Characterisation of *Necator americanus* excretory/secretory products

Jayden Anthony Logan

BSc (Hons)

This thesis is presented for the degree of Doctor of Philosophy from James Cook University, College of Public Health, Medical and Veterinary Sciences

Submitted October 2019

Acknowledgements

There are a multitude of people who deserve credit and have my sincerest gratitude for helping me along this journey. Firstly *Prof. Alex Loukas* – your unwavering optimism in the project, confidence in my abilities, infectious love of worms, and guidance throughout has made it all possible. Secondly to *Dr. Javier Sotillo* – you supported me on a daily basis throughout all aspects of my project and were always available to lend a hand. Thirdly to *Dr. Paul Giacomin* – your expertise and direction was indispensable, my only regret is not having you on the team from the beginning. Lastly to *Dr. Severine Navarro* – thank you for shaping me as a researcher, and helping me to hone my technical skills. The four of you got me across the line, you went above and beyond and I couldn't have done it without you.

The team at AITHM, the lab and administration team, each deserve individual page long mentions but unfortunately this is not the place for that. Not only for the technical advice, but for making coming in each day an absolute pleasure. A special mention to *Phill Walsh*, I couldn't have asked for a better, more understanding manager. Working for you made my whole experience a great deal better.

Thank you to the Australian Federal Government for funding my Australian Postgraduate Award (APA). To James Cook University for awarding me the APA, as well as for my college Top-Up scholarship.

Finally, thank you to my wife *Meg*, for your loving support and continuous encouragement. You deserve a PhD out of this too.

Statement of contributions

This thesis, the work it entails and the written components includes no previously published material except where due reference has been made. The contributions of others has been clearly stated below. The content of my thesis is a result of work that was commenced at the beginning of my candidature. It does not include a substantial component of work that has been submitted to any other university or tertiary institution to qualify for another award or degree of any kind.

I acknowledge that the University library must receive an electronic copy of my thesis which is subject to the policy and procedures of James Cook University. This thesis, in accordance with the Copyright Act 1968, will be made available, unless an embargo is approved by the Graduate School.

Every reasonable effort has been made to gain permission and acknowledge the owners of copyright material.

Nature of Assistance	Contribution	Contributors
Intellectual support	Experimental design Statistics support Editorial assistance	Prof. Alex Loukas Dr. Javier Sotillo Dr. Paul Giacomin Dr. Severine Navarro
Intellectual support	Proteomics support Genome reannotation assistance Diagnostics guidance	Dr. Srikanth Manda Dr. Young-Jun Choi Dr. Mark Pearson

Data collection	Research assistance	Dr. Javier Sotillo Dr. Paul Giacomin
Material	Whole ES mixture used for chapter 2	Prof. Ricardo Fujiwara
Project costs	Entire project	Prof. Alex Loukas

Abstract

The burden of chronic inflammatory diseases, including inflammatory bowel diseases and asthma, is increasing around the world. An inverse relationship between early life exposure to microbes and larger pathogens and the prevalence of these chronic diseases is well supported. The human hookworm, *Necator americanus*, is one such pathogen which infects hundreds of millions of people worldwide. The hygiene and 'old friends' hypotheses propose that by removing hookworms and other pathogens we have eliminated an important factor in the normal development of the immune system. Many parasites have evolved to regulate host immunity through the production of excretory/secretory (ES) products. I hypothesized that the *N. americanus* ES proteome contains proteins which could be exploited *ex vivo* to limit inflammatory responses.

Herein I describe, for the first time, the ES proteome of a human hookworm and utilize this data to carry out the first proteogenomic analysis of any parasitic helminth. This analysis significantly improved the available genome annotation, resulting in 3,425 fewer genes and increased numbers of introns and exons, total gene length and the percentage of the genome covered by genes. One hundred and ninety-eight ES proteins were identified using liquid chromatography tandem mass spectrometry with the dominant families found to be SCP/TAPS and 'hypothetical proteins'. In terms of sequence homology and families present, the findings compared well with closely related species including *Ancylostoma caninum*, *Nippostrongylus brasiliensis* and *Heligmosomoides polygyrus*.

Using the *Pichia pastoris* expression system I expressed 4 of the SCP/TAPS proteins and tested their diagnostic screening potential as well as their therapeutic efficacy in the trinitrobenzenesulfonic acid (TNBS) model of murine colitis and the house dust mite (HDM) model of murine asthma. Vehicle-treated mice that received TNBS experienced weight loss and significant colon pathology, and HDM-sensitized and -challenged mice had marked immune cell infiltration into the lung as well as increased

serum IgE compared with controls. None of the recombinant SCP/TAPS proteins were found to be therapeutic when administered intra-peritoneally in the TNBS colitis or HDM asthma models.

Ex vivo profiling of the recombinant ES proteins was carried out using peripheral blood mononuclear cells (PBMCs) isolated from the blood of healthy human volunteers. The hookworm recombinant protein NP17 consistently suppressed tumor necrosis factor (TNF)-α production by PBMCs that were stimulated with either LPS or PMA/ionomycin. A second protein, NP4, elicited pro-inflammatory responses in PBMCs in the form of elevated interleukin (IL)-6, IL-8 and TNF-α production. To provide insight into the potential binding partners of these proteins I used a human protein microarray. The results generated by probing of this array supported an immune regulatory role for at least some of these proteins in their interactions with host cells.

The description of the *N. americanus* proteome offers a valuable source of information on the proteins in the *N. americanus* secretome, and at least one of the recombinant proteins shows promise as a novel biologic for the treatment of human inflammatory disorders. The future implications of this data are varied and may be useful for understanding the molecular basis of host-parasite interactions. It also provides a long list of potential vaccine candidates, diagnostic targets, and immunomodulatory biologics. Indeed, by profiling the impact of just four of these proteins on human PBMCs I have shown that at least one protein selectively suppressed the production of inflammatory cytokines. This highlights the importance of characterizing ES proteins and the potential usefulness of this research.

Table of contents

List of tables

List of figures

List of abbreviations

AEC – airway epithelial cells

AHR – airway hyperresponsiveness

CD – Crohn's disease

cDCs – conventional dendritic cells

DALYs – disability adjusted life years

DSS - dextran sulfate sodium

DTT – dithiothreitol

DUF – domain of unknown function

emPAI – exponential modified abundance index

ES – excretory/secretory

FDR – false discovery rate

FoR – frequency of recognition

FT – Flow through

GO – gene ontology

GOI – gene of interest

GR – glucocorticoid receptor

GSSPs – genome search specific peptides

HDM – house dust mite

IAM – iodoacetamide

IBD – inflammatory bowel disease

IgE – immunoglobulin E

IL – interleukin

LPS – lipopolysaccharides

MM – mastermix

NF-κB - nuclear factor kappa-light-chain-enhancer of activated B cells

NK – natural killer

ORF – open reading frame

OVA – ovalbumin

PBMCs – peripheral blood mononuclear cells

PMA/ion - phorbol myristate acetate and ionomycin

PRR – pattern recognition receptors

RO – reverse osmosis

S/N – supernatant

SCP/TAPS - sperm-coating protein/Tpx/antigen 5/pathogenesis related-1/Sc7

SIT – specific immunotherapy

TGF-β – transforming growth factor β

Th – T-helper type

TLR – toll-like receptor

TNBS - 2,4,6-Trinitrobenzene sulphonic acid

TNF-α – tumor necrosis factor alpha

Tregs – regulatory T cells

UC – ulcerative colitis

1 Chapter 1 - Introduction

1.1 Inflammatory diseases

Inflammation is a key immune response that occurs in defence against pathogens, cellular damage, irradiation and toxin exposure [1]. Typically, it acts to limit or remove damage in these circumstances and initiate the healing process [1]. However, in cases of chronic inflammation and dysregulated immunity, a number of immune-mediated diseases can occur. Immune diseases, encompassing allergy and autoimmunity, are on the rise in developed nations [2]. This has been attributed to a number of factors including improved diagnosis, lifestyle and genetic factors. Of particular note for their high prevalence and burden of disease are asthma and inflammatory bowel diseases (IBD).

Atopic dermatitis, allergic rhinitis and asthma are common inflammatory diseases involving an inappropriate immune response to often harmless stimuli [3]. Asthma is a chronic, non-communicable, inflammatory disease characterised by repeated episodes of wheezing, chest tightening, coughing and shortness of breath [4]. Asthma affects people of all ages and demographics, and if left uncontrolled, these symptoms impose moderate to severe limitations on a person's quality of life. While it usually begins in childhood, it has been shown to appear at any stage of life [5]. Asthma is often a lifelong condition. Half of the children with persistent asthma and three-quarters with severe asthma at age 6 were shown to still be symptomatic at age 50 [6]. Given the global morbidity and health-care costs associated with asthma, the disease represents an area of great need. A number of studies have highlighted the association between asthma and IBD including both Crohn's disease (CD) and ulcerative colitis (UC) [7].

CD and UC are both idiopathic, immune-mediated inflammatory conditions affecting the gastrointestinal tract [8]. IBD typically begins in a person's twenties or thirties with a majority of afflicted individuals experiencing relapses and chronic disease [8]. The main feature of UC is varying

degrees of mucosal inflammation extending proximally from the rectum [9]. This inflammation often results in severe superficial mucosal ulceration, fistulas, stenosis, and intestinal granulomas [10]. CD is similar in that it involves inflammatory driven lesions and ulceration however, it can manifest throughout the entire gastrointestinal tract [9]. The clinical manifestations of IBD can include abdominal pain, weight loss, haemorrhagic diarrhoea, anorexia and tenesmus [11].

1.2 Epidemiology

Asthma is the most common chronic disease in children and the most common respiratory disease worldwide, estimated to affect 334 million people [12,13]. Australia leads the world in terms of asthma prevalence with 1 in 9 of the population reporting asthma in 2014-15 [14]. Asthma affects twice as many children as adults and is one of the leading causes of missed school and work days, emergency department visits and hospitalisations [15]. While asthma prevalence is consistent or reducing in many developed nations, developing countries are seeing drastic increases as they become progressively westernised [16]. Migration studies reveal the importance of environmental factors in these global trends. Immigrants moving from low to high prevalence asthma countries begin with lower prevalence, increasing to similar proportions as duration of residence increased [17]. This effect was generally greater in second generation migrants and applied to multiple allergic diseases [17].

IBD is primarily a disorder of the developed world. It afflicts more than five million people worldwide, including 1.4 million in the United States and three million in Europe [18]. The greatest increase in incidence however is being seen in traditionally low-prevalence regions including Asia and South-America [19]. While some studies have reported gender distribution bias for IBD, others have described no such effect [20,21]. Although IBD onset has typically been middle-aged, the number of reported cases of paediatric IBD is increasing in northern Europe [22]. Further to this, prevalence gradients have been observed from north-south and west-east with higher rates in north-western countries [23,24]. This effect suggests the importance of environmental factors in the development of IBD.

1.3 Aetiology and immunology

Asthma

Asthma entails a range of heterogeneous phenotypes for both children and adults afflicted. These phenotypes differ in presentation, pathophysiology and aetiology. Understanding the complex nature of the disease pathogenesis is key for the progress of treatment strategies as well as preventative measures. Research into asthma development highlights numerous genetic, host-environment interactions, *in utero* exposures, and immunological factors that contribute to the onset of asthma [16].

From studies of twins, family history, familial aggregation and segregation, it is widely accepted that genetic factors contribute significantly to the development of asthma [25]. However, the exact contribution of genes when compared with environmental risk factors is significantly disputed. Some studies have reported heritability accounts for just 35% of the risk, whereas others claim the figure to be as high as 95% [25]. This disparity in the literature may be explained by the fact that asthma results from multiple gene interactions. While some of these genes are protective, others contribute to disease pathogenesis, with each gene having its own propensity to be influenced by environmental factors [25]. Genome-wide association studies have identified polymorphisms from both adults and children implicated in abnormalities in adaptive immunity as well as the epithelium barrier. The most notable of these are *tim1, dpp10, opn3, ormdl3, phf11, gpra* and *pde4d* [26]. While each of these genes will not be discussed individually in this review, their importance should not be overlooked. Asthma concordance in monozygotic twins is notably only 50%, highlighting the importance of other factors such as early-life exposures in asthma development [27]. These early life exposures, and the failure to account for them appropriately, is another likely reason for the inconsistencies in heritability risk percentages.

A multitude of risk factors in the prenatal period have been implicated in asthma. These include maternal smoking, diet and nutrition, stress, antibiotic use and mode of delivery [28]. Prenatal

maternal smoking impaired lung function and was a consistent predictor for wheeze in the first year of life [29]. While the link between asthma and stress is less clear than maternal smoking, stress of a caregiver and infant immunoglobulin (Ig)E levels have been positively correlated [30]. Although roughly half of all infants experience wheezing, only 10-15% have diagnosed asthma by age 6 [31]. Exploring risk factors in childhood is essential for understanding this trend as well as potentially limiting disease frequency. Having multiple older siblings, consuming water with higher overall microbe content, attending early day care, growing up in close proximity with farm animals or cats and dogs, and living in a predominantly agrarian economy have all been reported to limit or reduce the risk of developing asthma [32-37]. Across these studies the impact was greatest in the first year of life and the reduction in the risk of sensitization observed was not allergen specific.

When considering the immune responses involved in asthma pathogenesis, two distinct paradigms (adaptive and innate) are consistently reported. Both paradigms are similar in that they involve an aberrant immune response to non-pathogenic stimuli [38]. Early life allergen exposure is widely accepted as a key risk factor for subsequent development of allergic asthma [39]. When an allergen enters the airways it can directly interact with both airway epithelial cells (AECs) and conventional dendritic cells (cDCs) through pattern-recognition receptor (PRR) binding (Figure 1.1). Upon activation, AECs produce the chemokines CCL2 and CCL20, which attract immature pre-cDCs [39]. An array of signaling cytokines are also produced by activated epithelial cells to support the maturation of these immature cDCs to CD11b+ cDCs including granulocyte-macrophage colony-stimulating factor (GM-CSF), IL-25, IL-1α, thymic stromal lymphopoietin (TSLP) and IL-33 [39]. These activated CD11b+ cDCs migrate to the draining mediastinal lymph nodes in an IL-13-dependent manner where they induce T cell differentiation and subsequent T helper type 2 (Th2) and T helper type 17 (Th17) responses [39]. Th2 cells secrete an array of pro-allergic cytokines including IL-3, IL-4, IL-5, IL-9, IL-13 and GM-CSF. These cytokines elicit responses that are dominated by eosinophils, but also contain mast cells, basophils, neutrophils, monocytes and macrophages [4]. Following sensitization, repeated exposure of

the airways to allergens results in an early-type bronchoconstrictor response (EAR) which is driven by mast cells and typically lasts between 5-90 minutes. Specifically, IgE binds to mast cells to induce release of histamine, leukotriene C_4 and prostaglandin D_2. The late phase response (LAR) follows the EAR 3-12 hours post allergen exposure and is characterized by infiltration and activation of eosinophils, additional LTC4 generation, release of Th2 cytokines from mast cells and T cells, and an overall increase in airway responsiveness. While studies into EAR and LAR do provide a mechanism by which allergen-induced exacerbations occur, they are not sufficient to explain the persistent inflammation seen in most instances of asthma. This Th2 type inflammation takes place in more than 80% of children and in a majority of adults with sensitization to pollen, pet dander, dust mites and fungi [40-42].

