Teladorsagia circumcincta Galectin-Mucosal Interactome in Sheep

By

Nur Nasuha Hafidi

Bachelor of Animal and Veterinary Bioscience (Hons)

Submitted to La Trobe University in total fulfilment of the requirements for the degree of Master of Science

31 December 2020

School of Life Sciences

College of Science, Health and Engineering

La Trobe University,

Victoria, Australia

Table of Contents

Acknowledgements

My deepest and most sincere thanks to my project supervisors Prof. Travis Beddoe and Prof. Michael Stear for their endless support and guidance throughout my candidature. Thank you, Travis, for your boundless patience and constant encouragement during such a tough time. To Jaclyn Swan, you have taught me everything I know these past two years and I would never have made it this far without you. To Dr Timothy Cameron, thank you for always being there to help me troubleshoot when the science went wrong, and to reach for things on the top shelf when I couldn't. Massive shoutout to the Beddoe lab group for being the best bunch of misfits a girl could ever hope to work with. It has been such a joy getting to know all of you.

I am so very thankful for my support systems, most importantly my Super Smash Bros – Alex Knox and Gemma Zerna. Thank you for keeping me sane and for allowing me to be the truest version of myself. No one else can ever make me laugh as hard as you two. Special mention to Molly Spaulding and Kimberly Lim for going on this crazy journey with me. I am so proud of what we have accomplished together. Thanks also to Yasmin Basir, Adriana Gan and Cassandra Scicluna for being my shoulders to cry on. I am eternally grateful.

No words can express my gratitude towards my family. I can't list everyone, but you have all helped me get to where I am today. To Ibu and Abi, I will be forever indebted to your never-ending love, support, and sacrifice. To my aunt, Dr Hamimah Ahmat, you are the strongest woman I know, and my biggest inspiration. To Nenek and Atok, I love you with all my heart. This is dedicated to all the ones I've lost during this journey – Nenek Amu, Baba Mok, Uncle Zeki, Papa Yusoff and Sasha Subash. I wish you were still here to see me. I hope I have made you proud.

Abbreviations

ACE	Angiotensin-converting enzyme
ACN	Acetonitrile
AGC	Automatic gain control
AID	Activation-induced cytidine deaminase
AUD	Australian dollars
BMMY	Buffered methanol-complex media
CD	Cluster of differentiation
CHB	Canaria Hair Breed
CLCA1	Calcium activated chloride channel
ConA-HRP	Concanavalin A-horseradish peroxidase conjugate
cRAP	common Repository of Adventitious Proteins
CRD	Carbohydrate recognition domain
DOC	Sodium deoxycholate
EDTA	Ethylenediaminetetraacetic acid
epg	Eggs per gram
ER	Endoplasmic reticulum
E/S	Excretory/secretory
FCGBP	Fc fragment of IgG binding protein
FDR	False discovery rate
FECRT	Faecal egg count reduction test
GI	Gastrointestinal
GO	Gene ontology
HCD	Higher-energy collisional dissociation
Hco-gal-1	*Haemonchus contortus* galectin 1

Ig	Immunoglobulins
IL-4	Interleukin-4
IPM	Integrated pest management
ITLN	Intelectin
L1	First stage larvae
L2	Second stage larvae
L3	Third stage larvae
L4	Fourth stage larvae
LFQ	Label free quantification
LGALS11	Ruminant galectin 11
LGALS14	Ruminant galectin 14
LTU-CPP	La Trobe University – Comprehensive Proteomics Platform
MCAM	Melanoma cell adhesion molecule
MePBS	Phosphate-buffered saline plus β-mercaptoethanol
MHC	Major histocompatibility complex
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MUC	Mucin
NCBI	National Center for Biotechnology Information
NHS	N-hydroxysuccinamide
Ni-NTA	Nickel-nitrilotriacetic acid
PBS	Phosphate-buffered saline
PEG	Polyethylene glycol
PRIDE	Proteomics identification database
PSM	Peptide spectrum match
PVDF	Polyvinylidene difluoride

rIL-5	recombinant Interleukin-5
RIPA	Radioimmunoprecipitation assay
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEA	Sea urchin sperm, enterokinase and agrin
ST	Abomasal scrape tissue
tBLASTn	translated Basic Local Alignment Search Tool
TCEP	Tris-2-carboxyethyl-phosphine
Tci-gal-1	*Teladorsagia circumcincta* galectin 1
Tci-gal-2	*Teladorsagia circumcincta* galectin 2
Tco-gal-2	*Trichostrongylus colubriformis* galectin 2
TFA	Trifluoroacetic acid
TFF	Trefoil factor family
Th1	T helper type 1
Th2	T helper type 2
Tl-gal	*Toxascaris leonina* galectin
Tsgal	*Trichinella spiralis* galectin
UniProtKb	Universal Protein Resource Knowledgebase
WT	Abomasal whole tissue
YEPD	Yeast extract peptone dextrose

Statement of Authorship

This thesis consists primarily of work by the author. Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis accepted for the award of any other degree or diploma. No other person's work has been used without due acknowledgment in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution. This work was supported by an Australian Government Research Training Program Scholarship.

Nur Nasuha Hafidi

31 December 2020

Abstract

Teladorsagia circumcincta is the most important gastrointestinal parasite in the livestock industry in temperate regions around the world, causing great economic losses. The infective third-stage larvae (L3) of this nematode secretes a large number of excretory-secretory (E/S) molecules that are presented at the host-parasite interface, some of which are likely to play critical roles in modulating the host immune response. One of the most abundant E/S molecules is a protein of unknown function termed Tci-gal-1, which has similarity to mammalian galectins. Galectins are a family of carbohydrate-binding molecules, with characteristic domain organization and affinity for β-galactosides, that mediates a variety of important cellular functions including inflammation and immune responses. To understand the role of Tci-gal-1 at the host-parasite interface, the ligands from sheep abomasal scrapes and whole tissue were identified by galectin-affinity chromatography and mass spectrometry. A total of 135 unique proteins were identified from whole abomasal tissue samples, while 89 proteins were isolated from abomasal scrape samples. Of these proteins, 63 were present in both samples. Many of the host proteins identified, such as trefoil factors and mucin-like proteins, play critical roles in the host response and act in concert to effectively eliminate nematodes during infection. The identification of Tci-gal-1 binding partners could provide new insights on host-parasite interactions, and potentially lead to the development of new interventions.

Chapter 1 – Literature Review

1.1 Introduction

1.1.1 Livestock industry

The livestock industry is of great economic importance worldwide, as domesticated ruminants such as cattle, sheep and goats provide the agricultural sector with vast resources to support the growing global population. As human populations have multiplied over the years, there is an increased demand and intensification of livestock production. In France alone, an estimated one billion land animals are culled annually for consumption, with global numbers significantly increasing over the past 50 years (Bonnet *et al.* 2020). In 2011, the east coast of Australia alone produced about 20% (5.7 million heads) of all cattle in the country, along with 1.7 million sheep. The Australian dairy industry produces approximately 9 billion litres of raw milk, of which only 38% is being exported (Gollnow *et al.* 2014; Marshall *et al.* 2017). Additionally, Australia is the largest exporter of goat meat in the world despite only producing less than 2 million goats per year. The extent of this industry makes it important to ensure the welfare of livestock animals are well managed to prevent significant production and economic losses.

1.1.2 Livestock diseases

Livestock are exposed to a variety of pathogens such as viruses, bacteria, protozoa, and helminths that naturally exist in the environment. Individuals are often co-infected with multiple pathogens of different species, with some pathogens having the potential to cause or exacerbate the development of diseases. Microbes also exist ubiquitously and can have symbiotic relationships with other organisms like helminths, be it commensalism, mutualism or parasitism (Sinnathamby *et al.* 2018). Being grazing ruminants, livestock in both small- and large-scale farms are highly susceptible to gastrointestinal parasites spread via the faecal-oral route. Infection by gastrointestinal parasites can cause parasitic gastroenteritis, hindering productivity and may lead to mortality (Koinari *et al.* 2013).

In 2014, Huang *et al.* discovered almost 87% of dairy cattle across Taiwan were infected with gastrointestinal parasites, nearly 82% of which are caused by protozoa. The remaining 8% encompasses helminth infections, mainly nematodes (roundworms), trematodes (flatworms) and cestodes (tapeworms). Calves under the age of one were noticeably more susceptible to disease, and often act as reservoirs due to their immature immune system against the parasites (Huang *et al.* 2014). Similarly, another study based in Ireland surveying 88 beef and dairy cattle farms from 19 counties revealed 49-67% of herds to be infected by gastrointestinal parasites (Murphy *et al.* 2006). In Papua New Guinea, Koinari *et al.* (2012) studied the gastrointestinal parasites of 165 sheep and goats and found 78% were infected with one or more types of parasites, with strongyle parasites being the most prevalent. Sinnathamby and colleagues (2018) studied the bacterial microbiome associated with nematode infections, specifically in the eggs, adult and third-stage larvae (L3) of *Haemonchus contortus*. They discovered that the symbionts of *H. contortus* are mainly commensals or mutualists that are part of the nematode's metabolism. Other microbes present in the host's faeces or soil have been known to be food sources for the nematodes, particularly for the free-living larval stages.

Ruminants, particularly in the dairy industry, are also highly susceptible to mastitis. Mastitis is one of the most common bovine infections. Blockage or trauma to the mammary glands make them susceptible to secondary bacterial infection, most commonly by organisms such as *Escherichia coli*, *Staphylococcus aureus* or *Streptococcus agalactiae* (Ruegg 2017). Infection with mastitis can be contagious and causes significant declines in milk yield (AUD $200/cow/year) (Ismail 2017).

1.1.3 Gastrointestinal nematode infections in livestock

Ruminants are highly susceptible to nematode infections through ingestion of infective stages of the parasites that reside on pasture and cause parasitic gastroenteritis. Irrigated pastures provide ideal environments for parasites to thrive and proliferate. The prevalence and distribution of nematode species is also dependent on climatic factors like rainfall and temperature, with the survival of some parasite life stages preferring warm tropical climates and others preferring cooler temperate climates (Roeber and Kahn 2014; McRae *et al.* 2015). Susceptibility of an animal to parasitic infections is also dependent on the age and health status of an animal, with young or old animals, immunocompromised animals, as well as infection with large larval loads predisposing an individual to severe and intense infections (Zajac 2006).

Nematodes are slender and elongated with a bilaterally symmetrical structure, as well as a tubular digestive system with openings at both ends (Finlay *et al.* 2014). The majority of gastrointestinal nematodes are monoxenous, requiring only one host to complete their lifecycle (Fig. 1) (Roeber *et al.* 2013). Eggs are passed onto the pasture in the animal's faeces where the first (L1) and second (L2) larval stages develop and later become infective third-stage larvae (L3). As the ruminant feeds on infected pasture, the parasite is ingested at the L3 life stage, which exsheaths and moves to the abomasum, developing into fourth-stage larvae (L4) and then maturing into adults (Demeler 2005; Roeber *et al.* 2013).

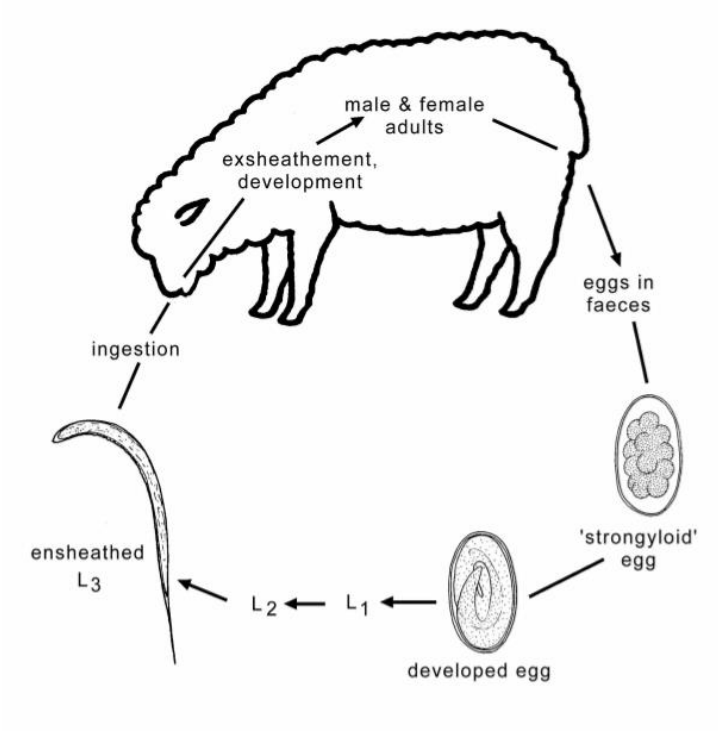


Fig. 1. General life cycle of strongyle nematodes. Eggs are passed onto the pasture in the animal's faeces and hatch into first-stage larvae (L1). They then develop into second-stage larvae (L2) before migrating onto pasture as infective third-stage larvae (L3). Upon ingestion, L3 exsheaths, penetrates and resides within the abomasum or intestines where they continue to develop into fourth-stage larvae (L4) and then mature into adults. Modified from Demeler (2005).

Despite the presence of many nematode species, only few pose a significant threat and cause disease. Most animals would also experience concurrent helminth infections. Common signs of infection with gastrointestinal nematodes are anaemia, scouring, decreased appetite, weight loss, decreased meat and milk yield, a decline in wool quality and low reproductive rates, making them one of the major causes of production losses across the globe. In 2014, an estimated AUD $436 million per annum in losses was recorded in Australia alone, with helminth infections accounting for approximately AUD $260 million (Roeber and Kahn 2014; Lane *et al.* 2015; Emery *et al.* 2016). This number has reduced compared to previous years, with the improvement likely attributed to the decline in animal numbers, as well as the development of new and improved detection and treatment methods, allowing for the

early detection of parasites in the herd and immediate administration of treatment for improved overall animal welfare (Roeber and Kahn 2014).

1.2 Trichostrongyloid nematode infections

The most important disease-causing gastrointestinal nematodes in most regions are part of the order Strongylida belonging to the Trichostrongyloidea superfamily (Table 1). Trichostrongyloid nematodes typically infect the gastrointestinal tracts of grazing ruminants, severely impacting the animals' health as well as their production profitability. This family of nematodes include *Haemonchus* spp. and *Cooperia* spp. which are prevalent in tropical regions, *Teladorsagia/Ostertagia* spp. in temperate regions, and *Trichostrongylus* spp. present throughout both. These nematodes are also commonly termed bursate nematodes due to the males characteristically possessing a copulatory bursa, whilst females of this family produces ovoid eggs that are released into the environment via the host's faeces (Roeber *et al.* 2013).

H. contortus is one of the largest and most pathogenic species of nematodes that infect sheep, goats and cattle, accounting for a large proportion of yearly production losses (Kebeta *et al.* 2020). Their high fecundity, with female worms producing between 5000 to 15,000 eggs per day, coupled with a very short pre-patent period of only 15 days, allows *H. contortus* to establish infection very quickly within a host (Emery *et al.* 2016). *H. contortus* is often referred to as the Barber's pole worm based on its visual appearance, with its characteristic coiled red pseudocoelomic fluid and white-coloured uterus (Roeber *et al.* 2013). Although not a primary cause of scouring, adults of this species feed on the blood of their host and infections at high intensity can cause fatal anaemia (Zajac 2006).

Table 1. Morphological features, host species, location in host and egg morphology of common gastrointestinal nematodes of goats, sheep, and cattle in Australasia (Isenstein 1971; Suarez *et al.* 1993; Sommer 1996). Table adapted from Roeber *et al.* (2013).

Species	**Host**	**Location in host**	**Length (mm)**	**Features**	**Egg morphology**
Haemonchus contortus	Cattle, sheep, goats	Abomasum	M: 10-20 F: 18-30	Red pseudocoelomic fluid and white coiled ovaries giving the appearance of a barber's pole. Presence of vulvar flaps depends on strain.	Curved sides with slightly thicker shell
Teladorsagia circumcincta	Sheep, goats	Abomasum	M: 7-8 F: 10-12	Small head and buccal cavity. Vulvar flaps can be present in females.	Curved sides
Ostertagia ostertagi	Cattle	Abomasum	M: 5-7 F: 6-8	Rounded head. Equal length spicules with tapered ends. Thorn-like cervical papillae in females.	Curved sides
Trichostrongylus axei	Cattle, sheep, goats	Abomasum	M: 2-6 F: 3-8	Dissimilar spicules of unequal length.	Elongated and pointed at one or both ends
T. colubriformis	Sheep, goats	Anterior small intestine	M: 4-8 F: 5-9	Equal length spicules with triangular tip.	Elongated and pointed at one or both ends
T. vitrinus	Sheep, goats	Anterior small intestine	M: 4-7 F: 5-8	Equal length spicules with sharp tips.	Elongated and pointed at one or both ends
T. rugatus	Sheep, goats	Small intestine	M: 4-7 F: 6-7	Dissimilar spicules with foot-like appearance.	Elongated and pointed at one or both ends
Cooperia curticei	Sheep, goats	Small intestine	M: 4-5 F: 5-6	Transverse striation of cuticle in all species. Watch-spring-like body posture and presence of a small cephalic vesicle are characteristic.	Parallel sides
C. oncophora	Cattle	Small intestine	M: 5-9 F: 6-8	Transverse striation of cuticle in all species. Coiled body posture. Large bursa present in male worms.	Parallel sides

Another important member of the trichostrongylid family are nematodes from the genus *Teladorsagia* or *Ostertagia,* also commonly known as brown stomach worms. Formerly known as *Ostertagia circumcincta,* the focal species of this study, *Teladorsagia circumcincta,* is of particular importance to the Australian livestock industry. Highly prevalent in temperate regions, *T. circumcincta* is the most common cause of ovine parasitic gastroenteritis in sheep and goats around the world (Turnbull *et al.* 2019). Unlike *H. contortus, T. circumcincta* does not cause disease by feeding on the host's blood but rather the developing larvae burrow into the abomasal crypts, leading to the formation of nodules (Roeber *et al.* 2013; Venturina *et al.* 2013). Larval burrowing also injures parietal cells and the mucosal lining, causing mucosal hyperplasia which impairs the ruminant host's ability to digest food. The damage also causes important digestive enzymes and proteins to leak from the stomach across the epithelium, which in turn leads to hypoproteinemia (Venturina *et al.* 2013). Other signs of Teladorsagiasis include intermittent diarrhea, dehydration, reduced appetite and weight (Craig *et al.* 2006; McNeilly *et al.* 2009). Adult worms are relatively small, growing between 7 to 12 mm in length, making them harder to detect than the larger *H. contortus*. Similar in pathology, *Ostertagia ostertagi* is prevalent in temperate regions, but mainly cause disease in cattle due to the simultaneous re-activation of arrested larvae in the abomasum (Newton *et al.* 1997).

Trichostrongylus spp., often known as Black scour worms, infects the small intestines of a wide range of herbivores worldwide through ingestion of infective larvae on pasture or contaminated water. In Australasia, *T. colubriformis, T. vitrines* and *T. rugatus* are the most prevalent species and are most pathogenic in young animals but also cause significant decrease in productivity in older sheep (Roeber *et al.* 2013). Species of *Trichostrongylus* develop by burrowing in the intestinal villi for 10 to 12 days and emerge as immature adults, causing enteritis due to major damage of the intestinal walls. Clinical signs include atrophy of the large intestinal mucosa microvilli, hyperplasia, hypertrophy and an influx of leukocytes, which significantly compromises nutrient uptake by the host system. Black scours occur as a result of haemorrhage in the intestines during severe infections

(> 2000 epg). At low intensities (100-500 epg), it may be difficult to differentiate a *Trichostrongylus* infection from malnutrition (Roeber *et al.* 2013).

Cooperia onchophora also commonly infects cattle in temperate regions and are more detrimental to calves due to their under-developed immune system (Kloosterman *et al.* 1984; Armour *et al.* 1987). Although of low pathogenicity, this parasite can contribute to parasitic gastroenteritis during mixed infections with other species such as *O. ostertagi*. *C. oncophora* resides in the intestines causing luminal damage and possibly anaemia, as well as causing up to a 13.5% decrease in weight gain (Li and Gasbarre 2009). In sheep, *Cooperia cuticei* is the dominant species commonly found in the small intestine, although of low pathogenicity in the absence of co-infection (Zajac 2006).

1.3 Parasite control and anthelmintic resistance

Parasite control is generally achieved using anthelmintics which allows for the rapid improvement of animal welfare and productivity, especially in combination with pasture management plans. Common broad-spectrum drugs include benzimidazoles, imidazothiazoles (levamisole) or macrocyclic lactones (avermectins and milbemycins) (Sutherland *et al.* 2003; Sargison *et al.* 2005; McNeilly *et al.* 2009; Roeber *et al.* 2013). The mechanism of benzimidazoles are the most well characterised amongst commonly used anthelmintics. This class of drug targets tubulin, an essential structural protein in eukaryotic cells, disrupting the formation of microtubules and hence diminishing uptake of glucose and protein secretion (Turnbull *et al.* 2019). This in turn leads to starvation and death of the parasite (Keegan *et al.* 2017; Turnbull *et al.* 2019).

Macrocyclic lactones, particularly ivermectin, have been extensively used since the 1980s to treat endo- and ectoparasitic infections in animals and humans (Whittaker *et al.* 2017). This class of drug

acts as a ligand-gated ion channel agonist and irreversibly binds to glutamate-gated chloride (GluCl) channels, causing an influx of chloride ions (Cl^-) in the nerve and muscle cells of nematodes. This hyperpolarisation effect inhibits the normal neuromuscular and physiological processes in the worms, leading to paralysis and starvation (Whittaker *et al.* 2017). Similarly, levamisole is an agonist for nicotinic acetylcholine-gated cation channels which causes the constant activation of cation channels. The continuous depolarisation of the nematode neuromuscular system can lead to paralysis, death or expulsion of the nematode from the host (Whittaker *et al.* 2017).

The excessive and improper use of therapeutic drugs has led to the escalation of resistance against these drugs, growing into a serious problem worldwide. Resistance is defined as the ability of a parasite to survive a usually lethal dose of anthelmintic drug, leading to parasitism of the infected host and the continuation of their life cycle (Besier and Love 2003). Highly fecund species like *H. contortus* are more likely to rapidly develop resistance through intergenerational selection as only worms that survive chemical treatments are able to reproduce, producing drug resistant progeny (Emery *et al.* 2016). The efficacy of a drug can be estimated via several methods. Most commonly, a faecal egg count reduction test (FECRT) is performed to diagnose the extent of anthelmintic resistance in parasites. The FECRT calculates the percentage of decline in number of nematode eggs before and after treatment. In sheep and goats, resistance is considered present if there is less than a 95% reduction in egg count 14 days post-treatment (Flanagan *et al.* 2011). The lack of development of new drugs, coupled with the continual selection for resistance in livestock, will lead to the rise of significant issues in future. This stresses the need for alternative parasite control methods.

