	2.27	4.79E-06
	Developmental process	4,593	46	20	2.26	5.31E-06
	Multicellular organismal process	4,750	46	21	2.19	1.62E-05
	Single-multicellular organism process	4,634	45	21	2.19	2.50E-05
	Multicellular organismal development	4,338	43	19	2.24	3.66E-05
	Anatomical structure development	3,855	40	17	2.34	4.48E-05
	Metabolic process	4,939	45	22	2.06	1.87E-04
	Cellular process	5,864	48	26	1.85	1.49E-03
	Protein metabolic process	1,771	23	8	2.93	2.88E-03
	Organic substance metabolic process	3,614	35	16	2.19	2.91E-03
	Macromolecule metabolic process	2,828	30	13	2.4	3.55E-03
	Primary metabolic process	3,507	34	16	2.19	4.33E-03
	Embryo development	3,049	31	14	2.3	5.51E-03
	Nematode larval development	2,013	24	9	2.69	7.21E-03
	Larval development	2,015	24	9	2.69	7.34E-03
	Post-embryonic development	2,032	24	9	2.67	8.47E-03
MF	Unclassified	12,036	25	53	-0.47	0.00E+00
	Molecular_function	8,512	66	38	1.75	1.51E-06
	Ion binding	2,936	34	13	2.61	3.58E-05
	Catalytic activity	3,840	37	17	2.18	8.04E-04
	Structural molecule activity	460	11	2	> 5	5.27E-03
	Binding	5,100	42	23	1.86	6.63E-03

Table 8-10: Enriched gene ontology terms for 52 predicted miRNA targets of Hco-miR-228-5p.

	GO term	Frequency			Fold Enrichment	p-value
		Background	Sample	Expected		
BP	Unclassified	10,443	16	44	-0.36	0.00E+00
	Biological_process	10,105	71	43	1.66	5.48E-07
	Single-organism process	7,476	58	32	1.83	1.36E-05
	Single-organism developmental process	4,579	43	19	2.22	3.62E-05
	Developmental process	4,593	43	19	2.21	3.98E-05
	Multicellular organismal process	4,750	43	20	2.14	1.11E-04
	Single-multicellular organism process	4,634	42	20	2.14	1.75E-04
	Multicellular organismal development	4,338	40	18	2.18	2.79E-04
	Metabolic process	4,939	43	21	2.06	3.58E-04
	Anatomical structure development	3,855	37	16	2.27	4.02E-04
	Organic substance metabolic process	3,614	34	15	2.22	2.65E-03
	Macromolecule metabolic process	2,828	29	12	2.42	4.03E-03
	Primary metabolic process	3,507	33	15	2.22	4.09E-03
	Protein metabolic process	1,771	22	8	2.93	4.72E-03
	Cellular process	5,864	45	25	1.81	6.38E-03
MF	Unclassified	12,036	25	51	0.49	0.00E+00

	Molecular_function	8,512	62	36	1.72	1.41E-05
	Ion binding	2,936	32	12	2.57	1.34E-04
	Catalytic activity	3,840	35	16	2.15	2.09E-03

Table 8-11: Enriched gene ontology terms for 68 predicted miRNA targets of Hco-miR-235-3p.

GO term		Frequency			Fold Enrichment	p-value
		Background	Sample	Expected		
BP	Unclassified	10,443	10	26	-0.38	0.00E+00
	Biological_process	10,105	42	26	1.64	3.72E-03
	Anatomical structure development	3,855	24	10	2.46	8.67E-03

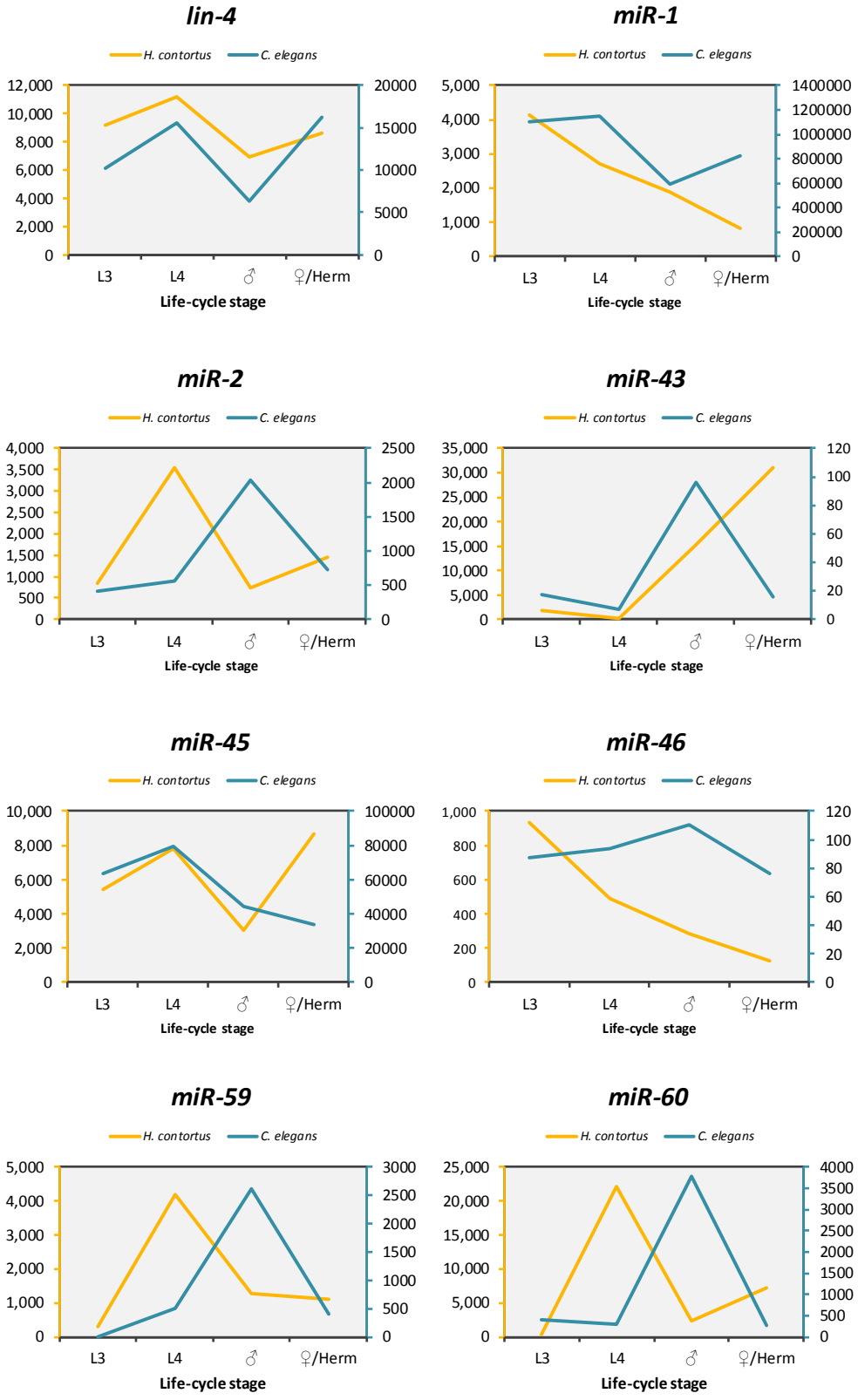
Table 8-12: Enriched gene ontology terms for 29 predicted miRNA targets of Hco-miR-5885-3p.

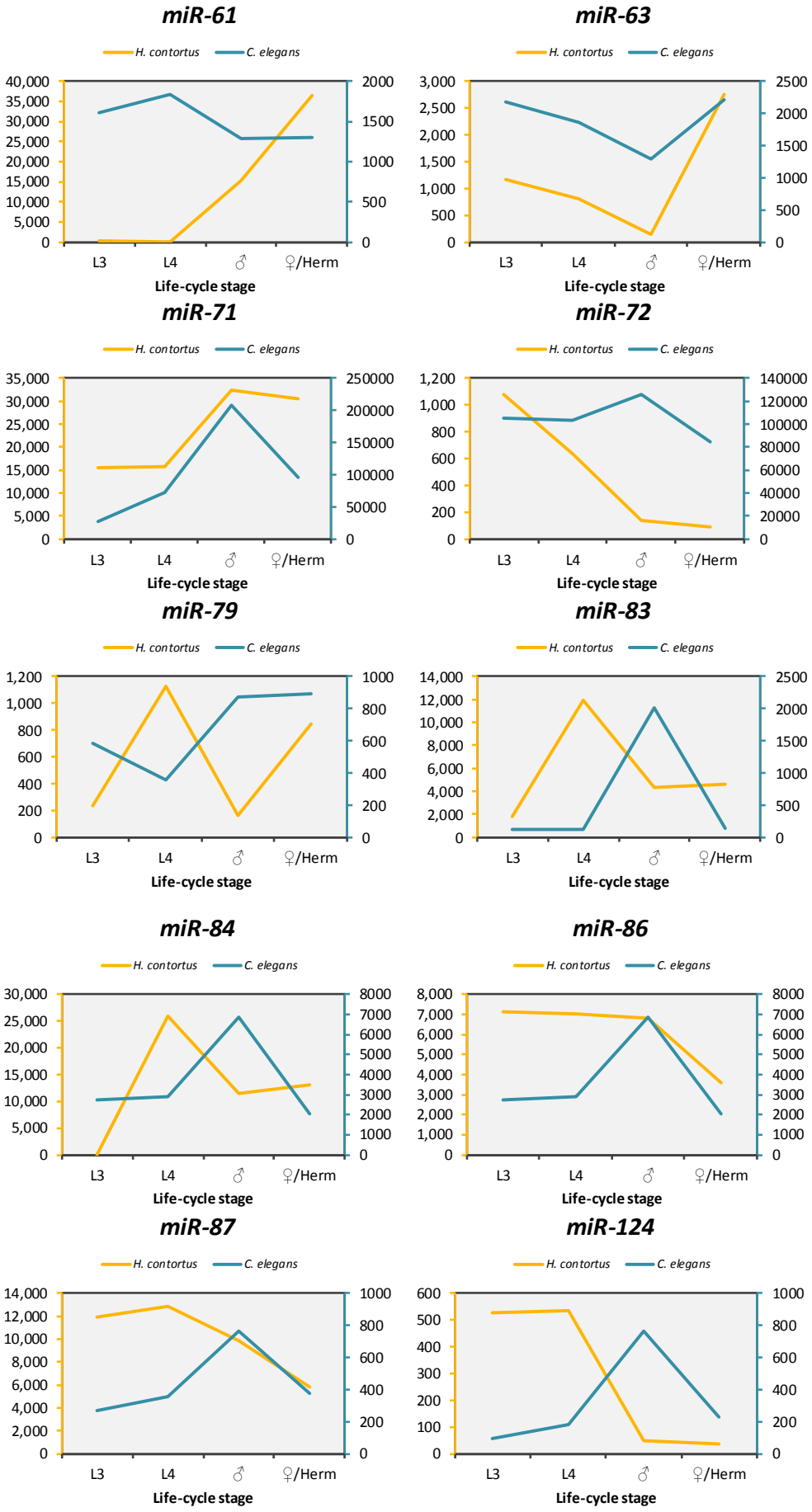
GO term		Frequency			Fold enrichment	p-value
		Background	Sample	Expected		
MF	Unclassified	12036	7	15.82	-0.44	0.00E+00
	oxidoreductase activity, acting on CH or CH2 groups	3	2	0	> 5	6.11E-03
BP	Unclassified	13903	12	18.27	-0.66	0.00E+00
CC	Unclassified	10443	7	13.72	-0.51	0.00E+00

Table 8-13: Enriched gene ontology terms for 38 predicted miRNA targets of Hco-miR-5885-5p.