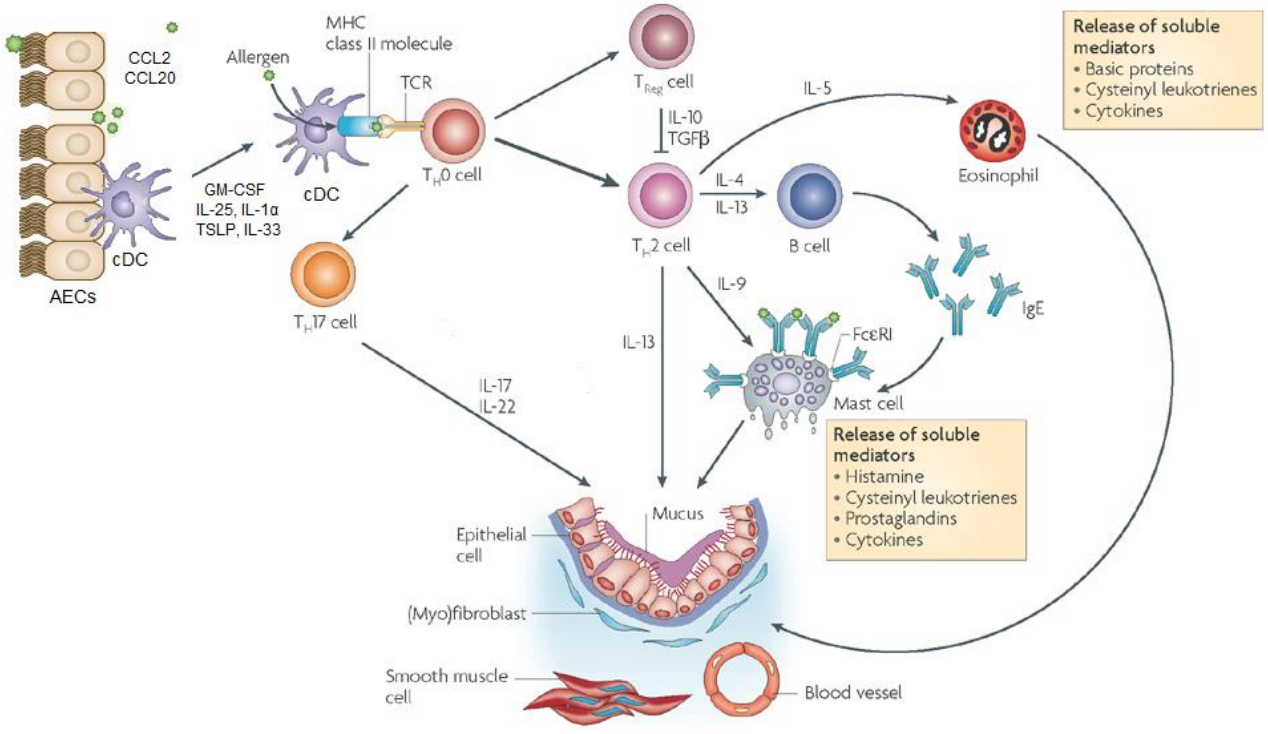


Figure 1.1: Summary of epithelial cell and DC interactions and subsequent responses during sensitisation and challenge of experimental asthma. Airway epithelial cells and cDCs express pattern-recognition receptors, directly activated by allergens which leads to the production of chemokines CCL2 and CCL20. Activated cDCs induce T cells to differentiate to Th2 cells, producing an array of pro-allergic

cytokines including IL-4, IL-5, IL-9, IL-13 and GM-CSF. These cytokines lead to EAR and LAR characteristic of allergic asthma. Adapted from: [43] and [44]

Asthma disease severity and duration are directly proportional to airway wall thickening. This airway remodeling is a key component of asthma pathogenesis to which numerous factors contribute. An increase in airway smooth muscle was first reported in 1922 as a key clinical feature in fatal asthma [45]. Since then a plethora of studies have reported on the varying contributors to airway remodeling and the underlying causes for this effect. Sub-epithelial basement membrane thickening in asthmatics results from the deposition of collagens type I, III, V, and VI along with fibronectin, tenascin, osteopontin and periostin [46,47]. These collagens are produced by a sheath of sub-epithelium myofibroblasts [4]. As a result of repeated injury from inflammation, an epithelial-mesenchymal trophic unit is established between the epithelial and smooth muscle layers of the airways [4]. Within this unit the epithelium is a significant source of platelet-derived growth factors, fibroblast growth factors, transforming growth factor-β (TGF-β) family members and functionally active periostin [48,49]. Furthermore, epidermal growth factors are produced by the epithelium, which play a key role in both smooth muscle proliferation and fibrosis [50]. Along with angiogenic factors and neurotrophin, these growth factors promote airway remodeling as a direct result of epithelial injury and impeded repair [4]. It is also important to note that defective resolution of inflammation pathways also play a role. By failing to effectively downregulate the inflammatory responses seen in chronic asthma, states of chronic inflammation contribute significantly to disease severity [51].

Inflammatory Bowel Disease

The aetiology of IBD is similar to asthma in that it is still poorly understood. While the triggers and underlying mechanisms are largely unknown, research points towards major contributing roles for individual genetic susceptibility, environmental factors, immune responses and gut microbiota.

The number of gene loci associated with IBD is currently 163, with 110 linked to both CD and UC, 23 to UC specifically and 30 to CD [52]. These findings highlight the significant overlap between the two major forms of IBD which is particularly important for understanding their shared pathogenesis. Some of the most notable of these genes are nucleotide-binding oligomerization domain containing 2 (*NOD2*), autophagy related 16 like 1 (*ATG16L1)*, and Immunity Related GTPase M (*IGRM1)*, which play important roles in autophagy and immune responses [53-55]. Defects in the genes whose products are necessary for normal detection of the gut microbiota lead to an increased risk of IBD. In particular, NOD2 for bacterial sensing along with caspase recruitment domain-containing protein 9 (CARD-9) for fungal detection [56,57]. *IL23R*, the gene which encodes the receptor for the pro-inflammatory cytokine IL-23, has been clearly linked with IBD for the role it plays in generation of Th17 cells [58,59]. Dysfunction in IL-10, a key immune regulatory cytokine, has also been linked with UC and CD [60]. IBD genetic research highlights a few key lessons in relation to the underlying disease mechanism. One is that there is an increasing number of gene loci implicated in IBD which emphasizes the important role genes play in disease pathogenesis. However, in terms of heritability, these variants only account for 20-25% in the aforementioned studies. This gap in explainable susceptibility loci is true for other polygenetic diseases such as asthma, and has been termed 'missing heritability' [61]. While many variants remain to be identified, considering gene-gene, gene-pathway and gene-environment interactions may help fill this gap [62].

Several environmental factors have been closely linked with an augmented risk of developing IBD including smoking, drugs, diet, stress and pollution. Of these factors, smoking is the most extensively explored prompter for IBD. A number of studies have shown a counter-intuitive protective effect of heavy smoking for UC as well as lower relapse rates, whereas for CD, smoking was found to increase the risk of development [63-66]. Prolonged, high-dose use of nonsteroidal anti-inflammatory drugs (NSAIDs) was linked with an increased risk of CD and UC as was the use of antibiotics [67,68]. Antibiotic

use was particularly detrimental for risk of IBD when use occurred within the first year of life [69]. This effect of antibiotics points to the importance of considering gut microbiota in IBD pathogenesis.

The human intestinal microbiome consists of more than 1,000 bacterial species with an individual person having approximately 160 species [70]. Differences, particularly reduced biodiversity, have been consistently reported between healthy individuals and IBD patients [71,72]. For CD these differences have been detected before the onset of disease, removing the possibility of IBD treatment-induced variation [73]. While there is no definitive microbial community for someone with IBD, a number of species have been found to be comparatively reduced or increased [74,75]. Some of these differences include reductions in Bifidobacteriaceae and Erysipelotrichaceae families, with increases in *Escherichia* and *Fusobacterium* genera [76,77]. When inflamed and non-inflamed sections of colon were compared in people with IBD, abnormal flora was also present in many cases. This suggests that inflammation is not the only driver of microbial dysbiosis [78].

The chronic inflammation experienced by patients with IBD is a hallmark of the disease and the main driver of clinical manifestations. The immunological mechanisms driving this inflammation, while poorly understood, vary across IBD types and occur in cycles of relapse and remission. In a healthy person, intestinal homeostasis is regulated by a careful balance of Th1, Th2, Th17 and regulatory responses. In IBD, dysfunction occurs at two key points: (1) impairment of the mucosal epithelial barrier and (2) inappropriate host acquired and innate immune responses. Although complex and often dissimilar across patients with IBD, the immunopathogenesis can be broken down into three distinct phases: (1) luminal contents penetrate into underlying tissue due to mucosal barrier defects, (2) limited clearance of this material due to inappropriate immune responses and (3) an acquired compensatory immune response which results in chronic inflammation and characteristic IBD associated lesions.

As the intestinal lumen contains a plethora of microorganisms, one of the key roles of the mucosal epithelium is to act as a physical barrier. Numerous studies have implicated disruption of the intestinal

epithelial barrier in the initiation and worsening of IBD [79]. Indeed, the apical junction complex, consisting of tight and adherens junctions, is compromised in people with IBD [80]. As a result, many of these people experience increased gut permeability [81,82]. Moreover, mucin production by intestinal goblet cells is reduced in people with IBD meaning the protective mucus layer is suboptimal [83]. Innate immune cells, including DCs, macrophages, and epithelial cells, use PRRs to identify and respond to stimuli appropriately [84]. Loss of function polymorphisms in the bacterial sensing gene NOD2 (a PRR) is one of the strongest genetic associations linked with IBD [57]. When tolerance of self-antigens is defective in the intestinal mucosa, either by genetic susceptibility or injury, it can contribute to IBD development [85]. These defects in the epithelium predispose people with IBD to increased antigen uptake, chronic immune activation and the resultant mucosal inflammation.

With increased permeabilisation, the mucosal immune system and commensal flora come in contact much more frequently. Bacterial clearance was shown to be limited in these interactions due to defective cytokine production by macrophages [86]. When the host immune system is overexposed to commensal bacteria, a breakdown of tolerance can occur. DCs provide the interface been intestinal epithelial cells and T lymphocytes. They ensure tolerance to commensals is maintained by promoting T cell differentiation to regulatory phenotypes. Overactive DCs have been observed in IBD at sites of inflammation [87]. These DCs induce differentiation of effector lymphocytes as well as natural killer (NK) and NK T cells while eliminating regulatory responses [88]. A key contributor to overactive DCs is the abnormalities in PRRs such as the NOD receptors previously mentioned [88]. Indeed, 15% of patients with CD are heterozygous or homozygous carriers of a major *NOD2* mutation [57]. These susceptibilities combine with a number of innate and adaptive immune alterations contributing to overall pathogenesis.

In active IBD there is an imbalance between regulatory and effector Th cells and the respective cytokines [89]. In a healthy mucosa, macrophages are characterized by hyporesponsiveness to Toll-like receptors (TLRs) and reduced capacity for priming adaptive responses when compared with peripheral

monocytes [90]. The acute phase of IBD involves dramatic increases in intestinal mucosa macrophages. In contrast to normal macrophages, these cells were shown to express a pro-inflammatory phenotype marked by increased production of IL-1, IL-6, IL-8, IL-12, TGF-β and TNF-α. TGF-β is a chemoattractant for additional macrophages, as well as neutrophils, and therefore increases recruitment of these cells [91]. IL-12 and TNF-α are potent inducers of pathogenic effector T cells [92]. Notably, the severity of UC and CD have both been correlated with increasing levels of serum TNF-α [93]. TNF-α augments production of other pro-inflammatory cytokines including IL-1β, IL-6, and IL-33, and has been shown to be involved in numerous pathological processes including increased epithelial permeability, granuloma formation and neutrophil recruitment [94-96]. For this reason it has become a key target for the treatment of both UC and CD with numerous anti-TNF drugs on the market [97].

1.4 Treatments and limitations

Asthma

Currently there is no preventative or cure for asthma itself. Instead, clinicians attempt to treat and prevent the symptoms to give patients the best quality of life possible. The first line of treatment for asthma is inhaled corticosteroids. Corticosteroids are the most efficient anti-inflammatory treatment available for numerous inflammatory diseases, including asthma [98]. Corticosteroids such as Budesonide work through the binding of glucocorticoid receptors (GR) found in the cytosol of most cells in the body [99]. Upon ligand binding and activation, the GRs are released from chaperone proteins (heat shock proteins) which then has three key anti-inflammatory effects within the cell [100]. Firstly the corticosteroid-receptor complex travels to the nucleus where it forms a homodimer and binds to the promotor region of DNA sequences called glucocorticoid response elements [100,101]. This binding trans-activates genes encoding anti-inflammatory proteins [101]. The second effect involves an interaction between the GR complex and other transcription factors – namely nuclear factor-kB (NF-kB), CREB-binding protein and histone deacetylase 2, which deacetylates histones to

suppress gene transcription of inflammatory molecules [100,101]. The final effect again takes place in the nucleus where the GR increases the expression of tristetraprolin, which binds to mRNAs encoding inflammatory cytokines, destabilizing them and therefore reducing their expression [100].

Although highly effective at treating the inflammation associated with asthma, glucocorticoids are not curative or disease limiting, even when started early in childhood [44]. Furthermore, there are significant limitations associated with long-term or high-dose glucocorticoid use. Some of the more serious effects include adrenal atrophy, hypertension, psychoses, immunosuppression, delayed wound healing, and foetal growth retardation [101]. Additionally, inhaled corticosteroids offer limited to no efficacy for virus-induced exacerbations or for asthmatics who smoke [44].

Corticosteroids are commonly used with β2-adrenoceptor agonists (β-agonists) which upon inhalation, mediate smooth muscle relaxation. β-agonists work by binding to β2 adrenergic receptors (β2AR) [102]. The receptor is coupled with stimulatory G protein and adenylyl cyclase. Activation results in the formation of cyclic adenosine monophosphate which stimulates phosphate kinase A to phosphorylate myosin light chain kinase in bronchial smooth muscle cells, ultimately ending in relaxation of the airways [102].

One of the major limitations of β-agonists is their short half-life, meaning continual drug re-administration to maintain a therapeutic effect [44,98]. To combat their short acting nature, formoterol and salmeterol, two long acting β-agonists (LABAs) can be used. However, a large systematic review of mortality associated with formoterol in 2009 found that patients taking this drug are at a 2.0-3.2-fold increased risk of death due to asthma [103]. Other studies have highlighted that subsensitivity (tolerance) occurs with prolonged use of LABAs, meaning higher doses are required to achieve the same efficacy [104,105]. Further, mono-treatment with LABAs is not recommended as it may mask worsening inflammation [106].

Comparatively recent treatment for asthma has emerged in the form of monoclonal antibodies (mAb). The only currently approved and available mAb for treatment of asthma is Omalizumab, which was first registered in Australia in 2002 [107]. Omalizumab is an anti-IgE mAb which disrupts the interaction between IgE and its receptor FcεRI, whilst also decreasing the expression of FcεRI on mast cells and basophils [107]. The result is inhibition of mast cell degranulation and prevention of histamine release. This in turn precludes bronchoconstriction and mucous production which would have otherwise taken place [107]. Other mAbs have been developed to target multiple Th2 and inflammatory pathways involved in asthma including anti-IL4, anti-IL5, anti-IL13, anti-TSLP and anti-TNF-α. The large number of clinical trials involving these mAbs have found them to be less successful than Omalizumab [108].

Allergen specific immunotherapy (SIT) is an alternative method for treatment of allergic rhinitis, some drug allergies, venom hypersensitivity and bronchial asthma [44]. SIT induces immunological tolerance rather than sensitisation, and stimulates the production of blocking IgG4 antibodies through repeated exposures to allergens [44]. Despite being quite effective for a majority of asthmatics, induction of tolerance with SIT usually takes several months and in some cases even longer. A key drawback of SIT is the reporting that some people experience systemic, life threatening side-effects and even death [109-111]. This has caused some nations (e.g the United Kingdom) to restrict the availability of SIT for asthma [112].