The reliance on anthelmintics for controlling nematode infections has been eased for certain species through the development of vaccines. Barbervax® is currently the only vaccine that is commercially available in Australia against *H. contortus* infections in sheep, providing between 75 to 95%

protection in young sheep and periparturient ewes (Britton *et al.* 2020; Kebeta *et al.* 2020). The vaccine uses a galactose-containing glycoprotein complex (H-gal-GP), a hidden antigen from the gut lining of *H. contortus* (Yanming *et al.* 2007; Broomfield *et al.* 2020). Immunological protection is gained after the parasite ingests host antibodies during blood feeding which results in binding to the hidden antigens and disruption of the worm's gut function (Broomfield *et al.* 2020). Although the vaccine does not immediately kill off adult worms, the egg output is significantly reduced, therefore decreasing parasite contamination on pasture (Broomfield *et al.* 2020).

Integrated pest management (IPM) is a synergetic control strategy to effectively minimise parasite numbers in a herd. IPM involves the merging of multiple control methods to manage parasite numbers both on pasture and within the animals. This usually involves the combination of chemical treatment with pasture management, nutritional supplementation, and breeding for resistance (Britton *et al.* 2020). Chemical treatment is not always possible due to concerns of chemical residues remaining within the animal and becoming a public health risk when distributed in the food chain. In these instances, biological control methods may be more favourable and eliminates the need for a withholding period before distribution of animal produce. Invertebrates such as earthworms and dung beetles are commonly found on pasture and help to break down and remove faecal matter, which in turn considerably reduces the number of free-living nematode stages. However, the presence and activity of these invertebrates are highly dependent on weather conditions which makes them difficult to rely on (Waller 2006).

Nematophagous fungi have been successfully used as a parasite control method in livestock. Specifically, *Duddingtonia flagrans* is often used as a feed additive due to the availability of spore material on a commercial scale, as well as its ability to survive the gut passage of the host, showing reduced numbers of infective nematode larvae in goats and horses upon treatment (Larsen *et al.* 1995;

Braga and de Araújo 2014; Buzatti *et al.* 2015; Vilela *et al.* 2020). This predacious fungi feeds on nematode larvae after trapping them in microscopic nets within the faeces (Peña *et al.* 2002). *D. flagrans* is a successful control method as it can grow very rapidly in faeces, effectively killing the parasite and inhibiting the continuation of the helminth's life cycle. Aside from having no detrimental effects on the environment, *D. flagrans* has also been described to be a clonal population worldwide with minimal genetic variability between isolates from different areas and are naturally occurring at low levels in the soil (Waller 2006).

1.4 Immune response against infection

1.4.1 Host-pathogen interactions

In response to infection with GI nematodes, the host elicits a T helper type 2 cell (Th2) immune response, triggering the expression of innate immune cells (McNeilly *et al.* 2009; Menzies *et al.* 2010; Maizels *et al.* 2012; Pemberton *et al.* 2012). Exposure to parasitic E/S antigens are associated with increased levels of Th2 cytokines, granular and globular leukocytes, along with parasite-specific antibodies which include the isotyopes IgA, IgG1 and IgE (McRae *et al.* 2015). Elevated IgA levels in sheep is positively correlated with *T. circumcincta* resistance, with worm length and fecundity being significantly diminished (Strain *et al.* 2002; McRae *et al.* 2015). IgA is predominantly expressed in host mucosa and has been known to recognise and mediate activity against a range of antigens from parasite L4 stages in the host intestine (McRae *et al.* 2015). IgA from the gastrointestinal tract can bind to parasites or their E/S antigens, with resistant sheep shown to carry higher levels of parasite-specific serum IgA in the bloodstream, which is positively correlated to mucosal IgA production in the intestines (McRae *et al.* 2014).

The establishment of a Th2 cytokine environment during infection, specifically interleukin-4 (IL-4), leads to the secretion of IgE and IgG1 following the activation of B cells (Yasuda and Nakanishi

2018). IgE is strongly associated with immune responses against parasitic infections, and can act in conjunction with larvae-specific IgG1 to mediate helminth expulsion (Negrão-Corrêa 2001). A study successfully showed that the administration of a combination of IgG1 and IgE to activation-induced cytidine deaminase (AID)-deficient mice, incapable of switching IgM to other isotypes, infected with *Nippostrongylus brasiliensis* promoted accelerated expulsion of the nematodes in a dose-dependent manner (Matsumoto *et al.* 2013). Increased IgG1 and IgE levels are also associated with reduced faecal egg count in sheep (McRae *et al.* 2015).

A recent study by Hernández *et al*. (2019) discovered that the vaccination of recombinant interleukin-5 (rIL-5) significantly lowers the blood eosinophil count in sheep, leaving them susceptible to subsequent experimental infection with *H. contortus*. IL-5 is an important Th2-associated cytokine involved in the differentiation and activation of eosinophilic responses against parasites, and increases the secretion of IgA from B cells (Yasuda and Nakanishi 2018). Vaccination of naturally nematode-resistant Canaria Hair Breed (CHB) sheep with rIL-5 negated the effects of naturally produced IL-5, hindering the stimulation of eosinophils and resulted in higher worm burden in the immunised group compared to the control group (Hernández *et al.* 2019). This study concluded that IL-5 and eosinophilic activation is required for the expulsion of *H. contortus* and targeting IL-5 production in sheep could potentially be an effective control method to decline nematode physiology and fecundity (Hernández *et al.* 2019).

Intelectins are also highly up-regulated during Th2 immune responses and have been identified as a hallmark protein of GI nematode infections, potentially playing an important role in parasite expulsion (Artis 2006). Intelectins are galactose-binding lectins expressed by mucus neck cells in the abomasum (Artis 2006; French *et al.* 2009; Pemberton *et al.* 2011). Although their functional importance has yet to be fully determined, intelectins are believed to be involved in anti-nematode

responses through direct and indirect interactions. A study by Artis (2006) suggested that intelectins may directly impair nematode fitness in the GI tract by binding to carbohydrate residues, making the parasites more susceptible to elimination. Three homologues of intelectin have been described in sheep: intelectin 1 (ITLN1), intelectin 2 (ITLN2) and intelectin 3 (ITLN3) (French *et al.* 2009). ITLN1 is widely reported to be critical in the host's innate defence against microbial infections (Komiya *et al.* 1998) while ITLN2 is the dominant intelectin that is upregulated around the period of worm expulsion (Artis 2006). Little is known about the function of ITLN3, however sequencing performed by French *et al*. (2009) revealed an 86% and 91% homology to ITLN1 and ITLN2, respectively. Although the mode of action is unknown, intelectins bind galactofuranosyl-containing residues in bacterial cell wall, therefore ITLN3 may interact with similar residues in parasitic nematodes and interfere with worm functions (Artis 2006).

Mucus is a gel-like substance that coats the GI epithelium, which physically separates and prevents the establishment of pathogens in the host (van Putten and Strijbis 2017; Sharpe *et al.* 2018). Although previously thought to only function as a physical barrier, recent studies have shown that mucus possesses additional functions that include presentation of specific ligands to trap pathogens (Sharpe *et al.* 2018), preventing parasitic nematodes like *T. circumcincta* from penetrating host abomasal glands (Pemberton *et al.* 2012). Mucus is mainly comprised of high molecular weight (> 1 MDa), heavily O-glycosylated glycoproteins known as mucins (Hasnain *et al.* 2011; Pemberton *et al.* 2011; Sharpe *et al.* 2018). Mucins can be categorised into two main subtypes: secreted mucins and membrane-bound or transmembrane mucins (Sharpe *et al.* 2018). Secreted mucins are responsible for the gel-like and viscous property of mucus while transmembrane mucins have both barrier and signalling functions (van Putten and Strijbis 2017).

Several mucin variants are of importance in the GI tract during infection. MUC6 is highly expressed in the normal stomach. Although the functions of this mucin are not adequately described in sheep, MUC6 has been known to co-localise with a trefoil factor family (TFF) peptide, TFF2, in mucus neck cells (Kjellev 2009; Scott *et al.* 2017). Surface mucus cells produce and secrete MUC5AC, a gel-forming mucin, together with mucin 2 (MUC2) (Hasnain *et al.* 2011; Scott *et al.* 2017). MUC5AC has been described to be involved in anti-nematode responses across various species, where it is upregulated shortly before worm expulsion (Hasnain *et al.* 2011). A previous study in mice has revealed that MUC5AC-deficient mice had a higher T helper type 1 (Th1) instead of Th2-type response, which was associated with chronic *T. muris* infection (Hasnain *et al.* 2011). The same study showed that the absence of MUC5AC resulted in more porous mucus, allowing for increased worm viability. Different to MUC2, MUC5AC is not easily degraded by nematode E/S-derived proteases, making MUC5AC a potential therapeutic target to induce parasite expulsion (Sharpe *et al.* 2018). Intelectins may indirectly aid nematode expulsion by altering the properties of mucus glycoproteins (Artis 2006; French *et al.* 2008). Previous studies on toads have shown that a homologous lectin in *Xenopus laevis* eggs (XL35) hardens the egg coat protein by interacting with mucus-like proteins to prevent polyspermy (Nishihara *et al.* 1986; Chang *et al.* 2004). Artis (2006) postulates that a similar cross-linking interaction may occur with nematode intelectins, forming a 'glycoprotein cement' on the nematode exterior to aid expulsion. Further exploration would be beneficial to determine if intelectins directly interact with nematodes, and if the cross-linking of mucus on the parasite surface interferes with motility.

Trefoil factors are cysteine-rich proteins that are secreted with mucus, forming complex structures with mucins that can influence mucus viscosity (Poulsen *et al.* 2005; Sharpe *et al.* 2018). The trefoil factor family (TFF) consists of three peptide variants: TFF1, TFF2 and TFF3 (Taupin and Podolsky 2003; Poulsen *et al.* 2005). This family of proteins are typically expressed in response to mucosal

damage (Taupin and Podolsky 2003). Previous findings have shown that all three members of the TFF are rapidly induced upon inflammatory injury, with TFF2 being upregulated as quickly as 30 mins post-injury (Taupin and Podolsky 2003). TFF2 has two trefoil domains, compared to one domain in the other variants, and is also more compact in structure, producing the highest viscosity amongst the three variants upon interaction with mucus (Poulsen *et al.* 2005). TFF2 is believed to be involved in aiding nematode expulsion. A study in mice infected with the hookworm *Nippostrongylus brasiliensis* showed a higher worm burden in TFF2 deficient mice, along with significantly more eggs deposited in the faeces (Wills-Karp *et al.* 2012). Mucins and TFFs act in conjunction with calcium activated chloride channels (CLCA1), which regulates mucus hydration via osmotic fluid transfer, to alter mucus viscosity (Rowe *et al.* 2009). Overall, it is evident that host molecules act in concert to prevent the establishment of parasites and effectively eliminate them from the host system.

1.4.2 Parasite immune evasion

Helminths modulate the host immune response, allowing them to remain in the host for long periods of time (Hoerauf *et al.* 2005; Hewitson *et al.* 2009). This is achieved through the stage-specific release of excretory/secretory (E/S) products, such as proteins, by the parasite to maintain a state of active suppression (Craig *et al.* 2006; Hewitson *et al.* 2009). These E/S products are fundamental for parasite survival in the host, playing a role in host penetration, feeding and evasion of the host's anti-parasitic immune response (Knox 2000). Each species of helminth secretes unique immunomodulatory molecules to target different aspects of the host immune response (Finlay *et al.* 2014). Disrupting the activity of these E/S proteins may lead to quicker elimination of the parasite in a host (Craig *et al.* 2006).

It has been described that parasitic stages of different species of nematodes have different levels of effectiveness at modulating host immunity. A previous study determined that the E/S products from

O. ostertagi L4 are able to suppress T cell proliferation *in vitro* but the E/S products of the L3 stages are not capable of active suppression (Gómez-Muñoz *et al.* 2004). Likewise, E/S products from *T. circumcincta* L4 can suppress T cell activation in sheep *in vitro* through the upregulation of the transcription factor Foxp3 which is involved T cell regulation (McNeilly *et al.* 2013). In contrast, E/S products from the L4 stages of other species like *Nematodirus battus* and *T. vitrinus* have no suppressive effects (McNeilly and Nisbet 2014).

A proteomic analysis performed by Smith *et al.* (2009) identified E/S products released by *T. circumcincta* larvae between 1-5 days post infection (dpi), showing that L3 secrete an astacin-like metalloprotease upon exsheathing, similar to that found in *O. ostertagi*, along with cathepsin F which is known to be abundant in *T. circumcincta* L4 (Smith *et al.* 2009). A lectin-binding protein known as galectin was also found to be highly expressed in the infective stages of *T. circumcincta* and is of particular interest due to its potential role in the modulation of host response.

1.5 Galectin

1.5.1 Galectin expression and functions

Galectins are a family of β-galactoside-binding lectins expressed by nearly all immune cells that bind to N-acetyllactosamine-containing glycans, especially lactose (Greenhalgh *et al.* 1999; Rabinovich and Toscano 2009). Galectins are characterised by the presence of at least one carbohydrate recognition domain (CRD) consisting of about 135 amino acid residues (Hughes 1999). This family of proteins can be categorised into three groups: prototype galectins, tandem repeat-type galectins, and chimera-type galectins (Fig. 2). Prototype galectins can exist as monomers or dimers and have a single CRD; tandem repeat-type galectins are bivalent in nature and possess two CRDs with different specificities; chimera-type galectin 3 contains one CRD that is connected to a non-lectin amino-terminal region (Rabinovich and Toscano 2009; Popa *et al.* 2018; Xu *et al.* 2018). At least 15 subtypes

of galectins have been identified in mammals thus far and have been named in order of their discovery (Xu *et al.* 2018). On the other hand, non-mammalian galectins are named according to the organism they were discovered in (Vasta 2009).

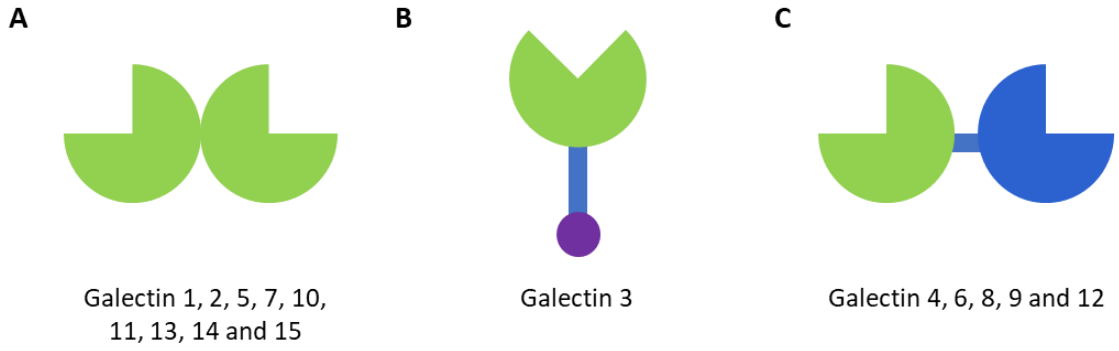


Fig. 2. Structure of mammalian galectins. (A) Dimeric prototype galectin with a single carbohydrate recognition domain (CRD). (B) Chimera-type galectin with a single CRD and non-lectin amino-terminal region. (C) Tandem repeat-type galectin with two CRDs of different binding specificities (Vasta 2009).

Studies have shown that galectins may have differing, and occasionally opposing roles inside and outside the cell, and are extensively involved in the innate and adaptive immune responses of mammals (Naqvi *et al.* 2020). Although their exact functions are uncertain, galectins seem to be involved in a variety of biological and immune processes such as host-pathogen recognition, cell adhesion, cell growth regulation, T cell polarisation, and apoptosis (Rabinovich and Gruppi 2005; Craig *et al.* 2006; Hewitson *et al.* 2009). The differential effects of galectins on immune responses may be influenced by a variety of intrinsic and extrinsic factors (Rabinovich and Toscano 2009).

Galectin function has been shown to be influenced by the galectin-receptor interaction on its target cell (Yuan *et al.* 2015; Naqvi *et al.* 2020). Galectins are also thought to be involved in cell signalling and activation, where some subtypes amplify immune responses while others inhibit the spread of inflammation (Rabinovich *et al.* 2002). Several studies have revealed that activated immune cells under certain inflammatory conditions can induce apoptosis in effector T cells at the end of an immune response by secreting an excess of galectin 1 (Rabinovich *et al.* 2002; Rabinovich and Toscano 2009). Galectin 9 induces the chemotaxis, activation and degranulation of eosinophils (Vasta 2009). On the other hand, galectin 3 appears to be anti-apoptotic and instead stimulates cell proliferation, DNA synthesis and growth of fibroblasts (Rabinovich *et al.* 2002).

1.5.2 Galectins in small ruminants

Galectin 11 (LGALS11) and 14 (LGALS14) localise in the abomasum of small ruminants and are upregulated during parasite infections (Young and Meeusen 2002; Ekwemalor *et al.* 2018). Exclusively found in ovine eosinophils, LGALS14 has been shown to only be secreted following the recruitment and activation of eosinophils during allergic reactions or parasitic infections (Young *et al.* 2009). LGALS14 binds to terminal N-acetyllactosamine residues and has the highest affinity for lacto-N-neotetraose which is expressed by helminths and can drive a Th2 immune response in the host. This prototype galectin is expressed in the gut and lung mucosa, and has been shown to bind to laminin and can cross-link mucins to change the rheological properties of mucus (Young *et al.* 2009). LGALS14 is also a known orthologue of galectin-10, which is also associated with eosinophilic responses to helminth infections and allergic reactions in humans (Young *et al.* 2012).

Similarly, LGALS11 in the gut has been found to be upregulated during parasitic infections as well as pregnancy (Lewis *et al.* 2007; Hoorens *et al.* 2011; Sakthivel *et al.* 2020). Also prototypical in structure, LGALS11 localises in the cytoplasm and nucleus of epithelial cells and is secreted by

abomasal crypts into the mucus layer during infection (Hoorens *et al.* 2011). LGALS11 is associated with eosinophilic inflammatory responses against parasite larvae, and is also involved in both innate and adaptive immune responses as it crosslinks mucus glycans to form a physical barrier against parasitic nematodes (Vasta 2009). A study in sheep suggested that upregulation of LGALS11 by parasites is restricted to the GI tract and does not occur at other mucosal surfaces (Hoorens *et al.* 2011). Additionally, induction of LGALS11 was not stimulated by the E/S products of *O. ostertagia*, but rather is an indirect response to the parasite (Hoorens *et al.* 2011). Expression of LGALS11 can be induced by specific parasite life stages, specifically the infective L3 stages of parasitic nematodes as they possess major surface antigens that are absent in the L4 and adult stages (Preston *et al.* 2015). The same study found that LSGALS11 bound to the pharynx region of L4 and adult *H. contortus*, disrupting feeding and development (Preston *et al.* 2015).

1.5.3 Galectins in gastrointestinal nematodes

Genes encoding for tandem repeat-type galectins have been isolated from the ovine parasitic gastrointestinal nematodes *T. circumcincta* (Tci-gal-1 and Tci-gal-2), *H. contortus* (Hco-gal-1), and *Trichostrongylus colubriformis* (Tco-gal-2), and have been hypothesized to be involved in parasite invasion and immunomodulation (Greenhalgh *et al.* 1999; Xu *et al.* 2018). For example, Hco-gal-1 binds with glycoproteins which are essential to cell migration, phagocytosis and secretion of cytokines (Xu *et al.* 2018). Another parasite galectin is Tsgal, a tandem repeat-type galectin isolated from *Trichinella spiralis* that is speculated to mimic the eosinophil chemokinetic functions of host galectin 9 to evade detection by the host immune system, allowing for parasite establishment (Xu *et al.* 2018). Overall, very little is known about parasite galectins and their roles in the host during infection, warranting the need for further investigation.

1.6 Conclusion

The parasitic nematode *T. circumcincta* is a major disease-causing pathogen of sheep, contributing to economic losses of up to AUD $260 million annually. Current parasite control relies on IPM, involving the combination of drugs and pasture management, however the exponential increase in anthelmintic resistance warrants the urgent need for more sustainable control options such as vaccines. Galectin plays a major role in regulating the host immune response, interacting with a wide range of host proteins of varying functions. As such, the role of *T. circumcincta* galectin (Tci-gal-1) in host-parasite interactions require further investigation. A greater understanding of Tci-gal-1 binding partners could potentially be used as a vaccine or therapeutic targets to control this parasite.

References

Armour, J, Bairden, K, Holmes, PH, Parkins, JJ, Ploeger, H, Salman, SK, McWilliam, PN (1987) Pathophysiological and parasitological studies on *Cooperia oncophora* infections in calves. *Research in Veterinary Science* **42**, 373-381.

Artis, D (2006) New weapons in the war on worms: identification of putative mechanisms of immune-mediated expulsion of gastrointestinal nematodes. *International Journal for Parasitology* **36**, 723-733.

Besier, RB, Love, SCJ (2003) Anthelmintic resistance in sheep nematodes in Australia: the need for new approaches. *Australian Journal of Experimental Agriculture* **43**, 1383-1391.

Bonnet, C, Bouamra-Mechemache, Z, Réquillart, V, Treich, N (2020) Viewpoint: Regulating meat consumption to improve health, the environment and animal welfare. *Food Policy* **97**, 1-11.

Braga, FR, de Araújo, JV (2014) Nematophagous fungi for biological control of gastrointestinal nematodes in domestic animals. *Applied microbiology and biotechnology* **98**, 71-82.

Britton, C, Emery, DL, McNeilly, TN, Nisbet, AJ, Stear, MJ (2020) The potential for vaccines against scour worms of small ruminants. *International Journal for Parasitology* **50**, 533-553.

Broomfield, MA, Doyle, EK, Kahn, LP, Smith, WD, Walkden-Brown, SW (2020) A simplified Barbervax® vaccination regimen in lambs to evoke immunological protection to *Haemonchus contortus*. *Veterinary Parasitology* **287**, 109243.

Buzatti, A, de Paula Santos, C, Fernandes, MAM, Yoshitani, UY, Sprenger, LK, dos Santos, CD, Molento, MB (2015) *Duddingtonia flagrans* in the control of gastrointestinal nematodes of horses. *Experimental Parasitology* **159**, 1-4.

Chang, BY, Peavy, TR, Wardrip, NJ, Hedrick, JL (2004) The *Xenopus laevis* cortical granule lectin: cDNA cloning, developmental expression, and identification of the eglectin family of lectins. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **137**, 115-129.

Craig, H, Wastling, JM, Knox, DP (2006) A preliminary proteomic survey of the in vitro excretory/secretory products of fourth-stage larval and adult *Teladorsagia circumcincta*. *Parasitology* **132**, 535-543.

Demeler, J (2005) The physiological site of action and the site of resistance to the macrocyclic lactone anthelmintics in sheep parasitic trichostrongyloid nematodes. University of Veterinary Medicine Hanover.

Ekwemalor, K, Adjei-Fremah, S, Asiamah, E, Eluka-Okoludoh, E, Osei, B, Worku, M (2018) Systemic expression of galectin genes in periparturient goats. *Small Ruminant Research* **168**, 60-68.

Emery, DL, Hunt, PW, Le Jambre, LF (2016) *Haemonchus contortus*: the then and now, and where to from here? *International Journal for Parasitology* **46**, 755-769.