GO term		Frequency			Fold enrichment	p-value
		Background	Sample	Expected		
MF	Unclassified	12036	13	20.5	-0.63	0.00E+00
BP	regulation of multicellular organismal process	734	10	1.25	> 5	3.86E-04
	regulation of biological process	3074	17	5.24	3.25	3.95E-03
	regulation of protein modification process	87	4	0.15	> 5	2.14E-02
	regulation of developmental process	690	8	1.18	> 5	2.38E-02
	biological regulation	3584	17	6.1	2.78	3.28E-02
CC	Unclassified	13903	15	23.68	-0.63	0.00E+00
	cell part	4261	18	7.26	2.48	2.41E-02
	cell	4272	18	7.28	2.47	2.50E-02

8.5 Miscellaneous




miR-61
H. contortus
C. elegans
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
2000
1500
1000
500
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-63
H. contortus
C. elegans
3,000
2,500
2,000
1,500
1,000
500
0
2500
2000
1500
1000
500
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-71
H. contortus
C. elegans
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
250000
200000
150000
100000
50000
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-72
H. contortus
C. elegans
1,200
1,000
800
600
400
200
0
140000
120000
100000
80000
60000
40000
20000
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-79
H. contortus
C. elegans
1,200
1,000
800
600
400
200
0
1000
800
600
400
200
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-83
H. contortus
C. elegans
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
2500
2000
1500
1000
500
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-84
H. contortus
C. elegans
30,000
25,000
20,000
15,000
10,000
5,000
0
8000
7000
6000
5000
4000
3000
2000
1000
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-86
H. contortus
C. elegans
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
8000
7000
6000
5000
4000
3000
2000
1000
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-87
H. contortus
C. elegans
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
1000
800
600
400
200
0
L3
L4
♂
♀/Herm
Life-cycle stage
miR-124
H. contortus
C. elegans
600
500
400
300
200
100
0
1000
800
600
400
200
0
L3
L4
♂
♀/Herm
Life-cycle stage

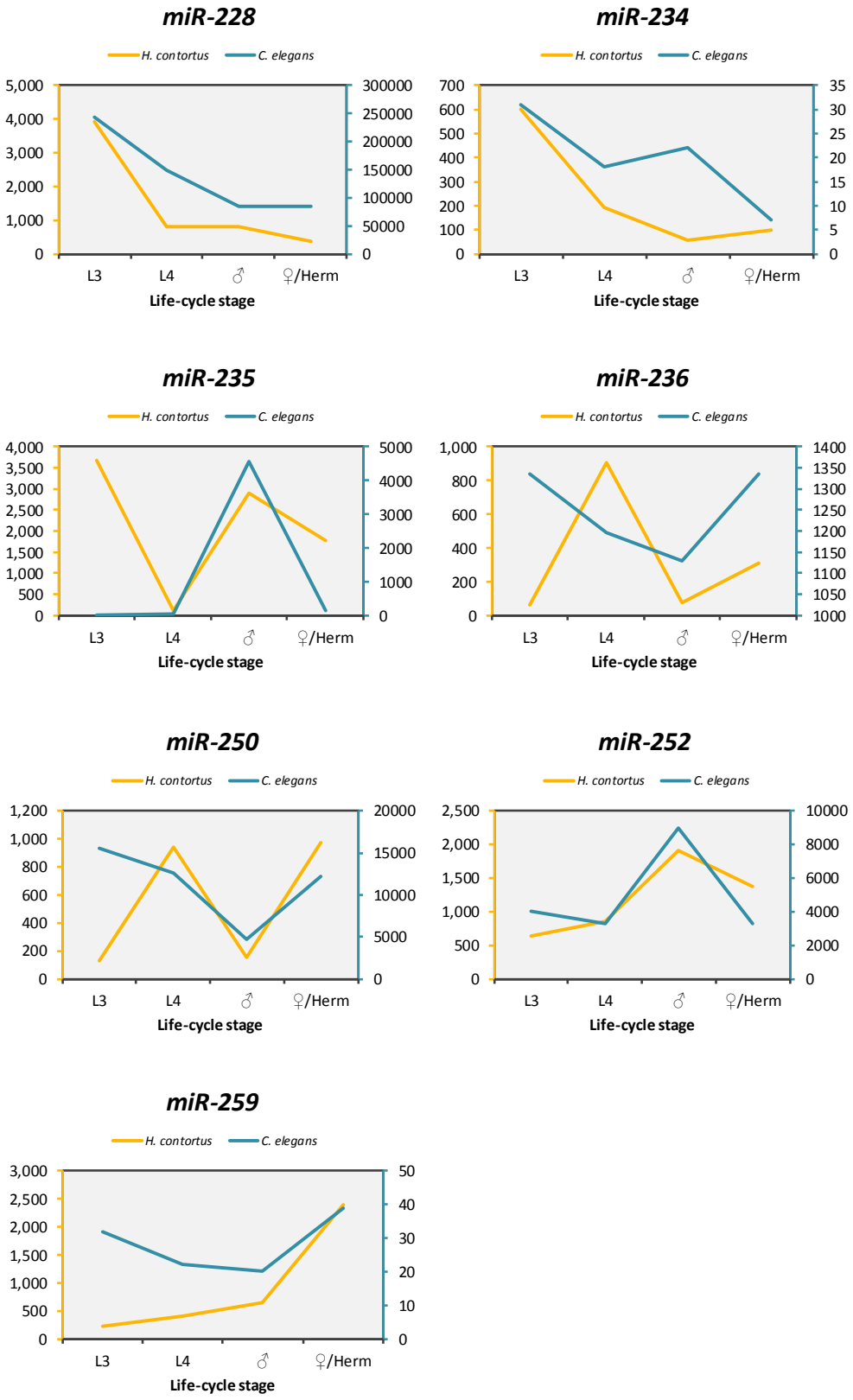


Fig. 8-1: Expression of miRNAs conserved in *C. elegans* and *H. contortus*. The *H. contortus* data (yellow) was obtained using microarrays (see Chapter 3), while the *C. elegans* data (blue) was generated by Kato *et al.* (2009) using RNAseq.

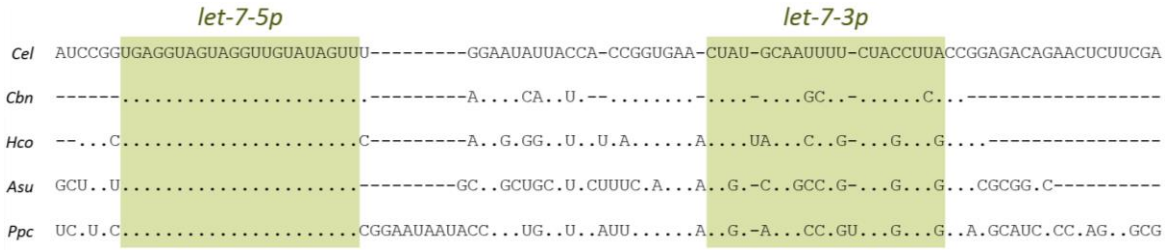


Fig. 8-2: Alignment of *let-7* miRNA stem-loops. Sequences from five nematode species (two free-living and three parasitic) aligned in Geneious (version 5) using BLASTn. *Cel* = *C. elegans*, *Cbn* = *Caenorhabditis brenneri*, *Hco* = *H. contortus*, *Asu* = *Ascaris suum*, *Ppc* = *Pristionchus pacificus*. Dots (.) represent bases identical the *C. elegans* sequence, while dashes (-) represent gaps. The 5' and 3' arms are represented by green squares.

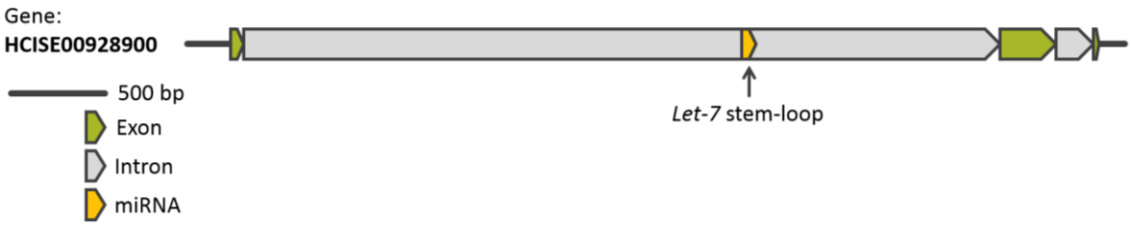


Fig. 8-3: Position of the *H. contortus let-7* miRNA stem-loop. The stem-loop is within the first intron of putative gene HCISE00928900 which comprises three exons and two introns over 4,659. Together, the exons generate a product of 438 bp (146 aa).

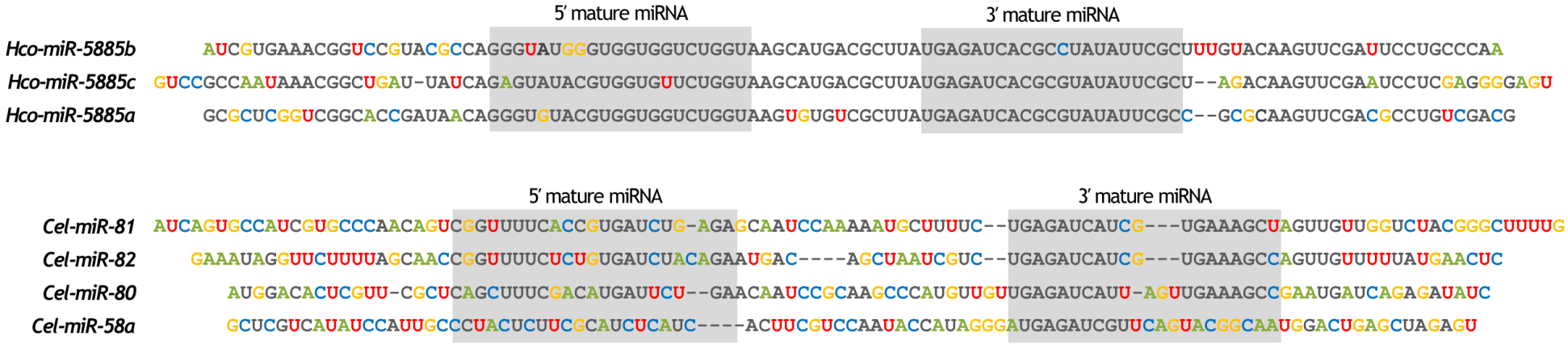


Fig. 8-4: Alignment of *miRNAs* in the *Bantam* family in *H. contortus* and *C. elegans*. A: Alignment of *Hco-miR-5885a/b/c* stem-loops. B: Alignment of Cel-miR-58a, miR-80, miR-81 and miR-82. All sequences taken from miRBase and aligned in Geneious (version 5) using Clustalw. Nucleotides in the majority at a specific position are coloured black, otherwise they are colour-coded: A = green, C = blue, U = red, G = yellow.

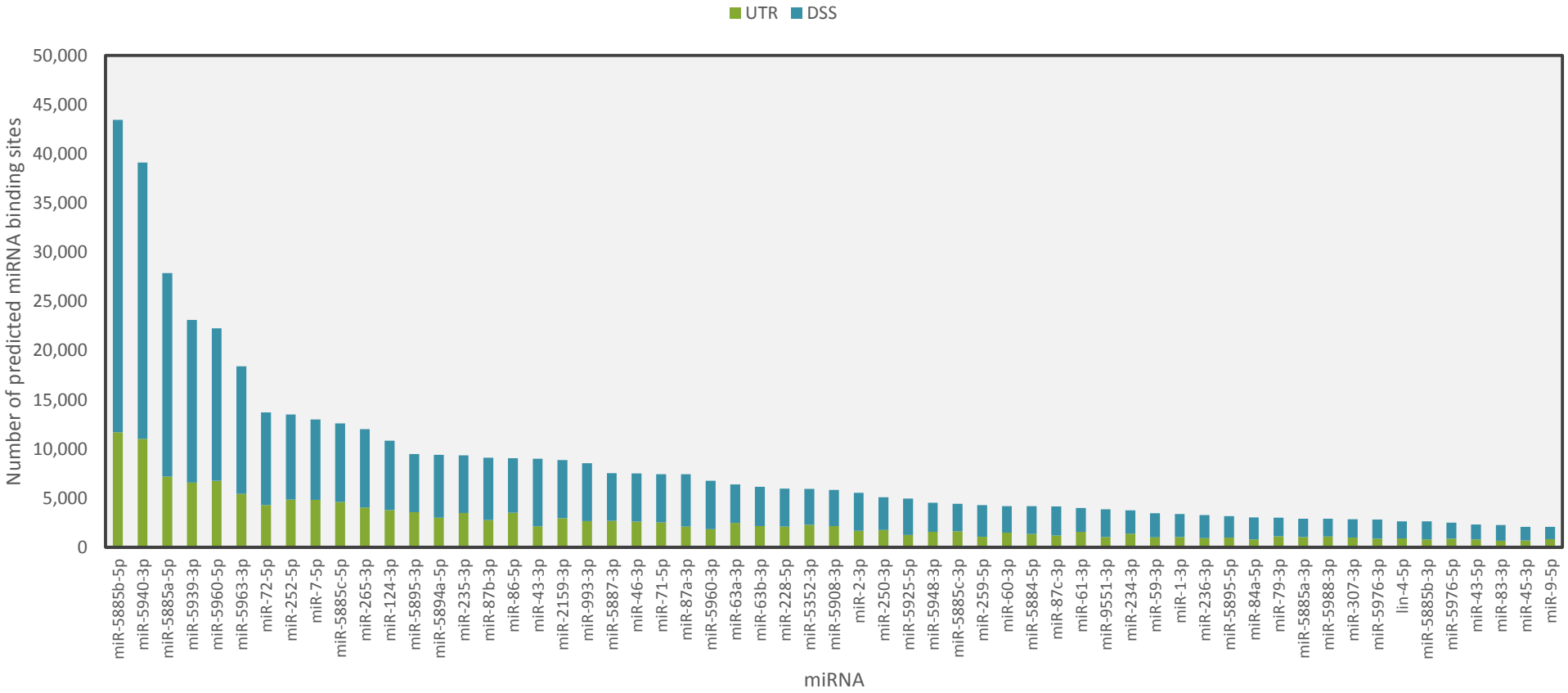


Fig. 8-5: *H. contortus* miRNA binding sites predicted *in silico*, broken down by sequence database. Two databases were interrogated for binding sites: 3' UTR (untranslated region) and 3' DSS (downstream sequence). Three algorithms were used to predict binding sites (miRanda, PITA and RNAhybrid). All 60 miRNAs displayed significant temporal or spatial variation in expression across L3, L4, adult male and adult female stages as well as adult female gut tissue.