Inflammatory Bowel Disease

Much like asthma, IBD has no cure. Instead treatment is aimed at reducing inflammation, dampening down the immune system, and inducing remission. The first line of treatment for UC is often aminosalicylates, while glucocorticoids are used for CD. When these prove unsuccessful, clinicians move on to other immunosuppressants and biologics such as antibody treatments. Treatment choices vary across both UC and CD with some drugs having been shown to be efficacious for one but not the

other [93]. For example, 5-aminosalicylates (5-ASAs) are effective at maintaining remission in UC by reducing pro-inflammatory cytokine production as well as blocking neutrophil recruitment and inactivating NF-κB [113,114]. While they work well for people with UC, they have little effect in CD [115]. Corticosteroids are a first line treatment for CD and are commonly used to induce remission through similar effects as 5-ASAs [116]. Other classical immunosuppressive drugs used in IBD are methotrexate and cyclosporine-A which suppress inflammatory cytokines including TNF-α and induce apoptosis [117-119]. Glucocorticoids, while often effective at inducing remission in people with IBD, come with numerous side effects. The GR receptor is located on many cell types which results in non-specific coverage, increased patient susceptibility to infections, and the other aforementioned effects as seen in asthma treatment.

Given that a few key cytokines such as TNF-α and IL-1β play such key pathogenic roles in IBD progression, therapy targeting these molecules has been explored. Indeed, some mAbs have been shown to be highly successful for treatment of IBD. Specifically, anti-TNF agents such as adalimumab and infliximab work well in maintaining remission in both UC and CD [120,121]. By blocking the TNF receptor, these drugs limit the production of other pro-inflammatory cytokines, prevent regulatory T cell (Treg) apoptosis and promote wound healing through alternatively activated macrophages (AAMacs) [122]. Although anti-TNF agents have drastically improved outcomes for people with IBD, they are not effective for everyone [123]. Moreover, many people taking these drugs see a reduction in efficacy over time while others become intolerant [124,125]. When anti-TNF stops working it isn't as simple as targeting other pro-inflammatory cytokines. Among other adverse effects, anti-IFN-γ and anti-IL-17A antibodies can also unexpectedly aggravate CD [126-128]. This attests to the complexity of the immune responses that result in IBD. Some traditionally pro-inflammatory cytokines may, in fact, have multiple functions.

There is no question that treatments for patients with IBD have improved over the past few decades. New treatments such as anti-TNF have combined with immunosupressants for greater remission rates

and improved quality of life. However, 30-40% of patients with CD and 20-30% with UC will still require surgery for their condition at some point [129]. This attests to the need for new therapeutics, particularly those that are better targeted and safer in the long term.

1.5 Hygiene hypothesis

The 'hygiene hypothesis', first proposed in 1989, came about in an attempt to explain the drastic rise in autoimmune and allergic diseases over the past few decades. The theory was based on observations from Strachan who found differences in the rates of atopy in children from rural versus urban settings, and from smaller vs larger families [130]. From this he concluded that those who grow up in relatively sterile environments were more likely to develop immune disorders [131].

Much of the support for the hygiene hypothesis comes from studies comparing rates of allergic diseases among populations in urban vs rural environments. Multiple studies have found that children who grew up in a microbe-rich environment (such as a farm) had reduced risk of allergic diseases [132,133]. Among these children, the effect was magnified for those who were around livestock specifically. Further, protection was notably higher when that exposure occurred prenatally or within the first year of life [134].

When looking globally at autoimmune disorders and allergy, there tends to be an inverse relationship between these afflictions and rates of infectious diseases [135]. Over the past few decades the prevalence of asthma, Crohn's disease and type 1 diabetes have increased by more than 300% in developed settings while many infectious diseases have been completely ablated by vaccines and improved sanitation [135]. Given that the rise of immune diseases has occurred almost exclusively in developed or quickly developing countries, it raised the question of whether we were becoming too clean for our own good.

In support of this, diseases such as asthma and IBD are more common in people exposed to antibiotics early in life [136-138]. Similarly, as aforementioned, early life exposures to environments higher in microbes (i.e. farmland and larger families) are linked to lower rates of allergy [132]. However, there is conflicting evidence which now suggests the hygiene hypothesis may be too simplistic. While lower socioeconomic nations tend to have higher rates of infection and lower rates of allergy and autoimmunity, it does not mean all infections are beneficial. For instance, repeated respiratory tract infections are associated with higher rates of asthma later in life [139].

Although epidemiological data does not provide a mechanism, one explanation for these findings is that early life infections are important for the normal development of the immune system. A refinement of the hygiene hypothesis was offered by Rook in 2003, called the 'old friends' hypothesis [140]. Rook suggested that regular and early exposure to a wide range of harmless microorganisms he termed 'old friends' is necessary for training the immune system to respond appropriately [140]. The problem is not just about numbers of infections but rather the diversity, timing, and nature of exposures. The human immune system evolved in a time before antibiotics, sterile drinking water, effective waste removal and sanitation. It was therefore frequently exposed to numerous microbes and parasites which, with recent urbanization and industrialization, have drastically been reduced. It is likely that lack of exposure to these 'old friends' has removed necessary educational material for appropriate immune function [141]. In support of this, numerous studies have now shown that the microbiota is implicated in a range of health problems including diabetes, allergy, obesity and even mental health [142-144].

1.6 Helminths, modified immunity, and regulatory T cells

Despite being the largest organisms that infect humans, most species of helminths are capable of surviving for years without provoking an immune response that is sufficient to eliminate all of the parasites. While the human immune system is often capable of expelling a parasite, the scale of the

response required would likely be more detrimental to the host than tolerating light to moderate worm burdens. For this reason, regulatory mechanisms have evolved, either parasite or host-induced, which limit parasite expulsion while protecting the host from excessive immunopathology [145,146]. Indeed, it has been hypothesised that the Type 2 arm of the immune response evolved to combat helminth infections, as well as to limit and repair the damage they inflict during migration and feeding [147]. This immune response is characteristic of helminth infections and involves production of the cytokines IL-4, IL-5, IL-9, IL-13, IL-25 and IL-33 [148]. Through various signaling cascades, these cytokines coordinate Th2 cell differentiation, B-cell class-switching to IgE, and inflammatory cell recruitment, all in an effort to kill or expel the parasite [149].

The Type 2 immune response to helminths and the immune response seen in allergy share significant similarities. Like a helminth infection, atopic asthma involves a Type 2 response, but results in mucus production in the lung, bronchoconstriction, smooth-muscle contraction and airway remodeling [43]. Logically, helminth infections — notably those where larval worms navigate through the pulmonary vasculature en route to the gut — should therefore augment asthmatic responses due to the similar nature of the immune responses they elicit; however, the opposite scenario may be true in some instances. This is likely because helminth infections promote a modified Type 2 response and concurrent regulatory immune responses that induce a state of hypo-responsiveness that prevents the development of allergic disease [149].

Tregs are an integral part of the immune system, responsible for regulating pro-inflammatory immune responses via the production of IL-10, IL-35 and TGF-β [150]. Studies have shown that Tregs from asthmatics are dysfunctional, which contributes to the exaggerated immune response to allergens [151,152]. Helminths on the other hand, have been shown to increase the number and enhance the functional capacity of Tregs during infection. For example, in cross-sectional human studies, chronic infection with schistosomes and filarial nematodes were associated with higher levels of IL-10 and augmented Treg activity [153,154]. Similarly, lymphocytes from children in areas hyper-endemic for

the intestinal parasites *Trichuris trichiura* and *Ascaris lumbricoides* were found to constitutively secrete more IL- 10 and TGF-β [150]. The opposite effect has also been reported with anthelmintic treatment of humans, which resulted in a reduction in Treg cell numbers [155]. Evidence from animal models provides a similar picture. For example, frequencies and absolute numbers of Foxp3 + Tregs increase early in an infection with *Heligmosomoides polygyrus*, a rodent model of gastrointestinal human helminth infection, and this is associated with protection against lung inflammation [156-158]. This induction and persistence of Tregs is important for long-term parasite survival, as Treg depletion leads to a protective immune response and parasite expulsion [145]. There have also been a number of clinical trials utilising experimental helminth infections to treat immune diseases which will be discussed later.

1.7 Hookworms

Hookworms, along with whipworms and roundworms, make up three of the most pertinent and life-limiting parasitic infections worldwide [159]. The *Necator* and *Ancylostoma* species of human hookworm infect more than 400 million people in tropical regions of Asia, Africa and South America [146]. The pathology caused by hookworm infection is predominantly due to blood loss caused by feeding adult worms. This blood loss is tightly correlated with worm burden. Chronic infection with *N. americanus* results in fatigue, abdominal pain, diarrhoea, weight loss and anaemia [160]. In children, anaemia can cause growth retardation and impairments in cognitive development [161]. In pregnant women, these infections lead to poor birth outcomes including low birth weight, increased perinatal morbidity as well as increased mortality [162,163]. Moreover, hookworm infections result in 3.2 million disability-adjusted life years (DALYs) lost every year [164]. Due to these effects, and as a result of their ubiquitous nature, these infections contribute significantly to widespread poverty in the majority of the places they are found [165].

The life cycle of *N. americanus* is direct, with no intermediate hosts involved. Eggs are passed out of the body in the faeces and, under favourable conditions, hatch releasing rhabditiform larvae. Parasites develop to infective L3, non-feeding filariform larvae that can penetrate the skin to infect a human host [166]. Upon infection the parasites migrate through the circulatory system to the lungs, where they move up the trachea, are eventually swallowed, and ultimately reach the small intestine where they can live for up to 10 years [146,166]. It is here in the small intestine where blood loss begins to occur, predominantly as a result of leakage around the feeding site of the worm [146].

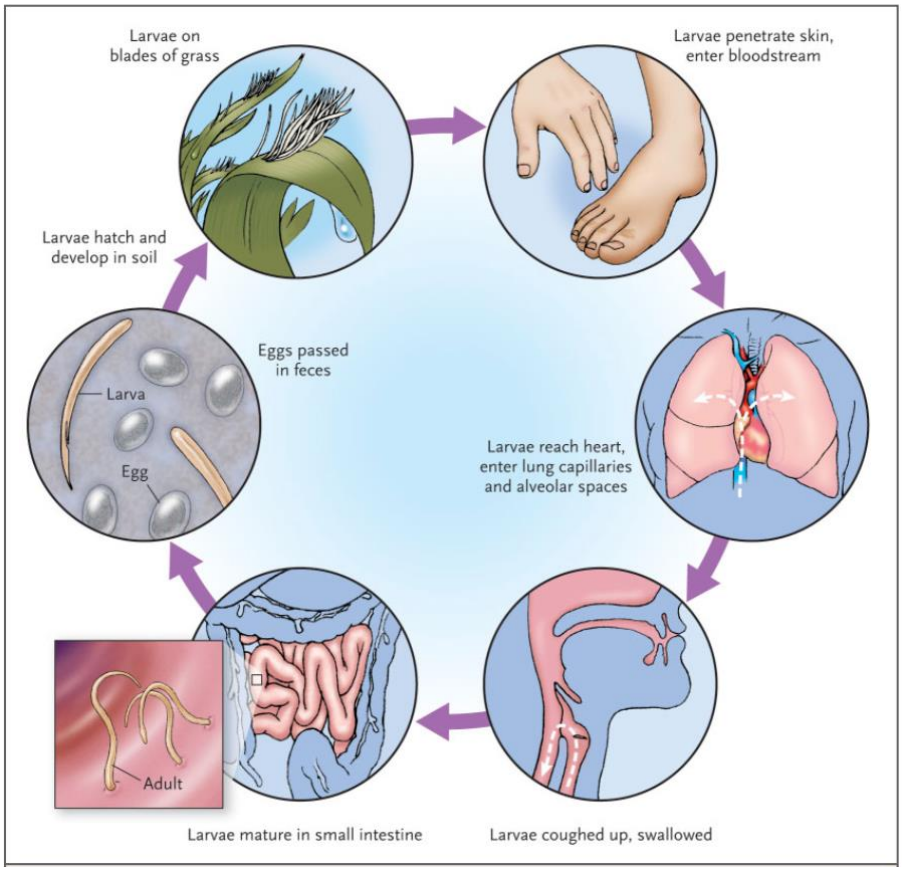


Figure 1.2: Hookworm life cycle. Image sourced from [160] with no modifications.

Diagnosis of hookworms are important to confirm infection for clinical purposes as well as mapping for transmission and targeted interventions. Examining faecal samples (molecular or microscopic) is key to diagnose infections. Given the above lifecycle, the faeces of infected persons include eggs, and often parts or whole larval stage parasites. For microscopic purposes, these must be processed fresh within 24 hours before the eggs completely hatch and migrate out of the stool [146]. Diagnostics are particularly important in mass drug administration (MDA) campaigns. In this context it shows the effectiveness of these programs while also dictating whether other measures will become necessary. Microscope-based methods, while cheaper and more readily available, are inadequate in their ability to detect very low level infections and consequently fail to inform MDA action appropriately [167].

Molecular based examinations are far more sensitive and are able to provide egg-per-gram measurements on frozen or fresh samples. While more costly and difficult to implement in the field, they are able to better differentiate hookworm species and obtain data on other parasitic infections simultaneously [168]. These tools are particularly useful for informing MDA approaches based on the risk of morbidity from infection. This kind of information is crucial for children and childbearing women who are most at risk of developing the iron deficiency anemia resultant from infections. The availability of 'omics' data, offers new tools for improving potential diagnostics as parasite antigens become increasing known [169].

1.7.1 Host-parasite interactions

Parasites have evolved alongside humans for millions of years. As a result, they have developed mechanisms to evade, subvert and ultimately survive their host's immune response. The ways by which parasites have achieved this immune evasion varies between parasites. One of the classic examples of parasite immune evasion is antigenic variation, which is employed by the unicellular parasites *Plasmodium falciparum* and *Trypanosoma brucei* [170,171]. Another example is seen in the

multicellular blood fluke *Schistosoma*, where host cell surface proteins are adsorbed onto the parasite's surface to make it undetectable by circulating immune cells [172-174].

In gastrointestinal helminth infections, long-term infections with limited pathology are common. This fact combined with the finding that infected populations exhibit less allergic diseases has led to studies attempting to explain how these infections limit inflammation and associated disease. The three most common human gastrointestinal helminths, *N. americanus, A. lumbricoides* and *T. trichiura*, are strongly linked with immune hyporesponsiveness measured as changes in circulating cytokines and regulatory cells. Specifically, the regulatory cytokines IL-10 and TGF-β1 have been reported to be elevated in patients in direct proportion to burdens of infection [150,175].

Despite the high prevalence of *N. americanus* infections globally, data on immune responses in human naturally acquired infections is limited. Nevertheless, experimental human infections with *N. americanus* and mouse studies with hookworm-like nematodes have been able to shed light on the immunopathological responses to these parasites. Initially, host immune responses are aimed at targeting invading parasites. These initial responses manifest as systemic and localized eosinophilia and mastocytosis [176,177]. The predominant response in experimental human infections is increased mucosal expression of Th1 cytokines such as IFN-γ, IL-2 and IL-15, and Th2 cytokines such as IL-4, IL-5 and IL-13 [177]. Even with these immune responses, adult hookworms are rarely ousted from the human gut. Additionally, despite this distinct immune response caused by hookworms, infected individuals do not show signs of allergy to the parasite [146]. Conversely, infected individuals have been shown to be afforded protection from developing allergies to other antigens [178].

Mouse immune responses to *Nippostrongylus brasiliensis* have been extensively explored because of its similarity to *N. americanus*. In this model there is significant evidence of host acquired immunity as well as clearance of the nematode from the lungs and gut. IgE-armed basophils have been shown to trap larvae in the skin, limiting further migration and subsequent development [179]. In the gut, *N.*

brasiliensis elicits Th2 cell responses in the form of IL-4 and IL-13 as well as goblet cell hyperplasia which is effective in causing worm expulsion [180]. In contrast with this model, there is little evidence of immunological clearance of *N. americanus* from humans despite similar robust Th2 responses. In experimental human infection, some immunity has been described with both polyclonal and parasite-specific IgE found to offer protection against hookworm infection [181]. Furthermore, a negative association between parasite-specific IL-5 concentrations and the likelihood of hookworm reinfection following treatment has been demonstrated [182]. While immunity to hookworm is not evident at a population level, this suggests an important role for eosinophils in the mediation of defence against these parasites [182].