Finlay, CM, Walsh, KP, Mills, KHG (2014) Induction of regulatory cells by helminth parasites: exploitation for the treatment of inflammatory diseases. *Immunological Reviews* **259**, 206-230.

Flanagan, A, Edgar, HW, Gordon, A, Hanna, RE, Brennan, GP, Fairweather, I (2011) Comparison of two assays, a faecal egg count reduction test (FECRT) and a coproantigen reduction test (CRT), for the diagnosis of resistance to triclabendazole in *Fasciola hepatica* in sheep. *Veterinary Parasitology* **176**, 170-176.

French, AT, Knight, PA, Smith, WD, Brown, JK, Craig, NM, Pate, JA, Miller, HR, Pemberton, AD (2008) Up-regulation of intelectin in sheep after infection with *Teladorsagia circumcincta*. *International Journal for Parasitology* **38**, 467-475.

French, AT, Knight, PA, Smith, WD, Pate, JA, Miller, HR, Pemberton, AD (2009) Expression of three intelectins in sheep and response to a Th2 environment. *Veterinary Research* **40**, 1-13.

Gollnow, S, Lundie, S, Moore, AD, McLaren, J, van Buuren, N, Stahle, P, Christie, K, Thylmann, D, Rehl, T (2014) Carbon footprint of milk production from dairy cows in Australia. *International Dairy Journal* **37**, 31-38.

Gómez-Muñoz, MT, Canals-Caballero, A, Almeria, S, Pasquali, P, Zarlenga, DS, Gasbarre, LC (2004) Inhibition of bovine T lymphocyte responses by extracts of the stomach worm *Ostertagia ostertagi*. *Veterinary Parasitology* **120**, 199-214.

Greenhalgh, CJ, Loukas, A, Newton, SE (1999) The organization of a galectin gene from *Teladorsagia circumcincta*. *Molecular and Biochemical Parasitology* **101**, 199-206.

Hasnain, SZ, Evans, CM, Roy, M, Gallagher, AL, Kindrachuk, KN, Barron, L, Dickey, BF, Wilson, MS, Wynn, TA, Grencis, RK, Thornton, DJ (2011) Muc5ac: a critical component mediating the rejection of enteric nematodes. *The Journal of Experimental Medicine* **208**, 893-900.

Hernández, JN, Meeusen, E, Rodríguez, F, Piedrafita, D, González, JF (2019) Increased susceptibility to *Haemonchus contortus* infection by interleukin-5 modulation of eosinophil responses in sheep. *Parasite Immunology* **42**,

Hewitson, JP, Grainger, JR, Maizels, RM (2009) Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. *Molecular and Biochemical Parasitology* **167**, 1-11.

Hoerauf, A, Satoguina, J, Saeftel, M, Specht, S (2005) Immunomodulation by filarial nematodes. *Parasite Immunology* **27**, 417-429.

Hoorens, P, Rinaldi, M, Mihi, B, Dreesen, L, Grit, G, Meeusen, E, Li, RW, Geldhof, P (2011) Galectin-11 induction in the gastrointestinal tract of cattle following nematode and protozoan infections. *Parasite Immunology* **33**, 669-678.

Huang, CC, Wang, LC, Pan, CH, Yang, CH, Lai, CH (2014) Investigation of gastrointestinal parasites of dairy cattle around Taiwan. *Journal of Microbiology, Immunology, and Infection* **47**, 70-4.

Hughes, RC (1999) Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochimica et Biophysica Acta* **1473**, 172-185.

Isenstein, RS (1971) Hybridization of Two Species of Nematodes Parasitic in Ruminants, *Cooperia oncophora* (Railliet, 1898) Ransom, 1907, and *Cooperia pectinata* Ransom, 1907. *The Journal of Parasitology* **57**, 320-326.

Ismail, ZB (2017) Mastitis vaccines in dairy cows: Recent developments and recommendations of application. *Veterinary World* **10**, 1057-1062.

Kebeta, MM, Hine, BC, Walkden-Brown, SW, Kahn, LP, Doyle, EK (2020) Evaluation of Barbervax® vaccination for lambing Merino ewes. *Veterinary Parasitology* **283**, 109187.

Keegan, JD, Good, B, de Waal, T, Fanning, J, Keane, OM (2017) Genetic basis of benzimidazole resistance in *Teladorsagia circumcincta* in Ireland. *Irish Veterinary Journal* **70**, 8.

Kjellev, S (2009) The trefoil factor family - small peptides with multiple functionalities. *Cellular and Molecular Life Sciences* **66**, 1350-1369.

Kloosterman, A, Albers, GAA, Van Den Brink, R (1984) Negative interactions between *Ostertagia ostertagi* and *Cooperia oncophora* in calves. *Veterinary Parasitology* **15**, 135-150.

Knox, DP (2000) Development of vaccines against gastrointestinal nematodes. *Parasitology* **120**, 43-61.

Koinari, M, Karl, S, Ryan, U, Lymbery, AJ (2013) Infection levels of gastrointestinal parasites in sheep and goats in Papua New Guinea. *Journal of Helminthology* **87**, 409-15.

Komiya, T, Tanigawa, Y, Hirohashi, S (1998) Cloning of the novel gene intelectin, which is expressed in intestinal paneth cells in mice. *Biochemical and Biophysical Research Communications* **251**, 759-762.

Lane, J, Jubb, T, Shephard, R, Webb-Ware, J, Fordyce, G (2015) Priority list of endemic diseases for the red meat industries. Meat & Livestock Australia Limited No. B.AHE.0010, Australia.

Larsen, M, Nansen, P, Wolstrup, J, Grønvold, J, Henriksen, SA, Zorn, A (1995) Biological control of trichostrongyles in calves by the fungus *Duddingtonia flagrans* fed to animals under natural grazing conditions. *Veterinary Parasitology* **60**, 321-330.

Lewis, SK, Farmer, JL, Burghardt, RC, Newton, GR, Johnson, GA, Adelson, DL, Bazer, FW, Spencer, TE (2007) Galectin 15 (LGALS15): A gene uniquely expressed in the uteri of sheep and goats that functions in trophoblast attachment. *Biology of Reproduction* **77**, 1027-1036.

Li, RW, Gasbarre, LC (2009) A temporal shift in regulatory networks and pathways in the bovine small intestine during *Cooperia oncophora* infection. *International Journal for Parasitology* **39**, 813-24.

Maizels, RM, Hewitson, JP, Smith, KA (2012) Susceptibility and immunity to helminth parasites. *Current Opinion in Immunology* **24**, 459-466.

Marshall, NA, Taylor, BM, Heyenga, S, Butler, D (2017) Vulnerability of the livestock industry in eastern Australia. *Sustainability Science* **13**, 393-402.

Matsumoto, M, Sasaki, Y, Yasuda, K, Takai, T, Muramatsu, M, Yoshimoto, T, Nakanishi, K (2013) IgG and IgE collaboratively accelerate expulsion of *Strongyloides venezuelensis* in a primary infection. *Infection and Immunity* **81**, 2518-2527.

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

McNeilly, TN, Nisbet, AJ (2014) Immune modulation by helminth parasites of ruminants: implications for vaccine development and host immune competence. *Parasite* **21**, 51.

McNeilly, TN, Rocchi, M, Bartley, Y, Brown, JK, Frew, D, Longhi, C, McLean, L, McIntyre, J, Nisbet, AJ, Wattegedera, S, Huntley, JF, Matthews, JB (2013) Suppression of ovine lymphocyte activation by *Teladorsagia circumcincta* larval excretory-secretory products. *Veterinary Research* **44**, 1-18.

McRae, KM, Good, B, Hanrahan, JP, Glynn, A, O'Connell, MJ, Keane, OM (2014) Response to *Teladorsagia circumcincta* infection in Scottish Blackface lambs with divergent phenotypes for nematode resistance. *Veterinary Parasitology* **206**, 200-207.

McRae, KM, Stear, MJ, Good, B, Keane, OM (2015) The host immune response to gastrointestinal nematode infection in sheep. *Parasite Immunology* **37**, 605-613.

Menzies, M, Reverter, A, Andronicos, N, Hunt, P, Windon, R, Ingham, A (2010) Nematode challenge induces differential expression of oxidant, antioxidant and mucous genes down the longitudinal axis of the sheep gut. *Parasite Immunology* **32**, 36-46.

Murphy, TM, Fahy, KN, McAuliffe, A, Forbes, AB, Clegg, TA, O'Brien, DJ (2006) A study of helminth parasites in culled cows from Ireland. *Preventive Veterinary Medicine* **76**, 1-10.

Naqvi, MA, Memon, MA, Jamil, T, Naqvi, SZ, Aimulajiang, K, Gadahi, JA, Xu, L, Song, X, Li, X, Yan, R (2020) Galectin domain containing protein from *Haemonchus contortus* modulates the immune functions of goat PBMCs and regulates CD4+ T-helper cells in vitro. *Biomolecules* **10**,

Negrão-Corrêa, D (2001) Importance of immunoglobulin E (IgE) in the protective mechanism against gastrointestinal nematode infection: looking at the intestinal mucosae. *Journal of the São Paulo Institute of Tropical Medicine* **43**, 291-299.

Newton, SE, Monti, JR, Greenhalgh, CJ, Ashman, K, Meeusen, EN (1997) cDNA cloning of galectins from third stage larvae of the parasitic nematode *Teladorsagia circumcincta*. *Molecular and Biochemical Parasitology* **86**, 143-153.

Nishihara, T, Wyrick, RE, Working, PK, Chen, YH, Hedrick, JL (1986) Isolation and characterization of a lectin from the cortical granules of *Xenopus laevis* Eggs. *Biochemistry* **25**, 6013-6020.

Pemberton, AD, Brown, JK, Craig, NM, Pate, J, McLean, K, Inglis, NF, Knox, D, Knight, PA (2012) Changes in protein expression in the sheep abomasum following trickle infection with *Teladorsagia circumcincta*. *Parasitology* **139**, 375-385.

Pemberton, AD, Verdon, B, Inglis, NF, Pearson, JP (2011) Sheep intelectin-2 co-purifies with the mucin Muc5ac from gastric mucus. *Veterinary Research* **91**, 53-57.

Peña, MT, Miller, JE, Fontenot, ME, Gillespie, A, Larsen, M (2002) Evaluation of *Duddingtonia flagrans* in reducing infective larvae of *Haemonchus contortus* in feces of sheep. *Veterinary Parasitology* **103**, 259-265.

Popa, SJ, Stewart, SE, Moreau, K (2018) Unconventional secretion of annexins and galectins. *Seminars in Cell & Developmental Biology* **83**, 42-50.

Poulsen, SS, Kissow, H, Hare, K, Hartmann, B, Thim, L (2005) Luminal and parenteral TFF2 and TFF3 dimer and monomer in two models of experimental colitis in the rat. *Regulatory Peptides* **126**, 163-171.

Preston, SJ, Beddoe, T, Walkden-Brown, S, Meeusen, E, Piedrafita, D (2015) Galectin-11: A novel host mediator targeting specific stages of the gastrointestinal nematode parasite, *Haemonchus contortus*. *International Journal for Parasitology* **45**, 791-796.

Rabinovich, GA, Gruppi, A (2005) Galectins as immunoregulators during infectious processes: from microbial invasion to the resolution of the disease. *Parasite Immunology* **27**, 103-114.

Rabinovich, GA, Rubinstein, N, Toscano, MA (2002) Role of galectins in inflammatory and immunomodulatory processes. *Biochimica et Biophysica Acta* **1572**, 274-284.

Rabinovich, GA, Toscano, MA (2009) Turning 'sweet' on immunity: galectin–glycan interactions in immune tolerance and inflammation. *Nature Reviews Immunology* **9**, 338-352.

Roeber, F, Jex, AR, Gasser, RB (2013) Advances in the diagnosis of key gastrointestinal nematode infections of livestock, with an emphasis on small ruminants. *Biotechnology Advances* **31**, 1135-52.

Roeber, F, Kahn, L (2014) The specific diagnosis of gastrointestinal nematode infections in livestock: larval culture technique, its limitations and alternative DNA-based approaches. *Veterinary Parasitology* **205**, 619-28.

Rowe, A, Gondro, C, Emery, D, Sangster, N (2009) Sequential microarray to identify timing of molecular responses to *Haemonchus contortus* infection in sheep. *Veterinary Parasitology* **161**, 76-87.

Ruegg, PL (2017) A 100-Year Review: Mastitis detection, management, and prevention. *Journal of Dairy Science* **100**, 10381-10397.

Sakthivel, D, Preston, S, Gasser, RB, Costa, T, Hernandez, JN, Shahine, A, Shakif-Azam, MD, Lock, P, Rossjohn, J, Perugini, MA, Gonzalez, JF, Meeusen, E, Piedrafita, D, Beddoe, T (2020) The oligomeric assembly of galectin-11 is critical for anti-parasitic activity in sheep (*Ovis aries*). *Communications Biology* **3**, 464.

Sargison, ND, Jackson, F, Bartley, DJ, Moir, ACP (2005) Failure of moxidectin to control benzimidazole-, levamisole- and ivermectin-resistant *Teladorsagia circumcincta* in a sheep flock. *Veterinary Record* **156**, 105-109.

Scott, I, Umair, S, Savoian, MS, Simpson, HV (2017) Abomasal dysfunction and cellular and mucin changes during infection of sheep with larval or adult *Teladorsagia circumcincta*. *PLoS One* **12**, 1-24.

Sharpe, C, Thornton, DJ, Grencis, RK (2018) A sticky end for gastrointestinal helminths; the role of the mucus barrier. *Parasite Immunology* **40**, 1-10.

Sinnathamby, G, Henderson, G, Umair, S, Janssen, P, Bland, R, Simpson, H (2018) The bacterial community associated with the sheep gastrointestinal nematode parasite *Haemonchus contortus*. *PLoS One* **13**, e0192164.

Smith, SK, Nisbet, AJ, Meikle, LI, Inglis, NF, Sales, J, Beynon, RJ, Matthews, JB (2009) Proteomic analysis of excretory/secretory products released by *Teladorsagia circumcincta* larvae early post-infection. *Parasite Immunology* **31**, 10-9.

Sommer, C (1996) Digital image analysis and identification of eggs from bovine parasitic nematodes. *Journal of Helminthology* **70**, 143-151.

Strain, SAJ, Bishop, SC, Henderson, NG, Kerr, A, McKellar, QA, Mitchell, S, Stear, MJ (2002) The genetic control of IgA activity against *Teladorsagia circumcincta* and its association with parasite resistance in naturally infected sheep. *Parasitology* **124**, 545-552.

Suarez, VH, Durette-Desset, MC, Cabaret, J (1993) Description of *Ostertagia ostertagi* and *Ostertagia leptospicularis* hybrids in experimentally infected sheep. *The Journal of Parasitology* **79**, 874-878.

Sutherland, IA, Brown, AE, Leathwick, DM, Bisset, SA (2003) Resistance to prophylactic treatment with macrocyclic lactone anthelmintics in *Teladorsagia circumcincta*. *Veterinary Parasitology* **115**, 301-309.

Taupin, D, Podolsky, DK (2003) Trefoil factors: initiators of mucosal healing. *Nature Reviews Molecular Cell Biology* **4**, 721-732.

Turnbull, F, Devaney, E, Morrison, AA, Laing, R, Bartley, DJ (2019) Genotypic characterisation of monepantel resistance in historical and newly derived field strains of *Teladorsagia circumcincta*. *International Journal for Parasitology: Drugs and Drug Resistance* **11**, 59-69.

van Putten, JPM, Strijbis, K (2017) Transmembrane mucins: signaling receptors at the intersection of inflammation and cancer. *Journal of Innate Immunity* **9**, 281-299.

Vasta, GR (2009) Roles of galectins in infection. *Nature Reviews Immunology* **7**, 424-438.

Venturina, VM, Gossner, AG, Hopkins, J (2013) The immunology and genetics of resistance of sheep to *Teladorsagia circumcincta*. *Veterinary Research Communications* **37**, 171-81.

Vilela, VLR, Feitosa, TF, Braga, FR, Santos, Ad, Bezerra, RA, Silva, GLL, Athayde, ACR, Araújo, JV (2020) Use of *Duddingtonia flagrans* in the control of gastrointestinal nematodes of feedlot goats. *Semina: Ciências Agrárias* **41**, 915-924.

Waller, PJ (2006) Sustainable nematode parasite control strategies for ruminant livestock by grazing management and biological control. *Animal Feed Science and Technology* **126**, 277-289.

Whittaker, JH, Carlson, SA, Jones, DE, Brewer, MT (2017) Molecular mechanisms for anthelmintic resistance in strongyle nematode parasites of veterinary importance. *Journal of Veterinary Pharmacology and Therapeutics* **40**, 105-115.

Wills-Karp, M, Rani, R, Dienger, K, Lewkowich, I, Fox, JG, Perkins, C, Lewis, L, Finkelman, FD, Smith, DE, Bryce, PJ, Kurt-Jones, EA, Wang, TC, Sivaprasad, U, Hershey, GK, Herbert, DR (2012) Trefoil factor 2 rapidly induces interleukin 33 to promote type 2 immunity during allergic asthma and hookworm infection. *The Journal of Experimental Medicine* **209**, 607-622.

Xu, J, Yang, F, Yang, DQ, Jiang, P, Liu, RD, Zhang, X, Cui, J, Wang, ZQ (2018) Molecular characterization of *Trichinella spiralis* galectin and its participation in larval invasion of host's intestinal epithelial cells. *Veterinary Research* **49**, 1-15.

Yanming, S, Ruofeng, Y, Muleke, CI, Guangwei, Z, Lixin, X, Xiangrui, L (2007) Vaccination of goats with recombinant galectin antigen induces partial protection against *Haemonchus contortus* infection. *Parasite Immunology* **29**, 319-326.

Yasuda, K, Nakanishi, K (2018) Host responses to intestinal nematodes. *International Immunology* **30**, 93-102.

Young, AR, Barcham, GJ, Kemp, JM, Dunphy, JL, Nash, A, Meeusen, EN (2009) Functional characterization of an eosinophil-specific galectin, ovine galectin-14. *Glycoconjugate Journal* **26**, 423-432.

Young, AR, Barcham, GJ, McWilliam, HE, Piedrafita, DM, Meeusen, EN (2012) Galectin secretion and binding to adult *Fasciola hepatica* during chronic liver fluke infection of sheep. *Veterinary Immunology and Immunopathology* **145**, 362-367.

Young, AR, Meeusen, EN (2002) Galectins in parasite infection and allergic inflammation. *Glycoconjugate Journal* **19**, 601-606.

Yuan, C, Zhang, H, Wang, W, Li, Y, Yan, R, Xu, L, Song, X, Li, X (2015) Transmembrane protein 63A is a partner protein of *Haemonchus contortus* galectin in the regulation of goat peripheral blood mononuclear cells. *Parasites & Vectors* **8**, 211.

Zajac, AM (2006) Gastrointestinal nematodes of small ruminants: life cycle, anthelmintics, and diagnosis. *Veterinary Clinics of North America: Food Animal Practice* **22**, 529-541.

Chapter 2 – *Teladorsagia circumcincta* galectin-mucosal interactome in sheep

This manuscript has been written for submission to International Journal for Parasitology. N. Hafidi performed sample collection, expression and purification of recombinant proteins, isolation of glycopeptide material, and mass spectrometry protein annotation and characterisation. P. Faou, R. Lowe and H. Rajapaksha at the La Trobe University – Comprehensive Proteomics Platform (LTU-CPP, La Trobe University, Melbourne, Victoria) performed mass spectrometry analysis. N. Hafidi wrote the paper with editing by T. Beddoe, J. Swan, C. Cairns and M. Stear.

Teladorsagia circumcincta galectin-mucosal interactome in sheep

Nur Nasuha Hafidi[1,2], Jaclyn Swan[1,2], Pierre Faou[3], Rohan Lowe[3], Harinda Rajapaksha[3], Callum Cairns[1], Michael Stear[1,2] and Travis Beddoe[1,2*]

[1]Department of Animal, Plant and Soil Science and Centre for AgriBioscience (AgriBio), La Trobe University, Victoria 3086, Australia.

[2]Centre for Livestock Interactions with Pathogens (CLiP), La Trobe University, Victoria 3086, Australia.

[3]Department of Biochemistry & Genetics, La Trobe Institute for Molecular Science La Trobe University, Victoria 3086, Australia.

*Corresponding author

Email: t.beddoe@latrobe.edu.au; m.stear@latrobe.edu.au

Abstract

Teladorsagia circumcincta is the most important gastrointestinal parasite in the livestock industry in temperate regions around the world, causing great economic losses. The infective third-stage larvae (L3) of this nematode secretes a large number of excretory-secretory (E/S) molecules that are presented at the host-parasite interface, some of which are likely to play critical roles in modulating the host immune response. One of the most abundant E/S molecules is a protein of unknown function termed Tci-gal-1, which has similarity to mammalian galectins. Galectins are a family of carbohydrate-binding molecules, with characteristic domain organization and affinity for β-galactosides, that mediates a variety of important cellular functions including inflammation and immune responses. To understand the role of Tci-gal-1 at the host-parasite interface, the ligands from sheep abomasal scrapes and whole tissue were identified by galectin-affinity chromatography and mass spectrometry. A total of 135 unique proteins were identified from whole abomasal tissue samples, while 89 proteins were isolated from abomasal scrape samples. Of these proteins, 63 were present in both samples. Many of the host proteins identified, such as trefoil factors and mucin-like proteins, play critical roles in the host response and act in concert to effectively eliminate nematodes during infection. The identification of Tci-gal-1 binding partners could provide new insights on host-parasite interactions, and potentially lead to the development of new interventions.

Keywords

Galectin; Glycoproteins; Parasite-host interaction, Mass spectrometry

1. Introduction

Gastrointestinal (GI) nematodes cause significant production and economic losses to livestock industries worldwide (Hoerauf *et al.* 2005; McNeilly *et al.* 2009; Finlay *et al.* 2014). Infection with GI nematodes such as *Haemonchus contortus*, *Teladorsagia circumcinta* and *Trichostrongylus* spp. results in significant production losses (AUD $260 million in 2011) and increased costs for prevention and control of the disease, with anthelmintics estimated to cost tens of billions of dollars annually worldwide (Roeber *et al.* 2013; Roeber and Kahn 2014). Current control methods are becoming less effective due to the rapid emergence of anthelmintic resistance to commonly used broad-spectrum drugs like benzimidazoles, levamisole or macrocyclic lactones (Sutherland *et al.* 2003; Sargison *et al.* 2005; McNeilly *et al.* 2009). Sustainable control solutions such vaccines and breeding for resistance are required, however increased knowledge of the host-parasite interface is required.

Of particular importance to the sheep industry is the GI nematode *Teladorsagia circumcincta*, the most significant cause of ovine parasitic gastroenteritis in temperate regions around the world (Craig *et al.* 2006; McNeilly *et al.* 2009). Commonly known as the brown stomach worm, this parasite follows a direct life cycle and does not require an intermediate host for larval development (Demeler 2005), with the third (L3) and fourth (L4) larval stages of this parasite being important for establishment of infection. In response to infection with *T. circumcincta*, the host's system elicits a T helper type 2 (Th2) immune response, stimulating the expression of innate defence molecules including cytokines, immunoglobulins, and eosinophils (McNeilly *et al.* 2009; Menzies *et al.* 2010; Maizels *et al.* 2012; Pemberton *et al.* 2012). Establishment of infection within the abomasal glands by L3 may cause clinical signs such as weight loss, reduced appetite, profuse scouring, and occasionally death (Craig *et al.* 2006; McNeilly *et al.* 2009).