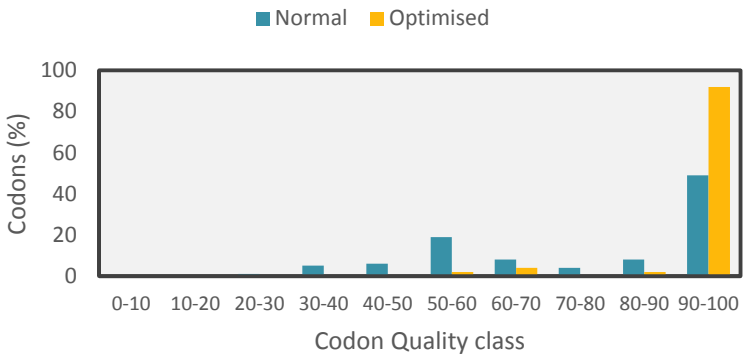


Fig. 8-6: Difference in codon quality of *Hco-alg-1* sequence before (normal) and after GeneArt® codon optimisation. The quality value of the most frequently used codon for a given amino acid in *E. coli* is set to 100, the remaining codons are scaled accordingly

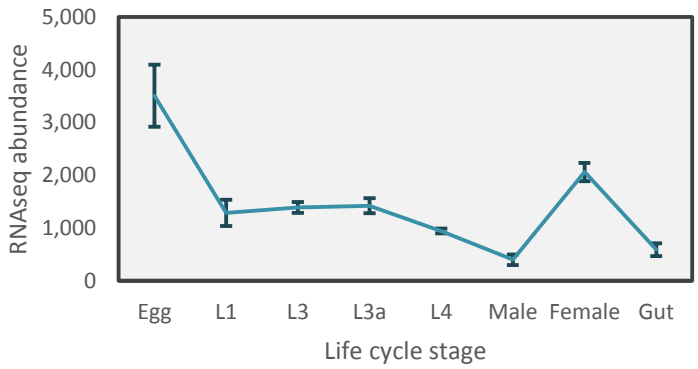


Fig. 8-7: RNAseq expression data for *Hco-alg-1*. Normalised abundance values generated from triplicate RNAseq data provided by R. Laing. Vertical bars represent ± 1 standard deviation from the mean.

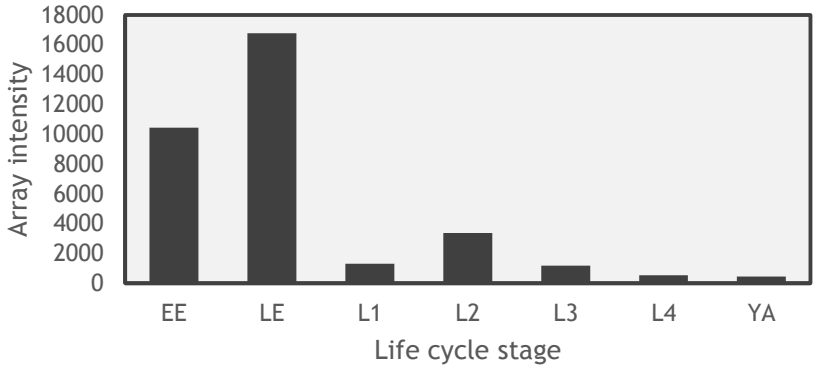


Fig. 8-8: Expression of *Cel-hbl-1*. Data obtained from tiling arrays performed by Spencer *et al.* *(*2011) and available online at WormViz (http://www.vanderbilt.edu/wormdoc/wormmap/WormViz.html). EE = early embryo. LE = late embryo. YA = young adult hermaphrodite.

Bibliography

Abbott AL, **Alvarez-Saavedra E**, **Miska EA**, **Lau NC**, **Bartel DP**, **Horvitz HR** and **Ambros V (2005)** The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. *Dev. Cell*. 9(3): 403–14.

Abbott EM, **Parkins JJ** and **Holmes PH (1988)** Influence of dietary protein on the pathophysiology of haemonchosis in lambs given continuous infections. *Res. Vet. Sci.* 4541–49.

Abrahante JE, **Daul AL**, **Li M**, **Volk ML**, **Tennessen JM**, **Miller EA** and **Rougvie AE (2003)** The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. *Dev. Cell*. 4(5): 625–37.

Ai L, **Xu MJ**, **Chen MX**, **Zhang YN**, **Chen SH**, **Guo J**, **Cai YC**, **Zhou XN**, **Zhu XQ** and **Chen JX (2012)** Characterization of microRNAs in Taenia saginata of zoonotic significance by Solexa deep sequencing and bioinformatics analysis. *Parasitol. Res.* 110(6): 2373–8.

Akat KM, Moore-McGriff D, Morozov P, Brown M, Gogakos T, Correa Da Rosa J, Mihailovic A, Sauer M, Ji R & other authors (2014) Comparative RNA-sequencing analysis of myocardial and circulating small RNAs in human heart failure and their utility as biomarkers. *Proc. Natl. Acad. Sci. U. S. A.* 11111151–6.

Alexiou P, **Maragkakis M**, **Papadopoulos GL**, **Reczko M** and **Hatzigeorgiou AG (2009)** Lost in translation: An assessment and perspective for computational microrna target identification. *Bioinformatics*.

Alkharouf NW, **Klink VP**, **Chouikha IB**, **Beard HS**, **MacDonald MH**, **Meyer S**, **Knap HT**, **Khan R** and **Matthews BF (2006)** Timecourse microarray analyses reveal global changes in gene expression of susceptible Glycine max (soybean) roots during infection by Heterodera glycines (soybean cyst nematode). *Planta*. 224(4): 838–852.

Alvarez-Saavedra E and **Horvitz HR (2010)** Many families of C. elegans microRNAs are not essential for development or viability. *Curr. Biol.* 20(4): 367–73. editor & translator, Elsevier Ltd.

Ambros V and **Horvitz H (1984)** Heterochronic mutants of the nematode Caenorhabditis elegans. *Science (80-.).* 226(4673): 409–416.

Ambros V (2011) MicroRNAs and developmental timing. *Curr. Opin. Genet. Dev.* 21(4): 511–517. editor & translator, Elsevier Ltd.

Anonymous (2014) Monepantel resistance reported on Dutch sheep farms. *Vet. Rec.* 175(17): 418.

Antebi A, **Yeh WH**, **Tait D**, **Hedgecock EM** and **Riddle DL (2000)** daf-12 encodes a nuclear receptor that regulates the dauer diapause and developmental age in C. elegans. *Genes Dev.* 14(12): 1512–27.

Ashrafi K, **Chang FY**, **Watts JL**, **Fraser AG**, **Kamath RS**, **Ahringer J** and **Ruvkun G (2003)** Genome-wide RNAi analysis of Caenorhabditis elegans fat regulatory genes. *Nature*. 421268–272.

Baker N, **Alsford S** and **Horn D (2011)** Genome-wide RNAi screens in African trypanosomes identify the nifurtimox activator NTR and the eflornithine transporter AAT6. *Mol. Biochem. Parasitol.* 17655–57.

Banks DJD, **Singh R**, **Barger IA**, **Pratap B** and **Le Jambre LF (1990)** Development and survival of infective larvae of Haemonchus contortus and Trichostrongylus colubriformis on pasture in a tropical environment. *Int. J. Parasitol.* 20(2): 155–160.

Bannister SC, **Tizard ML V**, **Doran TJ**, **Sinclair AH** and **Smith CA (2009)** Sexually dimorphic microRNA expression during chicken embryonic gonadal development. *Biol. Reprod.* 81(1): 165–76.

Barger I a and **Cox HW (1984)** Wool production of sheep chronically infected with Haemonchus contortus. *Vet. Parasitol.* 15(2): 169–75.

Barger IA and **Le Jambre LF (1979)** The role of inhibited larvae in the epidemiology of ovine haemonchosis. *Aust. Vet. J.* 55(12): 580–3.

Barger I and **Southcott W (1978)** Parasitism and production in weaner sheep grazing alternately with cattle. *Aust. J. Exp. Agric.* 18(92): 340.

Bargmann CI and **Horvitz HR (1991)** Control of larval development by chemosensory neurons in Caenorhabditis elegans. *Science*. 251(4998): 1243–1246.

Bassetto CC, **Silva MRL**, **Newlands GFJ**, **Smith WD**, **Ratti Júnior J**, **Martins CL** and **Amarante AFT (2014a)** Vaccination of grazing calves with antigens from the intestinal membranes of haemonchus contortus: Effects against natural challenge with haemonchus placei and haemonchus similis. *Int. J. Parasitol.* 44(10): 697–702. editor & translator, Australian Society for Parasitology Inc.

Bassetto CC, **Picharillo MÉ**, **Newlands GFJ**, **Smith WD**, **Fernandes S**, **Siqueira ER** and **Amarante a. FT (2014b)** Attempts to vaccinate ewes and their lambs against natural infection with Haemonchus contortus in a tropical environment. *Int. J. Parasitol.* 44(14): 1049–1054. editor & translator, Australian Society for Parasitology Inc.

Berberian JF and **Mizelle JD (1957)** Developmental Studies on Haemonchus contortus Rudolphi (1803). *Am. Midl. Nat.* 57(2): 421.

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.

Besier RB and **Dunsmore JD (1993)** The ecology of Haemonchus contortus in a winter rainfall region in Australia: the development of eggs to infective larvae. *Vet. Parasitol.* 45(3-4): 275–292.

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.

Bettinger JC, **Lee K** and **Rougvie AE (1996)** Stage-specific accumulation of the terminal differentiation factor LIN-29 during

Caenorhabditis elegans development. *Development*. 122(8): 2517–27.

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.

Boag B and **Thomas RJ (1977)** Epidemiological studies on gastro-intestinal nematode parasites of sheep: the seasonal number of generations and succession of species. *Res. Vet. Sci*. 2262–67.

Boisvenue RJ, **Stiff MI**, **Tonkinson L V.**, **Cox GN** and **Hageman R (1992)** Fibrinogen-degrading proteins from Haemonchus contortus used to vaccinate sheep. *Am. J. Vet. Res.* 531263–1265.

Boomker J, **Horak IG** and **De Vos V (1986)** The helminth parasites of various artiodactylids from some South African nature reserves. *Onderstepoort J. Vet. Res.* 53(2): 93–102.

Borchert GM, **Lanier W** and **Davidson BL (2006)** RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* 13(12): 1097–101.

Bouasker S and **Simard MJ (2012)** The slicing activity of miRNA-specific Argonautes is essential for the miRNA pathway in C. elegans. *Nucleic Acids Res.* 40(20): 10452–10462.

Brennecke J, **Hipfner DR**, **Stark A**, **Russell RB** and **Cohen SM (2003)** bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell*. 113(1): 25–36.

Brennecke J, **Stark A**, **Russell RB** and **Cohen SM (2005)** Principles of microRNA-target recognition. In *PLoS Biol.*, pages 0404–0418.

Brenner JL, **Jasiewicz KL**, **Fahley AF**, **Kemp BJ** and **Abbott AL (2010)** Loss of individual microRNAs causes mutant phenotypes in sensitized genetic backgrounds in C. elegans. *Curr. Biol.* 20(14): 1321–5. editor & translator, Elsevier Ltd.

Brenner S (1974) The genetics of Caenorhabditis elegans. *Genetics*. 77(1): 71–94.

Britton C, **Samarasinghe B** and **Knox DP (2012)** Ups and downs of RNA interference in parasitic nematodes. *Exp. Parasitol.* 13256–61.

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.*

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.* 189299–307.

Burke SL, **Hammell M** and **Ambros V (2015)** Robust Distal Tip Cell Pathfinding in the Face of Temperature Stress Is Ensured by Two Conserved microRNAs in Caenorhabditis elegans. *Genetics*. 200(4): 1201–1218.

Butcher R a, **Fujita M**, **Schroeder FC** and **Clardy J (2007)** Small-molecule pheromones that control dauer development in Caenorhabditis elegans. *Nat. Chem. Biol.* 3(7): 420–2.

Butz H, **Szabó P**, **Khella H**, **Nofech-Mozes R**, **Patocs A** and **Yousef G (2015)** miRNA-target network reveals miR-124as a key miRNA contributing to clear cell renal cell carcinoma aggressive behaviour by targeting CAV1 and FLOT1. *Oncotarget*. 6(14): 12543–57.

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.

Calura E, **Martini P**, **Sales G**, **Beltrame L**, **Chiorino G**, **D'Incalci M**, **Marchini S** and **Romualdi C (2014)** Wiring miRNAs to pathways: A topological approach to integrate miRNA and mRNA expression profiles. *Nucleic Acids Res.* 42(11): .

Camon E, Magrane M, Barrell D, Binns D, Fleischmann W, Kersey P, Mulder N, Oinn T, Maslen J & other authors (2003) The Gene Ontology Annotation (GOA) project: Implementation of GO in SWISS-PROT, TrEMBL, and interpro. *Genome Res.*

Carlsen AL, Schetter AJ, Nielsen CT, Lood C, Knudsen S, Voss A, Harris CC, Hellmark T, Segelmark M & other authors (2013) Circulating microRNA expression profiles associated with systemic lupus erythematosus. *Arthritis Rheum.* 65(5): 1324–34.

Cassada RC and **Russell RL (1975)** The dauerlarva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans. *Dev. Biol.* 46(2): 326–342.

Cerutti L, **Mian N** and **Bateman A (2000)** Domains in gene silencing and cell differentiation proteins: tTe novel PAZ domain and redefinition of the Piwi domain. *Trends Biochem. Sci*.

Chalfie M (1981) Mutations that lead to reiterations in the cell lineages of C. elegans. *Cell*. 24(1): 59–69.

Chang S, **Johnston RJ**, **Frøkjaer-Jensen C**, **Lockery S** and **Hobert O (2004)** MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature*. 430(7001): 785–9.