1.7.2 Live infections and ES products

A number of clinical trials have explored using either *N. americanus* or *Trichuris suis* to treat inflammatory disease. These helminths, in some clinical trials at least, have been found to limit disease in patients with IBD. One such study found that *T. trichiura* led to UC remission and an increase of IL-22+ CD4+ T cells of a single self-infected volunteer [183]. Another study showed promising results using *T. suis* ova (TSO) in the treatment of people with UC [184]. However, when large randomised controlled trials were carried out to confirm these findings, TSO therapy showed no statistical improvements for patients with either CD or UC [185,186].

There has been mixed support for helminthic treatment for allergic rhinitis and asthma. Some studies have reported improvements in airway responsiveness and a reduction in the need for medication, while others found no improvement in any tested parameters [187,188]. More recently, experimental infection with *N. americanus* was found to restore gluten tolerance in people with coeliac disease. While the precise mechanism for this improvement is unclear, duodenal biopsy post-infection revealed intra-epithelial upregulation of CD3+ CD4+ Foxp3+ Tregs with increasing gluten exposure [189]. Upregulation of Tregs in combination with reductions in pro-inflammatory IL-17A- and IFN-γ-expressing

T cells provides a potential mechanism by which *N. americanus* parasites limit disease outcomes. Multiple sclerosis (MS) patients infected with gastrointestinal helminths were found to have fewer relapses, lower disability scores and reduced MRI activity compared with uninfected MS participants. Following anthelmintic treatment, absolute numbers of CD4+ CD25+ Foxp3+ Tregs markedly decreased as did the number of IL-10- and TGF-β-secreting cells, and MS outcomes worsened [190].

While these studies show that helminth infection shows promise for suppressing immunopathology in some inflammatory diseases at least, a stigma exists around infections with a live pathogen to treat disease. This has led to more studies examining the molecular interface between host and parasite, with particular emphasis on the parasite secretome and its immunomodulatory properties.

A number of studies have shown that crude excretory/secretory (ES) products from the dog hookworm, *Ancylostoma caninum*, protected mice against chemically induced colitis. In these studies adult stage hookworms were removed from their hosts and cultured to generate crude ES products which was subsequently injected into mice. ES-treated mice had decreased macroscopic inflammation scores, less inflammation and reduced myeloperoxidase activity [191-193] after administration of colitis-inducing chemicals. The ES products of *A. caninum* and *N. brasiliensis* have already been characterized, revealing 313 and 315 proteins respectively [194,195]. These proteomic analyses not only validated these parasites as suitable model species for studying the human hookworm *N. americanus* but also identified a suite of proteins of unknown function. These proteins of unknown function represent an array of potentially immunomodulatory molecules which could be used to treat inflammatory disease. Characterisation of the *N. americanus* ES products, most likely due to the difficulty in acquiring sufficient material, is yet to be reported. A large proportion of the human hookworm proteins remain unknown in terms of identification and function. It is likely the profile of the ES products will be similar to the closely related aforementioned species. Of particular relevance in these publications has been the overwhelming dominance of SCP/TAPS proteins.

SCP/TAPS have also been labelled SCP/Ag5/PR-1/Tpx-1/Sc7 (Pfam: PF00188) and are a part of the cysteine-rich secretory protein (CRISP) “superfamily” [196]. SCP/TAPS proteins share a common primary structure usually including a signal peptide followed by the SCP domain [197]. SCP/TAPS homologues have been found in numerous invertebrates, particularly roundworms, flatworms, arthropods and plants. Despite their prevalence, attempts to define specific role classification for these proteins has been limited. In parasitic nematodes, Ancylostoma-secreted proteins (ASPs) were initially characterized from hookworms and later from related strongylid nematodes [197]. ASPs were found to be highly abundant in the ES products of L3 larvae. This finding has led to the belief that they play an important role in infection and transition from free-living to the parasitic adult stage of development [198,199]. Due to their immunogenic properties, ASPs have been and should continue to be explored for their potential as vaccine candidates [200,201].

Defining SCP/TAPS has typically been based on the number and features of the SCP domains. Three classifications have broadly emerged namely: ‘double domain SCP/TAPS’ with two separate SCP-domains, and ‘N-type’ and ‘C-type’ single domain SCP/TAPS [197]. Double domain and C-type single domain SCP/TAPS have been the dominant identified proteins in a range of parasitic nematodes [197]. Given the importance and abundance of SCP/TAPS in a wide range of eukaryotes, greater classification and functional information is needed. In particular functional data may aid approaches into control of infectious diseases (e.g. vaccines) or therapies based on immune regulatory roles.

A number of publications have begun to elucidate the functions of specific hookworm proteins by producing recombinant versions. Recombinant *A. caninum* NIF (neutrophil inhibitory factor), a SCP/TAPS protein, inhibited Th2 cytokine production, eosinophil recruitment and infiltration, mucous secretion, and goblet cell hyperplasia in an OVA-induced mouse model of asthma [202]. In 2016, Navarro and colleagues showed that a single recombinant protein from the ES of *A. caninum*, termed AIP-2, suppressed airway inflammation in a mouse model of asthma [203]. AIP-2 was revealed to have multiple immunomodulatory effects including suppressing the expression of costimulatory markers on

human DCs and inducing the expansion of mouse mesenteric CD103+ DCs which resulted in generation of Tregs that localize to the mucosa [203]. More recently, an ES protein from *H. polygyrus* (coined *H. polygyrus* TGF-beta mimic (Hp-TGM) with no homology to any TGF-β family member was shown to mimic the biological and functional effects of TGF-β [204]. This included binding to mammalian TGF-β receptors, and induction of both mouse and human Treg activity [204,205]. Another ES protein, *H. polygyrus* Alarmin Release Inhibitor (HpARI), was found to abrogate IL-33, eosinophilia, and group 2 innate lymphoid cell responses to allergen *in vivo [206]*. It also limited eosinophilic responses to *N. brasiliensis*, augmenting parasite burden. HpARI was shown to bind directly to human IL-33, inhibiting release and thereby preventing type 2 responses [206]. These findings in particular highlight the potential for specific hookworm ES products to be harnessed as anti-inflammatory therapeutics.

1.7.3 Hookworm ES products as vaccine candidates

There is no natural protective immune response to hookworm infections in the human host [207]. Instead, infections can persist for many years due to the immunoregulatory capacity of these parasites. The plausibility of a human hookworm vaccine is supported by the canine hookworm vaccine which utilised radiation-attenuated infective larvae [208]. This was first marketed in the 1970s and while highly effective, was discontinued due to cost of production and a lack of scaling feasibility [208]. Given the enormous burden of disease the human hookworm poses, a vaccine would need to utilize a low cost expression system and show scaling potential [207]. The ES products of hookworm L3 larvae and adults have been explored for antigens to meet these necessities. Indeed, recombinant antigens have been identified from both stages of the parasite which have been used in animal models and show protection against challenge infections [207]. Development of *Na*-ASP-2, the leading L3-stage antigen against *N. americanus*, was discontinued following generalised urticaria in chronically infected subjects post vaccination [209]. Following this outcome, efforts have shifted towards targeting the blood-feeding stage of the adult hookworm.

Using gut-expressed antigens as successful vaccine targets has been previously described for the ruminant parasite *Haemonchus contortus* [210]. In the human hookworm, two adult stage antigens that are currently being developed are the aspartic protease haemoglobinase (*Na*-APR-1) and the glutathione S-transferase (*Na*-GST-1) [211]. Both of these have shown efficacy in laboratory canines with *Na*-APR-1 having significant homology to APR-1 from *A. caninum* [211]. *Na*-GST-1 has since undergone a safety study in naïve and infected populations in Brazil with no adverse responses and significant antigen-specific IgG responses [212]. However, there are still a number of clinical trials and efficacy studies before the vaccine can be widely distributed.

If the current leading antigens do not result in the desired protection against hookworms in their target population, other candidates will become necessary. For these means, proteomics analyses of ES products such as those carried out in *N. brasiliensis* and *H. polygyrus* (mentioned above) become highly valuable. ES products from the adult stage of the human hookworm, as will be presented in this thesis, hold a plethora of potential vaccine candidates which play important roles for parasite survival. Targeting one or many of these in future studies could provide additional efficacy for a human hookworm vaccine.

1.8 Hypothesis underpinning this thesis

The *N. americanus* hookworm has co-evolved with humans over thousands of years, specifically adapting to regulate host immunity to ensure its longevity. There is a large body of evidence to suggest that the ES proteins at the parasite-host interface represent the primary means by which the parasite does this. The secretomes of a number of parasitic helminths have been characterized, with some individual molecules found to possess immunomodulatory capacity. I hypothesise that some of the human hookworm ES proteins, through similar capacity, will be able to improve outcomes in inflammatory disease models. This thesis aims to define the human hookworm secretome, produce

select proteins in recombinant forms, and subsequently explore their potential for regulating host immune responses.

2 Chapter 2 – Proteomic analysis of ES products

2.1 Introduction

Despite their potential biotechnological utility, only a limited number of *N. americanus* ES proteins have been described to date, and most of them have been identified as cDNAs based on their homology to proteins from more readily accessible hookworm species that parasitise animals, such as *Ancylostoma caninum* [213]. In terms of vaccine antigens, a handful of *N. americanus* ES products including glutathione-S-transferases, aspartic proteases and sperm-coating proteins/Tpx-1/Ag5/PR-1/Sc7 (SCP/TAPS), have been identified at the cDNA level, and vaccine efficacy of recombinant proteins assessed in animal models and phase 1 clinical trials [212,214,215]. SCP/TAPS, also referred to as venom allergen-like (VAL) or Activation-associated Secreted Proteins (ASPs) (Pfam accession number no. PF00188) have been reported from many helminths, but appear to be significantly expanded in the genomes and secreted proteomes of gut-resident clade IV and V parasitic nematodes, including the hookworms [169,194,195,216]

The relative paucity of functional information on *N. americanus* proteins can be attributed, at least in part, to the difficulty in obtaining parasite material and the absence until 2014 of the published *N. americanus* draft genome. Analysis of the 244 Mb draft genome and the predicted 19,151 genes [169] provided important information about the molecular mechanisms and pathways by which *N. americanus* interacts with its human host. In agreement with published transcriptomes of *N. americanus* and genomes/transcriptomes of other hookworm species, a select number of protein families were over-represented, including SCP/TAPS proteins and different mechanistic classes of proteases with various functions including haemoglobinolysis and tissue penetration [217-222]. Of the >19,000 predicted genes reported in the draft *N. americanus* genome, 8,176 genes had no known

InterPro domain. Additionally, more than half of the total proteins had either no blast homology to any gene from the NCBI database (10,771 proteins) or shared identity with a 'hypothetical' protein (3,043 proteins). These results highlight the importance of further annotation and refinement of the *N. americanus* genome [223].

In general, the practicality of genomic sequence data is dependent on the accuracy of gene annotation as well as the availability of functional, expression and localization information [224]. The high-throughput methods used when annotating a genome are prone to errors, therefore, to validate the predicted protein-coding genes, an analysis of the proteome is essential. Mass spectrometry provides useful data that can be used in a proteogenomic approach to improve genome annotation and identify novel peptides containing predicted protein sequences [225]. Methods including multiple fractionation techniques (in-gel and in-solution combination methods) improve the overall sensitivity and accuracy of our profiling [226].

In this chapter I perform the first proteogenomic analysis of a parasitic helminth, while also significantly improving the genome annotation and comprehensively characterizing the ES proteome of adult *N. americanus*. The findings provide valuable information on important families of proteins with both known and unknown functions that could be instrumental in host-parasite interactions, including protein families that might be key for parasite survival and protection of the host against excessive immunopathology. Characterization of these proteins will be useful for the identification of vaccine and drug targets and diagnostics for hookworm infection.

2.2 Methods

2.2.1 Ethics

Ethical approval for hamster animal experimentation to obtain *N. americanus* adult worms was obtained from the Federal University of Minas Gerais, Brazil (Protocol# 51/2013). Ethical approval for human experimental infection with *N. americanus* and subsequent culturing of L3 was obtained from the James Cook University Human Research Ethics Committee (ID# H5936).

2.2.2 Parasite material

Adult *N. americanus* were manually isolated from the intestines of experimentally infected golden hamsters (*Mesocricetus auratus*) upon euthanasia. For the isolation and purification of ES products, the worms were washed 3 times in phosphate-buffered saline (PBS) before being cultured overnight in a humidified incubator at 37°C, 5% CO_2, in RPMI 1640 (100 U/ml penicillin, 100 µg/ml streptomycin sulphate, 0.25 µg/ml amphotericin B). The supernatant was collected the following day and debris was removed by centrifuging the concentrated samples at 1,500 *g* for 3 minutes in a benchtop microfuge. ES products were concentrated with a 3 kDa cut-off Centricon filter membrane (Merck Millipore) and samples were stored at -80°C until use. For somatic adult extracts, data was used from Tang *et al* [169]. This data was the mass spec results for the adult *N. americanus* whole worm digest that was later used in section 2.2.5 for database searching. In brief, whole worms were ground under liquid nitrogen and solubilized using lysis buffer (1.0% (v/v) Triton X-100 in 40 mM Tris, 0.1% (w/v) SDS, pH 7.4). The extract was filtered through a 20 µm filter before fractionation.

For isolation of *N. americanus* L3, stool samples were collected from infected human volunteers and cultured as follows. dH_2O water was added to the stool sample until a thick paste was formed. This paste was then distributed onto moistened filter paper (VWR, Standard Grade, 110 mm) in Petri dishes (Sarstedt, 150 mm) and placed in a 25°C incubator for 8 days. Following incubation, the edges of each

plate were gently rinsed with RO water to obtain clean L3 preparations. Somatic extracts were prepared by adding 100 µl lysis buffer (3 M Urea, 0.2% SDS, 1% Triton X, 50 mM Tris-HCl) to approximately 6,000 larvae before repeated vortexing and sonication (4°C, probe sonicator, pulse setting) to digest the larvae. The extract was passed through a 0.45 µm filter (Millipore) before in-gel fractionation.

2.2.3 Fractionation and digest

SDS-PAGE fractionation and in-gel trypsin digestion

A total of 30 µg of *N. americanus* adult ES products was buffer exchanged into 50 mM NH_4HCO_3, freeze-dried and resuspended in Laemmli buffer (2% SDS, 10% glycerol, 0.002% bromophenol blue, 0.0625 M Tris-HCl, 5% dithiothreitol (DTT). The sample was boiled at 95°C for 5 minutes and electrophoresed on a 12% SDS-PAGE gel for 40 minutes at 150 V. The gel was stained with Coomassie Blue and 20 pieces (approximately 1 mm thick) were cut and placed into Eppendorf tubes. For the in-gel digestion, slices were de-stained and freeze-dried before incubating them at 65°C for 1 h in reduction buffer (20 mM DTT, Sigma, 50 mM NH_4HCO_3). Samples were then alkylated in 50 mM iodoacetamide (IAM, Sigma), 50 mM NH_4HCO_3 for 40 minutes at 37°C, washed three times with 25 mM NH_4HCO_3 and digested with 20 µg/ml of trypsin (Sigma) by incubating them for 16 h at 37°C with gentle agitation. Digestion was stopped and peptides were released from the slices by adding 0.1% TFA, 70% acetonitrile. This step was repeated 3 times with pooling of the corresponding supernatants to maximise peptide recovery for each sample. Finally, each sample was desalted with a ZipTip (Merck Millipore) and stored at -80°C until use.