Modulation of the host immune response by *T. circumcincta* allows the parasite to persist for long periods of time within the ruminant host (Hoerauf *et al.* 2005; Hewitson *et al.* 2009). This is achieved through the stage-specific release of excretory/secretory (E/S) proteins such as proteases, venom allergen-like proteins, lectins and other enzymes by the parasite to maintain a state of active suppression (Craig *et al.* 2006; Hewitson *et al.* 2009). These E/S proteins are fundamental for parasite survival in the host, playing a role in host tissue penetration, feeding and evasion of the host's anti-parasitic immune response (Knox 2000). Each species of helminth secretes unique immunomodulatory molecules capable of targeting different aspects of the host immune response, establishing favourable conditions for parasite survival (Finlay *et al.* 2014). Disrupting the activity of these E/S proteins may lead to quicker elimination of the parasite in a host (Craig *et al.* 2006). An earlier *in vitro* proteomic study on *T. circumcincta* E/S products revealed that infective L3 and L4 larval stages of *T. circumcincta* produced high levels of galectin (Craig *et al.* 2006).

Galectins are a family of β-galactoside-binding lectins expressed by nearly all immune cells that bind to N-acetyllactosamine-containing glycans, especially lactose (Greenhalgh *et al.* 1999; Rabinovich and Toscano 2009; Shi *et al.* 2018). Galectins are characterised by the presence of at least one carbohydrate recognition domain (CRD) consisting of approximately 135 amino acid residues (Hughes 1999). This family of proteins can be categorised into three groups: prototype galectins (single CRD), tandem repeat-type galectins (two CRDs with different glycan specificities), and chimera-type galectins (single CRD connected to a non-lectin amino-terminal region) (Rabinovich and Toscano 2009; Popa *et al.* 2018; Xu *et al.* 2018). Previous studies have demonstrated the differing, and occasionally opposing, roles of galectin inside and outside the cell (Vasta 2009). Although their exact functions are uncertain, galectins appear to be involved in a variety of biological and immune processes such as host-pathogen recognition, cell adhesion, cell growth regulation, T cell polarisation, and apoptosis (Rabinovich and Gruppi 2005; Craig *et al.* 2006; Hewitson *et al.* 2009).

Genes encoding tandem repeat-type galectins have been isolated from the ovine parasitic gastrointestinal nematodes *T. circumcincta* (Tci-gal-1 and Tci-gal-2), *Haemonchus contortus* (Hco-gal-1), and *Trichostrongylus colubriformis* (Tco-gal-2) (Greenhalgh *et al*. 1999), and have been hypothesized to be involved in parasite invasion and immunomodulation (Xu *et al.* 2018). For example, Hco-gal-1 binds with glycoproteins which are essential for cell migration, phagocytosis and secretion of cytokines (Xu *et al.* 2018). Another tandem repeat-type parasite galectin is Tsgal from *Trichinella spiralis* that is speculated to mimic host galectins, as it shows homology to host galectins and can carry out similar functions, enabling the parasite to evade detection by the host immune system (Xu *et al.* 2018). Tsgal may also play a role in *T. spiralis* larval invasion by binding to host surface proteins that modulate the host's immune response (Xu *et al.* 2018). *Toxascaris leonina* also produces a tandem repeat-type galectin (Tl-gal) that is a homologue of human galectin 9, capable of inhibiting host inflammatory reactions through the suppression of cytokine production (Hwang *et al.* 2016).

Tci-gal-1 is one of the most abundant E/S proteins secreted upon infection with a largely unknown function. There is limited knowledge about this parasite galectin and the effects it could mediate within the infected host. This study is the first to explore Tci-gal-1-binding molecules by using recombinant Tci-gal-1 to identify galectin glycoconjugates within the host abomasum in order to identify sheep glycoproteins in an endeavour to better understand host-parasite interactions.

2. Materials and Methods

2.1 Glycoprotein preparation

Three sheep abomasa collected from a local abattoir were cut along the greater curvature and thoroughly rinsed with MilliQ water to remove the contents. No helminth infections were apparent in

all three abomasa, and no obvious morphological changes was observed. Whole cuts of tissue from the folds of each abomasum as well as mucosal scrapes were prepared. The abomasal whole tissue (WT) and scrape (ST) samples were then washed with phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, pH 7.4), and 1 g aliquots of each sample were set up in duplicate in 2 ml microfuge tubes and frozen at -80 °C for 20 mins. Two 3 mm glass beads (Qiagen, Germany) were added to each tube and samples were homogenised using the Qiagen TissueLyser II (Qiagen, Germany) at 30 Hz for two rounds of 5 mins. The homogenised tissue was resuspended in an equal volume of 1% (w/v) sodium deoxycholate (DOC) dissolved in PBS and centrifuged at 16,000 x *g* for 20 mins. Supernatant was collected and stored at -80 °C until required.

2.2 Tci-gal-1 expression and purification

The Tci-gal-1 (NCBI accession number: U67147.1) gene was commercially synthesised and cloned into the pPICZα vector (Invitrogen, USA) by Bioneer Pacific (South Korea) via the *Pst* I and *Not* I restriction enzyme sites. Subsequent pPICZα-Tci-gal-1 plasmid translates an N-terminal alpha-factor signal sequence and a C-terminal hexahistidine tag flanking the *Pichia pastoris* codon optimised Tci-gal-1 gene. The pPICZα-Tci-gal-1 plasmid was linearised with *Sac* I digest (New England BioLabs Inc., USA) and chemically transformed into *P. pastoris* X33 using the *Pichia* EasyComp™ Transformation Kit according to manufacturer's instructions (Invitrogen, USA). Transformants were plated onto YEPD agar plates (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose, 2% (w/v) agar) containing Zeocin™ (100 µg/ml) and incubated at 28 °C for 4 days. Single colonies were grown in 10 ml of YEPD broth to make a glycerol stock.

Starter culture of Tci-gal-1 *P. pastoris* cells was grown from a glycerol stock inoculated into 50 ml conical tubes containing 10 ml of YEPD and incubated at 28 °C for 48 h in a shaking incubator (180 rpm) (NB-205LF, N-BIOTEK, Korea). Starter culture was use at 1:40 to inoculate 400 ml of

fresh buffered methanol-complex (BMMY) medium (1% (w/v) yeast extract, 2% (w/v) peptone, 1% (w/v) yeast nitrogen base, 100 mM potassium phosphate pH 6.0, 100 µg/ml Zeocin™, and 0.5% (v/v) methanol) in a 2 L baffled flask. Cultures were incubated for 96 h at 28 °C whilst shaking (160 rpm), with 0.5% (v/v) methanol added every 24 h. Cells were pelleted at 6000 x *g* for 30 mins and the supernatant was dialysed using membrane tubing with a 12 kDa molecular weight cut-off into starter buffer (5 mM NaH_2PO_4 pH 7.6, 50 mM NaCl and 2 mM imidazole) at 4 °C for 48 h. Dialysis incubations were repeated at least three times with a minimum of 4 h between each exchange. Dialysed media was concentrated with the application of polyethylene glycol (PEG) 8000 to the outside of the tubing and kept overnight at 4 °C (Astral Scientific, Australia).

His-tagged Tci-gal-1 was purified from dialysed supernatant using nickel-nitrilotriacetic acid (Ni-NTA) affinity chromatography. Briefly, 2 ml of Ni-NTA agarose resin (His60 Ni Superflow Resin, Takara Bio Inc.) was added to a purification column and equilibrated with 10 bed volumes of starter buffer. Concentrated culture supernatants were added to the column and allowed to flow through by gravity at a rate of 1 ml/min. The resin was then washed with 2 column volumes of wash buffer (starter buffer containing 20 mM imidazole). Bound proteins were eluted with 10 ml of elution buffer (starter buffer containing 250 mM imidazole), followed by 5 ml of elution buffer containing 500 mM imidazole. The elution fractions were pooled and buffer-exchanged using Amicon® 3K Ultra-15 centrifugal filter units into storage buffer (25 mM NaH_2PO_4 and 250 mM NaCl, pH 7.6) and stored at 4 °C. Successful expression and purification of recombinant Tci-gal-1 was determined by 12% (w/v) SDS-PAGE, followed by circular dichroism spectroscopy (Kelly *et al.* 2005) and mass spectrometry.

2.3 Determination of lactose binding affinity

The sugar binding affinity of Tci-gal-1 was determined using an assay similar to that described by Greenhalgh and Newton (1999). Briefly, 250 µg of Tci-gal-1 was added to 25 µl of lactose-Sepharose resin (α-Lactose-Agarose, Sigma-Aldrich, USA) equilibrated with PBS containing 4 mM β-mercaptoethanol (MePBS). The suspension was incubated on a rotating wheel at room temperature for 1 h and washed twice with 500 µl MePBS before eluting with 300 µl PBS containing 500 mM lactose. The elution was separated by 12% (w/v) SDS-PAGE and visualised using Coomassie brilliant blue R staining.

2.4 Conjugation of Tci-gal-1 onto N-hydroxysuccinamide (NHS)-activated sepharose

Active Tci-gal-1 was conjugated to NHS-activated Sepharose beads (GE Healthcare Life Science, USA) as described by Swan et al. (2019). Briefly, Tci-gal-1 (21 mg) was added to 8 ml of NHS-activated Sepharose beads and allowed to couple on a rotating wheel for 16 h at 4 °C, followed by 2 h at room temperature. The remaining active sites on the resin was blocked for 2 h at room temperature with 100 mM Tris-HCl pH 8.5, 500 mM NaCl and 10 mM tris-2-carboxyethyl-phosphine (TCEP). Then, the resin was washed three times for 10 mins at room temperature with 10 volumes of two alternating wash buffers. Wash buffer 1 contained 100 mM Tris-HCl pH 8.8, 500 mM NaCl and 10 mM TCEP. Wash buffer 2 contained 100 mM HEPES-HCl pH 6.8, 500 mM NaCl and 10 mM TCEP. Galectin-conjugated Sepharose was stored in 100 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.05% (w/v) sodium azide at 4 °C.

2.5 Sodium periodate treatment

An aliquot of each lysate preparation was treated with 20 mM sodium periodate dissolved in PBS and 50 mM sodium acetate buffer (pH 4.5), as described by Schallig and van Leeuwen (1996). Glycan modification by periodate treatment was confirmed by conducting a lectin blot. Briefly, 20 µl of

500 µg/ml of periodate treated and untreated whole tissue and scrape tissue were dotted onto a polyvinylidene difluoride (PVDF) membrane and probed with horseradish peroxidase conjugated Concanavalin A (ConA-HRP) lectin (Sigma-Aldrich, USA).

2.6 Isolation of Tci-gal-1-binding ligands from sheep abomasal tissue

A batch-binding technique was used to capture Tci-gal-1-binding ligands from WT and ST lysates as well as sodium periodate treated WT and ST lysates (*n*=3). Approximately 500 µg of each lysate was incubated with 400 µg of galectin conjugated Sepharose beads at room temperature for 3 h on a rotating wheel. Unbound proteins in the supernatant were removed by centrifugation at 500 x *g* for 1 min, followed by three washes with 600 µl RIPA dialysis buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 0.05% (v/v) Nonidet P-40, 0.01% (w/v) DOC and 1% (v/v) Triton X-100) for 10 mins each. Captured ligands were eluted with 600 µl of elution buffer (250 mM lactose, 20 mM Tris-HCl pH 8.0 and 100 mM NaCl) for 30 mins at room temperature. The washes and eluted proteins were analysed by 12% (w/v) SDS-PAGE and visualised by silver staining.

Briefly, protein bands on the gels were fixed with 40% (v/v) methanol and 13.5% (v/v) formalin for 10 mins immediately after electrophoresis. Gels were then washed twice in MilliQ water for 5 mins each and soaked in 0.02% (w/v) sodium thiosulfate for 1 min. After rinsing again with water, gels were soaked for 10 mins in 0.01% (w/v) silver nitrate and then in developing solution (3% (w/v) sodium carbonate, 0.05% (v/v) formalin and 1.6 x 10^{-5}% (v/v) sodium thiosulfate) until protein bands were adequately intensified. The reaction was stopped with 2.3 M citric acid. Once successful isolation of Tci-gal-1 glycoconjugates was confirmed, the eluted ligands were assessed using mass spectrometry (La Trobe University – Comprehensive Proteomics Platform (LTU-CPP), La Trobe University, Melbourne, Victoria).

2.7 Mass spectrometry

Eluted proteins were precipitated after adjustment to 0.02% (w/v) deoxycholate and 25% (v/v) trichloroacetic acid. Precipitated protein was washed in cold acetone before reconstitution in urea (8 M urea, 25 mM Tris-HCl, pH 8.0). Disulphide bonds were reduced by addition of TCEP to 2 mM for 60 mins, followed by addition of iodoacetamide to 38 mM for 45 mins in the dark to alkylate the reduced thiols. The sample was then diluted with 20 mM Tris-HCl to reduce the urea concentration below 1 M before addition of sequencing grade trypsin (Promega, USA) to achieve a 1:50 ratio compared to the original protein amount. Trypsin digestion was completed overnight at 37 °C. Tryptic peptides were desalted and concentrated using StageTips according to the published protocols by Rappsilber et al (2007).

Peptides were reconstituted in 0.1% (v/v) trifluoroacetic acid (TFA) and 2% (v/v) acetonitrile (ACN), and 500 µg of peptides were loaded onto C_{18} PepMap 100 µm ID × 2 cm trapping column (Thermo-Fisher Scientific, USA) at 5 µl/min for 6 mins, and washed for 6 mins before switching the pre-column in line with the analytical column (Acquity BHE C_{18}, 1.7 µm, 130 Å and 75 µm ID × 25 cm, Waters). The separation of peptides was performed at 250 nl/min using a linear ACN gradient of buffer A (0.1% (v/v) formic acid, 2% (v/v) ACN) and buffer B (0.1% (v/v) formic acid, 80% (v/v) ACN), starting at 5% buffer B to 35% over 90 mins, then 50% B in 15 mins followed by 95% B in 5 mins. The column was then cleaned for 5 mins at 95% B following a 5 mins equilibration step. Data were collected on a Thermo Orbitrap Eclipse (Thermo-Fisher Scientific) in Data Dependent Acquisition mode using m/z 350–1500 as MS scan range, HCD MS/MS spectra were collected in the orbitrap using a cycle time of 3 seconds per MS scan at 30 000 resolution. Dynamic exclusion parameters were set as follows: exclude isotope on, duration 60 s and using the peptide monoisotopic peak determination mode. Other instrument parameters for the instrument were: MS scan at 120 000

resolution, injection time Auto, AGC target Standard, HCD collision energy 30%, injection time Auto with AGT target at Standard. The isolation window of the quadrupole for the precursor was 1.6 m/z.

2.8 Protein identification and quantification

Raw files consisting of high-resolution MS/MS spectra were processed with PEAKS Studio 10 (build 20190129) software program (Bioinformatics Solutions Inc., Waterloo, ON, Canada) (Tran *et al.* 2019). Data were searched against an *Ovis aries* database (UniProt, June 2020). Spectra were additionally searched against the common contaminants database (common Repository of Adventitious Proteins (cRAP) – https://www.thegpm.org/crap/). Briefly, signature MS/MS spectra were searched using PEAKS DB algorithms using carbamidomethylation of cysteine as a fixed modification, and methionine oxidation as well as protein N-termini acetylation set as a variable modification with up to three modifications allowed per peptide. The maximum number of missed cleavages by trypsin digestion was set to two. Mass tolerances were set to ±10 ppm for parent ions and ± 0.5 Da for fragment ions. Minimum peptide length was set to 7, with a maximum mass of 4600 Da. Minimum and maximum peptide length for unspecific cleavage was set at 8 and 25 amino acids respectively. Protein and peptide spectrum matches (PSM) were reported at a false discovery rate (FDR) of 0.01%. Label free quantification (LFQ) of proteins was achieved using Andromeda, a built-in search engine within MaxQuant (Tyanova *et al.* 2016). Label Free quantification was done with 'Match between runs' using a match window of 0.7 min. Large LFQ ratios were stabilized to reduce the sensitivity for outliers, and data was normalised using the 'Quantile' method. Redundant proteins were identified and removed from the raw mass spectrometry dataset, including captured proteins that were present in the negative controls, as well as proteins that appeared in only one replicate. The average LFQ intensity of each remaining predicted protein was calculated and used to compare protein abundances.

2.9 Protein annotation and glycosylation analysis

Unique proteins were identified by searching the accession number of each protein for matches within the Universal Protein Resource Knowledgebase (UniProtKB) consortium database (*https://www.uniprot.org/*). The FASTA format protein sequences of uncharacterised proteins were scanned for conserved motifs against the InterPro 76.0 protein signature databases (*http://www.ebi.ac.uk/interpro/*), using the InterProScan tool. Additionally, the translated Basic Local Alignment Search Tool (tBLASTn) was used to search for regions of similarity of uncharacterised proteins between sequences in the National Center for Biotechnology Information (NCBI, USA) database. Functional annotations and cellular locations of each protein was predicted by assessing the domains and gene ontology (GO) terms of homologous proteins within UniProt and InterPro. Number of transmembrane domains and signal peptides were noted. Potential N- and O-glycosylation sites were predicted using the NetNGlyc 1.0 Server (*http://www.cbs.dtu.dk/services/NetNGlyc/*) and NetOGlyc 4.0 Server (*http://www.cbs.dtu.dk/services/NetOGlyc/*) respectively.

3. Results

3.1 Expression and purification of Tci-gal-1

Recombinant Tci-gal-1 was expressed in *P. pastoris* and purified by immobilized metal-ion affinity chromatography. The purified protein was analysed by SDS–PAGE and a single protein band at the expected molecular mass of 32.5 kDa was seen, showing that Tci-gal-1 was successfully expressed and purified as a soluble protein (Fig. 1A). Mass spectrometry was performed on purified recombinant protein to confirm its identity.

To confirm if the recombinant Tci-gal-1 was functional, lactose affinity chromatography was conducted (Greenhalgh and Newton 1999). Tci-gal-1 was shown to bind to the lactose-Sepharose beads and was eluted with lactose, indicating Tci-gal-1 was showing lectin-binding activity (Fig. 1B).

It was apparent that Tci-gal-1 was seen in the wash fraction which may be due to the galectin overloading the resin binding capacity. Circular dichroism spectroscopy revealed a predominantly beta-sheet secondary structure, conforming to the spectra in previous literature (Supplementary Fig. S1) (Arata *et al.* 1997).

3.2 Tci-gal-1-binding glycoproteins

The elution profile of galectin-affinity chromatography from WT and ST showed minimal difference between replicates. Treatment of lysates with 20 mM sodium periodate successfully altered the glycan structures as demonstrated by the lectin dot blot (Fig. 2), where the periodate-treated lysates were substantially less recognised by the lectin ConA. This confirms that the bands visualised on the silver stained SDS-PAGE showed non-specific binding of protein extracts to the resin complex in the negative controls (Fig. 3). These non-specific proteins found in the periodate-treated controls were identified across the WT and ST dataset and excluded.

Analysis by MS/MS revealed a total of 990 and 821 proteins captured from the WT and ST lysates respectively, from a great diversity of cellular locations and with varying functions. The WT pull-down assay identified 135 unique proteins (Fig. 4A) which bound specifically to galectin conjugated Sepharose beads, while 855 proteins were identified as non-specific (Supplementary Table S1). The major cellular localisation of unique proteins was in the cell membrane and cytoplasm, with 31 proteins in each subcellular location. Proteins in these cellular locations perform various functions, most dominantly being involved in regulatory functions as well as metabolic processes within the cell. Transmembrane domains were identified in 27 of the 31 defined membrane proteins. The remaining proteins were found in the mitochondria (22/135), extracellular space (12/135) nucleus (11/135), endoplasmic reticulum (ER) (9/135), cytoskeleton (7/135), Golgi apparatus (5/135), ribosome (4/135), and the lysosome (3/135) (Fig. 4A). Eleven of the 135 unique proteins

demonstrated no predicted glycosylation sites (Supplementary Table S1). For example, the most abundant WT protein, Solute carrier family 25 member 31, had no predicted glycosylation sites and was a mitochondrial protein. The top 50 most abundant WT proteins are displayed in Table 1.

From the ST pull-down, 89 unique proteins were isolated (Supplementary Table S2), with 732 proteins identified as non-specific. Similar to the WT data, a large proportion of the identified ST proteins were derived from the cell membrane (24/89) and cytoplasm (18/89). Transmembrane domains were identified in 22 of 24 membrane proteins. This is followed by proteins in the extracellular space (9/89), mitochondria (9/89), ER (7/89), nucleus (6/89), cytoskeleton (6/89), Golgi apparatus (4/89), ribosome (3/89), and lysosome (3/89) (Fig. 4B). Seven of the 89 unique proteins demonstrated no predicted glycosylation sites (Supplementary Table S2). Non-glycosylated proteins in both datasets consisted of mainly ribonuclear transport proteins. The top 50 most abundant ST proteins are displayed in Table 2.

The pull-down assays from both tissue types showed a large proportion of overlapping identified proteins, with 63 proteins observed to be present in both datasets (Fig. 5). The extracellular immune protein trefoil factor 2 (TFF2) presented as the most abundant protein in ST samples and second most abundant in WT samples, alongside other important immune proteins such as integrins, immunoglobulins, major histocompatibility complex (MHC) domains, and an angiotensin-converting enzyme (ACE). Another notable protein is the sea urchin sperm, enterokinase and agrin (SEA) domain-containing proteins which are present in both datasets and are related to mucins, along with an Fc fragment of IgG binding protein (FCGBP) in the WT sample.

4. Discussion

T. circumcincta secretes a range of proteins to extract nutrients and modulate the host immune response to ensure survival within its ruminant host. Tci-gal-1 has been identified as a galectin that has the potential to modulate the host immune response. However, the function of this parasite galectin remains to be elucidated. This study demonstrates that functional Tci-gal-1 can be expressed and purified from a yeast expression system, allowing for the identification of ligands from sheep abomasum whole and scrape tissue using mass spectrometry. A total of 135 whole abomasal tissue proteins and 89 scrape tissue proteins were found to bind specifically to Tci-gal-1-conjugated resin in a carbohydrate-dependant manner. Of these proteins, 63 were found to be present in both tissue types. Investigation of these ovine abomasal proteins focussed on extracellular and membrane proteins, considering galectin is secreted into the extracellular space and is therefore unlikely to naturally interact with the subcellular proteins in the mitochondria, cytoskeleton, ER and ribosome during infection (Vasta 2009). This study indicates that Tci-gal-1 interacts with an array of host glycoproteins, potentially identifying parasitic defensive interactions.