Chen H, **Lan H-Y**, **Roukos DH** and **Cho WC (2014a)** Application of microRNAs in diabetes mellitus. *J. Endocrinol.* 222(1): R1–R10.

Chen MX, Ai L, Xu MJ, Chen SH, Zhang YN, Guo J, Cai YC, Tian LG, Zhang LL & other authors (2011) Identification and characterization of microRNAs in Trichinella spiralis by comparison with Brugia malayi and Caenorhabditis elegans. *Parasitol. Res.* 109(3): 553–8.

Chen X, Li Q, Wang J, Guo X, Jiang X, Ren Z, Weng C, Sun G, Wang X & other authors (2009) Identification and characterization of novel amphioxus microRNAs by Solexa sequencing. *Genome Biol.* 10(7): R78.

Chen X, **Li Z-Y**, **Maleewong W**, **Maleewong P**, **Liang J**, **Zeng X**, **Zheng H**, **Wu Z-D** and **Sun X (2014b)** Serum aca-mir-146a is a potential biomarker for early diagnosis of Angiostrongylus cantonensis infection. *Parasitol. Res.* 113(9): 3221–7.

Chen X-P, **Chen Y-G**, **Lan J-Y** and **Shen Z-J (2014c)** MicroRNA-370 suppresses proliferation and promotes endometrioid ovarian cancer chemosensitivity to cDDP by negatively regulating ENG. *Cancer Lett.* 353(2): 201–210.

Chi SW, **Zang JB**, **Mele A** and **Darnell RB (2009)** Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*. 460(7254): 479–486.

Clark PM, **Loher P**, **Quann K**, **Brody J**, **Londin ER** and **Rigoutsos I (2014)** Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types. *Sci. Rep.* 45947.

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.

Cohen R, **Eastoe R**, **Hotson I** and **Smeal M (1972)** Sheep production on the north coast of New South Wales. 1. Effect of anthelmintics and grazing management on production and helminthosis of Merino wethers. *Aust. J. Exp. Agric.* 12(56): 247–251.

Coronnello C, Hartmaier R, Arora A, Huleihel L, Pandit K V., Bais AS, Butterworth M, Kaminski N, Stormo GD & other authors (2012) Novel Modeling of Combinatorial miRNA Targeting Identifies SNP with Potential Role in Bone Density. *PLoS Comput. Biol.* 8(12): .

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.

Cox DN, **Chao A**, **Baker J**, **Chang L**, **Qiao D** and **Lin H (1998)** A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.* 12(23): 3715–3727.

Crook M (2014) The dauer hypothesis and the evolution of parasitism: 20years on and still going strong. *Int. J. Parasitol.*

Crosby ME, **Kulshreshtha R**, **Ivan M** and **Glazer PM (2009)** MicroRNA regulation of DNA repair gene expression in hypoxic stress. *Cancer Res.* 69(3): 1221–9.

Cucher M, **Prada L**, **Mourglia-Ettlin G**, **Dematteis S**, **Camicia F**, **Asurmendi S** and **Rosenzvit M (2011)** Identification of Echinococcus granulosus microRNAs and their expression in different life cycle stages and parasite genotypes. *Int. J. Parasitol.* 41(3-4): 439–48.

Cucher M, **Macchiaroli N**, **Kamenetzky L**, **Maldonado L**, **Brehm K** and **Rosenzvit M (2015)** High-throughput characterization of Echinococcus spp. metacestode miRNomes. *Int. J. Parasitol.* 45(4): 253–267. editor & translator, Australian Society for Parasitology Inc.

Curry E, **Safranski TJ** and **Pratt SL (2011)** Differential expression of porcine sperm microRNAs and their association with sperm morphology and motility. *Theriogenology*. 76(8): 1532–9.

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.

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A & other authors (2012) The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* 221775–1789.

Devaney E, **Winter AD** and **Britton C (2010)** microRNAs: a role in drug resistance in parasitic nematodes? *Trends Parasitol.* 26(9): 428–33. editor & translator, Elsevier Ltd.

Diaz SA, **Brunet V**, **Lloyd-Jones GC**, **Spinner W**, **Wharam B** and **Viney M (2014)** Diverse and potentially manipulative signalling with ascarosides in the model nematode C. elegans. *BMC Evol. Biol.* 1446.

Dilda F, **Gioia G**, **Pisani L**, **Restelli L**, **Lecchi C**, **Albonico F**, **Bronzo V**, **Mortarino M** and **Ceciliani F (2012)** Escherichia coli lipopolysaccharides and Staphylococcus aureus enterotoxin B differentially modulate inflammatory microRNAs in bovine monocytes. *Vet. J.* 192(3): 514–6.

Ding X, **Ye J**, **Wu X**, **Huang L**, **Zhu L** and **Lin S (2015)** Deep sequencing analyses of pine wood nematode Bursaphelenchus xylophilus microRNAs reveal distinct miRNA expression patterns during the pathological process of pine wilt disease. *Gene*. 555(2): 346–356. editor & translator, Elsevier B.V.

Dinnik, J.A. and Dinnik NN (1958) Observations on the development of Haemonchus contortus larvae under field conditions in the Kenya highlands. *Bull. Epizoot. Dis. Africa*. 611–21.

Dobson RJ, **Hosking BC**, **Besier RB**, **Love S**, **Larsen JWA**, **Rolfe PF** and **Bailey JN (2011)** Minimising the development of anthelmintic resistance, and optimising the use of the novel anthelmintic monepantel, for the sustainable control of nematode parasites in Australian sheep grazing systems. *Aust. Vet. J.* 89(5): 160–6.

Doench JG and **Sharp PA (2004)** Specificity of microRNA target selection in translational repression. *Genes Dev.* 18(5): 504–511.

Donald AD (1968) Ecology of the free-living stages of nematode parasites of sheep. *Aust. Vet. J.* 44(4): 139–144.

Drudge JH, **Leland SE** and **Wyant ZN (1957)** Strain variation in the response of sheep nematodes to the action of phenothiazine. II. Studies on pure infections of Haemonchus contortus. *Am. J. Vet. Res.* 18(67): 317–325.

Du Z, **He F**, **Yu Z**, **Bowerman B** and **Bao Z (2015)** E3 ubiquitin ligases promote progression of differentiation during C. elegans embryogenesis. *Dev. Biol.* 398267–279.

Dye-Holden L and **Walker RJ (1990)** Avermectin and avermectin derivatives are antagonists at the 4-aminobutyric acid (GABA) receptor on the somatic muscle cells of Ascaris; is this the site of anthelmintic action? *Parasitology*. 101(02): 265.

Ecsedi M, **Rausch M** and **Großhans H (2015)** The let-7 microRNA Directs Vulval Development through a Single Target. *Dev. Cell*. 32(3): 335–344. editor & translator, Elsevier Inc.

Elbashir SM, **Harborth J** and **Lendeckel W (2001)** Duplexes of 21 ± nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*. 4111–5.

Elling AA, Mitreva M, Recknor J, Gai X, Martin J, Maier TR, McDermott JP, Hewezi T, McK Bird D & other authors (2007) Divergent evolution of arrested development in the dauer stage of Caenorhabditis elegans and the infective stage of Heterodera glycines. *Genome Biol.* 8(10): R211.

Elmén J, Lindow M, Schütz S, Lawrence M, Petri A, Obad S, Lindholm M, Hedtjärn M, Hansen HF & other authors (2008) LNA-mediated microRNA silencing in non-human primates. *Nature*. 452896–899.

Eulalio A and **Mano M (2015)** MicroRNA Screening and the Quest for Biologically Relevant Targets. *J. Biomol. Screen.*

Fang Z and **Rajewsky N (2011)** The impact of miRNA target sites in coding sequences and in 3???UTRs. *PLoS One*. 6(3): .

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L & other authors (2014) Pfam: The protein families database. *Nucleic Acids Res.*

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*. 391806–811.

Frank F, **Sonenberg N** and **Nagar B (2010)** Structural basis for 5’-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature*. 465(7299): 818–822.

Fraser AG, **Kamath RS**, **Zipperlen P**, **Martinez-Campos M**, **Sohrmann M** and **Ahringer J (2000)** Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. *Nature*. 408325–330.

Friedländer MR, **Chen W**, **Adamidi C**, **Maaskola J**, **Einspanier R**, **Knespel S** and **Rajewsky N (2008)** Discovering microRNAs from deep sequencing data using miRDeep. *Nat. Biotechnol.* 26(4): 407–415.

Friedman RC, **Farh KKH**, **Burge CB** and **Bartel DP (2009)** Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19(1): 92–105.

Fu Y, Lan J, Wu X, Yang D, Zhang Z, Nie H, Hou R, Zhang R, Zheng W & other authors (2013) Identification of Dirofilaria immitis miRNA using illumina deep sequencing. *Vet. Res.* 44(1): 3. editor & translator, Veterinary Research.

Fujiwara T and **Yada T (2013)** miRNA-target prediction based on transcriptional regulation. *BMC Genomics*. 14 Suppl 2(Suppl 2): S3.

Fukushige T, **Siddiqui ZK**, **Chou M**, **Culotti JG**, **Gogonea CB**, **Siddiqui SS** and **Hamelin M (1999)** MEC-12, an alpha-tubulin required for touch sensitivity in C. elegans. *J. Cell Sci.* 112 (Pt 3395–403.

Garcia DM, **Baek D**, **Shin C**, **Bell GW**, **Grimson A** and **Bartel DP (2011)** Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat. Struct. Mol. Biol.*

Geldhof P, **Visser A**, **Clark D**, **Saunders G**, **Britton C**, **Gilleard J**, **Berriman M** and **Knox D (2007)** RNA interference in parasitic helminths: current situation, potential pitfalls and future prospects. *Parasitology*. 134609–619.

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.* 36801–810.

Gibson TE and **Everett G (1976)** The ecology of the free-living stages on Haemonchus contortus. *Br. Vet. J.* 132(1): 50–9.

Gibson T (1973) Recent advances in the epidemiology and control of parasitic gastroenteritis in sheep. *Vet. Rec.* 92469–473.

Gitlin L, **Karelsky S** and **Andino R (2002)** Short interfering RNA confers intracellular antiviral immunity in human cells. *Nature*. 418430–434.

Glazov EA, **Kongsuwan K**, **Assavalapsakul W**, **Horwood PF**, **Mitter N** and **Mahony TJ (2009)** Repertoire of bovine miRNA and miRNA-like small regulatory RNAs expressed upon viral infection. *PLoS One*. 4(7): e6349.

Golden JW and **Riddle DL (1982)** A pheromone influences larval development in the nematode Caenorhabditis elegans. *Science (80-.).* 218(4572): 578–80.

Golden JW and **Riddle DL (1984)** The Caenorhabditis elegans dauer larva: developmental effects of pheromone, food, and temperature. *Dev. Biol.* 102(2): 368–78.

Gomez IG, Mackenna D, Roach AM, Johnson BG, Xia T-H, Kaimal V, Borza D-B, Zhang J, Liu S & other authors (2013) Anti-miR21 Protects Collagen 4A3 Defi cient Mice from Progression of Alport Disease by Decreasing Oxidative Stress. In *Kidney Week*.

Goodrich JA and **Kugel JF (2006)** Non-coding-RNA regulators of RNA polymerase II transcription. *Nat. Rev. Mol. Cell Biol.* 7612–616.

Griffiths-Jones S, **Hui JHL**, **Marco A** and **Ronshaugen M (2011)** MicroRNA evolution by arm switching. *EMBO Rep.* 12(2): 172–177.

Grimson A, **Farh KKH**, **Johnston WK**, **Garrett-Engele P**, **Lim LP** and **Bartel DP (2007)** MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Mol. Cell*. 27(1): 91–105.

Grosshans H, **Johnson T**, **Reinert KL**, **Gerstein M** and **Slack FJ (2005)** The temporal patterning microRNA let-7 regulates several transcription factors at the larval to adult transition in C. elegans. *Dev. Cell*. 8(3): 321–30.

Guo L, **Zhao Y**, **Yang S**, **Zhang H** and **Chen F (2014)** An integrated analysis of miRNA, lncRNA, and mRNA expression profiles. *Biomed Res. Int.* 2014.

Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp AC & other authors (2010) Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell*. 141(1): 129–141.

Hamilton AJ and **Baulcombe DC (1999)** A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*.

286950–952.

Hamilton B, **Dong Y**, **Shindo M**, **Liu W**, **Odell I**, **Ruvkun G** and **Lee SS (2005)** A systematic RNAi screen for longevity genes in C. elegans. *Genes Dev.*

Hammell CM, **Karp X** and **Ambros V (2009)** A feedback circuit involving let-7-family miRNAs and DAF-12 integrates environmental signals and developmental timing in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* 106(44): 18668–73.

Hammer M, **Mages J**, **Dietrich H**, **Servatius A**, **Howells N**, **Cato ACB** and **Lang R (2006)** Dual specificity phosphatase 1 (DUSP1) regulates a subset of LPS-induced genes and protects mice from lethal endotoxin shock. *J. Exp. Med.* 203(1): 15–20.

Hammond SM, **Boettcher S**, **Caudy a a**, **Kobayashi R** and **Hannon GJ (2001)** Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science*. 293(5532): 1146–50.

Haussecker D and **Kay MA (2010)** miR-122 continues to blaze the trail for microRNA therapeutics. *Mol. Ther.* 18240–242.