In-solution trypsin digestion and off-gel fractionation

A total of 70 µg of each *N. americanus* extract (L3 somatic and adult ES products) was buffer exchanged into 50 mM NH_4HCO_3 before adding DTT to 20 mM and incubating for 10 minutes at 65°C. Alkylation was carried out by adding IAM to 55 mM and incubating for 45 minutes at room temperature in the

dark. Samples were digested with 2 µg of trypsin by incubating for 16 h at 37°C with gentle agitation. Following trypsin digestion, peptides were fractionated using a 3100 OFFGEL Fractionator (Agilent Technologies) according to the manufacturer's protocol with a 24-well format as described previously [195]. In brief, Immobiline DryStrip pH 3-10 (24 cm) gel strips were rehydrated in rehydration buffer in the assembled loading frame. Digested peptides were diluted in dilution buffer to a volume of 3.6 ml and loaded equally across the 24 well cassette. The sample was run at a current of 50 µA until 50 kilovolt hours (kVh) had elapsed. Upon completion of the fractionation, samples were collected, desalted with ZipTip and stored at -80°C until use.

2.2.4 Mass spectrometry

Both L3 and adult extracts from both fractionation methods were analyzed by LC-MS/MS on a Shimadzu Prominence Nano HPLC (Japan) coupled to a Triple Tof 5600+ mass spectrometer (SCIEX, Canada) equipped with a nano electrospray ion source. Fifteen (15) µl of each extract was injected onto a 50 mm x 300 µm C18 trap column (Agilent Technologies, Australia) at 60 µl/min. The samples were de-salted on the trap column for 6 minutes using 0.1% formic acid (aq) at 60 µl/min. The trap column was then placed in-line with the analytical nano HPLC column, a 150 mm x 100 µm 300SBC18, 3.5 µm (Agilent Technologies, Australia) for mass spectrometry analysis. For peptide elution and analysis the nano-HPLC pump was initially held at 2% solvent B for 6 minutes followed by a linear gradient of 2-40% solvent B over 80 minutes at 500 nl/minute flow rate and then a steeper gradient from 40% to 80% solvent B in 10 minutes was applied. Solvent B was held at 80% for 5 minutes for washing the column and returned to 2% solvent B for equilibration prior to the next sample injection. Solvent A consisted of 0.1% formic acid (aq) and solvent B contained 90/10 acetonitrile/ 0.1% formic acid (aq). The ionspray voltage was set to 2,200 V, declustering potential (DP) 100V, curtain gas flow 25, nebulizer gas 1 (GS1) 12 and interface heater at 150°C. The mass spectrometer acquired 250 ms full scan TOF-MS data followed by 20 by 250 ms full scan product ion data in an Information Dependent Acquisition (IDA)

mode. Full scan TOF-MS data was acquired over the mass range 300-1600 and for product ion ms/ms 80-1600. Ions observed in the TOF-MS scan exceeding a threshold of 150 counts and a charge state of +2 to +5 were set to trigger the acquisition of product ion, MS/MS spectra of the resultant 20 most intense ions. The data was acquired using Analyst TF 1.6.1 (SCIEX, Canada).

2.2.5 Proteogenomics

The mass spectrometry raw data was searched against the *N. americanus* protein database [169] using SequestHT algorithm in Proteome Discoverer 2.1 (Thermo Scientific, Bremen, Germany). Trypsin was used as the protease, allowing a maximum of two missed cleavages. Carbamidomethylation of cysteine was specified as a fixed modification, and oxidation of methionine was included as variable modifications. The minimum peptide length was specified as 6 amino acids. The mass error of parent ions was set to 10 ppm, and for fragment ions it was set to 0.05 Da. Protein inference was based on the rule of parsimony and required one or more unique peptides. False Discovery Rate (FDR) of 1% at peptide and protein levels was applied. The unassigned spectra from the protein database search were further searched against the six-frame translated genome database of *N. americanus* [169]. The genome sequences were downloaded from WormBase ParaSite release 9 (http://parasite.wormbase.org/Necator_americanus_prjna72135/Info/Index) in FASTA format and translated in all six frames using in-house python scripts. All the sequences greater than 10 amino acids between any two stop codons were added to the database. The search was again performed using SequestHT with precursor mass tolerance of 50 ppm and fragment ion tolerance of 1 Da. Carbamidomethylation of cysteine was specified as a fixed modification, and oxidation of methionine was included as variable modifications. Results were obtained at 1% FDR for both protein and peptide level.

The identified peptides in the search against the six-frame translated genome were mapped back to the protein database using standalone BLAST. Any sequences that mapped 100% to the protein

database were discarded. The filtered peptides were mapped to the *N. americanus* genome using the standalone tblastn program [169]. Peptide identifications that unambiguously mapped to a single region in the genome, also known as Genome Search Specific Peptides (GSSPs), were considered to perform proteogenomics-based annotation of novel coding regions. The known gene annotation GFF3 was downloaded from WormBase ParaSite for *N. americanus* [169]. Using in-house scripts, GSSPs were characterized into various categories: intergenic, intronic, exon-extension, alternative frame, N-terminal extensions or repeat regions with respect to the known regions from GFF3. Additionally, for all the intergenic peptides, a search was undertaken for potential open reading frames (ORFs) within the stretch of amino acids. Each of the spectra was further manually validated.

2.2.6 Genome annotation

The genome annotation from *N. americanus* [169] was updated using the MAKER pipeline v2.31.8 [227] in collaboration with Dr Makedonka Mitreva and colleagues at Washington Univertsity of St Louis. The genome assembly (GenBank assembly accession: GCA_000507365.1) was softmasked for repetitive elements with RepeatMasker v4.0.6 using a species-specific repeat library created by RepeatModeler v1.0.8, RepBase repeat libraries [228] and a list of known transposable elements provided by MAKER [227]. From the NCBI Sequence Read Archive (SRA), *N. americanus* RNA-Seq data [169] (Adult: SRR609895, SRR831085, SRR892200; L3: SRR609894, SRR831091, SRR89220) were obtained. After adapter and quality trimming using Trimmomatic v0.36 [229], RNA-Seq reads were aligned to the genome using HISAT2 v2.0.5 [230] with the --dta option and subsequently assembled using StringTie v1.2.4 [231]. The resulting alignment information and transcript assembly were used by BRAKER [232] and MAKER pipelines, respectively, as extrinsic evidence data. In addition, protein sequences from SwissProt UniRef100 [233] and WormBase ParaSite WS258 [234] (*Ancylostoma ceylanicum* PRJNA231479, *Brugia malayi* PRJNA10729, *Caenorhabditis elegans* PRJNA13758, *Onchocerca volvulus* PRJEB513, *Pristionchus pacificus* PRJNA12644, *Trichinella spiralis* PRJNA12603 and *Strongyloides ratti*

PRJEB125) were provided to MAKER as protein homology evidence. Following the developer's recommendation [235], the protein-coding gene models of Tang *et al.* [169] were passed to MAKER as pred_gff to update the models by adding new 3' and 5' exons, additional UTRs, and merging split models. This method, however, cannot change internal exons nor create new annotations where evidence suggests a gene but no corresponding model is previously present. To address this shortcoming, additional *ab initio* gene predictions were generated using BRAKER v2.0.1 [232] and passed to MAKER so that the intron-exon model that best matched the evidence could be included in the final annotation set. Within the BRAKER pipeline, the gene prediction tools GeneMark [236] and AUGUSTUS [237] were trained utilizing the *N. americanus* RNA-Seq alignment and protein homology information from *C. elegans*. The GFFPs identified in the present study were used to confirm the validity of proposed annotation changes and resolve competing gene predictions through manual curation. Gene models with no evidence support were not included in the final annotation build to reduce false positives in the existing annotations. However, *ab initio* gene predictions that encoded Pfam domains as detected by InterProScan v5.19 [238] were rescued to enhance overall accuracy by balancing sensitivity and specificity [227,239]. The completeness of the annotated gene set was assessed using BUSCO v3.0 with Eukaryota-specific single copy orthologs (OrthoDB v9) [240].

2.2.7 Proteomics

Protein Identification

Mascot version 2.5 (Matrix Science), X!Tandem (The Global Proteome Machine Organisation) version Jackhammer and Comet v2014.02 rev.2 were used to analyze data from the mass spectrometer. Searches were carried out against a database comprised of either the updated genome annotation provided with this study, or the *N. americanus* genome [169], both appended to the common repository of adventitious proteins (cRAP; http://www.thegpm.org/crap/) database (to detect potential contamination including host material). The following parameters were used: enzyme,

trypsin; variable modifications, oxidation of methionine, carbamidomethylation of cysteine, deamidation of asparagine and glutamine; maximum missed cleavages, 2; precursor ion mass tolerance, 50 ppm; fragment ion tolerance 0.1 Da; charge states, 2+, 3+ 4+. A FDR of 0.1% was applied, and a filter of greater than 2 significant unique sequences was used to further improve the robustness of data. The mass spectrometry data have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD010669.

Bioinformatic Analysis of Proteomic Sequence Data

Gene ontology (GO) annotations were assigned using the program Blast2GO and Pfam analysis was performed using HMMER [241]. Pfam domains were detected at the $p<0.001$ threshold for the HMMER software. Putative signal peptides were predicted with SignalP v4.1, transmembrane domains with TMHMM v2.0, and non-classical protein secretion with the SecretomeP 2.0 server [242-244]. REViGO, an online tool, was used to summarise and plot GO terms [245]. UpSetR was used to group proteins based on whether they had a GO term with one, two or all of the GO categories (biological process, molecular function or cellular process) [246].

2.2.8 Comparing similar species

Similarity analysis

A similarity analysis was carried out based on the Parkinson and Blaxter method using an in-house script [221,247]. Given the difficulty of working with *N. americanus* specifically (in terms of accessibility to samples and establishment of life cycle in other hosts), other hookworms are frequently used to model this human parasite. Data from the secreted proteome of adult *N. americanus* (described herein) was compared to the published secreted proteomes from adult stages of related nematode species including *H. polygyrus*, *A. caninum* and *N. brasiliensis* [195,215,248].

Protein family similarity visualization

H. polygyrus, *A. caninum* and *N. brasiliensis* SCP/TAPS and protease protein sequences were obtained from their respective published secreted proteomes [194,195,248]. These sequences were aligned with adult *N. americanus* homologous protein sequences using BLAST. Significant sequence alignments were visualized using Circos [249].

Phylogenetic Analyses

SCP/TAPS proteins were identified from published secretomes of 6 species of parasitic helminths including *A. caninum, Ascaris suum, H. polygyrus, N. brasiliensis, Trichuris muris* and *Toxocara canis.* SCP/TAPS proteins were identified from *N. americanus* ES products using the proteome data generated in this study. Sequences were sorted into single and double SCP/TAPS domain proteins for individual phylogenetic analysis due to distinct differences described previously [197]. A list of the proteins and their respective sequences used for this analysis can be found in Supplemental Data 1. A multiple sequence alignment was carried out using the alignment program MUSCLE. Outliers with poor alignment (long unaligned regions) were detected and filtered out using ODseek. PhyML, a phylogeny software, was used for a maximum-likelihood (ML) phylogenetic analyses of SCP/TAPS amino acid sequences. The tree was visualized with The Interactive Tree of Life (iTOF) online phylogeny tool (https://itol.embl.de/) [250].

2.3 Results

The supplementary data for this chapter contains large datasets which are inappropriate for inclusion in this thesis. This has been uploaded to bioRxiv along with the rest of the data from this chapter and is available at: https://www.biorxiv.org/content/10.1101/406843v1.supplementary-material

2.3.1 Proteogenomic analysis and genome annotation

Different proteomes from two life stages (infective L3 somatic extract and adult somatic extracts and ES products) of *N. americanus* were analysed by mass spectrometry in a 5600+ ABSciex Triple Tof instrument to perform a proteogenomic analysis of the hookworm genome. After excluding peptides that mapped accurately to the existing protein database entry, a total of 218 novel peptides were identified that did not match any protein sequences of *N. americanus* contained in the annotations of Tang *et al* (11). Of these 218 novel peptides, 83 were found exclusively in adult somatic extracts, 50 exclusively in L3, and 67 exclusively in adult ES, while 10 peptides were found in both the adult somatic extract and ES products, and 8 were found in both adult and larval somatic extracts. No common peptides were identified from adult ES products and larval somatic extract.

Newly identified peptides could be grouped into the following six categories: (1) peptides mapping to intergenic regions; (2) peptides mapping to introns; (3) peptides mapping to alternative reading frames; (4) peptides extending gene boundaries (exon extensions); (5) peptides mapping to N-terminal extensions; and (6) peptides mapping to repeat regions (not gene) regions (Figure 2.1). Of the identified peptides, more than half belonged to group 1, 18% to group 2, 5% to group 3, 15% to group 4, 2% to group 5, and 6% to group 6 (Figure 2.1).

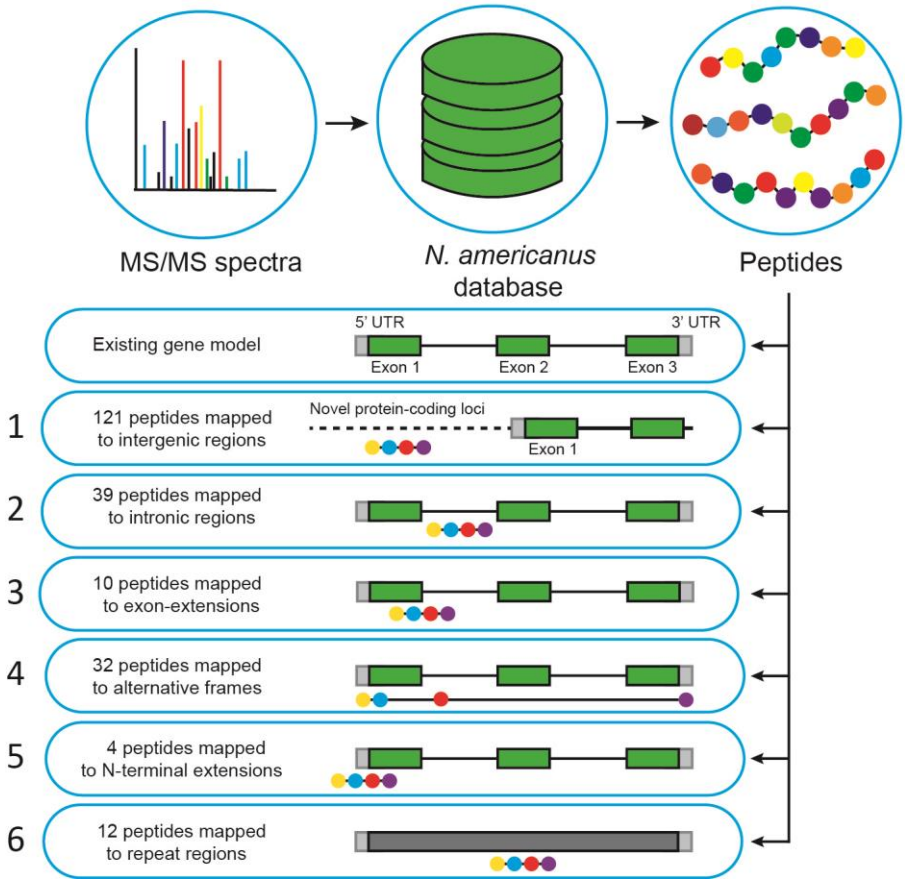


Figure 2.1: Depiction of the proteogenomic process as well as the types and numbers of peptide corrections identified.

Group 1 peptides were further analyzed for an ORF including whether or not they contained a methionine; see Supplemental data 2. From this data, one-third of group 1 peptides were found to overlap with a viable ORF, with the shortest having just 21 amino acids and the longest 381 amino acids. The translated protein sequences with their respective peptides, six-frame translated genome (SFG) ID and ORF length are provided in Supplemental data 3.