The membrane protein TFF2 was a highly abundant protein identified in both sample types. Trefoil factors are cysteine-rich proteins that are secreted in mucus, forming complex structures with mucins that can influence mucus viscosity (Poulsen *et al.* 2005; Sharpe *et al.* 2018). The trefoil factor family (TFF) consists of three peptide variants: TFF1, TFF2 and TFF3 (Taupin and Podolsky 2003; Poulsen *et al.* 2005). This family of proteins are typically expressed in response to mucosal damage (Taupin and Podolsky 2003). Previous findings have shown that all three members of the TFF are rapidly induced upon inflammatory injury, with TFF2 being upregulated as quickly as 30 mins post-injury (Taupin and Podolsky 2003). TFF2 has two trefoil domains, compared to one domain in the other variants, and is more compact in structure which produces a higher viscosity upon interaction with mucins (Poulsen *et al.* 2005). TFF2 is believed to be involved in aiding nematode expulsion as shown in a study in mice infected with the hookworm *Nippostrongylus brasiliensis* which showed a higher

worm burden in TFF2 deficient mice (Wills-Karp *et al.* 2012). Mucins and TFFs act in conjunction with calcium activated chloride channels (CLCA1) in the membrane which regulates mucus hydration via osmotic fluid transfer, ultimately altering mucus viscosity (Rowe *et al.* 2009).

Several immunoglobulin (Ig) domains were also identified as Tci-gal-1 interacting partners in the extracellular space, such as the FCGBP, several cluster of differentiation molecules, and an immunoglobulin-like cell adhesion molecule. GI nematode infections are correlated with an increased Th2-type immune response, indicated by increased levels of Th2 cytokines, granular and globular leukocytes, and parasite-specific antibodies which include IgA, IgG1 and IgE (McRae *et al.* 2015). Elevated IgA levels in sheep is widely known to be positively correlated with *T. circumcincta* resistance, with worm length and fecundity being significantly diminished (McRae *et al.* 2015). IgA from the gastrointestinal tract can bind to parasites or their E/S antigens, with higher levels of parasite-specific IgA found in resistant sheep (McRae *et al.* 2014). On the other hand, IgE is strongly associated with immune responses against parasitic infections, acting with larvae-specific IgG1 to mediate helminth expulsion, as demonstrated in Merino lambs resistant to *H. contortus* (Negrão-Corrêa 2001). A study successfully showed that the administration of a combination of IgG1 and IgE promoted accelerated expulsion of the nematodes in activation-induced cytidine deaminase (AID)-deficient mice, incapable of switching IgM to other isotypes, infected with *Nippostrongylus brasiliensis* (Matsumoto *et al.* 2013). Increased IgG1 and IgE levels are also associated with reduced faecal egg count in sheep (McRae *et al.* 2015). Therefore, the identification of Ig domains in the present study suggests the desire to diminish the potency of parasite specific antibodies through parasite galectin interactions.

A range of surface exposed cluster of differentiation (CD) molecules were also identified in both sample datasets. CD antigens are expressed by different subsets of precursor stem cells as the cells

differentiate along specific myeloid and lymphoid lineages. These glycoproteins are capable of many functions, such as cell signalling, transport or cell-cell interactions, and may also be immunoglobulins, complement factors or immune cell precursors (Belov *et al.* 2001). For example, CD47, CD54 and CD63 identified in this dataset are associated with and form complexes with integrins, affecting cell function and is essential in immune responses (Roebuck and Finnegan 1999; Brown 2001; Pols and Klumperman 2009). On the other hand, other CD antigens like CD34, CD36, CD46, and CD136 are immune precursor cells or cellular receptors (Gaggar *et al.* 2003; Silverstein and Febbraio 2009; Suga *et al.* 2009).

Several mucin-like proteins were identified in this study: a mucin 1 (MUC1)-like, mucin 13 (MUC13)-related protein and a mucin 18 (MUC18)-related protein. Mucus is a gel-like substance that coats the GI epithelium and can act as a physical barrier, which prevents the establishment of pathogens in the host (van Putten and Strijbis 2017; Sharpe *et al.* 2018). More recently, studies have shown that mucus possesses additional functions that include the presentation of specific ligands to trap pathogens (Sharpe *et al.* 2018), preventing parasitic nematodes like *T. circumcincta* from penetrating host abomasal glands (Pemberton *et al.* 2012). Mucus is mainly comprised of high molecular weight (> 1 MDa), heavily O-glycosylated glycoproteins known as mucins (Hasnain *et al.* 2011; Pemberton *et al.* 2011; Sharpe *et al.* 2018). Mucins can be categorised into two main subtypes: secreted mucins and membrane-bound or transmembrane mucins (Hoorens *et al.* 2011; Sharpe *et al.* 2018). Secreted mucins are responsible for the gel-like and viscous property of mucus while transmembrane mucins have both barrier and signalling functions (van Putten and Strijbis 2017).

In addition to mucus proteins, SEA domains related to MUC1 and MUC13 were identified in the datasets. SEA domains are hypothesized to be involved in reducing the impacts of mechanical stress, as they are located in extracellular regions of transmembrane mucins and break at their proteolytic

cleavage point upon force and reduce the effects of mechanical manipulation (Palmai-Pallag *et al.* 2005; Pelaseyed *et al.* 2013). Both MUC1 and MUC13 possess a single SEA domain on the C-terminal of their mucin domains (Pelaseyed *et al.* 2013). MUC1 is a major gastric transmembrane mucin with anti-inflammatory properties that is critical against enteropathogenic bacteria (Sheng *et al.* 2013; van Putten and Strijbis 2017). Glycosylated tandem repeats of MUC1 can extend up to 500 nm above the epithelial cell surface, forming a dense glycocalyx impermeable to bacterial pathogens (van Putten and Strijbis 2017). MUC1 is continually internalized by clathrin-mediated endocytosis and recycled back to the cell surface, expanding its carrying capacity (Sheng *et al.* 2013). A clathrin adaptor, a key component of clathrin-mediated endocytosis, was identified in this study. Conversely, MUC13 is a membrane-associated sialomucin typically expressed in the intestinal tract with more pro-inflammatory characteristics (Sheng *et al.* 2013). Despite being highly abundant, MUC13 is poorly described in ruminants but has been extensively reported to be a promising early marker in human cancer screening (Maher *et al.* 2011; Nishii *et al.* 2015; Filippou *et al.* 2018). The identification of both mucin-like proteins and the key SEA domain highlights a potential major role of Tci-gal-1 to alter the composition of mucus, potentially allowing *T. circumcincta* to advance through the protective barrier.

A melanoma cell adhesion molecule (MCAM) also known as MUC18 or CD146 was isolated from abomasal scrape samples. MUC18 is part of the immunoglobulin superfamily, consisting of five Ig domains, and is important in inflammatory responses (Shih 1999). Similar to MUC13, this membrane glycoprotein is associated with metastatic cancer and is used as a marker of tumour progression in humans (Sers *et al.* 1993). During infection, intelectins may indirectly aid nematode expulsion by altering the properties of mucus glycoproteins (Artis 2006; French *et al.* 2008). Previous studies on toads have shown a homologous lectin in *Xenopus laevis* eggs (XL35) hardens the egg coat protein by interacting with mucus-like proteins to prevent polyspermy (Nishihara *et al.* 1986; Chang *et al.*

2004). Artis (2006) postulates that a similar cross-linking interaction may occur with nematode intelectins, forming a 'glycoprotein cement' on the nematode exterior to aid expulsion.

Integrins were also amongst the top 50 most abundant proteins across both datasets, particularly ITGA1, ITGA3 and ITGB6. These cell surface receptors are comprised of an α and β subunit, with combinations of the two subunits resulting in unique binding and signalling specificities (Giancotti and Ruoslahti 1999). Involved in multiple signalling pathways, these glycoproteins interact with a variety of extracellular matrix proteins such as kinases, fibronectin, collagen and other molecules to mediate the activities in the extracellular matrix (Ginsberg *et al.* 2005). The interaction between Tci-gal-1 and integrins may hinder these signalling pathways to occur, impairing the host's natural responses.

In summary, parasite galectins can perform a wide range of functions although their role in host-parasite interactions remain to be elucidated. This study has identified glycoprotein ligands in two types of host-derived abomasal samples that bound to Tci-gal-1. For the first time, it is indicated that Tci-gal-1 interacts with host glycoproteins such as trefoil factors, immune proteins and integrins. Analysis by mass spectrometry identified that most host-parasite protein interactions occur in the extracellular space and on the cell surface. These data have led to the hypothesis that the expression of Tci-gal-1 by the parasite upon infection impairs the host's first line of defence, weakening immune effector cell responses and hence enabling *T. circumcincta* to advance through the GI tract. Tci-gal-1 also has strong affinity for mucins in combination with other mucus-associated proteins such as trefoil factors, limiting the efficiency of nematode expulsion by potentially altering mucus viscosity. In conclusion, this study has broadened the knowledge of the host-parasite interface with regard to *T. circumcincta* infection, as well as identified novel candidates for future study in parasite control strategies.

References

Arata, Y, Hirabayashi, J, Kasai, K (1997) The two lectin domains of the tandem-repeat 32-kDa galectin of the nematode *Caenorhabditis elegans* have different binding properties. Studies with recombinant protein. *The Journal of Biochemistry* **121**, 1002-1009.

Artis, D (2006) New weapons in the war on worms: identification of putative mechanisms of immune-mediated expulsion of gastrointestinal nematodes. *International Journal for Parasitology* **36**, 723-733.

Belov, L, de la Vega, O, dos Remedios, CG, Mulligan, SP, Christopherson, RI (2001) Immunophenotyping of leukemias using a cluster of differentiation antibody microarray. *the Journal of Cancer Research* **61**, 4483-4489.

Brown, E (2001) Integrin-associated protein (CD47) and its ligands. *Trends in Cell Biology* **11**, 130-135.

Chang, BY, Peavy, TR, Wardrip, NJ, Hedrick, JL (2004) The *Xenopus laevis* cortical granule lectin: cDNA cloning, developmental expression, and identification of the eglectin family of lectins. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **137**, 115-129.

Craig, H, Wastling, JM, Knox, DP (2006) A preliminary proteomic survey of the in vitro excretory/secretory products of fourth-stage larval and adult *Teladorsagia circumcincta*. *Parasitology* **132**, 535-543.

Demeler, J (2005) The physiological site of action and the site of resistance to the macrocyclic lactone anthelmintics in sheep parasitic trichostrongyloid nematodes. University of Veterinary Medicine Hanover.

Filippou, PS, Ren, AH, Korbakis, D, Dimitrakopoulos, L, Soosaipillai, A, Barak, V, Frenkel, S, Pe'er, J, Lotem, M, Merims, S, Molina, R, Blasutig, I, Bogdanos, DP, Diamandis, EP (2018) Exploring the potential of mucin 13 (MUC13) as a biomarker for carcinomas and other diseases. *Clinical Chemistry and Laboratory Medicine* **56**, 1945-1953.

Finlay, CM, Walsh, KP, Mills, KHG (2014) Induction of regulatory cells by helminth parasites: exploitation for the treatment of inflammatory diseases. *Immunological Reviews* **259**, 206-230.

French, AT, Knight, PA, Smith, WD, Brown, JK, Craig, NM, Pate, JA, Miller, HR, Pemberton, AD (2008) Up-regulation of intelectin in sheep after infection with *Teladorsagia circumcincta*. *International Journal for Parasitology* **38**, 467-475.

Gaggar, A, Shayakhmetov, DM, Lieber, A (2003) CD46 is a cellular receptor for group B adenoviruses. *Nature Medicine* **9**, 1408-1412.

Giancotti, FG, Ruoslahti, E (1999) Integrin signalling. *Science* **285**, 1028-1032.

Ginsberg, MH, Partridge, A, Shattil, SJ (2005) Integrin regulation. *Current Opinion in Cell Biology* **17**, 509-516.

Greenhalgh, CJ, Loukas, A, Newton, SE (1999) The organization of a galectin gene from *Teladorsagia circumcincta*. *Molecular and Biochemical Parasitology* **101**, 199-206.

Greenhalgh, CJ, Newton, SE (1999) Recombinant expression of a galectin from the sheep gastrointestinal parasite *Teladorsagia circumcincta*: its use in isolating galectin-glycoconjugates. *Protein Expression and Purification* **16**, 152-159.

Hasnain, SZ, Evans, CM, Roy, M, Gallagher, AL, Kindrachuk, KN, Barron, L, Dickey, BF, Wilson, MS, Wynn, TA, Grencis, RK, Thornton, DJ (2011) Muc5ac: a critical component mediating the rejection of enteric nematodes. *The Journal of Experimental Medicine* **208**, 893-900.

Hewitson, JP, Grainger, JR, Maizels, RM (2009) Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. *Molecular and Biochemical Parasitology* **167**, 1-11.

Hoerauf, A, Satoguina, J, Saeftel, M, Specht, S (2005) Immunomodulation by filarial nematodes. *Parasite Immunology* **27**, 417-429.

Hoorens, PR, Rinaldi, M, Li, RW, Goddeeris, B, Claerebout, E, Vercruysse, J, Geldhof, P (2011) Genome wide analysis of the bovine mucin genes and their gastrointestinal transcription profile. *BMC genomics* **12**,

Hughes, RC (1999) Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochimica et Biophysica Acta* **1473**, 172-185.

Hwang, EY, Jeong, MS, Park, SK, Ha, SC, Yu, HS, Jang, SB (2016) Structural basis for carbohydrate recognition and anti-inflammatory modulation by gastrointestinal nematode parasite *Toxascaris leonina* galectin. *The Journal of Biological Chemistry* **291**, 25326-25338.

Kelly, SM, Jess, TJ, Price, NC (2005) How to study proteins by circular dichroism. *Biochimica et Biophysica Acta* **1751**, 119-139.

Knox, DP (2000) Development of vaccines against gastrointestinal nematodes. *Parasitology* **120**, 43-61.

Maher, DM, Gupta, BK, Nagata, S, Jaggi, M, Chauhan, SC (2011) Mucin 13: structure, function, and potential roles in cancer pathogenesis. *Molecular Cancer Research* **9**, 531-537.

Maizels, RM, Hewitson, JP, Smith, KA (2012) Susceptibility and immunity to helminth parasites. *Current Opinion in Immunology* **24**, 459-466.

Matsumoto, M, Sasaki, Y, Yasuda, K, Takai, T, Muramatsu, M, Yoshimoto, T, Nakanishi, K (2013) IgG and IgE collaboratively accelerate expulsion of *Strongyloides venezuelensis* in a primary infection. *Infection and Immunity* **81**, 2518-2527.

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

McRae, KM, Good, B, Hanrahan, JP, Glynn, A, O'Connell, MJ, Keane, OM (2014) Response to *Teladorsagia circumcincta* infection in Scottish Blackface lambs with divergent phenotypes for nematode resistance. *Veterinary Parasitology* **206**, 200-207.

McRae, KM, Stear, MJ, Good, B, Keane, OM (2015) The host immune response to gastrointestinal nematode infection in sheep. *Parasite Immunology* **37**, 605-613.

Menzies, M, Reverter, A, Andronicos, N, Hunt, P, Windon, R, Ingham, A (2010) Nematode challenge induces differential expression of oxidant, antioxidant and mucous genes down the longitudinal axis of the sheep gut. *Parasite Immunology* **32**, 36-46.

Negrão-Corrêa, D (2001) Importance of immunoglobulin E (IgE) in the protective mechanism against gastrointestinal nematode infection: looking at the intestinal mucosae. *Journal of the São Paulo Institute of Tropical Medicine* **43**, 291-299.

Nishihara, T, Wyrick, RE, Working, PK, Chen, YH, Hedrick, JL (1986) Isolation and characterization of a lectin from the cortical granules of *Xenopus laevis* Eggs. *Biochemistry* **25**, 6013-6020.

Nishii, Y, Yamaguchi, M, Kimura, Y, Hasegawa, T, Aburatani, H, Uchida, H, Hirata, K, Sakuma, Y (2015) A newly developed anti-Mucin 13 monoclonal antibody targets pancreatic ductal adenocarcinoma cells. *International journal of oncology* **46**, 1781-1787.

Palmai-Pallag, T, Khodabukus, N, Kinarsky, L, Leir, SH, Sherman, S, Hollingsworth, MA, Harris, A (2005) The role of the SEA (sea urchin sperm protein, enterokinase and agrin) module in cleavage of membrane-tethered mucins. *FEBS Journal* **272**, 2901-11.

Pelaseyed, T, Zach, M, Petersson, AC, Svensson, F, Johansson, DG, Hansson, GC (2013) Unfolding dynamics of the mucin SEA domain probed by force spectroscopy suggest that it acts as a cell-protective device. *FEBS Journal* **280**, 1491-501.

Pemberton, AD, Brown, JK, Craig, NM, Pate, J, McLean, K, Inglis, NF, Knox, D, Knight, PA (2012) Changes in protein expression in the sheep abomasum following trickle infection with *Teladorsagia circumcincta*. *Parasitology* **139**, 375-385.

Pemberton, AD, Verdon, B, Inglis, NF, Pearson, JP (2011) Sheep intelectin-2 co-purifies with the mucin Muc5ac from gastric mucus. *Veterinary Research* **91**, 53-57.

Pols, MS, Klumperman, J (2009) Trafficking and function of the tetraspanin CD63. *Experimental Cell Research* **315**, 1584-1592.

Popa, SJ, Stewart, SE, Moreau, K (2018) Unconventional secretion of annexins and galectins. *Seminars in Cell & Developmental Biology* **83**, 42-50.

Poulsen, SS, Kissow, H, Hare, K, Hartmann, B, Thim, L (2005) Luminal and parenteral TFF2 and TFF3 dimer and monomer in two models of experimental colitis in the rat. *Regulatory Peptides* **126**, 163-171.

Rabinovich, GA, Gruppi, A (2005) Galectins as immunoregulators during infectious processes: from microbial invasion to the resolution of the disease. *Parasite Immunology* **27**, 103-114.

Rabinovich, GA, Toscano, MA (2009) Turning 'sweet' on immunity: galectin–glycan interactions in immune tolerance and inflammation. *Nature Reviews Immunology* **9**, 338-352.

Roeber, F, Jex, AR, Gasser, RB (2013) Impact of gastrointestinal parasitic nematodes of sheep, and the role of advanced molecular tools for exploring epidemiology and drug resistance - an Australian perspective. *Parasites & Vectors* **6**, 1-13.

Roeber, F, Kahn, L (2014) The specific diagnosis of gastrointestinal nematode infections in livestock: larval culture technique, its limitations and alternative DNA-based approaches. *Veterinary Parasitology* **205**, 619-28.

Roebuck, KA, Finnegan, A (1999) Regulation of intercellular adhesion molecule-1 (CD54) gene expression. *Journal of Leukocyte Biology* **66**, 876-888.

Rowe, A, Gondro, C, Emery, D, Sangster, N (2009) Sequential microarray to identify timing of molecular responses to *Haemonchus contortus* infection in sheep. *Veterinary Parasitology* **161**, 76-87.

Sargison, ND, Jackson, F, Bartley, DJ, Moir, ACP (2005) Failure of moxidectin to control benzimidazole-, levamisole- and ivermectin-resistant *Teladorsagia circumcincta* in a sheep flock. *Veterinary Record* **156**, 105-109.

Sers, C, Kirsch, K, Rothbächer, U, Riethmüller, G, Johnson, JP (1993) Genomic organization of the melanoma-associated glycoprotein MUC18: implications for the evolution of the

immunoglobulin domains. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8514-8518.

Sharpe, C, Thornton, DJ, Grencis, RK (2018) A sticky end for gastrointestinal helminths; the role of the mucus barrier. *Parasite Immunology* **40**, 1-10.

Sheng, YH, Triyana, S, Wang, R, Das, I, Gerloff, K, Florin, TH, Sutton, P, McGuckin, MA (2013) MUC1 and MUC13 differentially regulate epithelial inflammation in response to inflammatory and infectious stimuli. *Mucosal Immunology* **6**, 557-68.

Shi, W, Xue, C, Su, XZ, Lu, F (2018) The roles of galectins in parasitic infections. *Acta Tropica* **177**, 97-104.

Shih, IM (1999) The role of CD146 (Mel-CAM) in biology and pathology. *The Journal of Pathology* **189**, 4-11.

Silverstein, RL, Febbraio, M (2009) CD36, a Scavenger Receptor Involved in Immunity, Metabolism, Angiogenesis, and Behavior. *Science Signalling* **2**,

Suga, H, Matsumoto, D, Eto, H, Inoue, K, Aoi, N, Kato, H, Araki, J, Yoshimura, K (2009) Functional implications of CD34 expression in human adipose–derived stem/progenitor cells. *Stem Cells and Development* **18**, 1201-1210.

Sutherland, IA, Brown, AE, Leathwick, DM, Bisset, SA (2003) Resistance to prophylactic treatment with macrocyclic lactone anthelmintics in *Teladorsagia circumcincta*. *Veterinary Parasitology* **115**, 301-309.

Taupin, D, Podolsky, DK (2003) Trefoil factors: initiators of mucosal healing. *Nature Reviews Molecular Cell Biology* **4**, 721-732.

Tran, NH, Qiao, R, Xin, L, Chen, X, Liu, C, Zhang, X, Shan, B, Ghodsi, A, Li, M (2019) Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry. *Nature Methods* **16**, 63-66.

Tyanova, S, Temu, T, Cox, J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**, 2301-2319.

van Putten, JPM, Strijbis, K (2017) Transmembrane mucins: signaling receptors at the intersection of inflammation and cancer. *Journal of Innate Immunity* **9**, 281-299.

Vasta, GR (2009) Roles of galectins in infection. *Nature Reviews Immunology* **7**, 424-438.

Wills-Karp, M, Rani, R, Dienger, K, Lewkowich, I, Fox, JG, Perkins, C, Lewis, L, Finkelman, FD, Smith, DE, Bryce, PJ, Kurt-Jones, EA, Wang, TC, Sivaprasad, U, Hershey, GK, Herbert, DR (2012) Trefoil factor 2 rapidly induces interleukin 33 to promote type 2 immunity during allergic asthma and hookworm infection. *The Journal of Experimental Medicine* **209**, 607-622.

Xu, J, Yang, F, Yang, DQ, Jiang, P, Liu, RD, Zhang, X, Cui, J, Wang, ZQ (2018) Molecular characterization of *Trichinella spiralis* galectin and its participation in larval invasion of host's intestinal epithelial cells. *Veterinary Research* **49**, 1-15.