Hedgecock EM and **White JG (1985)** Polyploid tissues in the nematode Caenorhabditis elegans. *Dev. Biol.* 107(1): 128–133.

Hewezi T, **Howe P**, **Maier TR** and **Baum TJ (2008)** Arabidopsis small RNAs and their targets during cyst nematode parasitism. *Mol. Plant. Microbe. Interact.* 21(12): 1622–1634.

Hiszczyńska-Sawicka E, **Gatkowska JM**, **Grzybowski MM** and **Długońska H (2014)** Veterinary vaccines against toxoplasmosis. *Parasitology*. FirstView1–14.

Hoberg ER, **Lichtenfels JR** and **Gibbons L (2004)** Phylogeny for species of Haemonchus (Nematoda: Trichostrongyloidea): considerations of their evolutionary history and global biogeography among Camelidae and Pecora (Artiodactyla). *J. Parasitol.* 90(5): 1085–102.

Hon LS and **Zhang Z (2007)** The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression. *Genome Biol.* 8(8): R166.

Horwich MD and **Zamore PD (2008)** Design and delivery of antisense oligonucleotides to block microRNA function in cultured Drosophila and human cells. *Nat. Protoc.* 3(10): 1537–1549.

Hoy AM and **Buck AH (2012)** Extracellular small RNAs: what, where, why? *Biochem. Soc. Trans.* 40(4): 886–90.

Hoy AM, Lundie RJ, Ivens A, Quintana JF, Nausch N, Forster T, Jones F, Kabatereine NB, Dunne DW & other authors (2014a) Parasite-Derived MicroRNAs in Host Serum As Novel Biomarkers of Helminth Infection. *PLoS Negl. Trop. Dis.* 8(2): .

Hoy AM, Lundie RJ, Ivens A, Quintana JF, Nausch N, Forster T, Jones F, Kabatereine NB, Dunne DW & other authors (2014b) Parasite-Derived MicroRNAs in Host Serum As Novel Biomarkers of Helminth Infection. *PLoS Negl. Trop. Dis.* 8(2): e2701 (W. E. Secor, Ed.).

Hu W, **Wang T**, **Yue E**, **Zheng S** and **Xu J-H (2014)** Flexible microRNA arm selection in rice. *Biochem. Biophys. Res. Commun.* 447(3): 526–530.

Huang Q-X, **Cheng X-Y**, **Mao Z-C**, **Wang Y-S**, **Zhao L-L**, **Yan X**, **Ferris VR**, **Xu R-M** and **Xie B-Y (2010)** MicroRNA discovery and analysis of pinewood nematode Bursaphelenchus xylophilus by deep sequencing. *PLoS One*. 5(10): e13271.

Hudder A and **Novak RF (2008)** miRNAs: Effectors of Environmental Influences on Gene Expression and Disease. *Toxicol. Sci.* 103(2): 228–240.

Hur K, **Toiyama Y**, **Schetter AJ**, **Okugawa Y**, **Harris CC**, **Boland CR** and **Goel A (2015)** Identification of a Metastasis-Specific MicroRNA Signature in Human Colorectal Cancer. *JNCI J. Natl. Cancer Inst.* 107(3): dju492–dju492.

Hutvágner G, **McLachlan J**, **Pasquinelli a E**, **Bálint E**, **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*. 293(5531): 834–8.

Hutvagner G and **Simard MJ (2008)** Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* 9(1): 22–32.

Hydbring P and **Badalian-Very G (2013)** Clinical applications of microRNAs. *F1000Research*. 1–15.

Ibáñez-Ventoso C, **Yang M**, **Guo S**, **Robins H**, **Padgett RW** and **Driscoll M (2006)** Modulated microRNA expression during adult lifespan in Caenorhabditis elegans. *Aging Cell*. 5(3): 235–246.

Iqbal J, Shen Y, Huang X, Liu Y, Wake L, Liu C, Deffenbacher K, Lachel CM, Wang C & other authors (2015) Global microRNA expression profiling uncovers molecular markers for classification and prognosis in aggressive B-cell lymphoma. *Blood*. 125(7): 1137–1145.

Ithal N, **Recknor J**, **Nettleton D**, **Hearne L**, **Maier T**, **Baum TJ** and **Mitchum MG (2007)** Parallel genome-wide expression profiling of host and pathogen during soybean cyst nematode infection of soybean. *Mol. Plant. Microbe. Interact.* 20(3): 293–305.

Ivey KN and **Srivastava D (2015)** microRNAs as Developmental Regulators. *Cold Spring Harb. Perspect. Biol.* 7(7): .

Iyer A, **Zurolo E**, **Prabowo A**, **Fluiter K**, **Spliet WGM**, **van Rijen PC**, **Gorter JA** and **Aronica E (2012)** MicroRNA-146a: a key regulator of astrocyte-mediated inflammatory response. *PLoS One*. 7(9): e44789.

Izrayelit Y, **Srinivasan J**, **Campbell SL**, **Jo Y**, **von Reuss SH**, **Genoff MC**, **Sternberg PW** and **Schroeder FC (2012)** Targeted metabolomics reveals a male pheromone and sex-specific ascaroside biosynthesis in Caenorhabditis elegans. *ACS Chem. Biol.* 7(8): 1321–5.

Jackson RJ and **Standart N (2007)** How do microRNAs regulate gene expression? *Sci. STKE*. 2007(367): re1.

Jain S and **Parker R (2013)** Ten Years of Progress in GW/P Body Research. *New York*. 76823–43.

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 (80-.).* 344(6188): 1168–1173.

Jin W, **Dodson M V**, **Moore SS**, **Basarab JA** and **Guan LL (2010)** Characterization of microRNA expression in bovine adipose

tissues: a potential regulatory mechanism of subcutaneous adipose tissue development. *BMC Mol. Biol.* 1129.

Jin W, **Ibeagha-Awemu EM**, **Liang G**, **Beaudoin F**, **Zhao X** and **Guan LL (2014)** Transcriptome microRNA profiling of bovine mammary epithelial cells challenged with Escherichia coli or Staphylococcus aureus bacteria reveals pathogen directed microRNA expression profiles. *BMC Genomics*. 15181.

Jing Q, **Huang S**, **Guth S**, **Zarubin T**, **Motoyama A**, **Chen J**, **Di Padova F**, **Lin S-C**, **Gram H** and **Han J (2005)** Involvement of MicroRNA in AU-Rich Element-Mediated mRNA Instability. *Cell*. 120(5): 623–634.

John B, **Enright AJ**, **Aravin A**, **Tuschl T**, **Sander C** and **Marks DS (2004)** Human microRNA targets. *PLoS Biol.* 2(11): .

Johnson SM, **Lin SY** and **Slack FJ (2003)** The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Dev. Biol.* 259(2): 364–79.

Johnson TE, **Mitchell DH**, **Kline S**, **Kemal R** and **Foy J (1984)** Arresting development arrests aging in the nematode Caenorhabditis elegans. *Mech. Ageing Dev.* 28(1): 23–40.

Johnston M and **Hutvagner G (2011)** Posttranslational modification of Argonautes and their role in small RNA-mediated gene regulation. *Silence*.

Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, Kanapin A, Le Bot N, Moreno S & other authors (2003) Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. *Nature*. 421(6920): 231–237.

Kaminsky R (2003) Drug resistance in nematodes: a paper tiger or a real problem? *Curr. Opin. Infect. Dis.* 16(6): 559–64.

Kaminsky R, Ducray P, Jung M, Clover R, Rufener L, Bouvier J, Weber SS, Wenger A, Wieland-Berghausen S & other authors (2008) A new class of anthelmintics effective against drug-resistant nematodes. *Nature*. 452(7184): 176–80.

Kaplan F, Srinivasan J, Mahanti P, Ajredini R, Durak O, Nimalendran R, Sternberg PW, Teal PE a, Schroeder FC & other authors (2011) Ascaroside expression in Caenorhabditis elegans is strongly dependent on diet and developmental stage. *PLoS One*. 6(3): e17804.

Kaplan RM (2004) Drug resistance in nematodes of veterinary importance: a status report. *Trends Parasitol.* 20(10): 477–81.

Karbiener M, **Glantschnig C** and **Scheideler M (2014a)** Hunting the needle in the haystack: a guide to obtain biologically meaningful microRNA targets. *Int. J. Mol. Sci.* 15(11): 20266–89.

Karbiener M, **Glantschnig C** and **Scheideler M (2014b)** Hunting the Needle in the Haystack: A Guide to Obtain Biologically Meaningful MicroRNA Targets. *Int. J. Mol. Sci.* 15(11): 20266–20289.

Karp X and **Ambros V (2012)** Dauer larva quiescence alters the circuitry of microRNA pathways regulating cell fate progression in C. elegans. *Development*. 139(12): 2177–86.

Karp X, **Hammell M**, **Ow MC** and **Ambros V (2011)** Effect of life history on microRNA expression during C. elegans development. *RNA*. 17(4): 639–51.

Kasuga H, **Fukuyama M**, **Kitazawa A**, **Kontani K** and **Katada T (2013)** The microRNA miR-235 couples blast-cell quiescence to the nutritional state. *Nature*. 497(7450): 503–6. editor & translator, Nature Publishing Group.

Kato M, **de Lencastre A**, **Pincus Z** and **Slack FJ (2009)** Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during Caenorhabditis elegans development. *Genome Biol.* 10(5): R54.

Keller O, **Kollmar M**, **Stanke M** and **Waack S (2011)** A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics*. 27(6): 757–763.

Kertesz M, **Iovino N**, **Unnerstall U**, **Gaul U** and **Segal E (2007)** The role of site accessibility in microRNA target recognition. *Nat. Genet.* 39(10): 1278–84.

Khvorova A, **Reynolds A** and **Jayasena SD (2003)** Functional siRNAs and miRNAs exhibit strand bias. *Cell*. 115(2): 209–16.

Kim DH, **Saetrom P**, **Snøve O** and **Rossi JJ (2008)** MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 105(42): 16230–16235.

Kimble J and **Sharrock WJ (1983)** Tissue-specific synthesis of yolk proteins in Caenorhabditis elegans. *Dev. Biol.* 96(1): 189–196.

Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, Ishihara G, Kawaoka S, Sugano S & other authors (2014) A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature*. 509633–6.

Klass M and **Hirsh D (1976)** Non-ageing developmental variant of Caenorhabditis elegans. *Nature*. 260(5551): 523–525.

Kloosterman WP, **Wienholds E**, **de Bruijn E**, **Kauppinen S** and **Plasterk RHA (2006)** In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat. Methods*. 3(1): 27–29.

Kluiver J, **Slezak-Prochazka I**, **Smigielska-Czepiel K**, **Halsema N**, **Kroesen BJ** and **van den Berg A (2012)** Generation of miRNA sponge constructs. *Methods*. 58(2): 113–117.

Knox DP and **Jones DG (1990)** Studies on the presence and release of proteolytic enzymes (proteinases) in gastro-intestinal nematodes of ruminants. *Int. J. Parasitol.* 20243–249.

Knox D (2011) Proteases in blood-feeding nematodes and their potential as vaccine candidates. *Adv. Exp. Med. Biol.* 712155–176.

Kozomara A and **Griffiths-Jones S (2011)** MiRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39(SUPPL. 1): .

Krause RM, **Buisson B**, **Bertrand S**, **Corringer PJ**, **Galzi JL**, **Changeux JP** and **Bertrand D (1998)** Ivermectin: a positive allosteric effector of the alpha7 neuronal nicotinic acetylcholine receptor. *Mol. Pharmacol.* 53(2): 283–94.

Krützfeldt J, **Rajewsky N**, **Braich R**, **Rajeev KG**, **Tuschl T**, **Manoharan M** and **Stoffel M (2005)** Silencing of microRNAs in vivo with ‘antagomirs’. *Nature*. 438(7068): 685–9.

Kuchenbauer F, Mah SM, Heuser M, McPherson A, Ruschmann J, Rouhi A, Berg T, Bullinger L, Argiropoulos B & other authors (2011) Comprehensive analysis of mammalian miRNA* species and their role in myeloid cells. *Blood*. 118(12): 3350–3358.

Kudlow B a, **Zhang L** and **Han M (2012)** Systematic Analysis of Tissue-Restricted miRISCs Reveals a Broad Role for MicroRNAs in Suppressing Basal Activity of the C. elegans Pathogen Response. *Mol. Cell*. 46(4): 530–41. editor & translator, Elsevier Inc.

Kwak PB and **Tomari Y (2012)** The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat. Struct. Mol. Biol.* 19(2): 145–51.

Lacey E (1990) Mode of action of benzimidazoles. *Parasitol. Today*. 6(4): 112–5.

Lacey E and **Gill JH (1994)** Biochemistry of benzimidazole resistance. *Acta Trop.* 56(2-3): 245–62.

Lagendijk AK, **Moulton JD** and **Bakkers J (2012)** Revealing details: whole mount microRNA in situ hybridization protocol for zebrafish embryos and adult tissues. *Biol. Open*. 1(6): 566–569.

Laing R, Kikuchi T, Martinelli A, Tsai IJ, Beech RN, Redman E, Holroyd N, Bartley DJ, 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.

Larsen PL (1993) Aging and resistance to oxidative damage in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* 90(19): 8905–9.

Leathwick DM, **Ganesh S** and **Waghorn TS (2015)** Evidence for reversion towards anthelmintic susceptibility in Teladorsagia circumcincta in response to resistance management programmes. *Int. J. Parasitol. Drugs Drug Resist.* 59–15.