These results highlighted the need for improving the previously published gene models, so we updated the genome annotation of *N. americanus* using a new RNA-Seq based gene calling pipeline as outlined in the Materials and Methods (Table 2.1). The total number of predicted *N. americanus* genes decreased by 3,425, with substantial increases in both the number of exons and introns. The total

coding sequence (CDS) length increased by 2.22 Mb with the mean gene length (including introns and UTRs) nearly doubling from 4.3 kb to 8.1 kb. Furthermore, the percentage of the genome covered by genes increased by 18.3%, and the percentage of detected BUSCOs among the predicted genes increased from 95.7% to 97.4% (with 31% reduction in the number of fragmented BUSCOs). While these improvements are attributable in most part to the use of more sophisticated genome annotation methods utilizing RNA-Seq data and the inclusion of more extensive, up-to-date homologous protein databases, the peptide sequences generated in this study contributed directly to the refinement of 14 gene models. The newly annotated, improved gene models were used in subsequent proteomic analysis of ES products, and are publicly available on Nematode.net (56, 57).

	Original annotation	Updated annotation
Number of genes	19,153	15,728
Number of exons	122,849	148,780
Number of introns	103,696	133,052
Number of CDS	19,153	15,728
Overlapping genes	395	2,424
Contained genes	2	386
Total gene length (bp)	82,090,364	126,651,725
Total exon length (bp)	15,470,227	24,553,753
Total intron length (bp)	66,827,529	102,364,076
Total CDS length (bp)	15,461,420	17,683,227
Mean gene length (bp)	4,286	8,053
Mean exon length (bp)	126	165
Mean intron length (bp)	644	769
Mean CDS length (bp)	807	1,124
% of genome covered by genes	33.6	51.9
% of genome covered by CDS	6.3	7.2
Mean exons per mRNA	6.4	9.5
Mean introns per mRNA	5.4	8.5
Complete BUSCOs	82.84%	88.45%
Fragmented BUSCOs	12.87%	8.91%
Missing BUSCOs	4.29%	2.64%

Table 2.1: Summary of the updated genome annotation of human hookworm. CDS – coding DNA sequence; bp – base pairs; BUSCOs - Benchmarking Universal Single-Copy Orthologs.

2.3.2 Analysis of the ES products from *N. americanus* adult worms

A comprehensive analysis of the ES products from adult worms was carried out using in-gel and off-gel fractionation and the tryptic peptides were analyzed using LC-MS/MS. Mascot, X!Tandem and Comet searches were carried out against a database including the predicted proteins from the annotated *N. americanus* genome available in this study appended to the cRAP sequences available at http://www.thegpm.org/crap/. A total of 186 and 141 proteins were identified using Mascot and X!Tandem/Comet, respectively. All ES proteins were identified with at least two unique peptides, at a 99.0% probability and FRD of 0.1%. The two search methods combined found a total of 198 proteins (Supplemental data 4). These 198 proteins were obtained using the updated genome annotation generated as part of this study. In comparison, using the first version of the annotated genome sequence we identified 203 proteins using the same analytical software. Using the exponential modified protein abundance index (emPAI) and the newly annotated genome, the most abundant proteins in the ES products of *N. americanus* adult worms were ranked, and the top 30 are shown in Table 2.2. One emPAI value was listed for each of the digestion methods used (in-gel and off-gel) and their combined value was used to rank the proteins.

Accession	Blast2GO Description	MASCOT score		iProphet probabilty		NumP		SP	TD	emPAI		Pfam domain
		In-gel	Off-gel	In-gel	Off-gel	In-gel	Off-gel			In-gel	Off-gel	
NAME_05596	Hypothetical protein	19496	15287	1	1	11	11	N	0	867.19	308.05	No conserved domains
NAME_07794	Hypothetical protein	21210	39596	1	1	18	22	Y	0	47.88	199.86	Single SCP
NAME_11917	SCP	35077	43195	1	1	18	31	N	0	11.87	68.4	Double SCP
NAME_13724	Hypothetical protein	7701	23099	1	1	5	7	N	0	7.06	64.55	No conserved domains
NAME_15197	SCP	9446	12866	1	1	6	7	N	0	9.55	50.11	No conserved domains
NAME_14329	SCP	10911	15142	1	1	10	15	N	0	14.34	43.25	Single SCP
NAME_11146	SCP	21158	29296	1	1	24	27	Y	0	17.03	22.85	Double SCP
NAME_10941	Hypothetical protein	983	3870	1	1	3	8	Y	0	2.6	33.23	Single SCP
NAME_09098	SCP	6908	47379	1	1	13	25	Y	0	3.27	31.87	Double SCP
NAME_05595	SCP-like protein, partial	0	23369	0	1	0	10	Y	0	0	26.61	Single SCP
NAME_11145	SCP	19923	25558	1	1	18	22	N	0	6.35	16.62	Double SCP
NAME_02907	Glu Leu Phe Val dimerization	6645	6095	1	1	20	22	N	0	6.82	7.63	Single ELFV_dehydrog_N
NAME_10752	Nematode fatty acid retinoid	12416	3051	1	1	11	8	Y	0	9.76	4.65	Single Gp-FAR-1
NAME_05865	Globin	3566	5667	1	1	6	10	N	0	2.63	9.67	Single Globin
NAME_15231	SCP	117	1844	0	1	3	6	N	0	1.75	9.66	Single SCP
NAME_05592	Hypothetical protein	0	4150	0	1	0	4	Y	0	0	11.02	Single SCP
NAME_13132	Ferritin	1808	2833	1	1	7	8	N	0	3.39	7.31	Single Ferritin-like
NAME_05794	Hypothetical protein	2681	2271	1	1	7	9	N	0	5.18	5.21	No conserved domains
NAME_13809	SCP	8287	14385	1	1	7	9	Y	0	3.99	5.92	Single SCP
NAME_07942	Copper zinc superoxide dismutase	4241	4469	1	1	4	8	N	0	1.41	6.25	Single Copper/zinc superoxide
NAME_09181	Hypothetical protein	27774	68153	1	1	41	50	N	0	3.17	4.43	No conserved domains
NAME_10294	SCP	6557	11684	1	1	7	9	Y	0	2.58	4.8	Single SCP
NAME_02523	Histidine operon leader	23085	24957	1	1	66	63	Y	0	3.68	3.3	Many NPA
NAME_13379	Flavodoxin	0	1089	0.9985	0	0	4	N	0	0	6.89	No conserved domains

NAME_02090	SCP	3916	12638	1	1	8	14	Y	0	1.35	5.47	Double SCP
NAME_01069	SCP	3816	11201	1	1	10	12	Y	0	1.91	4.63	Double SCP
NAME_12806	SCP	8094	8066	1	1	6	8	Y	0	1.62	4.42	Single SCP
NAME_02548	Globin	887	1473	1	1	4	6	N	0	3.01	3.03	Single Globin
NAME_05396	SCP	1339	5930	1	1	4	8	Y	0	1.48	4.13	Single SCP
NAME_13850	SCP	3510	4006	1	0.9933	7	8	Y	0	1.97	3.07	Single SCP

Table 2.2: Top 30 most abundant proteins identified with in-gel and OFF-GEL fractionation and ranked using summed exponential modified protein abundance index (emPAI) score. Blast2GO was used to obtain descriptions of each protein. Abbreviations key: NumP – number of significant peptides; SP – signal peptide; TD – transmembrane domain; Mascot scores >28 indicate identity or extensive homology ($p < 0.05$).

The conserved Pfam domains of the 198 ES proteins identified were analysed. The most abundant protein family in the ES products was the SCP/TAPS family with 54/198 proteins containing a single or double cysteine-rich secretory protein family (CAP) domain (PF00188) (Figure 2.2B). The top 10 most abundant protein families are displayed in Figure 2.2B. Many of these proteins were described by Blast2GO as 'SCP-partial' or 'SCP-like', but for a more standardized annotation in my analysis they have been grouped and referred to as 'SCP/TAPS'. The second most frequently represented group (with 42 proteins) were proteins with one or more domains of unknown function (DUF). Other abundant families included Ancylostoma secreted protein related (ASPR) proteins and metalloproteases with 35 and 9 of each identified, respectively (Supplemental data 5). Despite some published reports classifying ASPRs as SCP/TAPS proteins, they are a diverse set of secreted cysteine rich proteins based on Pfam annotation and therefore have been grouped separately (28). Of the 198 identified ES proteins, 51% contained a predicted signal peptide. Supporting the accuracy of the new gene model, 96 of the identified adult ES proteins were predicted to contain a signal peptide compared to just 75 using the previous genome annotation. Of the remaining 49% which did not have a signal peptide, 45% were predicted to be non-classically secreted by SecretomeP.

The adult *N. americanus* ES proteins were annotated using Blast2GO (46). In total, 30 GO terms were identified (following the removal of parent child redundancy) belonging to one of the three GO database categories: biological processes, molecular function or cellular component (Figure 2.2A, for raw data see Supplemental data 6). Blast2GO returned biological process GO terms for 87/198 proteins, molecular function GO terms for 99/198 proteins and cellular component GO terms for 83/198 proteins (Figure 2.2A). The most prominent biological process term was proteolysis (Figure 2.2D), with 15% (29 proteins) of total ES products being involved in a proteolytic process. The most prominent single molecular function term was "hydrolase activity" followed by "peptidase activity" and "metal ion binding" (Figure 2.2C). Interestingly, 50/198 proteins did not return any GO terms (Figure 2.2A). Of these, SCP/TAPS proteins made up 64% with no known biological process, molecular function, or

cellular component. This highlights a significant knowledge gap surrounding SCP/TAPS produced by helminths.

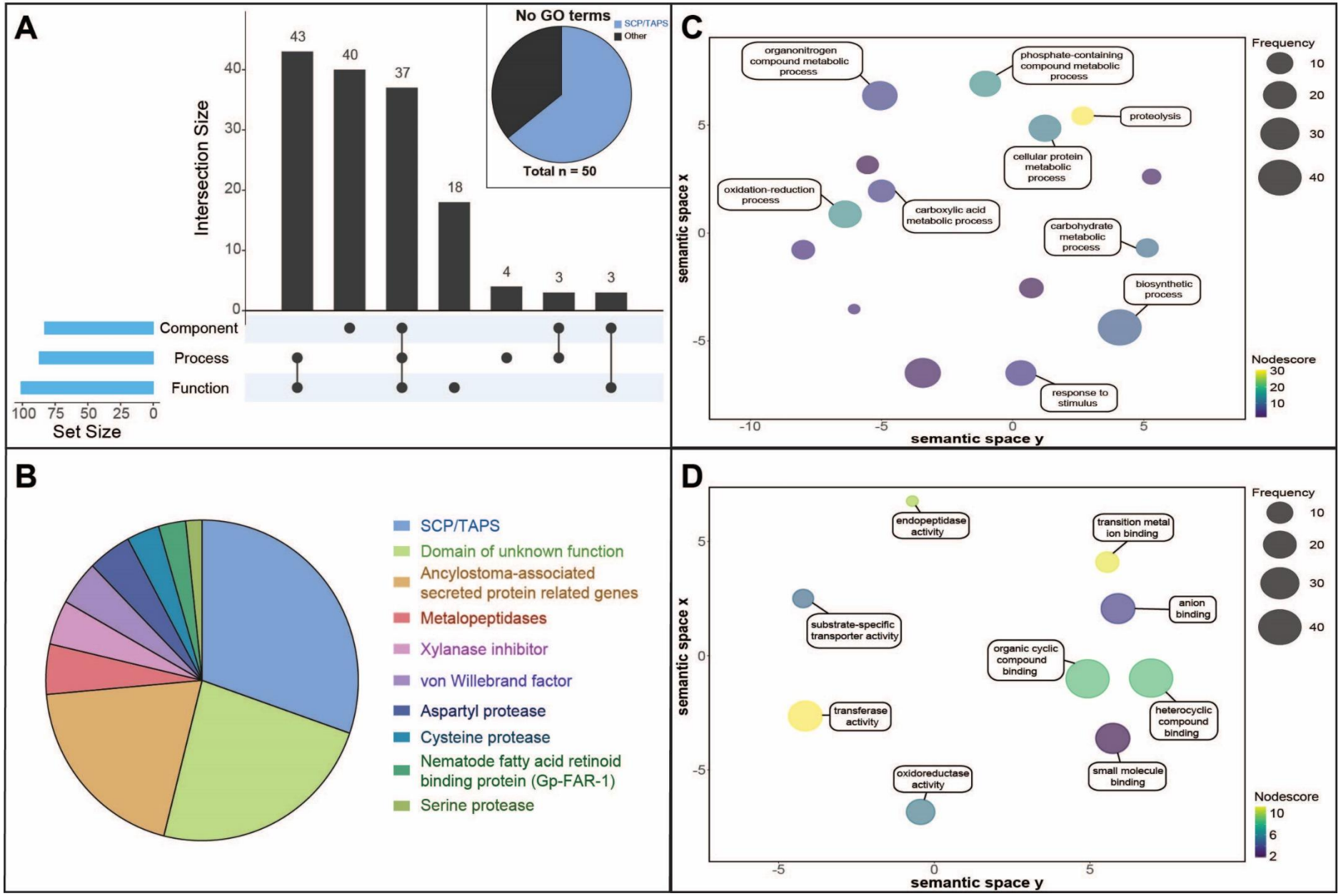

A
No GO terms
SCP/TAPS
Other
Total n = 50
Intersection Size
43
40
37
18
4
3
3
Component
Process
Function
100 75 50 25 0
Set Size
B
SCP/TAPS
Domain of unknown function
Ancylostoma-associated secreted protein related genes
Metalopeptidases
Xylanase inhibitor
von Willebrand factor
Aspartyl protease
Cysteine protease
Nematode fatty acid retinoid binding protein (Gp-FAR-1)
Serine protease
C
organonitrogen compound metabolic process
phosphate-containing compound metabolic process
proteolysis
cellular protein metabolic process
oxidation-reduction process
carboxylic acid metabolic process
carbohydrate metabolic process
biosynthetic process
response to stimulus
semantic space x
semantic space y
Frequency
10
20
30
40
Nodescore
30
20
10
D
endopeptidase activity
transition metal ion binding
anion binding
substrate-specific transporter activity
organic cyclic compound binding
heterocyclic compound binding
transferase activity
oxidoreductase activity
small molecule binding
semantic space x
semantic space y
Frequency
10
20
30
40
Nodescore
10
6
2

Figure 2.2: Gene ontology (GO) and protein families of adult *N. americanus* ES proteins. (A) UpSetR plot displaying the number of each gene ontology (GO) term categories (cellular component, biological process, or molecular function) available (Blast2GO) for adult *N. americanus* ES proteins. Proteins with no available GO terms are broken down in a pie chart into 'SCP/TAPS' proteins 'other'. **(B)** Top 10 most abundant protein families in the ES products of adult *N. americanus*. **(C)** Biological processes of adult *N. americanus* ES proteins ranked by nodescore (Blast2GO) and plotted using REViGO. Semantically similar GO terms plot close together, increasing heatmap score signifies increasing nodescore from Blast2GO, while circle size denotes the frequency of the GO term from the underlying database. **(D)** Molecular functions of adult *N. americanus* ES proteins ranked by nodescore (Blast2GO) plotted using REViGO.

2.3.3 Similarity analysis of the ES products from different gastrointestinal nematode species

A similarity analysis of ES proteomic data from *N. americanus* and three of the most commonly used animal models for human hookworms, *A. caninum*, *H. polygyrus* and *N. brasiliensis* was carried out (Figure 2.3). A total of 15, 10 and one *N. americanus* ES proteins had unique homology to ES proteins from *A. caninum*, *H. polygyrus* and *N. brasiliensis*, respectively. This included one SCP/TAPS protein (NAME_13724) which was similar only to an *H. polygyrus* protein, while 3 of the proteins were similar only to *A. caninum* ES proteins with domains of unknown function (NAME_07734, NAME_09181, NAME_09182). Seven proteins shared homology with only *A. caninum* and *H. polygyrus* proteins, 11 shared homology with only *A. caninum* and *N. brasiliensis* proteins and 7 shared homology with only *H. polygyrus* and *N. brasiliensis* proteins. One hundred and twenty-two *N. americanus* proteins shared different degrees of homology with proteins from *A. caninum*, *H. polygyrus* and *N. brasiliensis*.