Figures

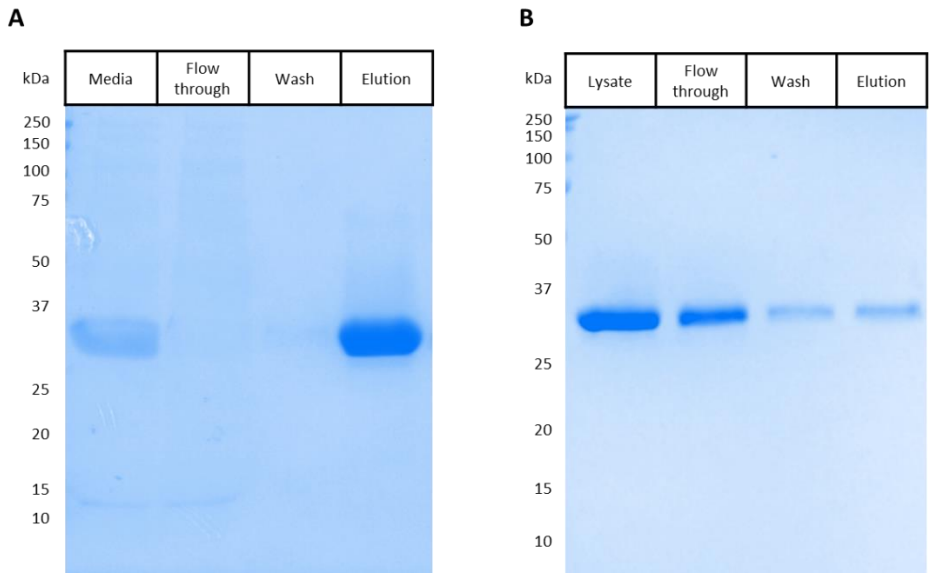


Figure 1. Purification and characterisation of *Teladorsagia circumcincta* galectin (Tci-gal-1). (A) Tci-gal-1 was recombinantly expressed in *Pichia pastoris* and purified via nickel affinity chromatography. (B) Recombinant Tci-gal-1 was shown to be functional by elution from a lactose-Sepharose column. Equal volumes (15 µl) of each sample was loaded onto the gel.

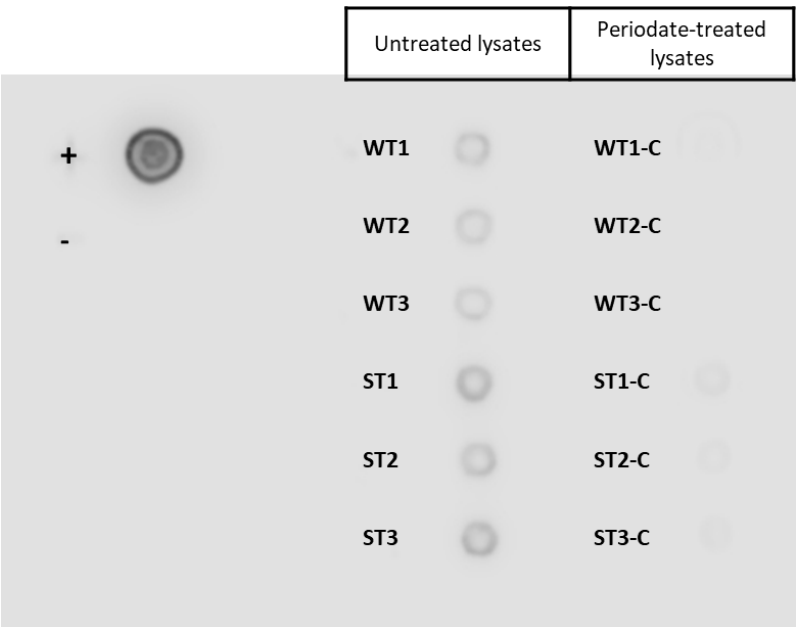


Figure 2. Confirmation of disruption of glycan structures by sodium periodate. ConA lectin dot blot confirming successful disruption of glycan structures after treatment of sheep abomasal whole tissue (WT) and scrape tissue (ST) with 20 mM sodium periodate. Approximately 20 µl of each sample were dotted onto a polyvinylidene difluoride (PVDF) membrane and probed with horseradish peroxidase conjugated Concanavalin A (ConA-HRP) lectin. (WT1-3) Biological replicates of abomasal whole tissue extracts. (ST1-3) Biological replicates of abomasal scrape extracts. (-C) Periodate-treated control samples. (+) *Fasciola hepatica* whole worm extract used as a positive control. (-) Periodate treated *Fasciola hepatica* whole worm extract used as negative control.

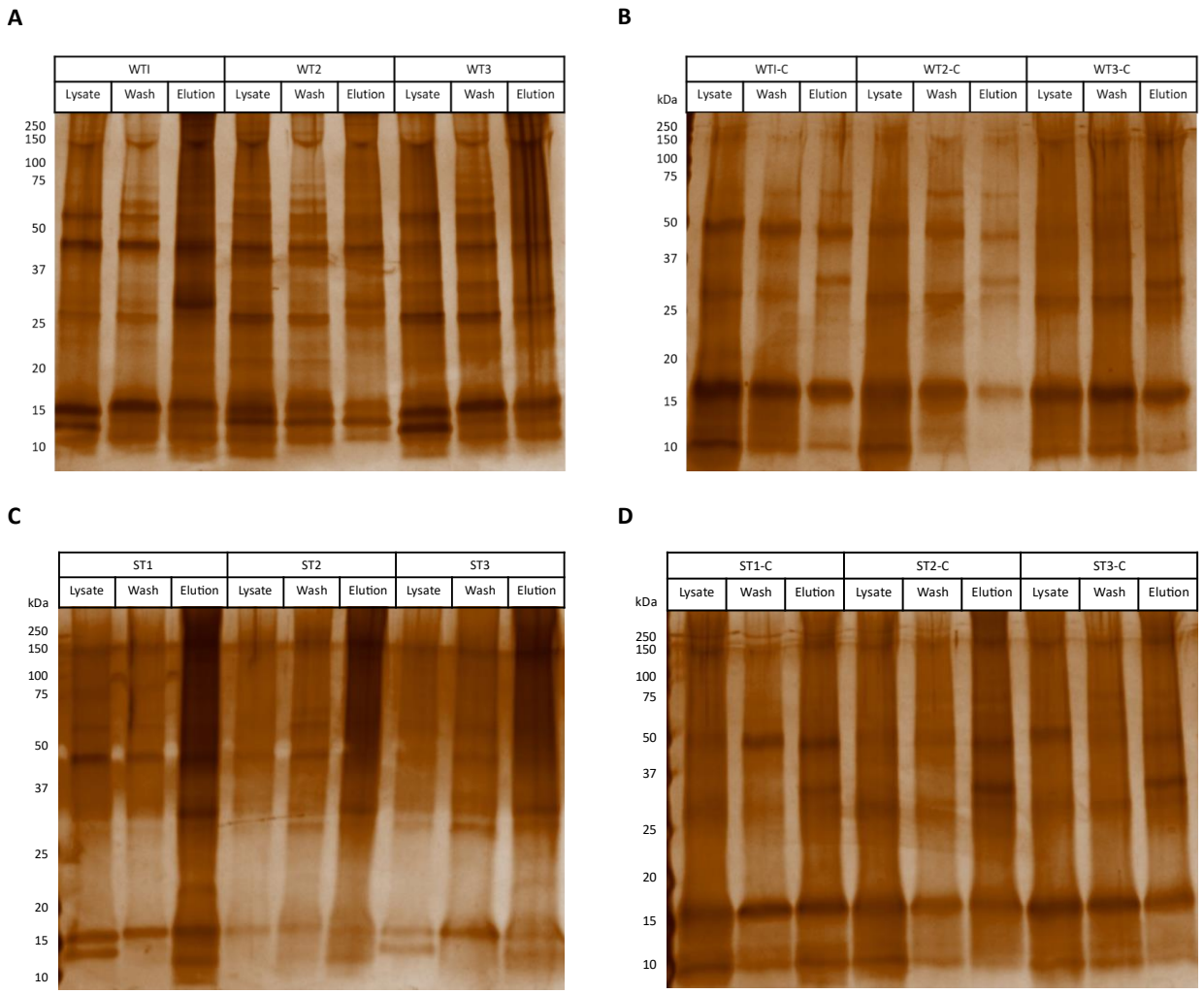
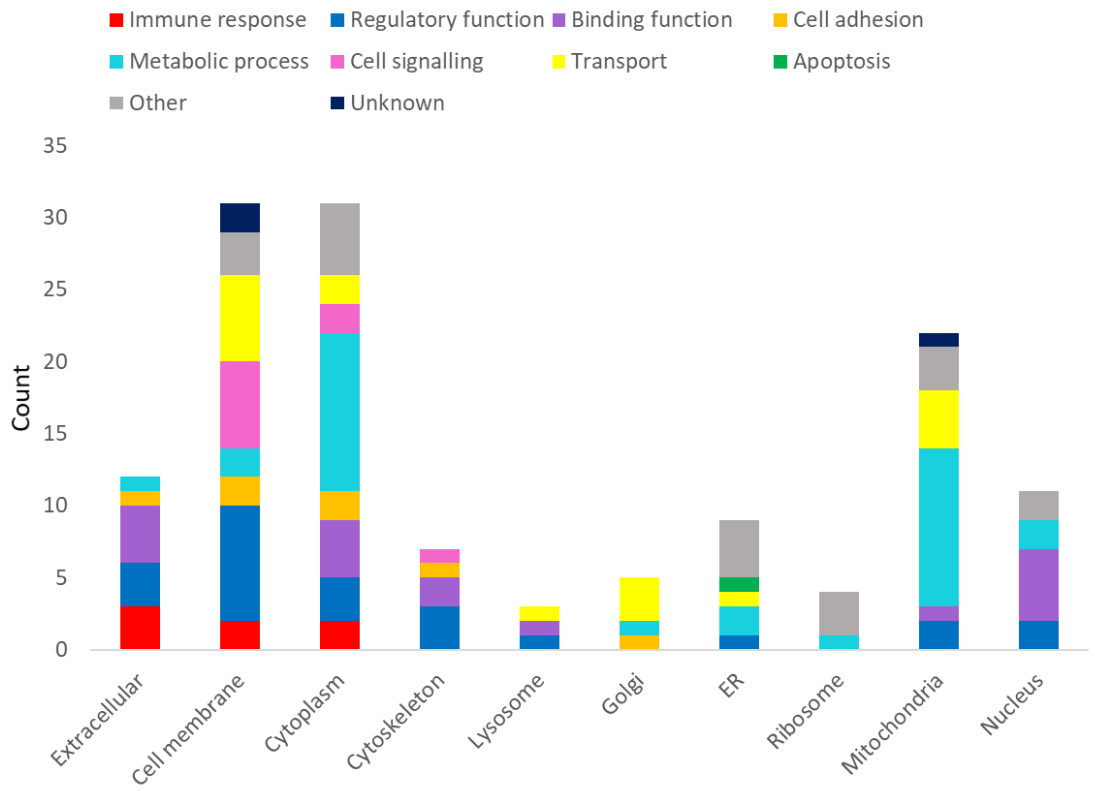


Figure 3. Protein profiles of abomasal whole tissue (WT) and abomasal scrape tissue (ST) extracts that bound to Tci-gal-1. Silver stained 12% (w/v) SDS-PAGE gels showing protein abomasal whole tissue and scrape tissue lysates bound to immobilised *Teladorsagia circumcincta* galectin (Tci-gal-1). (A) Lysate, wash and elution prepared from sheep abomasal whole tissue added to immobilised Tci-gal-1, performed in triplicate. (B) Lysate, wash and elution prepared from sheep abomasal whole tissue, treated with 20 mM sodium periodate, added to immobilised Tci-gal-1, performed in triplicate. (C) Lysate, wash and elution prepared from sheep abomasal scrape tissue added to immobilised Tci-gal-1, performed in triplicate. (D) Lysate, wash and elution prepared from sheep abomasal scrape tissue, treated with 20 mM sodium periodate, added to immobilised Tci-gal-1, performed in triplicate. Bound glycoproteins in all replicates were washed and eluted with 250 mM lactose. Approximately 0.5 mg/ml of starting lysate was added onto the gel in the 'Lysate' columns. Each well was loaded with 15 µl of sample.

A


Immune response
Regulatory function
Binding function
Cell adhesion
Metabolic process
Cell signalling
Transport
Apoptosis
Other
Unknown
Count
35
30
25
20
15
10
5
0
Extracellular
Cell membrane
Cytoplasm
Cytoskeleton
Lysosome
Golgi
ER
Ribosome
Mitochondria
Nucleus

B

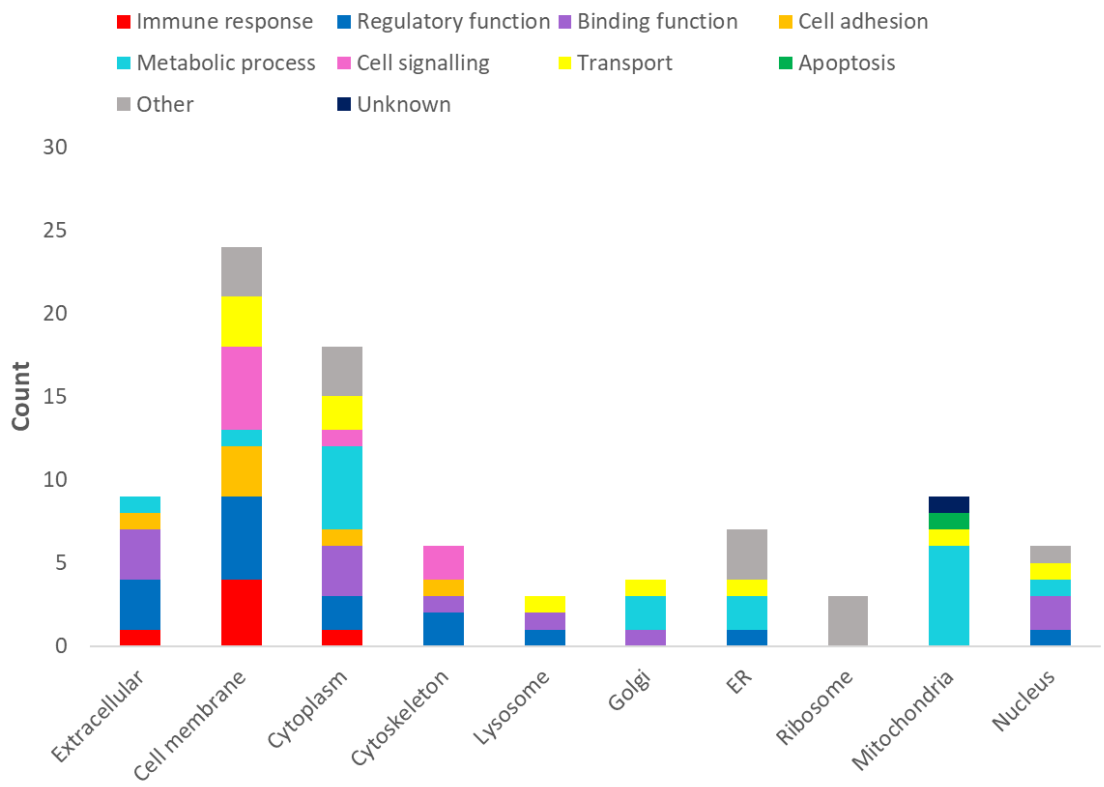

Immune response
Regulatory function
Binding function
Cell adhesion
Metabolic process
Cell signalling
Transport
Apoptosis
Other
Unknown
Count
30
25
20
15
10
5
0
Extracellular
Cell membrane
Cytoplasm
Cytoskeleton
Lysosome
Golgi
ER
Ribosome
Mitochondria
Nucleus

Figure 4. Characterisation of proteins in sheep abomasal tissue that interacted with *Teladorsagia circumcincta* galectin (Tci-gal-1). The profiles were categorised based on biological processes and cellular locations of Tci-gal-1-bound abomasal whole tissue (A) and Tci-gal-1-bound abomasal scrape tissue (B).

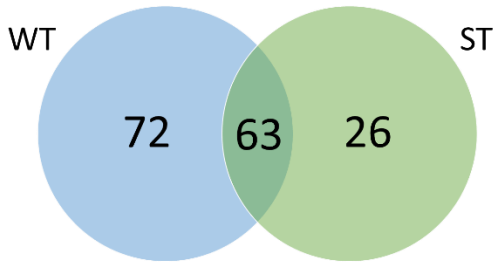


Figure 5. Venn diagram showing the distribution of sheep abomasal whole tissue (WT) and abomasal scrape tissue (ST) that specifically bound to *Teladorsagia circumcincta* galectin (Tci-gal-1).

Table 1. Identification by mass spectroscopy of top 50 most abundant sheep abomasal whole tissue (WT) proteins eluted from a Tci-gal-1 column.

Accession number	Gene ID	Protein annotation	Mol. weight (kDa)	Sequence length	Unique peptides	Average abundance	Number of predicted glycosylation sites		Signal peptides	Predicted cellular location	Predicted function
							N-glyc	O-glyc			
W5PCT5	SLC25A31	Solute carrier family 25 member 31	28.72	262	6	24.86	0	0	N	Mitochondria	Membrane transport
W5PLZ4	TFF2	Trefoil factor 2	14.31	131	2	24.57	0	1	Y	Extracellular	Regulatory function
W5P7F8	CD63	Tetraspanin	25.83	236	4	24.37	4	0	N	Cell membrane	Regulatory function
W5P0U4	MUC1	SEA domain-containing protein (Mucin 1 related)	58.65	588	4	23.98	5	158	Y	Cell membrane	Regulatory function
W5PBS4	LRP1	LDL receptor related protein 1	502.55	4526	14	23.56	33	122	N	Cell membrane	Regulatory function
W5QAH5	-	Transcription factor, GTP-binding domain	30.79	277	1	23.51	2	1	N	Nucleus	Other
W5QA42	MFAP4	Fibrinogen C-terminal domain-containing protein	21.42	193	2	23.38	1	0	Y	Extracellular	Metabolic function
W5PQ79	RPS15	Ribosomal protein S15	16.43	141	2	23.34	0	0	N	Ribosome	Other
W5PZK7	ACTA2	Actin alpha 2, smooth muscle	42.01	377	1	23.26	1	1	N	Cytoskeleton	Regulatory function
W5P5W1	ITGA3	Integrin subunit alpha 3	115.18	1039	10	23.09	11	4	Y	Cell membrane	Cell signalling
W5PEN0	ITGA1	Integrin subunit alpha 1	129.29	1166	9	23.09	19	15	Y	Cell membrane	Regulatory function
P00922	CA2	Carbonic anhydrase 2	29.21	260	2	23.06	0	1	N	Cell membrane	Other
W5PTZ9	LOC114116824	Histone H3	15.39	136	1	22.93	0	7	N	Nucleus	Binding function
W5P8R7	FCGBP	IgGFc-binding protein	271.26	2544	17	22.92	10	44	Y	Extracellular	Binding function
W5PNH6	FAM234A	Family with sequence similarity 234 member A-related	57.62	540	7	22.86	3	6	N	Cell membrane	Unknown
W5QBH1	CSRP1	Cysteine and glycine rich protein 1	18.72	175	2	22.59	3	7	N	Nucleus	Binding function
W5QB48	LOC101113624	ATP synthase, subunit F	10.84	93	2	22.59	0	3	N	Mitochondria	Metabolic function
W5NWN0	GSTZ1	Glutathione S-transferase zeta 1	26.70	240	4	22.54	4	1	N	Mitochondria	Metabolic function

Table 1 (continued)

Accession number	Gene ID	Protein annotation	Mol. weight (kDa)	Sequence length	Unique peptides	Average abundance	Number of predicted glycosylation sites		Signal peptides	Predicted cellular location	Predicted function
							N-glyc	O-glyc			
W5PI50	GLRX	Glutaredoxin domain-containing protein	11.81	106	2	22.53	1	0	N	Cytoplasm	Regulatory function
W5NY03	CD163	CD163, scavenger receptor cysteine rich domain	122.35	1136	6	22.45	7	13	Y	Cell membrane	Cell signalling
W5Q5W2	SCARB2	CD36, scavenger receptor class B member 2	49.78	439	5	22.38	9	0	N	Lysosome	Binding function
W5QCQ4	DHX9	DExH-box helicase 9	142.13	1289	9	22.36	7	29	N	Nucleus	Regulatory function
W5Q436	SLC25A5	Mitochondrial carrier protein	29.33	263	3	22.34	1	0	N	Mitochondria	Regulatory function
W5PPJ0	STT3A	STT3 oligosaccharyltransferase complex catalytic subunit A	90.02	789	5	22.34	1	1	N	ER	Metabolic function
W5QD93	LOC101119050	Dehydrogenase/reductase SDR member 4-like	29.60	279	3	22.31	1	3	N	Mitochondria	Metabolic function
W5NTX6	RIPK1	Receptor interacting serine/threonine kinase 1	87.21	770	1	22.26	2	38	N	Cytoplasm	Regulatory function
W5PQK3	PFKL	ATP-dependent 6-phosphofructokinase	82.65	752	5	22.23	2	1	N	Cytoplasm	Metabolic function
W5Q8K4	SLC3A2	Solute carrier family 3 member 2	63.34	577	5	22.22	3	7	N	Cytoplasm	Metabolic function
W5QFP1	PABPC4	Polyadenylate-binding protein	72.31	660	2	22.19	0	8	N	Cytoplasm	Binding function
W5PQS4	PROCR	MHC class 1-like antigen recognition domain	27.09	241	3	22.18	4	1	Y	Cytoskeleton	Cell signalling
W5QJ31	MTHFD1	Methylenetetrahydrofolate deh ydrogenase 1	105.89	977	6	22.17	0	4	N	Cytoplasm	Metabolic function
Q863C4	ITGB6	Integrin beta-6	85.75	787	3	22.17	7	17	Y	Cell membrane	Cell signalling
W5QFZ8	RPL13	Ribosomal protein L13	23.41	203	2	22.14	1	4	N	Ribosome	Other
W5PEE9	LAMP1	Lysosomal associated membrane protein 1	42.10	392	3	22.13	17	5	N	Lysosome	Regulatory function
W5Q9K1	QSOX1	Sulfhydryl oxidase	81.60	747	5	22.11	2	30	Y	Golgi	Metabolic function

Table 1 (continued)

Accession number	Gene ID	Protein annotation	Mol. weight (kDa)	Sequence length	Unique peptides	Average abundance	Number of predicted glycosylation sites		Signal peptides	Predicted cellular location	Predicted function
							N-glyc	O-glyc			
W5NQL2	OLA1	Obg-like ATPase 1	47.25	417	6	22.09	3	2	N	Nucleus	Binding function
W5PTV7	ACE	Angiotensin-converting enzyme	138.24	1206	5	22.08	6	12	N	Extracellular	Binding function
W5P375	TCP1	T-complex 1	60.52	558	6	22.07	0	3	N	Cytoplasm	Regulatory function
W5PVM5	LAMA5	Laminin subunit alpha 5	375.79	3464	5	22.07	13	173	N	Extracellular	Regulatory function
W5QC22	CD47	CD47 molecule	32.16	291	3	22.07	5	4	N	Extracellular	Immune response
W5QCP9	COL6A3	Collagen type VI alpha 3 chain, von Willebrand factor A domain	340.43	3154	6	22.00	6	86	Y	Extracellular	Binding function
W5PHN8	AKR	Aldo-keto reductase	36.67	323	2	21.99	1	0	N	Cytoplasm	Metabolic function
W5P061	ABCC3	ATP binding cassette subfamily C member 3	169.72	1526	5	21.97	6	13	N	Cell membrane	Ion transport
W5Q633	BLVRA	Biliverdin reductase A	33.61	296	4	21.94	0	0	N	Cytoplasm	Metabolic function
W5PEH7	SRP68	Signal recognition particle subunit SRP68	70.90	626	6	21.90	0	19	N	Cytoplasm	Binding function
W5QEY8	GFPT1	Glutamine--fructose-6-phosphate transaminase 1	67.81	603	4	21.86	1	1	N	Cytoplasm	Metabolic function
W5Q2E5	PTPRC	Protein tyrosine phosphatase receptor type C	139.33	1235	4	21.83	19	55	N	Cell membrane	Cell signalling
W5Q496	RPL36	60S ribosomal protein L36	12.21	105	3	21.81	0	4	N	Ribosome	Metabolic function
W5P6U5	LMAN1	Lectin, mannose binding 1	52.87	465	5	21.79	0	4	N	Golgi	Protein transport
W5PK26	LASP1	LIM and SH3 protein 1	29.70	260	4	21.74	0	9	N	Cytoplasm	Binding function

Table 2. Identification by mass spectroscopy of top 50 most abundant sheep abomasal scrape tissue (ST) proteins eluted from a Tci-gal-1 column.