Lee RC, **Feinbaum RL** and **Ambros V (1993)** The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*. 75(5): 843–54.

Lee SS (2003) DAF-16 Target Genes That Control C. elegans Life-Span and Metabolism. *Science (80-.).* 300(5619): 644–647.

Lee Y, **Jeon K**, **Lee J-T**, **Kim S** and **Kim VN (2002)** MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* 21(17): 4663–70.

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): 4051–60.

Leignel V, **Silvestre A**, **Humbert JF** and **Cabaret J (2010)** Alternation of anthelmintic treatments: A molecular evaluation for benzimidazole resistance in nematodes. *Vet. Parasitol.* 17280–88.

Leung AKL and **Sharp PA (2010)** MicroRNA Functions in Stress Responses. *Mol. Cell*.

Lewis BP, **Shih IH**, **Jones-Rhoades MW**, **Bartel DP** and **Burge CB (2003)** Prediction of Mammalian MicroRNA Targets. *Cell*. 115(7): 787–798.

Li SSL, **Yu SL**, **Kao LP**, **Zong YT**, **Singh S**, **Bo ZC**, **Ho BC**, **Liu YH** and **Yang PC (2009a)** Target identification of micrornas expressed highly in human embryonic stem cells. *J. Cell. Biochem.* 106(6): 1020–1030.

Li S-C, **Liao Y-L**, **Ho M-R**, **Tsai K-W**, **Lai C-H** and **Lin W (2012a)** miRNA arm selection and isomiR distribution in gastric cancer. *BMC Genomics*. 13(Suppl 1): S13.

Li X, **Wang X**, **Zhang S**, **Liu D**, **Duan Y** and **Dong W (2012b)** Identification of soybean microRNAs involved in soybean cyst nematode infection by deep sequencing. *PLoS One*. 7(6): e39650.

Li X, **Cassidy JJ**, **Reinke CA**, **Fischboeck S** and **Carthew RW (2009b)** A MicroRNA Imparts Robustness against Environmental Fluctuation during Development. *Cell*. 137(2): 273–282.

Li Z, **Chen X**, **Zen X**, **Liang J**, **Wei J**, **Lv Z**, **Sun X** and **Wu Z-D (2014)** MicroRNA expression profile in the third- and fourth-stage larvae of Angiostrongylus cantonensis. *Parasitol. Res.* 113(5): 1883–96.

Lichtenfels JR, **Pilitt** and **Jambre LF Le (1986)** Cuticular Ridge Patterns of Haemonchus contortus and Haemonchus placet (Nematoda: Trichostrongyloidea). *Proc. Helm. Soc. Wash.* 53(1): 94–101.

Lichtenfels JR, **Pilitt PA**, **Gibbons LM** and **Boomker JDF (2001)** Haemonchus horaki n. sp. (nematoda: trichostrongyloidea) from the grey rhebuck pelea capreolus in south africa 87(5): 1095–1103.

Lichtenfels JR, **Pilitt PA**, **Gibbons LM** and **Hoberg EP (2002)** Redescriptions of Haemonchus mitchelli and Haemonchus okapiae (Nematoda: Trichostrongyloidea) and description of a unique synlophe for the haemonchinae. *J. Parasitol.* 88(5): 947–60.

Lim LP, **Lau NC**, **Weinstein EG**, **Abdelhakim A**, **Yekta S**, **Rhoades MW**, **Burge CB** and **Bartel DP (2003)** The microRNAs of Caenorhabditis elegans. *Genes Dev.* 17(8): 991–1008.

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.

Lin H and **Spradling AC (1997)** A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development*. 1242463–2476.

Lin Q, **Gao Z**, **Alarcon RM**, **Ye J** and **Yun Z (2009)** A role of miR-27 in the regulation of adipogenesis. *FEBS J.* 276(8): 2348–2358.

Lin S, **Johnson SM**, **Abraham M**, **Vella MC**, **Pasquinelli A**, **Gamberi C**, **Gottlieb E** and **Slack FJ (2003a)** The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. *Dev. Cell*. 4(5): 639–50.

Lin S-Y, **Johnson SM**, **Abraham M**, **Vella MC**, **Pasquinelli A**, **Gamberi C**, **Gottlieb E** and **Slack FJ (2003b)** The C. elegans hunchback Homolog, hbl-1, Controls Temporal Patterning and Is a Probable MicroRNA Target. *Dev. Cell*. 4(5): 639–650.

Lingel A, **Simon B**, **Izaurralde E** and **Sattler M (2004)** Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat. Struct. Mol. Biol.* 11(6): 576–577.

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 C, **Voronin D**, **Poole CB**, **Bachu S**, **Rogers MB**, **Jin J**, **Ghedin E**, **Lustigman S**, **McReynolds LA** and **Unnasch TR (2015)** Functional analysis of microRNA activity in Brugia malayi. *Int. J. Parasitol.* 45(9-10): 579–583.

Liu Q, **Tuo W**, **Gao H** and **Zhu X-Q (2010)** MicroRNAs of parasites: current status and future perspectives. *Parasitol. Res.* 107(3): 501–7.

Lodygin D, **Tarasov V**, **Epanchintsev A**, **Berking C**, **Knyazeva T**, **Körner H**, **Knyazev P**, **Diebold J** and **Hermeking H (2008)** Inactivation of miR-34a by aberrant CpG methylation in multiple types of cancer. *Cell Cycle*. 72591–2600.

Long D, **Lee R**, **Williams P**, **Chan CY**, **Ambros V** and **Ding Y (2007)** Potent effect of target structure on microRNA function. *Nat. Struct. Mol. Biol.* 14(4): 287–294.

Lucanic M, **Graham J**, **Scott G**, **Bhaumik D**, **Benz CC**, **Hubbard A**, **Lithgow GJ** and **Melov S (2013)** Age-related micro-RNA abundance in individual C. elegans. *Aging (Albany. NY).* 5(6): 394–411.

Lucas MPD and **Lozano E (2011)** Cel-bantam miRNA family regulates TGF-b Sma/Mab pathway presented in. *Int. Worm Meet.*

Luteijn MJ and **Ketting RF (2013)** PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat. Rev. Genet.* 14(8): 523–34.

Macosko EZ, **Pokala N**, **Feinberg EH**, **Chalasani SH**, **Butcher RA**, **Clardy J** and **Bargmann CI (2009)** A hub-and-spoke circuit drives pheromone attraction and social behaviour in C. elegans. *Nature*. 458(7242): 1171–5.

Mahen EM, **Watson PY**, **Cottrell JW** and **Fedor MJ (2010)** mRNA secondary structures fold sequentially but exchange rapidly in vivo. *PLoS Biol.* 8(2): .

Majlessi M, **Nelson NC** and **Becker MM (1998)** Advantages of 2'-O-methyl oligoribonucleotide probes for detecting RNA targets. *Nucleic Acids Res.* 26(9): 2224–2229.

Marco a., **Ninova M**, **Ronshaugen M** and **Griffiths-Jones S (2013a)** Clusters of microRNAs emerge by new hairpins in existing transcripts. *Nucleic Acids Res.* 1–8.

Marco A, **MacPherson JI**, **Ronshaugen M** and **Griffiths-Jones S (2012)** MicroRNAs from the same precursor have different targeting properties. *Silence*. 3(1): 8.

Marco A, **Kozomara A**, **Hui JHL**, **Emery AM**, **Rollinson D**, **Griffiths-Jones S** and **Ronshaugen M (2013b)** Sex-Biased Expression of MicroRNAs in Schistosoma mansoni. *PLoS Negl. Trop. Dis.* 7(9): e2402 (M. K. Jones, Ed.).

Marín RM and **Vaníek J (2011)** Efficient use of accessibility in microRNA target prediction. *Nucleic Acids Res.* 39(1): 19–29.

Martin RJ, **Verma S**, **Levandoski M**, **Clark CL**, **Qian H**, **Stewart M** and **Robertson AP (2006)** Drug resistance and neurotransmitter receptors of nematodes: recent studies on the mode of action of levamisole. *Parasitology*. 131(S1): S71.

Matrajt M (2010) Non-coding RNA in apicomplexan parasites. *Mol. Biochem. Parasitol.* 174(1): 1–7. editor & translator, Elsevier B.V.

McKeown C, **Praitis V** and **Austin J (1998)** sma-1 encodes a betaH-spectrin homolog required for Caenorhabditis elegans morphogenesis. *Development*. 125(11): 2087–2098.

Mederos AE, **Banchero GE** and **Ramos Z (2014)** First report of monepantel Haemonchus contortus resistance on sheep farms in Uruguay. *Parasit. Vectors*. 7(1): 598.

Miretti S, **Martignani E**, **Accornero P** and **Baratta M (2013)** Functional effect of mir-27b on myostatin expression: a relationship in Piedmontese cattle with double-muscled phenotype. *BMC Genomics*. 14194.

Miska E a, **Alvarez-Saavedra E**, **Abbott AL**, **Lau NC**, **Hellman AB**, **McGonagle SM**, **Bartel DP**, **Ambros VR** and **Horvitz HR (2007)** Most Caenorhabditis elegans microRNAs are individually not essential for development or viability. *PLoS Genet.* 3(12): e215.

Moghal N and **Sternberg PW (2003)** Extracellular domain determinants of LET-23 (EGF) receptor tyrosine kinase activity in Caenorhabditis elegans. *Oncogene*. 22(35): 5471–5480.

Mogilyansky E and **Rigoutsos I (2013)** The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death Differ.* 20(12): 1603–1614.

Mondou E, **Dufort I**, **Gohin M**, **Fournier E** and **Sirard M-A (2012)** Analysis of microRNAs and their precursors in bovine early embryonic development. *Mol. Hum. Reprod.* 18(9): 425–34.

Monney T and **Hemphill A (2014)** Vaccines against neosporosis: What can we learn from the past studies? *Exp. Parasitol.*

Morris K V., **Santoso S**, **Turner AM**, **Pastori C** and **Hawkins PG (2008)** Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet.* 4(11): .

Moss EG, **Lee RC** and **Ambros V (1997)** The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. *Cell*. 88(5): 637–46.

Mossi R, **Jónsson ZO**, **Allen BL**, **Hardin SH** and **Hübscher U (1997)** Replication factor C interacts with the C-terminal side of proliferating cell nuclear antigen. *J. Biol. Chem.* 272(3): 1769–1776.

Munn EA (1977) A helical, polymeric extracellular protein associated with the luminal surface of Haemonchus contortus intestinal cells. *Tissue Cell*. 923–34.

Munn EA, **Greenwood CA** and **Coadwell WJ (1987)** Vaccination of young lambs by means of a protein fraction extracted from adult Haemonchus contortus. *Parasitology*. 94 (Pt 2)385–397.

Naifang S, **Minping Q** and **Minghua D (2013)** Integrative Approaches for microRNA Target Prediction: Combining Sequence Information and the Paired mRNA and miRNA Expression Profiles. *Curr. Bioinform.* 8(1): 37–45.

Nakanishi K, **Weinberg DE**, **Bartel DP** and **Patel DJ (2012)** Structure of yeast Argonaute with guide RNA. *Nature*. 1–9. editor & translator, Nature Publishing Group.

Newlands GFJ, **Skuce PJ**, **Nisbet AJ**, **Redmond DL**, **Smith SK**, **Pettit D** and **Smith WD (2006)** Molecular characterization of a family of metalloendopeptidases from the intestinal brush border of Haemonchus contortus. *Parasitology*. 133357–368.

Newton SE and **Munn EA (1999)** The development of vaccines against gastrointestinal nematode parasites, particularly Haemonchus contortus. *Parasitol. Today*.

Nicolas FE (2011) Experimental validation of microRNA targets using a luciferase reporter system. *Methods Mol. Biol.* 732139–152.

Nicolas FE, **Pais H**, **Schwach F**, **Lindow M**, **Kauppinen S**, **Moulton V** and **Dalmay T (2008)** Experimental identification of microRNA-140 targets by silencing and overexpressing miR-140. *RNA*. 14(12): 2513–2520.

Niiranen L, **Espelid S**, **Karlsen CR**, **Mustonen M**, **Paulsen SM**, **Heikinheimo P** and **Willassen NP (2007)** Comparative expression study to increase the solubility of cold adapted Vibrio proteins in Escherichia coli. *Protein Expr. Purif.* 52(1): 210–218.

Ninova M, **Ronshaugen M** and **Griffiths-Jones S (2014)** Conserved Temporal Patterns of MicroRNA Expression in Drosophila Support a Developmental Hourglass Model. *Genome Biol. Evol.* 6(9): 2459–2467.

Nix P, **Hammarlund M**, **Hauth L**, **Lachnit M**, **Jorgensen EM** and **Bastiani M (2014)** Axon regeneration genes identified by RNAi screening in C. elegans. *J. Neurosci.* 34629–45.

Ohler U, **Yekta S**, **Lim LP**, **Bartel DP** and **Burge CB (2004)** Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. *RNA*. 10(9): 1309–1322.

Oikonomou G and **Shaham S (2011)** The Glia of Caenorhabditis elegans. *Glia*. 59(9): 1253–1263.

Onyali IO, **Onwuliri CO** and **Ajayi JA (1990)** Development and survival of Haemonchus contortus larvae on pasture at Vom, Plateau State, Nigeria. *Vet. Res. Commun.* 14(3): 211–6.

Orban TI (2005) Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA*. 11(4): 459–469.