From the 54 SCP/TAPS proteins found in *N. americanus* ES products, one (NAME_13724) shared homology with a protein found in only *H. polygyrus*, while another SCP/TAPS protein (NAME_15177) shared homology with proteins from both *A. caninum* and *N. brasiliensis*. Fifty-one of 54 SCP/TAPS proteins were similar to all compared species, leaving a single SCP/TAPS protein (NAME_11218) which did not have homology to any SCP/TAPS protein from the compared species. Twenty-five *N. americanus* ES proteins did not have homology to any proteins in the secretome of *A. caninum*, *H. polygyrus* or *N.*

brasiliensis. Notable proteins among these 25 were NAME_01848 (aspartyl protease) and NAME_05081 (zinc metalloprotease).

Of the 26 proteases in the *N. americanus* ES products, 22 had homologs in all compared species, while one serine protease (NAME_06735) only had homologs in *H. polygyrus* and *A. caninum* and one metalloprotease (NAME_00535, peptidase family M1) only had homologs in the ES products of *A. caninum*, and *N. brasiliensis*. One zinc metalloprotease (NAME_05081) and one serine protease (NAME_01250) were only found in the ES products of *N. americanus* and therefore did not have any similarity to ES products from the other species.

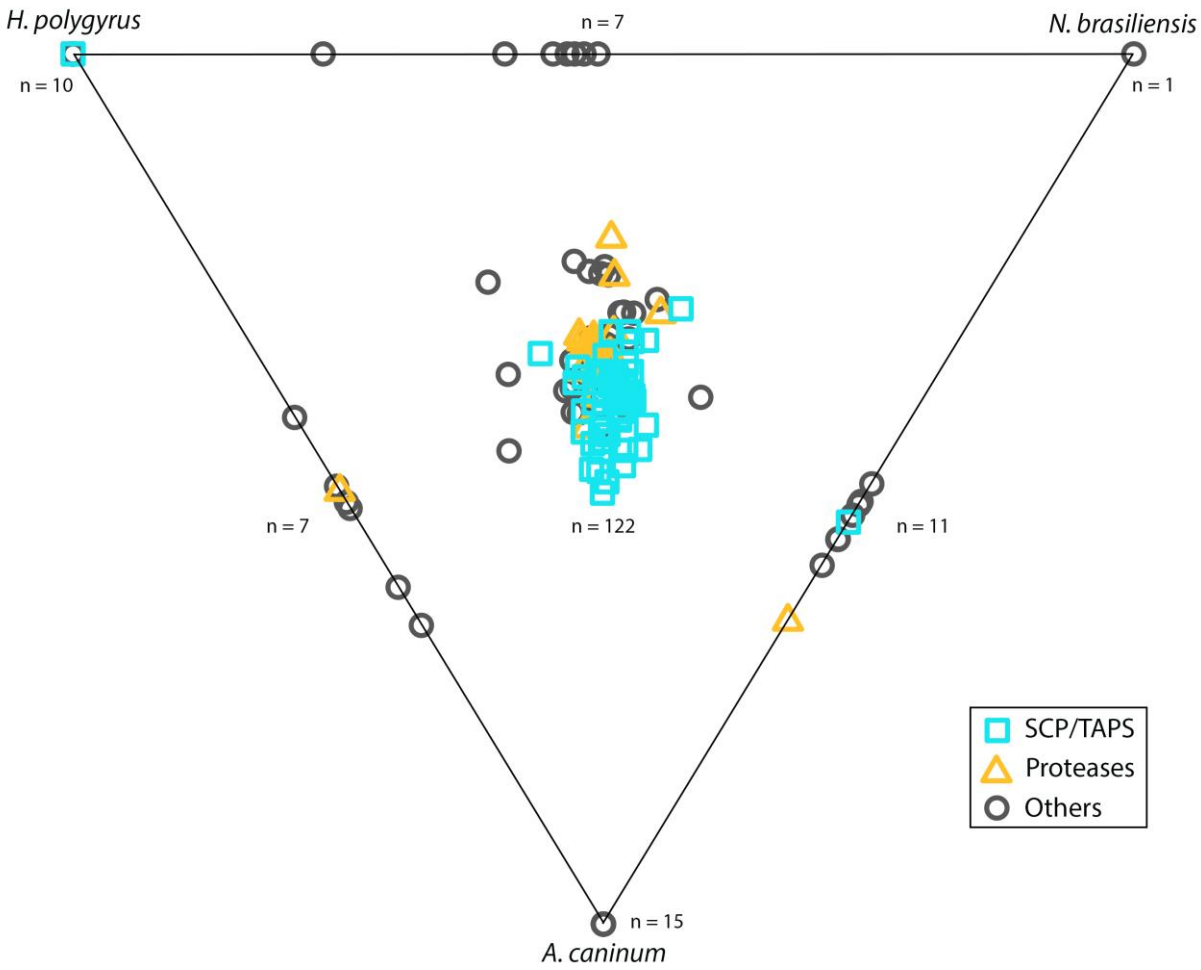


Figure 2.3: Similarity analysis of excretory/secretory (ES) proteins from helminths commonly used to model the human hookworm. Excretory/secretory (ES) products from *Ancylostoma caninum*,

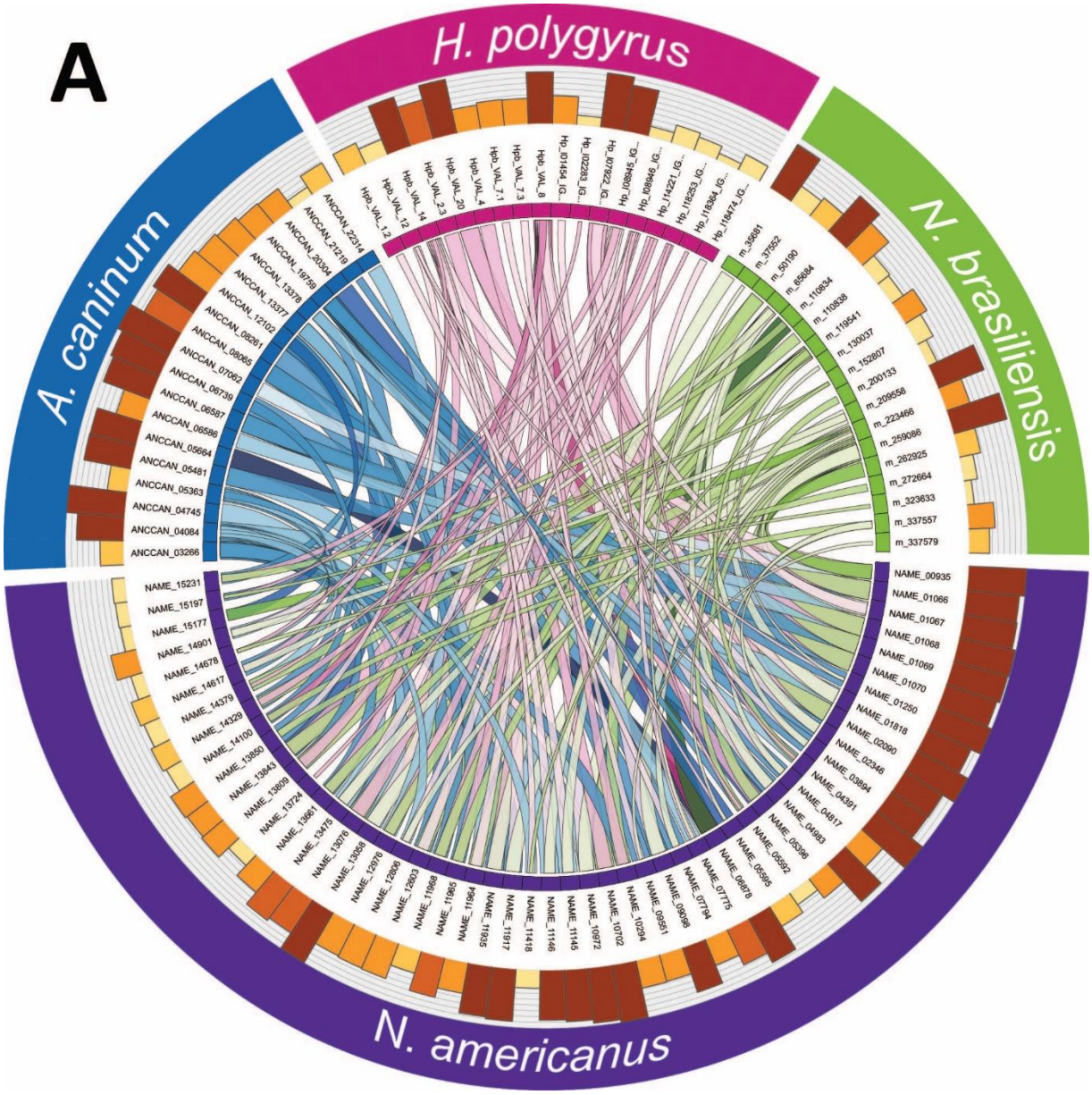
Heligmosomoides polygyrus, *and Nippostrongylus brasiliensis* were compared with the ES proteome of *Necator americanus* and displayed in a Simitri plot. SCP/TAPS are represented by a teal square, proteases by an orange triangle and any other protein by a grey circle. Points on the diagram triangle represent sequences which only had similarity to the labelled species. Points along the edges of the triangle are sequences which had similarity to two of the three species (given at the respective ends of the edge). Any sequence in the middle area of the triangle represents a sequence with similarity to all three compared species.

2.3.4 Homology analysis of SCP/TAPS and proteases in the ES products of *N. americanus*

Since SCP/TAPS proteins and proteases from *N. americanus* numerically dominate the ES protein dataset and likely play key roles in infection, migration and parasite establishment, I performed an in-depth comparison of these families of proteins from all four species of gastrointestinal nematodes. Adult *N. americanus* ES SCP/TAPS protein sequences were aligned with homologs from *H. polygyrus*, *A. caninum* and *N. brasiliensis* ES SCP/TAPS protein sequences using BLAST. Sequences which aligned with maximum scores >36 were visualized using Circos (Figure 2.4A). SCP/TAPS from *N. americanus* are more similar to *A. caninum* than to *H. polygyrus* or *N. brasiliensis*, as denoted by thicker, darker ribbons (Figure 2.4A). The sequences, their homologs and the corresponding blast scores are detailed in Supplemental data 7. As with the similarity analysis, NAME_11218 had no significant alignment to any of the compared species.

A similar analysis was performed for the proteases from each of the aforementioned species (Figure 2.4B). These proteases have been grouped together into their respective mechanistic classes: aspartyl (ASP), cysteine (CYS), metallo (MET) or serine (SER) proteases. In general, all three comparator species had high protease sequence homology to metalloproteases from *N. americanus*. Surprisingly, aspartyl protease sequences were more similar between *H. polygyrus*, *N. brasiliensis* and *N. americanus* than sequences from *A. caninum*. It is interesting to note that *N. americanus* contained more aspartyl proteases than the other nematode species analyzed (*N. americanus* – 8; *A. caninum* – 4; *H. polygyrus*

– 6; *N. brasiliensis* – 6). Serine proteases had the lowest sequence homology across all of the compared species and protease subclasses.



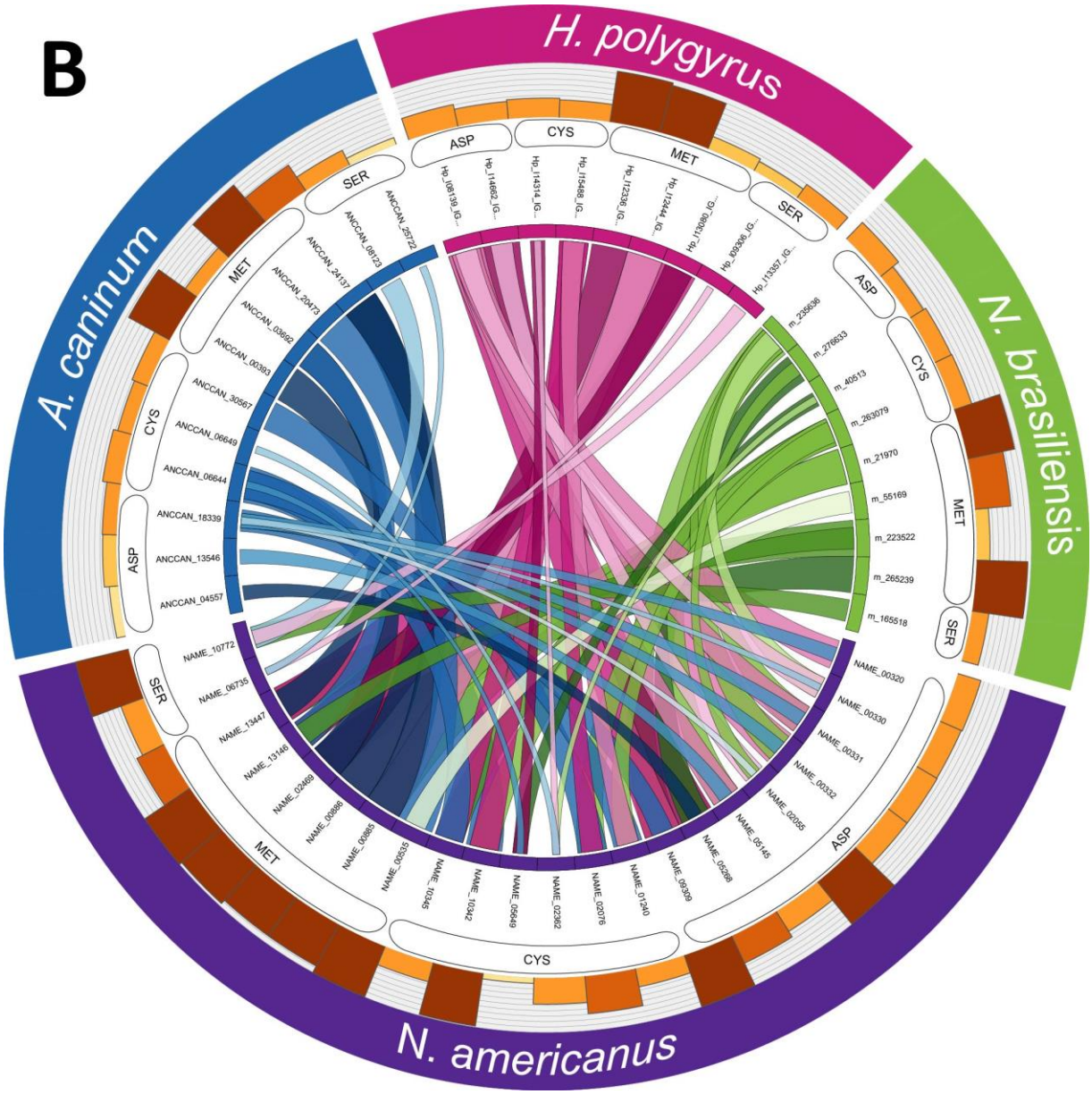


Figure 2.4: SCP/TAPS proteins in the excretory/secretory (ES) products of human hookworm are most closely related to SCP/TAPS proteins in the ES proteome of dog hookworm. SCP/TAPS **(A)** and protease **(B)** protein names are displayed in a circle with *Necator americanus* (purple), *Ancylostoma caninum* (blue), *Heligmosomoides polygyrus* (pink), and *Nippostrongylus brasiliensis* (green). Ribbon thickness is relative to the maximum score obtained in the BLAST search while darker ribbons denote higher sequence percent identity. The corresponding bars provide relative sequence length of each protein. Respective protease mechanistic classes: aspartic (ASP), cysteine (CYS), metallo (MET) or serine (SER).