Accession number	Gene ID	Protein annotation	Mol. weight (kDa)	Sequence length	Unique peptides	Average abundance	Number of predicted glycosylation sites		Signal peptides	Predicted cellular location	Predicted function
							N-glyc	O-glyc			
W5PLZ4	TFF2	Trefoil factor 2	14.31	131	2	24.21	0	1	Y	Extracellular	Regulatory function
W5QAH5	-	Transcription factor, GTP-binding domain	30.79	277	1	23.93	2	1	N	Nucleus	Other
W5PQ79	RPS15	Ribosomal protein S15	16.43	141	2	23.80	0	0	N	Ribosome	Other
W5P5W1	ITGA3	Integrin subunit alpha 3	115.18	1039	10	23.64	11	4	Y	Cell membrane	Cell signalling
W5QB48	LOC101113624	ATP synthase, subunit F	10.84	93	2	23.00	0	3	N	Mitochondria	Metabolic function
W5PBS4	LRP1	LDL receptor related protein 1	502.55	4526	14	22.89	33	122	N	Cell membrane	Regulatory function
W5PEN0	ITGA1	Integrin subunit alpha 1	129.29	1166	9	22.86	19	15	Y	Cell membrane	Regulatory function
W5P1J8	CAO	Copper amine oxidase	85.25	766	3	22.71	4	11	Y	Cytoplasm	Metabolic function
Q863C4	ITGB6	Integrin beta-6	85.75	787	3	22.58	7	17	Y	Cell membrane	Cell signalling
W5Q9K1	QSOX1	Sulfhydryl oxidase	81.60	747	5	22.46	2	30	Y	Golgi	Metabolic function
W5P5W6	NDRG1	N-myc downstream regulated 1	44.65	414	3	22.36	0	19	N	Cytoplasm	Regulatory function
W5NVK4	SPCS1	Signal peptidase complex subunit 1	18.11	163	1	22.34	1	2	N	ER	Other
W5NY03	CD163	CD163, scavenger receptor cysteine rich domain	122.35	1136	6	22.30	7	13	Y	Cell membrane	Cell signalling
W5P3H8	IGF2R	Insulin like growth factor 2 receptor, mannose-6-phosphate receptor	271.70	2463	7	22.28	16	13	N	Golgi	Binding function
W5PVR9	ERMP1	Endoplasmic reticulum metallopeptidase 1	100.02	905	4	22.28	2	5	N	ER	Metabolic function
W5QA42	MFAP4	Fibrinogen C-terminal domain-containing protein	21.42	193	2	22.21	1	0	Y	Extracellular	Metabolic function

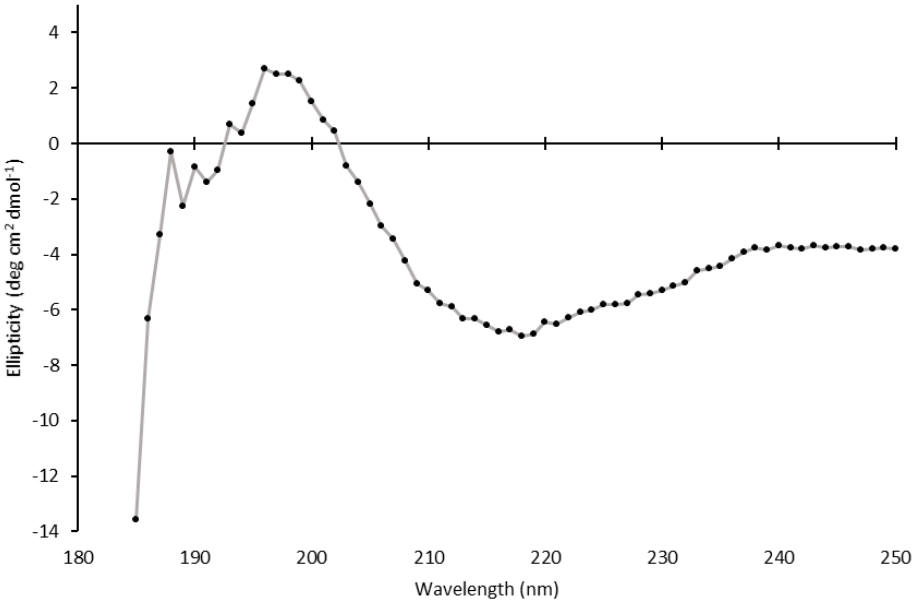
Table 2 (continued)

Accession number	Gene ID	Protein annotation	Mol. weight (kDa)	Sequence length	Unique peptides	Average abundance	Number of predicted glycosylation sites		Signal peptides	Predicted cellular location	Predicted function
							N-glyc	O-glyc			
W5QCH8	NPTN	Neuroplastin	42.43	376	6	22.21	4	3	N	Cell membrane	Cell adhesion
W5P6A5	SPINT2	Serine peptidase inhibitor, Kunitz type 2	27.96	250	2	22.17	2	9	Y	Cell membrane	Regulatory function
W5PEE9	LAMP1	Lysosomal associated membrane protein 1	42.10	392	3	22.15	17	5	N	Lysosome	Regulatory function
W5QC22	CD47	CD47 molecule	32.16	291	3	22.12	5	4	N	Cell membrane	Cell adhesion
W5NTX6	RIPK1	Receptor interacting serine/threonine kinase 1	87.21	770	1	22.10	2	38	N	Cytoplasm	Regulatory function
W5PTV7	ACE	Angiotensin-converting enzyme	138.24	1206	5	22.07	6	12	N	Extracellular	Binding function
W5P1H4	CTL4	Choline transporter-like protein 4	79.60	711	3	21.98	8	6	N	Cell membrane	Membrane transport
W5P180	MTX1	Metaxin	32.19	285	4	21.95	0	5	N	Mitochondria	Protein transport
W5Q0Z6	MCAM	Melanoma cell adhesion molecule (MUC18-related)	68.13	615	11	21.90	6	6	N	Cell membrane	Immune response
W5NYJ4	DKC1	PUA domain-containing protein	56.76	506	3	21.86	0	20	N	Nucleus	Regulatory function
W5QG70	MUC13	SEA domain-containing protein	53.93	508	2	21.79	9	62	Y	Cell membrane	Cell signalling
W5NYN6	RTN3	Reticulon	111.06	1028	1	21.78	1	151	N	ER	Other
W5P0R5	RPL21E	Ribosomal protein L21e	18.41	160	1	21.71	0	0	N	Ribosome	Other
W5Q8Y5	HDLBP	High density lipoprotein binding protein/Vigilin	141.91	1273	3	21.71	0	30	N	Nucleus	Binding function
W5PQ27	GATD3A	Glutamine Amidotransferase Like Class 1 Domain Containing 3A	28.69	274	4	21.66	1	8	Y	Mitochondria	Unknown
W5PK26	LASP1	LIM and SH3 protein 1	29.70	260	4	21.66	0	9	N	Cytoplasm	Binding function
W5Q501	GNAS	GNAS complex locus	111.42	1037	3	21.63	1	71	N	Cytoplasm	Cell signalling

Table 2 (continued)

Accession number	Gene ID	Protein annotation	Mol. weight (kDa)	Sequence length	Unique peptides	Average abundance	Number of predicted glycosylation sites		Signal peptides	Predicted cellular location	Predicted function
							N-glyc	O-glyc			
W5PQS4	PROCR	MHC class 1-like antigen recognition domain	27.09	241	3	21.61	4	1	Y	Cytoskeleton	Cell signalling
W5QJ31	MTHFD1	Methylenetetrahydrofolate dehydrogenase 1	105.89	977	6	21.61	0	4	N	Cytoplasm	Metabolic function
W5PVM5	LAMA5	Laminin subunit alpha 5	375.79	3464	5	21.60	13	173	N	Extracellular	Regulatory function
W5QFP1	PABPC4	Polyadenylate-binding protein	72.31	660	2	21.60	0	8	N	Cytoplasm	Binding function
W5PD64	ARL8B	ADP ribosylation factor-like GTPase 8B	20.41	176	3	21.57	2	0	N	Lysosome	Binding function
W5PDH7	NPC1	Niemann-Pick C type protein	141.53	1275	4	21.54	14	9	Y	Lysosome	Lipid transport
W5P066	CYB5B	Cytochrome b5 type B	16.98	153	2	21.49	0	1	N	Mitochondria	Metabolic function
W5P900	RPL29	60S ribosomal protein L29	17.13	156	2	21.48	2	0	N	Ribosome	Other
W5P246	TM9SF2	Nonaspanin	76.13	666	3	21.46	0	2	Y	Cell membrane	Membrane transport
W5P1G5	ERGIC1	Endoplasmic reticulum-golgi intermediate compartment 1	32.62	290	1	21.46	1	0	N	Nucleus	ER-Golgi transport
W5P906	DPP4	Dipeptidyl peptidase 4	88.44	765	3	21.45	9	7	Y	Cell membrane	Metabolic function
W5PWS0	CD46	Membrane cofactor protein CD46	39.43	363	3	21.35	3	10	Y	Cell membrane	Immune response
W5PZG7	PECAM1	Platelet and endothelial cell adhesion molecule 1 (Ig-like)	90.96	812	3	21.35	11	2	N	Extracellular	Immune response
W5QG66	ITGB5	Integrin beta-5	88.13	802	3	21.31	5	25	N	Cell membrane	Cell signalling
W5P8N8	IDH3B	Isocitrate dehydrogenase [NAD] subunit	42.50	385	4	21.27	0	6	N	Mitochondria	Metabolic function
W5PXR3	ENPP1	Ectonucleotide pyrophosphatase/phosphodiesterase 1	94.68	823	4	21.27	6	8	N	Extracellular	Regulatory function
W5Q3J8	HLA-DRB3	MHC class II antigen DRB3	27.99	244	2	21.23	1	1	Y	Cell membrane	Immune response

Supplementary figures



Supplementary Figure S1. Circular dichroism spectra of recombinant Tci-gal-1 expressed in *Pichia pastoris* showing tertiary structure to be β-sheets in accordance with spectra presented by Arata *et al.* (1997).

Table S1. Identification by mass spectroscopy of sheep abomasal whole tissue (WT) proteins eluted from a Tci-gal-1 column.

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
Immune response									
W5PI11	TAPBP	Ig-like domain-containing protein	2	19.74	1	6	1	Y	Cell membrane
W5PWS0	CD46	Membrane cofactor protein CD46	3	21.18	3	10	1	Y	Cell membrane
W5PH55	OAS	2'-5'-oligoadenylate synthase	4	21.54	5	16	0	N	Cytoplasm
W5QD49	PSME1	Proteasome activator subunit 1	2	20.94	1	1	0	N	Cytoplasm
W5PBE1	ISG15	Ubiquitin-like protein, interferon-stimulated gene 15	2	19.52	1	0	0	N	Extracellular
W5PZG7	PECAM1	Platelet and endothelial cell adhesion molecule 1	3	21.24	11	2	1	N	Extracellular
W5QC22	CD47	CD47 molecule	3	22.07	5	4	5	N	Extracellular
Cell signalling									
W5Q2E5	PTPRC	Protein tyrosine phosphatase receptor type C	4	21.83	19	55	1	N	Cell membrane
W5P5W1	ITGA3	Integrin subunit alpha 3	10	23.09	11	4	1	Y	Cell membrane
W5NY03	CD163	CD163, scavenger receptor cysteine rich domain	6	22.45	7	13	1	Y	Cell membrane
Q863C4	ITGB6	Integrin beta-6	3	22.17	7	17	1	Y	Cell membrane
W5QG66	ITGB5	Integrin beta-5	3	20.94	5	25	1	N	Cell membrane
W5QG70	MUC13	SEA domain-containing protein	2	21.04	9	62	1	Y	Cell membrane
W5Q438	GNB2/3	G protein subunit beta 2/3	3	21.48	0	2	0	N	Cytoplasm
W5Q501	GNAS	GNAS complex locus	3	21.45	1	71	0	N	Cytoplasm
W5PQS4	PROCR	MHC class 1-like antigen recognition domain	3	22.18	4	1	1	Y	Cytoskeleton
Cell adhesion									
W5PEE6	ITGA2	Integrin subunit alpha 2	2	19.42	11	12	1	N	Cell membrane
W5QCH8	NPTN	Neuroplastin	6	21.55	4	3	1	N	Cell membrane
W5PLW1	CTNND1	Catenin delta 1	3	20.38	7	73	0	N	Cytoplasm
W5NY24	CTNNB1	Catenin beta 1	2	20.57	5	16	0	N	Cytoplasm

Table S1 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
W5Q5C7	CTNNA1	Catenin alpha 1	2	21.42	0	20	0	N	Cytoskeleton
W5Q263	ICAM1/CD54	Ig-like domain-containing protein	2	21.18	7	1	1	Y	Extracellular
W5PEX4	FNDC3A	Fibronectin type III domain containing 3A	2	20.41	5	28	1	N	Golgi
Apoptosis									
W5QCK9	TMEM214	Transmembrane protein 214	4	20.20	4	19	1	N	ER
Transport									
W5Q867	SLC12A2	Solute carrier family 12 member 2	3	20.56	4	41	12	N	Cell membrane
W5P061	ABCC3	ATP binding cassette subfamily C member 3	5	21.97	6	13	14	N	Cell membrane
W5Q3Y9	ATP2A1/2	Calcium-transporting ATPase	3	21.05	3	0	8	N	Cell membrane
W5P1H4	CTL4	Choline transporter-like protein 4	3	21.66	8	6	10	N	Cell membrane
W5P246	TM9SF2	Nonaspanin	3	21.12	0	2	9	Y	Cell membrane
W5P9V5	PIGR	Polymeric immunoglobulin receptor	3	21.63	4	4	1	Y	Cell membrane
W5QHF2	AP2M1	Clathrin adaptor, mu subunit	4	20.59	2	6	0	N	Cytoplasm
W5NZ10	NUTF2	Nuclear transport factor 2	2	21.54	0	0	0	N	Cytoplasm
W5P7A1	SAR1A	Secretion associated Ras related GTPase 1A	1	21.38	1	1	0	N	ER
W5P6U5	LMAN1	Lectin, mannose binding 1	5	21.79	0	4	1	N	Golgi
W5P0R0	COPG1	Coatomer subunit gamma	3	20.55	1	10	0	N	Golgi
W5PDH7	NPC1	Niemann-Pick C type protein	4	21.36	14	9	9	Y	Lysosome
W5P180	MTX1	Metaxin	4	21.49	0	5	0	N	Mitochondria
W5P9Q8	NDUFC2	NADH dehydrogenase [ubiquinone] 1 subunit C2	2	20.46	0	3	1	N	Mitochondria
B2MVX2	SLC25A11	Mitochondrial 2-oxoglutarate/malate carrier protein	4	20.03	0	4	0	N	Mitochondria
W5PS50	MPC2	Mitochondrial pyruvate carrier	1	19.62	0	2	3	N	Mitochondria
W5PCT5	SLC25A31	Solute carrier family 25 member 31	6	24.86	0	0	0	N	Mitochondria

Table S1 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
Regulatory function									
W5P0U4	MUC1	SEA domain-containing protein (Mucin 1 related)	4	23.98	5	158	1	Y	Cell membrane
W5P7F8	CD63	Tetraspanin	4	24.37	4	0	4	N	Cell membrane
W5P8X9	CLIC4	Chloride intracellular channel protein	2	21.58	0	0	0	N	Cell membrane
W5PBS4	LRP1	LDL receptor related protein 1	14	23.56	33	122	1	N	Cell membrane
W5PEN0	ITGA1	Integrin subunit alpha 1	9	23.09	19	15	1	Y	Cell membrane
W5NUR4	ADAM9	ADAM metallopeptidase domain 9	2	19.74	5	18	1	Y	Cell membrane
W5P6A5	SPINT2	Serine peptidase inhibitor, Kunitz type 2	2	21.46	2	9	1	Y	Cell membrane
W5QAM4	CAMK2D	Protein kinase domain-containing protein	2	19.38	2	21	1	N	Cell membrane
W5PI50	GLRX	Glutaredoxin domain-containing protein	2	22.53	1	0	0	N	Cytoplasm
W5P375	TCP1	T-complex 1	6	22.07	0	3	0	N	Cytoplasm
W5NTX6	RIPK1	Receptor interacting serine/threonine kinase 1	1	22.26	2	38	0	N	Cytoplasm
W5PZK7	ACTA2	Actin alpha 2, smooth muscle	1	23.26	1	1	0	N	Cytoskeleton
W5Q8V2	LIMA1	LIM domain and actin binding 1	2	20.32	2	102	1	N	Cytoskeleton
W5PIU7	MTPN	Myotrophin	2	21.33	1	1	0	N	Cytoskeleton
W5QG21	PDIA5	Protein disulfide-isomerase A5	2	20.16	0	3	1	N	ER
W5PLZ4	TFF2	Trefoil factor 2	2	24.57	0	1	0	Y	Extracellular
W5PVM5	LAMA5	Laminin subunit alpha 5	5	22.07	13	173	0	N	Extracellular
W5PXR3	ENPP1	Ectonucleotide pyrophosphatase/phosphodiesterase 1	4	21.15	6	8	0	N	Extracellular
W5PEE9	LAMP1	Lysosomal associated membrane protein 1	3	22.13	17	5	1	N	Lysosome
W5Q436	SLC25A5	Mitochondrial carrier protein	3	22.34	1	0	1	N	Mitochondria
W5PGB0	SIRT3	NAD-dependent protein deacetylase	2	20.15	1	6	0	N	Mitochondria
W5QCQ4	DHX9	DExH-box helicase 9	9	22.36	7	29	0	N	Nucleus
W5NYJ4	DKC1	PUA domain-containing protein	3	20.69	0	20	0	N	Nucleus
Binding function									
W5PEH7	SRP68	Signal recognition particle subunit SRP68	6	21.90	0	19	0	N	Cytoplasm
W5PK26	LASP1	LIM and SH3 protein 1	4	21.74	0	9	0	N	Cytoplasm

Table S1 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
W5QFP1	PABPC4	Polyadenylate-binding protein	2	22.19	0	8	0	N	Cytoplasm
W5PBY5	RI	Ribonuclease inhibitor-like protein	3	21.31	0	0	0	N	Cytoplasm
W5PGT5	MYH11	Myosin heavy chain 11	1	20.25	10	52	0	N	Cytoskeleton
W5NYP8	DYNC1H1	Dynein cytoplasmic 1 heavy chain 1	6	21.59	7	86	0	N	Cytoskeleton
W5PTV7	ACE	Angiotensin-converting enzyme	5	22.08	6	12	0	N	Extracellular
W5QCP9	COL6A3	Collagen type VI alpha 3 chain, von Willebrand factor A domain	6	22.00	6	86	0	Y	Extracellular
W5P8R7	FCGBP	IgGFc-Binding Protein	17	22.92	10	44	0	Y	Extracellular
W5PSB6	ACE2	Angiotensin-converting enzyme 2	1	20.66	6	3	1	Y	Extracellular
W5P3H8	IGF2R	Insulin like growth factor 2 receptor, mannose-6-phosphate receptor	7	21.66	16	13	1	N	Golgi
W5Q5W2	SCARB1	CD36, Scavenger receptor class B member 1	5	22.38	9	0	1	N	Lysosome
W5PEC3	ANXA6	Annexin	4	21.67	1	14	0	N	Mitochondria
W5PTZ9	LOC114116824	Histone H3	1	22.93	0	7	0	N	Nucleus
W5NQL2	OLA1	Obg-like ATPase 1	6	22.09	3	2	0	N	Nucleus
W5Q1M9	DDX17	DEAD-box helicase 17	1	20.72	3	35	0	N	Nucleus
W5QBH1	CSRP1	Cysteine and glycine rich protein 1	2	22.59	3	7	0	N	Nucleus
W5QIV1	S100A11	S100 Calcium Binding Protein A11	3	21.47	1	2	0	N	Nucleus
Metabolic function									
W5P906	DPP4	Dipeptidyl peptidase 4	3	21.21	9	7	0	Y	Cell membrane
W5P0T6	FOLH1	Folate hydrolase 1 (Glutamate carboxypeptidase 2-like)	4	21.20	9	1	1	N	Cell membrane
W5PQK3	PFKL	ATP-dependent 6-phosphofructokinase	5	22.23	2	1	0	N	Cytoplasm
W5Q8K4	SLC3A2	Solute carrier family 3 member 2 (glycosyl hydrolase domain)	5	22.22	3	7	1	N	Cytoplasm
W5PHN8	AKR	Aldo-keto reductase	2	21.99	1	0	0	N	Cytoplasm
W5Q633	BLVRA	Biliverdin reductase A	4	21.94	0	0	0	N	Cytoplasm
W5PX34	LYPLA1	Lysophospholipase 1	2	20.73	0	0	0	N	Cytoplasm

Table S1 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
W5PSI1	PSMD3	Proteasome 26S subunit, non-ATPase 3 (PCI domain)	2	21.37	1	17	0	N	Cytoplasm
W5PJ69	CKM	Creatine kinase, M-type	2	20.86	1	3	0	N	Cytoplasm
W5QJ31	MTHFD1	Methylenetetrahydrofolate dehydrogenase 1	6	22.17	0	4	0	N	Cytoplasm
W5PPQ4	GNE	Glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase	2	20.06	3	0	0	N	Cytoplasm
W5QEY8	GFPT1	Glutamine--fructose-6-phosphate transaminase 1	4	21.86	1	1	0	N	Cytoplasm
W5PB30	GALE	UDP-glucose 4-epimerase (GT)	1	21.00	1	1	0	N	Cytoplasm
W5PPJ0	STT3A	STT3 oligosaccharyltransferase complex catalytic subunit A	5	22.34	1	1	13	N	ER
W5QH00	RETSAT	Retinol saturase	6	21.37	0	5	0	Y	ER
W5QA42	MFAP4	Fibrinogen C-terminal domain-containing protein	2	23.38	1	0	0	Y	Extracellular
W5Q9K1	QSOX1	Sulfhydryl oxidase	5	22.11	2	30	1	Y	Golgi
W5NWN0	GSTZ1	Glutathione S-transferase zeta 1	4	22.54	4	1	0	N	Mitochondria
W5QD93	LOC101119050	Dehydrogenase/reductase SDR member 4-like	3	22.31	1	3	0	N	Mitochondria
W5QBS0	ACAD10/11	Acyl-CoA dehydrogenase family	5	21.61	1	20	1	N	Mitochondria
W5NQ36	ECI2	Enoyl-CoA delta isomerase 2	5	21.56	1	5	0	N	Mitochondria
W5NR14	IDH3A	Isocitrate dehydrogenase [NAD] subunit	3	21.34	2	0	0	N	Mitochondria
W5PYA5	NDUFA8	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 8	4	21.22	0	3	0	N	Mitochondria
W5QB48	LOC101113624	ATP synthase, subunit F	2	22.59	0	3	0	N	Mitochondria
W5QFD8	PYROXD2	Pyridine nucleotide-disulphide oxidoreductase	2	21.47	2	0	0	N	Mitochondria
W5P8N8	IDH3B	Isocitrate dehydrogenase [NAD] subunit	4	20.98	0	6	0	N	Mitochondria
W5P066	CYB5B	Cytochrome b5 type B	2	20.79	0	1	1	N	Mitochondria
W5Q2E8	PKS	Polyketide synthase, enoylreductase domain	2	20.60	2	2	0	N	Mitochondria