Orellana EA and **Kasinski AL (2015)** MicroRNAs in Cancer: A Historical Perspective on the Path from Discovery to Therapy. *Cancers (Basel).* 7(3): 1388–405.

Pagano DJ, **Kingston ER** and **Kim DH (2013)** The mir-58 family of microRNAs regulates the tissue-specific expression of PMK-2 p38 MAPK that functions in host defense. *Int. Worm Meet.*

Parrish JZ, **Xu P**, **Kim CC**, **Jan LY** and **Jan YN (2009)** The microRNA bantam Functions in Epithelial Cells to Regulate Scaling Growth of Dendrite Arbors in Drosophila Sensory Neurons. *Neuron*. 63(6): 788–802.

Pasquinelli a E, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B & other authors (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*. 408(6808): 86–9.

Pasquinelli AE (2012) MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat. Rev. Genet.* 13(4): 271–282. editor & translator, Nature Publishing Group.

Peterson SM, **Thompson JA**, **Ufkin ML**, **Sathyanarayana P**, **Liaw L** and **Congdon CB (2014)** Common features of microRNA target prediction tools. *Front. Genet.*

Piedrafita D, **Preston S**, **Kemp J**, **de Veer M**, **Sherrard J**, **Kraska T**, **Elhay M** and **Meeusen E (2013)** The effect of different adjuvants on immune parameters and protection following vaccination of sheep with a larval-specific antigen of the gastrointestinal nematode, Haemonchus contortus. *PLoS One*. 8(10): e78357.

Pierce ML, **Weston MD**, **Fritzsch B**, **Gabel HW**, **Ruvkun G** and **Soukup GA (2008)** MicroRNA-183 family conservation and ciliated neurosensory organ expression. *Evol. Dev.* 10(1): 106–113.

Poole CB, **Gu W**, **Kumar S**, **Jin J**, **Davis PJ**, **Bauche D** and **McReynolds L a (2014a)** Diversity and expression of microRNAs in the filarial parasite, Brugia malayi. *PLoS One*. 9(5): e96498.

Poole CB, **Gu W**, **Kumar S**, **Jin J**, **Davis PJ**, **Bauche D** and **McReynolds LA (2014b)** Diversity and expression of microRNAs in the filarial parasite, Brugia malayi. *PLoS One*. 9(5): .

Poole RJ, **Bashllari E**, **Cochella L**, **Flowers EB** and **Hobert O (2011)** A Genome-Wide RNAi screen for factors involved in neuronal specification in Caenorhabditis elegans. *PLoS Genet.* 7.

Porter MY and **Koelle MR (2010)** RSBP-1 is a membrane-targeting subunit required by the Galpha(q)-specific but not the Galpha(o)-specific R7 regulator of G protein signaling in Caenorhabditis elegans. *Mol. Biol. Cell*. 21(2): 232–243.

Prichard R (2009) *Drug Resistance in Nematodes*. *Antimicrob. Drug Resist.*

Redman E, **Grillo V**, **Saunders G**, **Packard E**, **Jackson F**, **Berriman M** and **Gilleard JS (2008)** Genetics of mating and sex determination in the parasitic nematode Haemonchus contortus. *Genetics*. 180(4): 1877–87.

Redmond DL, **Knox DP**, **Newlands G** and **Smith WD (1997)** Molecular cloning and characterisation of a developmentally regulated putative metallopeptidase present in a host protective extract of Haemonchus contortus. *Mol. Biochem. Parasitol.* 8577–87.

Rehmsmeier M, **Steffen P**, **Hochsmann M** and **Giegerich R (2004)** Fast and effective prediction of microRNA/target duplexes. *RNA*. 101507–1517.

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): 901–6.

Rhoades MW, **Reinhart BJ**, **Lim LP**, **Burge CB**, **Bartel B** and **Bartel DP (2002)** Prediction of plant microRNA targets. *Cell*. 110(4): 513–520.

Riddle D and **Albert P (1997)** Genetic and environmental regulation of dauer larva development. In *C. elegans II*., pages 739–768.

Riddle DL, **Swanson MM** and **Albert PS (1981)** Interacting genes in nematode dauer larva formation. *Nature*. 290(5808): 668–671.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ & other authors (2007) Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell*. 1291311–1323.

Ritchie W, **Rasko JEJ** and **Flamant S (2013)** MicroRNA target prediction and validation. *Adv. Exp. Med. Biol*. 77439–53.

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.

Roberts TC (2014) The MicroRNA Biology of the Mammalian Nucleus. *Mol. Ther. Nucleic Acids*. 3(August): e188.

Rosa BA, **Jasmer DP** and **Mitreva M (2014)** Genome-Wide Tissue-Specific Gene Expression, Co-expression and Regulation of Co-expressed Genes in Adult Nematode Ascaris suum. *PLoS Negl. Trop. Dis.* 8(2): e2678 (A. R. Jex, Ed.).

Rose BYJH (1963) Observations on the free-living stages of the stomach worm Haemonchus contortus 469–481.

Rougvie AE and **Ambros V (1995)** The heterochronic gene lin-29 encodes a zinc finger protein that controls a terminal differentiation event in Caenorhabditis elegans. *Development*. 121(8): 2491–500.

Roush SF and **Slack FJ (2009)** Transcription of the C. elegans let-7 microRNA is temporally regulated by one of its targets, hbl-1. *Dev. Biol*. 334(2): 523–534.

Rowe JB, **Nolan J V**, **de Chaneet G**, **Teleni E** and **Holmes PH (1988)** The effect of haemonchosis and blood loss into the abomasum on digestion in sheep. *Br. J. Nutr*. 59(1): 125–39.

Ruby JG, **Jan C**, **Player C**, **Axtell MJ**, **Lee W**, **Nusbaum C**, **Ge H** and **Bartel DP (2006)** Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell*. 127(6): 1193–207.

Ruike Y, **Ichimura A**, **Tsuchiya S**, **Shimizu K**, **Kunimoto R**, **Okuno Y** and **Tsujimoto G (2008)** Global correlation analysis for micro-RNA and mRNA expression profiles in human cell lines. *J. Hum. Genet*. 53(6): 515–523.

Ruvinsky I and **Ruvkun G (2003)** Functional tests of enhancer conservation between distantly related species. *Development*. 130(21): 5133–5142.

Ruvkun G, **Wightman B** and **Ha I (2004)** The 20 years it took to recognize the importance of tiny RNAs. *Cell*. 116(2 Suppl): S93–6, 2 p following S96.

Sachs R, **Gibbons LM** and **Lweno MF (1973)** Species of Haemonchus from domestic and wild ruminants in Tanzania, East Africa, including a description of H. dinniki n.sp. *Z. Tropenmed. Parasitol.* 24(4): 467–75.

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): 51–59. editor & translator, Australian Society for Parasitology Inc.

Sargison ND, **Jackson F**, **Bartley DJ** and **Moir ACP (2005)** Failure of moxidectin to control benzimidazole-, levamisole- and ivermectin-resistant Teladorsagia circumcinda in a sheep flock. *Vet. Rec.* 156(4): 105–9.

Schirle NT and **MacRae IJ (2012)** The crystal structure of human Argonaute2. *Science* .336(6084):1037-40

Schnall-Levin M, **Zhao Y**, **Perrimon N** and **Berger B (2010)** Conserved microRNA targeting in Drosophila is as widespread in coding regions as in 3'UTRs. *Proc. Natl. Acad. Sci. U. S. A.* 107(36): 15751–6.

Scott I, **Pomroy WE**, **Kenyon PR**, **Smith G**, **Adlington B** and **Moss a. (2013)** Lack of efficacy of monepantel against *Teladorsagia circumcincta* and *Trichostrongylus colubriformis*. *Vet. Parasitol.* 198(1-2): 166–171. editor & translator, Elsevier B.V.

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.

Selkirk ME, **Huang SC**, **Knox DP** and **Britton C (2012)** The development of RNA interference (RNAi) in gastrointestinal nematodes.

Shaham S, **Reddien PW**, **Davies B** and **Horvitz HR (1999)** Mutational analysis of the *Caenorhabditis elegans* cell-death gene *ced-3*. *Genetics*. 153(4): 1655–71.

Shao C-C, **Xu M-J**, **Alasaad S**, **Song H-Q**, **Peng L**, **Tao J-P** and **Zhu X-Q (2014)** Comparative analysis of microRNA profiles between adult *Ascaris lumbricoides* and *Ascaris suum*. *BMC Vet. Res.* 10(1): 99.

Sharman PA, **Smith NC**, **Wallach MG** and **Katrib M (2010)** Chasing the golden egg: vaccination against poultry coccidiosis. *Parasite Immunol.* 32(8):590-8.

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.* 23:497–508.

Shkap V, **de Vos AJ**, **Zweygarth E** and **Jongejan F (2007)** Attenuated vaccines for tropical theileriosis, babesiosis and heartwater: the continuing necessity. *Trends Parasitol.* 23(9):420-6.

Shubber AH, **Lloyd S** and **Soulsby EJ (1981)** Infection with gastrointestinal helminths. Effect of lactation and maternal transfer of immunity. *Z. Parasitenkd.* 65(2): 181–9.

Siddle KJ, **Tailleux L**, **Deschamps M**, **Loh Y-HE**, **Deluen C**, **Gicquel B**, **Antoniewski C**, **Barreiro LB**, **Farinelli L** and **Quintana-Murci L (2015)** Bacterial infection drives the expression dynamics of microRNAs and their isomiRs. *PLOS Genet.* 11(3): e1005064.

Silva SS, **Lopes C**, **Teixeira AL**, **Sousa M. C de** and **Medeiros R (2015)** Forensic miRNA: potential biomarker for body fluids? *Forensic Sci. Int. Genet.* 141–10.

Simon JM and **Sternberg PW (2002)** Evidence of a mate-finding cue in the hermaphrodite nematode *Caenorhabditis elegans*. *Proc.*

Natl. Acad. Sci. U. S. A. 99(3): 1598–603.

Siomi MC, **Sato K**, **Pezic D** and **Aravin AA (2011)** PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* 12246–258.

Sivanathan S, **Duncan JL** and **Urquhart GM (1984)** Some factors influencing the immunisation of sheep with irradiated *Haemonchus contortus* larvae. *Vet. Parasitol.* 16(3-4): 313–23.

Skorobogata O, **Escobar-Restrepo JM** and **Rocheleau CE (2014)** An AGEF-1/Arf GTPase/AP-1 ensemble antagonizes LET-23 EGFR basolateral localization and signaling during *C. elegans* vulva induction. *PLoS Genet.* 10(10): e1004728.

Small EM and **Olson EN (2011)** Pervasive roles of microRNAs in cardiovascular biology. *Nature*. 469(7330): 336–42.

Smith KA and **Maizels RM (2014)** IL-6 controls susceptibility to helminth infection by impeding Th2 responsiveness and altering the Treg phenotype in vivo. *Eur. J. Immunol.* 44(1): 150–61.

Smith WD, **Jackson F** and **Jackson E (1982)** Changes in the flow and composition of gastric lymph in sheep given a primary infection of 10,000 *Haemonchus contortus*. *Res. Vet. Sci.* 32(2): 184–8.

Smith WD, **Smith SK** and **Murray JM (1994)** Protection studies with integral membrane fractions of *Haemonchus contortus. Parasite Immunol.* 16231–241.

Smith WD, **Smith SK**, **Pettit D**, **Newlands GFJ** and **Skuce PJ (2000)** Relative protective properties of three membrane glycoprotein fractions from *Haemonchus contortus. Parasite Immunol.* 2263–71.

Spencer WC, Zeller G, Watson JD, Henz SR, Watkins KL, McWhirter RD, Petersen S, Sreedharan VT, Widmer C & other authors (2011) A spatial and temporal map of *C. elegans* gene expression. *Genome Res.* 21(2): 325–341.

Starich T, **Sheehan M**, **Jadrich J** and **Shaw J (2001)** Innexins in *C. elegans*. *Cell Commun. Adhes.* 8(4-6): 311–314.

Stefani G, **Chen X**, **Zhao H** and **Slack FJ (2015)** A novel mechanism of LIN-28 regulation of let-7 microRNA expression revealed by in vivo HITS-CLIP in *C. elegans*. *RNA.* 21(5): 985–96.

Stiernagle T (2006) Maintenance of *C. elegans*. *WormBook.* (1999): 1–11.

Stricklin SL, **Griffiths-Jones S** and **Eddy SR (2005)** *C. elegans* noncoding RNA genes. *WormBook.* 1–7.

Suarez VH, **Cristel SL** and **Busetti MR (2009)** Epidemiology and effects of gastrointestinal nematode infection on milk productions of dairy ewes. *Parasite*. 16141–147.

Sulston JE, **Schierenberg E**, **White JG** and **Thomson JN (1983)** The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* 100(1): 64–119.

Takasaki S (2015) Roles of microRNAs in cancers and development, pages 375–413.

Taki F a, **Pan X** and **Zhang B (2014)** Chronic nicotine exposure systemically alters microRNA expression profiles during post-embryonic stages in *Caenorhabditis elegans*. *J. Cell. Physiol.* 229(1): 79–89.

Tang G (2005) siRNA and miRNA: An insight into RISCs. *Trends Biochem. Sci.*

Taylor M a., **Learmount J**, **Lunn E**, **Morgan C** and **Craig BH (2009)** Multiple resistance to anthelmintics in sheep nematodes and comparison of methods used for their detection. *Small Rumin. Res.* 86(1-3): 67–70.

Than MT, **Kudlow B a.** and **Han M (2013a)** Functional analysis of neuronal microRNAs in *Caenorhabditis elegans* dauer formation by combinational genetics and neuronal miRISC immunoprecipitation. *PLoS Genet.* 9(6): e1003592.

Todd KS, **Levine ND** and **Boatman PA (1976)** Effect of temperature on survival of free-living stages of *Haemonchus contortus*. *Am. J. Vet. Res.* 37(8): 991–2.

Tolia NH and **Joshua-Tor L (2007)** Slicer and the argonautes. *Nat. Chem. Biol.* 3(1): 36–43.

Torley KJ, **da Silveira JC**, **Smith P**, **Anthony R V**, **Veeramachaneni DNR**, **Winger QA** and **Bouma GJ (2011)** Expression of miRNAs in ovine fetal gonads: potential role in gonadal differentiation. *Reprod. Biol. Endocrinol.* 92.

Trionfini P, **Benigni A** and **Remuzzi G (2014)** MicroRNAs in kidney physiology and disease. *Nat. Rev. Nephrol.*

Tripurani SK, **Lee K-B**, **Wee G**, **Smith GW** and **Yao J (2011)** MicroRNA-196a regulates bovine newborn ovary homeobox gene (NOBOX) expression during early embryogenesis. *BMC Dev. Biol.* 1125.

Tritten L, **O'Neill M**, **Nutting C**, **Wanji S**, **Njouendoui A**, **Fombad F**, **Kengne-Ouaffo J**, **Mackenzie C** and **Geary T (2014a)** Loa loa and Onchocerca ochengi miRNAs detected in host circulation. *Mol. Biochem. Parasitol.* 198(1): 14–17. editor & translator, Elsevier B.V.

Tritten L, **Burkman E**, **Moorhead A**, **Satti M**, **Geary J**, **Mackenzie C** and **Geary T (2014b)** Detection of circulating parasite-derived microRNAs in filarial infections. *PLoS Negl. Trop. Dis.* 8(7): e2971.

Ule J, **Jensen KB**, **Ruggiu M**, **Mele A**, **Ule A** and **Darnell RB (2003)** CLIP identifies Nova-regulated RNA networks in the brain. *Science*. 302(5648): 1212–5.

Valinezhad Orang A, **Safaralizadeh R** and **Kazemzadeh-Bavili M (2014)** Mechanisms of miRNA-mediated gene regulation from common downregulation to mRNA-specific upregulation. *Int. J. Genomics*. 20141–15.

Vance V and **Vaucheret H (2001)** RNA silencing in plants--defense and counterdefense. *Science*. 2922277–2280.

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.

Vegh P, Foroushani ABK, Magee DA, McCabe MS, Browne JA, Nalpas NC, Conlon KM, Gordon S V, Bradley DG & other authors (2013) Profiling microRNA expression in bovine alveolar macrophages using RNA-seq. *Vet. Immunol. Immunopathol.* 155(4): 238–44.

Veljkovic E, **Stasiuk S**, **Skelly PJ**, **Shoemaker CB** and **Verrey F (2004)** Functional characterization of *Caenorhabditis elegans* heteromeric amino acid transporters. *J. Biol. Chem.* 279(9): 7655–7662.

Vella MC and **Slack FJ (2005)** *C. elegans* microRNAs. *WormBook*. 1–9.

Vercruysse J, **Knox DP**, **Schetters TPM** and **Willadsen P (2004)** Veterinary parasitic vaccines: pitfalls and future directions. *Trends Parasitol.* 20(10):488-92.

Voinnet O (2009) Origin, Biogenesis, and Activity of Plant MicroRNAs. *Cell*.136(4): 669-687.

Volovik Y, **Moll L**, **Marques FC**, **Maman M**, **Bejerano-Sagie M** and **Cohen E (2014)** Differential regulation of the heat shock factor 1 and DAF-16 by neuronal nhl-1 in the nematode *C. elegans*. *Cell Rep*. 92192–2205.

Wallace E, **Morrell N**, **Yang X**, **Long L**, **Stevens H**, **Nilsen M**, **Loughlin L**, **Mair K**, **Baker A** and **MacLean M (2015)** A sex-specific microRNA-96/5HT1B axis influences development of pulmonary hypertension. *Am. J. Respir. Crit. care*. 191(12):1432-42.

Wan WB, Migawa MT, Vasquez G, Murray HM, Nichols JG, Gaus H, Berdeja A, Lee S, Hart CE & other authors (2014) Synthesis, biophysical properties and biological activity of second generation antisense oligonucleotides containing chiral phosphorothioate linkages. *Nucleic Acids Res.* 42(22): 13456–13468.

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.* 211462–1477.

Wang XG, **Yu JF**, **Zhang Y**, **Gong DQ** and **Gu ZL (2012)** Identification and characterization of microRNA from chicken adipose tissue and skeletal muscle. *Poult. Sci.* 91(1): 139–49.

Wang Y, **Li X** and **Hu H (2011b)** Transcriptional regulation of co-expressed microRNA target genes. *Genomics*. 98(6): 445–452.

Watts JL and **Browse J (2002)** Genetic dissection of polyunsaturated fatty acid synthesis in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 99(9): 5854–5859.

Weaver BP, **Zabinsky R**, **Weaver YM**, **Lee ES**, **Xue D** and **Han M (2014)** CED-3 caspase acts with miRNAs to regulate non-apoptotic gene expression dynamics for robust development in *C. elegans*. *Elife*. 31–22.

Weber JA, **Baxter DH**, **Zhang S**, **Huang DY**, **How Huang K**, **Jen Lee M**, **Galas DJ** and **Wang K (2010)** The microRNA spectrum in 12 body fluids. *Clin. Chem.* 56(11): 1733–1741.

Weterman MA, **van Groningen JJ**, **Tertoolen L** and **van Kessel AG (2001)** Impairment of MAD2B-PRCC interaction in mitotic checkpoint defective t(X;1)-positive renal cell carcinomas. *Proc. Natl. Acad. Sci. U. S. A.* 98(24): 13808–13813.

White JG, **Southgate E**, **Thomson JN** and **Brenner S (1986)** The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. B Biol. Sci.* 314(1165): 1–340.

Wienholds E and **Plasterk RH a (2005)** MicroRNA function in animal development. *FEBS Lett.* 579(26): 5911–22.

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): 855–62.

Winter AD, **Weir W**, **Hunt M**, **Berriman M**, **Gilleard JS**, **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.

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): .

de Wit E, **Linsen SEV**, **Cuppen E** and **Berezikov E (2009)** Repertoire and evolution of miRNA genes in four divergent nematode species. *Genome Res.* 19(11): 2064–2074.

Witkos TM, **Koscianska E** and **Krzyzosiak WJ (2011)** Practical aspects of microRNA target prediction. *Curr. Mol. Med.* 11(2): 93–109.

Wu X, Fu Y, Yang D, Xie Y, Zhang R, Zheng W, Nie H, Yan N, Wang N & other authors (2013) Identification of neglected cestode *Taenia multiceps* microRNAs by illumina sequencing and bioinformatic analysis. *BMC Vet. Res.* 9162.

Xu D, **Gao Y**, **Huang L** and **Sun Y (2014)** Changes in miRNA expression profile of space-flown *Caenorhabditis elegans* during Shenzhou-8 mission. *Life Sci. Sp. Res.* 1(1): 44–52.

Xu M-J, **Ai L**, **Fu J-H**, **Nisbet AJ**, **Liu Q-Y**, **Chen M-X**, **Zhou D-H** and **Zhu X-Q (2012)** Comparative characterization of microRNAs from the liver flukes *Fasciola gigantica* and *F. hepatica*. *PLoS One*. 7(12): e53387.

Xu MJ, **Wang CR**, **Huang SY**, **Fu JH**, **Zhou DH**, **Chang QC**, **Zheng X** and **Zhu XQ (2013a)** Identification and characterization of microRNAs in the pancreatic fluke *Eurytrema pancreaticum*. *Parasit. Vectors*. 625.

Xu MJ, **Fu JH**, **Nisbet AJ**, **Huang SY**, **Zhou DH**, **Lin RQ**, **Song HQ** and **Zhu XQ (2013b)** Comparative profiling of microRNAs in male and female adults of *Ascaris suum*. *Parasitol. Res.* 112(3): 1189–95.

Yamada K, **Hirotsu T**, **Matsuki M**, **Butcher RA**, **Tomioka M**, **Ishihara T**, **Clardy J**, **Kunitomo H** and **Iino Y (2010)** Olfactory plasticity is regulated by pheromonal signaling in *Caenorhabditis elegans*. *Science*. 329(5999): 1647–50.

Yamaguchi H and **Miyazaki M (2014)** Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules*. 4(1): 235–51.

Yan KS, **Yan S**, **Farooq A**, **Han A**, **Zeng L** and **Zhou M-M (2003)** Structure and conserved RNA binding of the PAZ domain. *Nature*. 426(6965): 468–474.

Yan R, **Wang J**, **Xu L**, **Song X** and **Li X (2014)** DNA vaccine encoding *Haemonchus contortus* actin induces partial protection in goats. *Acta Parasitol.* 59(4): 698–709.

Yekta S (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science*. 304(5670): 594–596.

Yoneda T, **Benedetti C**, **Urano F**, **Clark SG**, **Harding HP** and **Ron D (2004)** Compartment-specific perturbation of protein handling activates genes encoding mitochondrial chaperones. *J. Cell Sci.* 117(Pt 18): 4055–4066.

Yoo AS and **Greenwald I (2005)** LIN-12/Notch activation leads to microRNA-mediated down-regulation of Vav in C. elegans. *Science.* 310(5752): 1330–1333.

Yoon CH, **Chang C**, **Hopper NA**, **Lesa GM** and **Sternberg PW (2000)** Requirements of multiple domains of SLI-1, a Caenorhabditis elegans homologue of c-Cbl, and an inhibitory tyrosine in LET-23 in regulating vulval differentiation. *Mol. Biol. Cell.* 11(11): 4019–4031.

Yoon J-H, **Abdelmohsen K** and **Gorospe M (2013)** Posttranscriptional gene regulation by long noncoding RNA. *J. Mol. Biol.* 4253723–30.

Yu L, **Liao Q**, **Chen X**, **Xu L**, **Zeng X**, **Lv Z**, **Sun X**, **Zhen H** and **Wu Z (2014)** Dynamic expression of miR-132, miR-212, and miR-146 in the brain of different hosts infected with Angiostrongylus cantonensis. *Parasitol. Res.* 113(1): 91–9.

Yu Z, **Jian Z**, **Shen SH**, **Purisima E** and **Wang E (2007)** Global analysis of microRNA target gene expression reveals that miRNA targets are lower expressed in mature mouse and Drosophila tissues than in the embryos. *Nucleic Acids Res.* 35(1): 152–164.

Yue D, **Liu H** and **Huang Y (2009)** Survey of Computational Algorithms for MicroRNA Target Prediction. *Curr. Genomics*. 10(7): 478–492.

Zhang B, **Pan X**, **Cobb GP** and **Anderson TA (2007a)** microRNAs as oncogenes and tumor suppressors. *Dev. Biol.* 302(1): 1–12.

Zhang L, **Ding L**, **Cheung TH**, **Dong M-Q**, **Chen J**, **Sewell AK**, **Liu X**, **Yates JR** and **Han M (2007b)** Systematic identification of C. elegans miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol. Cell.* 28(4): 598–613.

Zhao G-H, **Xu M-J** and **Zhu X-Q (2013)** Identification and characterization of microRNAs in Baylisascaris schroederi of the giant panda. *Parasit. Vectors*. 6(1): 216. editor & translator, Parasites & Vectors.

Zhao G, **Yan R**, **Muleke CI**, **Sun Y**, **Xu L** and **Li X (2012)** Vaccination of goats with DNA vaccines encoding H11 and IL-2 induces partial protection against Haemonchus contortus infection. *Vet. J.* 19194–100.

Zhong X, **Chung ACK**, **Chen H-Y**, **Meng X-M** and **Lan HY (2011)** Smad3-Mediated Upregulation of miR-21 Promotes Renal Fibrosis. *J. Am. Soc. Nephrol.* 221668–1681.

Zisoulis DG, **Lovci MT**, **Wilbert ML**, **Hutt KR**, **Liang TY**, **Pasquinelli AE** and **Yeo GW (2010a)** Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. *Nat. Struct. Mol. Biol.* 17(2): 173–179.

Zisoulis DG, **Lovci MT**, **Wilbert ML**, **Hutt KR**, **Liang TY**, **Pasquinelli AE** and **Yeo GW (2010b)** Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. *Nat. Struct. Mol. Biol.* 17(2): 173–9. editor & translator, Nature Publishing Group.

Zisoulis DG, **Yeo GW** and **Pasquinelli AE (2011)** Comprehensive Identification of miRNA Target Sites in Live Animals. *Methods Mol. Biol.* 732169–185.

Zorio DA, **Cheng NN**, **Blumenthal T** and **Spieth J (1994)** Operons as a common form of chromosomal organization in C. elegans. *Nature*. 372(6503): 270–2. editor & translator, Nature Publishing Group.

Zou Y, **Chiu H**, **Zinovyeva A**, **Ambros V**, **Chuang C-F** and **Chang C (2013)** Developmental decline in neuronal regeneration by the progressive change of two intrinsic timers. *Science*. 340(6130): 372–6.