Table S1 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
W5PM93	ACSS2	Acyl-CoA synthetase short chain family member 2	3	21.41	3	14	0	N	Nucleus
W5NVQ4	CAND1	Cullin associated and neddylation dissociated 1	3	21.25	1	9	0	N	Nucleus
W5Q496	RPL36	60S ribosomal protein L36	3	21.81	0	4	0	N	Ribosome
Other									
P00922	CA2	Carbonic anhydrase 2	2	23.06	0	1	0	N	Cell membrane
W5PGA9	NCSTN	Nicastrin	3	20.18	11	8	1	Y	Cell membrane
W5Q9C5	LNPEP	Leucyl and cystinyl aminopeptidase	2	20.18	15	6	1	N	Cell membrane
W5PUD2	RPS21	40S ribosomal protein S21	4	21.30	0	0	0	N	Cytoplasm
W5PZ25	CAPN8	Calpain 8	2	21.64	3	4	0	N	Cytoplasm
A7UHZ2	PSMD4	von Willebrand factor, type A domain	2	21.38	0	15	0	N	Cytoplasm
W5QFZ3	CCT4	T-complex protein 1 subunit delta	4	20.62	1	2	0	N	Cytoplasm
W5PJX0	CCT6B	Chaperonin containing TCP1 subunit 6B	3	20.96	2	1	0	N	Cytoplasm
W5Q7E0	CKAP4	Cytoskeleton associated protein 4	8	21.19	0	32	1	N	ER
W5PMD1	SPCS2	Signal peptidase complex subunit 2	4	21.37	0	0	2	N	ER
W5QDT3	DAD1	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1	2	20.25	0	0	3	N	ER
W5NVK4	SPCS1	Signal peptidase complex subunit 1	1	21.32	1	2	2	N	ER
W5Q341	NDUFA12	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12	3	20.91	1	10	0	N	Mitochondria
W5Q2P3	LETM1	Leucine zipper and EF-hand containing transmembrane protein 1	2	21.16	1	24	1	N	Mitochondria
W5P779	HNRNPL	Heterogeneous nuclear ribonucleoprotein L	5	21.68	2	9	0	N	Nucleus
W5QAH5	-	Transcription factor, GTP-binding domain	1	23.51	2	1	0	N	Nucleus
W5QFZ8	RPL13	Ribosomal protein L13	2	22.14	1	4	0	N	Ribosome
W5P0I6	RPS28	40S ribosomal protein S28	2	19.85	0	0	0	N	Ribosome
W5PQ79	RPS15	Ribosomal protein S15	2	23.34	0	0	0	N	Ribosome

Table S1 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
Unknown									
W5PNH6	FAM234A	Family with sequence similarity 234 member A-related	7	22.86	3	6	1	N	Cell membrane
W5Q8Z5	CA4	Carbonic anhydrase 4	6	20.69	4	5	0	Y	Cell membrane
W5PQ27	GATD3A	Glutamine amidotransferase-like class 1 domain containing 3A	4	21.35	1	8	0	Y	Mitochondria

Table S2. Identification by mass spectroscopy of sheep abomasal scrape tissue (ST) proteins eluted from a Tci-gal-1 column.

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
Immune response									
W5Q0Z6	MCAM	Melanoma cell adhesion molecule (MUC18-related)	11	21.90	6	6	1	N	Cell membrane
W5Q3J8	HLA-DRB3	MHC class II antigen DRB3	2	21.23	1	1	1	Y	Cell membrane
W5Q2V1	HLA-DQB1	MHC class II antigen DQB1	2	20.59	0	3	1	Y	Cell membrane
W5PWS0	CD46	Membrane cofactor protein CD46	3	21.35	3	10	1	Y	Cell membrane
W5QD49	PSME1	Proteasome activator subunit 1	2	20.13	1	1	0	N	Cytoplasm
Cell signalling									
W5P5W1	ITGA3	Integrin subunit alpha 3	10	23.64	11	4	1	Y	Cell membrane
Q863C4	ITGB6	Integrin beta-6	3	22.58	7	17	1	Y	Cell membrane
W5NY03	CD163	CD163, scavenger receptor cysteine rich domain	6	22.30	7	13	1	Y	Cell membrane
W5QG66	ITGB5	Integrin beta-5	3	21.31	5	25	1	N	Cell membrane
W5QG70	MUC13	SEA domain-containing protein	2	21.79	9	62	1	Y	Cell membrane
W5Q501	GNAS	GNAS complex locus	3	21.63	1	71	0	N	Cytoplasm
W5PF02	SRPRB	Signal recognition particle receptor, beta subunit	3	21.06	1	4	1	N	Cytoskeleton
W5PQS4	PROCR	MHC class 1-like antigen recognition domain	3	21.61	4	1	1	Y	Cytoskeleton
Cell adhesion									
W5QC22	CD47	CD47 molecule	3	22.12	5	4	5	N	Cell membrane
W5NTA5	CD34/PODXL	CD34/Podocalyxin	2	20.97	9	104	1	N	Cell membrane
W5QCH8	NPTN	Neuroplastin	6	22.21	4	3	1	N	Cell membrane
W5NY24	CTNNB1	Catenin beta 1	2	20.95	5	16	0	N	Cytoplasm
W5Q5C7	CTNNA1	Catenin alpha 1	2	21.17	0	20	0	N	Cytoskeleton
W5Q263	ICAM1/CD54	Ig-like domain-containing protein	2	21.19	7	1	1	Y	Extracellular
Apoptosis									
W5PD03	NDUFA13	GRIM-19	3	21.23	0	0	1	N	Mitochondria

Table S2 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
Regulatory function									
W5PEN0	ITGA1	Integrin subunit alpha 1	9	22.86	19	15	1	Y	Cell membrane
W5PBS4	LRP1	LDL receptor related protein 1	14	22.89	33	122	1	N	Cell membrane
W5P6A5	SPINT2	Serine peptidase inhibitor, Kunitz type 2	2	22.17	2	9	1	Y	Cell membrane
W5NUR4	ADAM9	ADAM metallopeptidase domain 9	2	20.52	5	18	1	Y	Cell membrane
W5QAM4	CAMK2D	Protein kinase domain-containing protein	2	20.49	2	21	1	N	Cell membrane
W5P5W6	NDRG1	N-myc downstream regulated 1	3	22.36	0	19	0	N	Cytoplasm
W5NTX6	RIPK1	Receptor interacting serine/threonine kinase 1	1	22.10	2	38	0	N	Cytoplasm
W5P889	SEPTIN9	Septin 9	3	20.67	0	49	0	N	Cytoskeleton
W5PIU7	MTPN	Myotrophin	2	20.23	1	1	0	N	Cytoskeleton
W5QG21	PDIA5	Protein disulfide-isomerase A5	2	20.29	0	3	1	N	ER
W5PLZ4	TFF2	Trefoil factor 2	2	24.21	0	1	0	Y	Extracellular
W5PVM5	LAMA5	Laminin subunit alpha 5	5	21.60	13	173	0	N	Extracellular
W5PXR3	ENPP1	Ectonucleotide pyrophosphatase/phosphodiesterase 1	4	21.27	6	8	0	N	Extracellular
W5PEE9	LAMP1	Lysosomal associated membrane protein 1	3	22.15	17	5	1	N	Lysosome
W5NYJ4	DKC1	PUA domain-containing protein	3	21.86	0	20	0	N	Nucleus
Transport									
W5P1H4	CTL4	Choline transporter-like protein 4	3	21.98	8	6	10	N	Cell membrane
W5P246	TM9SF2	Nonaspanin	3	21.46	0	2	9	Y	Cell membrane
W5P9V5	PIGR	Polymeric immunoglobulin receptor	3	20.67	4	4	1	Y	Cell membrane
Q9XT28	ATOX1	Copper transport protein ATOX1	3	20.15	0	0	0	N	Cytoplasm
W5NZ10	NUTF2	Nuclear transport factor 2	2	20.62	0	0	0	N	Cytoplasm
W5QGX9	SEC22B	SEC22 homolog B, vesicle trafficking protein	2	20.34	1	0	1	Y	ER
W5PLQ6	ARCN1	Coatomer subunit delta	4	21.09	0	18	0	N	Golgi
W5PDH7	NPC1	Niemann-Pick C type protein	4	21.54	14	9	9	Y	Lysosome
W5P180	MTX1	Metaxin	4	21.95	0	5	0	N	Mitochondria
W5P1G5	ERGIC1	Endoplasmic reticulum-Golgi intermediate compartment 1	1	21.46	1	0	2	N	Nucleus

Table S2 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
Binding function									
W5QFP1	PABPC4	Polyadenylate-binding protein	2	21.60	0	8	0	N	Cytoplasm
W5PK26	LASP1	LIM and SH3 protein 1	4	21.66	0	9	0	N	Cytoplasm
W5PBY5	RI	Ribonuclease inhibitor-like protein	3	20.88	0	0	0	N	Cytoplasm
W5PGT5	MYH11	Myosin heavy chain 11	1	20.94	10	52	0	N	Cytoskeleton
W5PTV7	ACE	Angiotensin-converting enzyme	5	22.07	6	12	0	N	Extracellular
W5NUA0	PRELP	Proline and arginine rich end leucine rich repeat protein (Prolargin)	4	20.63	2	4	0	Y	Extracellular
W5PSB6	ACE2	Angiotensin-converting enzyme 2	1	20.61	6	3	1	Y	Extracellular
W5P3H8	IGF2R	Insulin like growth factor 2 receptor, mannose-6-phosphate receptor	7	22.28	16	13	1	N	Golgi
W5PD64	ARL8B	ADP ribosylation factor-like GTPase 8B	3	21.57	2	0	0	N	Lysosome
W5Q8Y5	HDLBP	High density lipoprotein binding protein/Vigilin	3	21.71	0	30	0	N	Nucleus
W5QBH1	CSRP1	Cysteine and glycine rich protein 1 (LIM domain, zinc finger)	2	21.19	3	7	0	N	Nucleus
Metabolic function									
W5P906	DPP4	Dipeptidyl peptidase 4	3	21.45	9	7	0	Y	Cell membrane
W5P1J8	CAO	Copper amine oxidase	3	22.71	4	11	0	Y	Cytoplasm
W5Q4B6	PGLS	6-phosphogluconolactonase	2	20.35	0	1	0	N	Cytoplasm
W5QJ31	MTHFD1	Methylenetetrahydrofolate dehydrogenase 1	6	21.61	0	4	0	N	Cytoplasm
W5PJ69	CKM	Creatine kinase, M-type	2	21.05	1	3	0	N	Cytoplasm
W5PSI1	PSMD3	Proteasome 26S subunit, non-ATPase 3 (PCI domain)	2	20.76	1	17	0	N	Cytoplasm
W5QD45	UGT1A1	UDP-glucuronosyltransferase	2	20.90	1	4	0	Y	ER
W5PVR9	ERMP1	Endoplasmic reticulum metallopeptidase 1	4	22.28	2	5	8	N	ER
W5QA42	MFAP4	Fibrinogen C-terminal domain-containing protein	2	22.21	1	0	0	Y	Extracellular
W5P8L8	B4GALNT2	Beta-1,4 N-acetylgalactosaminyltransferase	2	20.61	3	2	1	N	Golgi
W5Q9K1	QSOX1	Sulfhydryl oxidase	5	22.46	2	30	1	Y	Golgi
W5PIZ6	MLYCD	Malonyl-CoA decarboxylase	1	20.98	0	4	0	N	Mitochondria
W5QB48	LOC101113624	ATP synthase, subunit F	2	23.00	0	3	0	N	Mitochondria

Table S2 (continued)

Accession number	Alternate gene ID	Protein annotation	Unique peptides	Average abundance	Number of predicted glycosylation sites		TM domains	Signal peptides	Predicted cellular location
					N-glyc	O-glyc			
W5QFD8	PYROXD2	Pyridine nucleotide-disulphide oxidoreductase	2	21.01	2	0	0	N	Mitochondria
W5P066	CYB5B	Cytochrome b5 type B	2	21.49	0	1	1	N	Mitochondria
W5P8N8	IDH3B	Isocitrate dehydrogenase [NAD] subunit	4	21.27	0	6	0	N	Mitochondria
W5Q2E8	PKS	Polyketide synthase, enoylreductase domain	2	20.83	2	2	0	N	Mitochondria
W5NVQ4	CAND1	Cullin associated and neddylation dissociated 1	3	20.35	1	9	0	N	Nucleus
Other									
P00922	CA2	Carbonic anhydrase 2	2	21.15	0	1	0	N	Cell membrane
W5PGA9	NCSTN	Nicastrin	3	20.74	11	8	1	Y	Cell membrane
W5Q9C5	LNPEP	Leucyl and cystinyl aminopeptidase	2	20.42	15	6	1	N	Cell membrane
W5NTZ9	AIMP2	Aminoacyl tRNA synthase complex-interacting multifunctional protein 2	2	20.25	1	5	0	N	Cytoplasm
W5QFZ3	CCT4	T-complex protein 1 subunit delta	4	20.00	1	2	0	N	Cytoplasm
W5PJX0	CCT6B	Chaperonin containing TCP1 subunit 6B	3	20.79	2	1	0	N	Cytoplasm
W5PZH6	MPDU1	Mannose-P-dolichol utilization defect 1 protein	2	20.99	0	0	4	N	ER
W5NYN6	RTN3	Reticulon	1	21.78	1	151	3	N	ER
W5NVK4	SPCS1	Signal peptidase complex subunit 1	1	22.34	1	2	2	N	ER
W5QAH5	-	Transcription factor, GTP-binding domain	1	23.93	2	1	0	N	Nucleus
W5P0R5	RPL21E	Ribosomal protein L21e	1	21.71	0	0	0	N	Ribosome
W5P900	RPL29	60S ribosomal protein L29	2	21.48	2	0	0	N	Ribosome
W5PQ79	RPS15	Ribosomal protein S15	2	23.80	0	0	0	N	Ribosome
Unknown									
W5PQ27	GATD3A	Glutamine amidotransferase-like class 1 domain containing 3A	4	21.66	1	8	0	Y	Mitochondria

Chapter 3 – Future directions

The above investigation gained valuable insight into the binding partners of parasitic galectin towards host tissue that did not have a distinguishable active *T. circumcincta* infection. Further research using experimentally infected animals would be able to gain further insight into the roles of galectins during an active infection. Future studies could look into the independent expression and purification of the N- and C-terminal CRDs of tandem-repeat Tci-gal-1, and conduct CRD affinity chromatography with each individual domain to determine their respective binding specificities. This was partially explored by Greenhalgh and Newton (1999), who showed differences in binding between full length Tci-gal-1 and its C-terminal domain when incubated with sheep sera. Repetition of the above experiment with both Tci-gal-1 terminals could potentially identify the varying roles of each CRD (Greenhalgh and Newton 1999). This has been successfully completed with galectin 9, revealing that its C-CRD induces T cell death while the N-CRD is more dominant in mediating innate immune cell activation (Li *et al.* 2011). Similar *in vitro* studies could be done with parasite galectin to explore their effects on sheep abomasal mucosa. Conducting galectin knock-out or knock-down experiments in *T. circumcincta* and experimentally infecting those strains in sheep could also help better understand galectin function. Mice have been widely used as a model for galectin knockout studies. For example, a 2017 study showed that galectin 3 deficiency in mice challenged with *Trypanosoma cruzi* led to greater parasite survival and replication *in vitro,* as well as increased systemic parasitaemia *in vivo* (da Silva *et al.* 2017).

Organoid explant systems have been used in past studies to investigate host-parasite interactions. A study in 2004 challenged goat and sheep abomasal tissue explants with exsheathed *T. circumcincta* L3 *in vitro* to explore the tissue association phase of larval development (Jackson *et al.* 2004). The results from this study show much higher larval establishment in naïve tissue compared to tissue from

previously infected sheep and goats, demonstrating a similarity to effector mechanisms in response to larval challenge *in vivo* (Jackson *et al.* 2004). Similar methods have been used to study the immediate immune response against *H. contortus* in *ex vivo* sheep abomasa. Kemp and colleagues (2009) explored rapid rejection of *H. contortus* L3 following repeated larval infection and observed an increase of mast cells and globular leukocytes as expected. Although there was no significant increase in tissue eosinophils, the eosinophil-specific galectin 14 was upregulated, suggesting a potential role in parasite immune exclusion (Kemp *et al.* 2009). Hypersensitisation to the parasite resulted the upregulation of these immune effector cells, leading to significant inhibition of parasite establishment. A more recent study discovered that galectin 1 is also upregulated in an *ex vivo* sheep abomasa segment model incubated with *H. contortus* L4 (El-Ashram *et al.* 2018). Glycan mapping or immunofluorescence studies can also be conducted on *T. circumcincta* to investigate if Tci-gal-1 is self-binding to parasite proteins, acting as a camouflage mechanism. Hco-gal-1 has been demonstrated to mimic the eosinophil chemokinetic activity of host galectin 9, modulating the host's immune response in favour of nematode development (Turner *et al.* 2008). This emphasised the need for further exploration of the interaction between parasite- and host-derived galectins.

Apart from the unknown binding nature of the captured glycoproteins, several limitations were identified in this study. Predominantly, none of the sheep sampled were infected with GI nematodes. Since sheep abomasa were obtained from an abattoir, the clinical history of the sheep host is unknown, such as the age or level of infection at the time of slaughter. Additionally, sheep were sampled in early autumn while *T. circumcincta* infections are generally more prolific in late winter where rainfall is consistent (O'Connor *et al.* 2006). The absence of infection most likely contributed to the lack of expected hallmark proteins during infection, such as intelectins and certain mucin variants, being captured. While it is interesting to determine the binding partners from the early onset of infection, it would be beneficial to conduct parallel experiments of known naïve and infected individuals to detect differences in galectin activity. Level of infection as well as co-infection with other species of

nematodes should be considered in subsequent studies by sampling during peak periods of infection, or experimentally infecting sheep with *T. circumcincta*.

Column chromatography using immobilised recombinant Tci-gal-1 was used in this study to isolate sheep abomasal proteins that interacted with the parasite galectin. Despite the ease and efficiency of this method, it is not without limitations. During the galectin-NHS Sepharose conjugation process, an excess of Tci-gal-1 was observed in the wash fraction. Although this could be due to overloading of the resin binding capacity, other factors such as non-optimal binding conditions may have also contributed to the low coupling efficiency. Recombinant galectin lost in the flow-through and wash fractions also made it difficult to determine the concentration of protein successfully coupled onto the Sepharose resin. Moreover, Tci-gal-1 affinity chromatography with sheep proteins may be affected if remaining active binding sites in the NHS Sepharose resin were not completely blocked with a suitable buffer, potentially allowing non-specific binding of host proteins to the resin. These factors combined may hinder the binding success of the sheep proteins to the immobilised parasite galectin, disrupting the specificity of the pull-down assay.

Furthermore, many proteins with predicted glycosylation sites lacked an ER signal peptide. Glycosylation occurs in the ER and Golgi apparatus, and signal peptides mediate protein entry into this secretory pathway (Faye *et al.* 2005). However, the absence of signal peptides could be due to incomplete protein sequences in the database which may have led to undetected signal peptides within the sequence. Repeating this experiment in a more controlled manner using experimentally infected sheep will allow for more definitive conclusions to be made about the potentially critical role galectins play in establishing infection.

References

da Silva, AA, Teixeira, TL, Teixeira, SC, Machado, FC, Dos Santos, MA, Tomiosso, TC, Tavares, PCB, Brigido, R, Martins, FA, Silva, NSL, Rodrigues, CC, Roque-Barreira, MC, Mortara, RA, Lopes, DS, Avila, VMR, da Silva, CV (2017) Galectin-3: A Friend but Not a Foe during *Trypanosoma cruzi* Experimental Infection. *Frontiers in Cellular Infection Microbiology* **7**, 1-9.

El-Ashram, S, Li, C, Abouhajer, F, Mehmood, R, Al Nasr, I, Zhang, Y, Lu, T, Yili, D, Suo, X, Haoji, Z, Li, Z, Huang, S (2018) An *ex vivo* abomasal ovine model to study the immediate immune response in the context of *Haemonchus contortus* larval-stage. *Veterinary Parasitology* **254**, 105-113.

Faye, L, Boulaflous, A, Benchabane, M, Gomord, V, Michaud, D (2005) Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* **23**, 1770-1778.

Greenhalgh, CJ, Newton, SE (1999) Recombinant expression of a galectin from the sheep gastrointestinal parasite *Teladorsagia circumcincta*: its use in isolating galectin-glycoconjugates. *Protein Expression and Purification* **16**, 152-159.

Jackson, F, Greer, AW, Huntley, J, McAnulty, RW, Bartley, DJ, Stanley, A, Stenhouse, L, Stankiewicz, M, Sykes, AR (2004) Studies using *Teladorsagia circumcincta* in an *in vitro* direct challenge method using abomasal tissue explants. *Veterinary Parasitology* **124**, 73-89.

Kemp, JM, Robinson, NA, Meeusen, ENT, Piedrafita, DM (2009) The relationship between the rapid rejection of *Haemonchus contortus* larvae with cells and mediators in abomasal tissues in immune sheep. *International Journal for Parasitology* **39**, 1589-1594.

Li, Y, Feng, J, Geng, S, Geng, S, Wei, H, Chen, G, Li, X, Wang, L, Wang, R, Peng, H, Han, G, Shen, B, Li, Y (2011) The N- and C-terminal carbohydrate recognition domains of galectin-9 contribute differently to its multiple functions in innate immunity and adaptive immunity. *Molecular Immunology* **48**, 670-677.

O'Connor, LJ, Walkden-Brown, SW, Kahn, LP (2006) Ecology of the free-living stages of major trichostrongylid parasites of sheep. *Veterinary Parasitology* **142**, 1-15.

Turner, DG, Wildblood, LA, Inglis, NF, Jones, DG (2008) Characterization of a galectin-like activity from the parasitic nematode, *Haemonchus contortus*, which modulates ovine eosinophil migration in vitro. *Veterinary Immunology and Immunopathology* **122**, 138-145.