2.3.5 Phylogenetic analysis of SCP/TAPS proteins

SCP/TAPS sequences of 6 parasitic nematodes were obtained from published secretomes and compared to *N. americanus* SCP/TAPS identified in this study. Sequences were grouped by whether they had a single or double domain and then aligned using MUSCLE and PhyML. From the 7 total species, 232 SCP/TAPS were reported with 134 single-domain proteins and 98 double-domain. For the single SCP/TAPS-domain proteins an unrooted tree was generated. The analysis identified seven main clades. Three of these clades consisted entirely of sequences from *N. americanus* and *A. caninum*, while sub-clades in 3 of the other main clades followed a similar trend. A majority of *T. muris* single-domain sequences formed a sub-clade with *A. suum* and *T. canis*, grouping together the non-clade V helminths. An unrooted tree was also generated for double domain sequences. The analysis identified five main clades. *N. americanus* SCP/TAPS clustered almost exclusively with sequences from *A. caninum* again, representing one main clade and several sub-clades. *H. polygyrus* and *N. brasiliensis* also formed a number of distinct sub-clades. *T. canis*, the only non-clade V helminth, was reported to only produce one double-domain SCP/TAPS protein in its ES products which was not closely related to any of the compared sequences. Interestingly, no double-domain SCP/TAPS were reported for *A. suum* or *T. muris*. Another trend across both single and double domain trees was the diversity of evolution between SCP/TAPS within a single species. For example, the single-domain tree included a sub-clade of 7 *N. americanus*-only sequences while other *N. americanus* sequences had more similarity to mouse hookworm sequences.

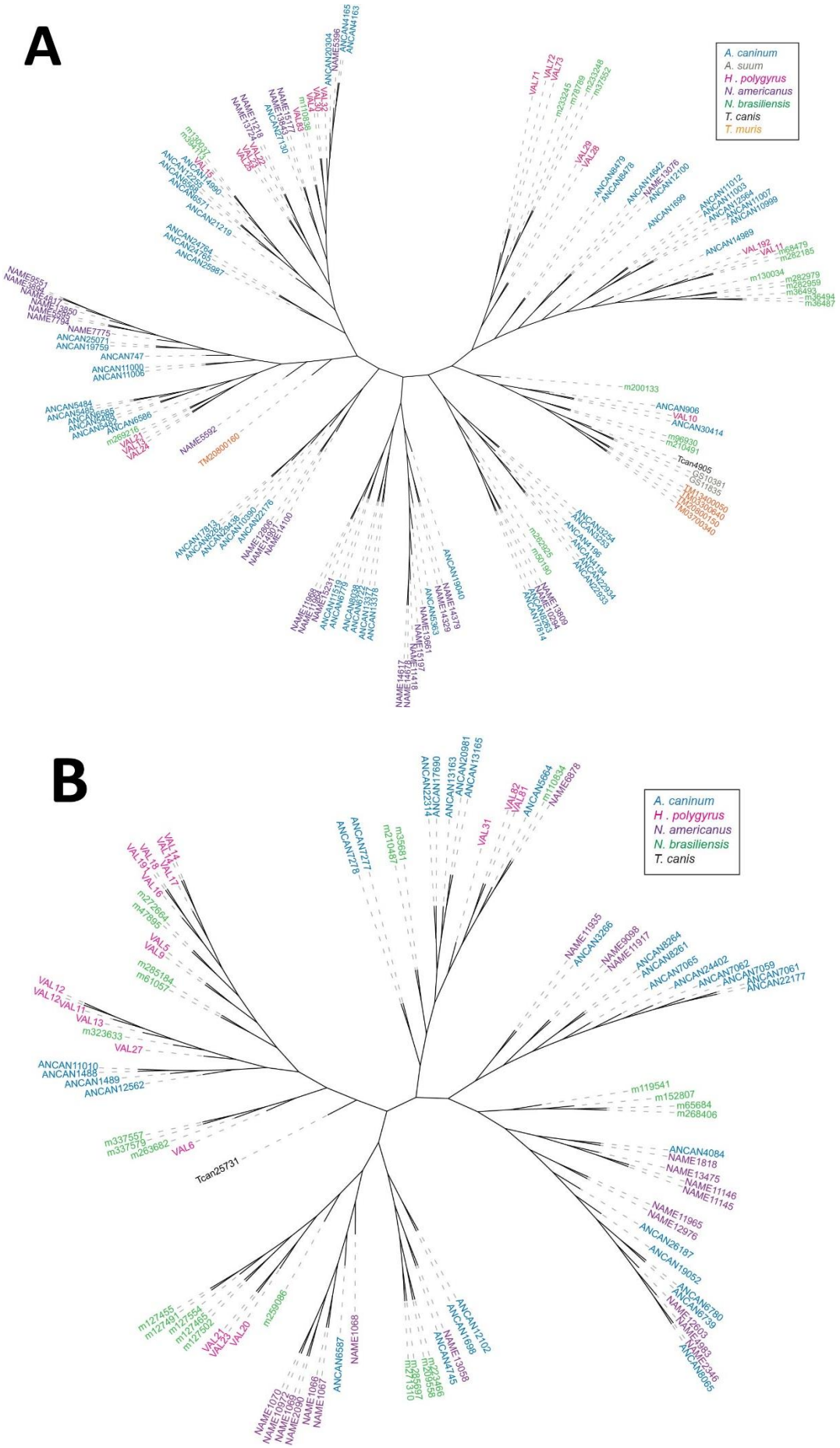

A
A. caninum
A. suum
H . polygyrus
N. americanus
N. brasiliensis
T. canis
T. muris
B
A. caninum
H . polygyrus
N. americanus
N. brasiliensis
T. canis

Figure 2.5: Phylogenetic analysis of single and double domain SCP/TAPS proteins from various species of helminths. Phylogenetic relationships of **(A)** single-domain and **(B)** double-domain SCP/TAPS proteins determined with MUSCLE alignment software. PhyML was used for a maximum-likelihood phylogenetic analysis and results were visualized with The Interactive Tree of Life (iTOF) online phylogeny tool. *Necator americanus* sequences are highlighted in purple with comparator species each denoted by a different color (see key).

2.4 Discussion

The genome of *N. americanus* was sequenced in 2014, providing an important dataset to facilitate efforts to combat hookworm disease; however, the tools available at that time for annotating genes from parasitic helminths were limited. For instance, the number of proteins with a top hit to a 'hypothetical protein' present in the original genome annotation was 3,043, corresponding to 15.8% of the total predicted proteins [169]. In addition, inferring gene and protein functions for parasitic nematodes is a major challenge as most species are genetically intractable and databases and algorithms are biased towards (free-living) model nematodes [223]. Proteogenomics is a relatively new approach in which proteomic data is used to improve genome annotation [225,251], and although it had never been applied to parasitic helminths (until now), its potential utility in this area has been suggested [252]. High-throughput sequencing and gene prediction tools are prone to false-negative and false-positive predictions which can lead to missed genes, false exons or exon boundaries and/or incorrect translational start/stop sites, so knowing the sequences of the proteins expressed by an organism will help to improve gene predictions.

The proteogenomic analysis carried out in this study addresses some of these challenges by improving the characterization of predicted proteins from the annotated *N. americanus* genome. This study, while primarily focused on the ES products, provided an opportunity to update the *N. americanus* genome. For this end, I included L3 stage material to provide more data to add to this annotation. Overall, 121 peptides were identified that mapped to intergenic regions in the first draft genome sequence for *N. americanus*. Peptides that map to intergenic regions are important as they can lead to identification of novel protein-coding genes or corrections of pre-existing models [225]. To investigate whether these peptides are likely to be new genes we checked for any potential ORFs where the peptides map. Thirty-two [227] of the peptides identified mapped to alternative ORFs than those described in the current gene model. While these peptides map to known coding regions, they highlight out-of-frame ORFs which is likely to, once correctly annotated, result in an entirely different protein. Of the total newly

identified peptides, 39 of them mapped to introns. Peptides mapping to introns can lead to identification of novel splice isoforms or amendments in gene structure. Peptides mapping to exon extensions and N-terminal extensions were less abundant, with 10 and 4 peptides respectively. These groups of peptides suggest a possible correction in reading frame or an incorrect start site annotation.

The decrease in gene number seen in this study is in line with other genome annotations. For example, the *Schistosoma mansoni* genome was originally thought to encode 11,809 genes, however further annotation has reduced this number to 10,772 [253]. Given that this is the first re-annotation of the *N. americanus* genome since the original draft was published, a substantial decrease in gene number was to be expected, and 15,728 genes is comparable with the predicted gene numbers of other nematode genomes [254]. Despite being a parasitic nematode, *N. americanus* has almost 5,000 fewer genes than its free-living relative *C. elegans* (https://parasite.wormbase.org/).

Of particular importance in the characterization of parasite-specific genes is the presence of a signal peptide. Of the secreted ES proteins from *N. americanus*, 51% were predicted to have a signal peptide; this is also in agreement with ES proteomes from other parasitic nematodes [194,213,248]. The presence of extracellular proteins without predicted signal peptides in the ES products of *N. americanus* could be due to one of three reasons: (a) the protein is secreted via an alternative pathway, including release of parasite extracellular vesicles [255-258] or non-classical secretory signals; (b) the lack of full-length RNA transcript sequence to confirm gene model accuracy resulted in an error in the predicted sequence (i.e. truncation or ORF shift); (c) the pre-set D-cutoff threshold of 0.33 in SignalP results in false-negative predictions. In support of this, 45% of the sequences without signal peptides were predicted to be non-classically secreted.

Helminth secretomes represent the molecular host-parasite interface [259], and provide useful insights into the biological strategies employed by these parasites to ensure longevity inside their respective hosts [259]. At this interface, ES products have been implicated in numerous roles from initial penetration/invasion and feeding to host immune regulation [259]. Obtaining sufficient ES products to

generate the proteome described in this study was time consuming due to the difficulty in culturing sufficient quantities of parasite material in hamsters. For this reason, human infection with *N. americanus* is frequently modelled using other hookworms and related nematodes that survive in rodents or larger animals, including *Ancylostoma* sp*., H. polygyrus* and *N. brasiliensis.* It should be noted that the *N. americanus* material employed in this study was obtained from parasites adapted to hamsters through many generations of selective pressure [260,261]. Adult stage *N. americanus* parasites obtained from hamster infection were smaller and less fecund than those from human hosts [262]. These phenetic differences may be associated with differences in transcriptional patterns. To date there have been no studies directly comparing the genomes or secretomes of *N. americanus* maintained in hamsters and those obtained directly from human infections. Such data would enhance the validity of the data obtained in this study.

The protein family analysis of *N. americanus* ES products revealed a diverse number of known and unknown domains (391 domains total), which attests to the many biological functions of ES proteins as well as to the lack of information on these proteomes. The ES products of the three comparator species used here had similar protein family profiles with 458 (*A. caninum*), 434 (*H. polygyrus*) and 628 (*N. brasiliensis*) unique domains present. To assess the overall usefulness of these models, a similarity analysis was carried out on their ES products. This analysis relates relative protein sequence similarities in a plot where each of the identified *N. americanus* ES proteins is compared to the ES proteome of the comparator species [194,247]. The majority of the *N. americanus* ES proteins (173/198), including 51/54 SCP/TAPS proteins, had homologs in the ES products of all three comparators, highlighting the relevance and usefulness of all three models. A total of 25 proteins did not have similarity to ES proteins from the other 3 nematodes analyzed. Since *A. caninum*, *H. polygyrus* and *N. brasiliensis* are animal parasites, these 25 *N. americanus* proteins might have evolved to such an extent that they target human-specific pathways. Two unique proteins of interest are the aspartyl protease NAME_01848 (PF00026) and the zinc metalloprotease NAME_05081 (PF01546). Metalloproteases and aspartyl

proteases play crucial roles in host tissue penetration and parasite feeding, and as such, proteolytic enzymes from helminths may be of particular interest as vaccine and/or drug targets [263,264].

The most abundantly represented protein family (54/198; 27%) in *N. americanus* ES products was proteins containing a single or double SCP/TAPS domain (PF00188). This finding aligns with previous work that highlighted the abundance of this protein family in nematode ES products in particular [213,265]. For instance, a total of 45, 90 and 25 SCP/TAPS proteins were found in the ES products from *N. brasiliensis* (45/313; 14%), *A. caninum* (90/315; 29%) and *H. polygyrus* respectively (25/374; 12%). Interestingly, this family of proteins is also abundant in the extracellular vesicles secreted by different nematodes [255,258]. Given that SCP/TAPS proteins from *N. brasiliensis* are almost exclusively secreted by the adult developmental stage [194], these proteins are likely coordinating specific roles in the gastrointestinal tract of the host. In fact, it has also been shown that SCP/TAPS are overrepresented at the transcript level in *N. americanus* adult worms [169]. Serum-activated *A. caninum* L3 larvae were also shown to upregulate SCP/TAPS in culture [266]. While relatively little functional information is available for SCP/TAPS proteins, neutrophil inhibitory factor (NIF), an SCP/TAPS protein in the ES of *A. caninum*, was reported to abrogate neutrophil adhesion to the endothelium [267]. However, a *N. americanus* homolog of this protein was not detected in the current study, despite the presence of a NIF-encoding gene in the draft genome [169]. SCP/TAPS proteins are thought to play numerous and diverse roles at the host-parasite interface, from defense mechanisms, normal body formation and lifespan [197]. The diverse nature of *N. americanus* SCP/TAPS sequences is evidenced by their phylogenetic relationships (Figure 5A and B). Blast analyses presented in the Circos plot (Figure 4A) revealed varying degrees of sequence homology to SCP/TAPS proteins from *A. caninum*, *H. polygyrus* or *N. brasiliensis*. These SCP/TAPS proteins should be further explored to understand their roles in *N. americanus*-human host interactions. Given the limited availability of information regarding the function of helminth SCP/TAPS proteins in general, it was unsurprising that GO analyses revealed no known molecular function or biological process for 32/54 of the SCP/TAPS from *N. americanus* ES

products, and more studies should be performed to characterize the properties of this intriguing family of proteins.

Despite the lack of functional information on the SCP/TAPS proteins in parasitic helminths, of the species we compared in this study, *N. americanus* proteins were generally most similar to those from *A. caninum*, which could simply be a reflection of the phylogenetic similarity between the two species. (Figure 5A and 5B). The trees generated in this study highlight strong clade-specific similarities between SCP/TAPS in the ES products of the compared species. In support of this, the vast majority of the SCP/TAPS proteins came from the four clade V species, while *A. suum*, *T. muris* and *T. canis* had only 2, 5 and 2 SCP/TAPS proteins respectively. The clustering of *A. caninum* with *N. americanus* and *H. polygyrus* with *N. brasiliensis* supports the notion of host-specific roles for SCP/TAPS. Another trend across both single and double domain trees was the diversity of evolution between SCP/TAPS within a single species. For example, the single-domain tree included a sub-clade of 7 *N. americanus*-only sequences indicating an important human-specific role for this evolutionary cluster. This compares well with a number of other sub-clades that included proteins from the four predominant species. These SCP/TAPS proteins are more likely to share a common function, potentially in host-infection or parasite development within the mammalian host.

The phylogenetic analysis strongly attests to the preferred use of *A. caninum* for studying hookworm molecular biology and ES products in general [268]. This finding is reinforced by the similarity analysis represented in the Circos plot of SCP/TAPS (Figure 4A). Not only was there a greater number of SCP/TAPS homologs in the ES products of *A. caninum* but these proteins also had relatively higher blast scores (denoted by link ribbon thickness) and higher percent identity scores (denoted by ribbon darkness). This type of analysis can prove useful since it also reveals which species to consider for investigating a specific *Necator* SCP/TAPS protein. For example, NAME_13850 has significant sequence homology to a SCP/TAPS protein from each of the compared species; however, the species with the highest sequence homology is *A. caninum* (ANCCAN_19759), making this protein the most relevant to

study as a model for NAME_13850. Prior to updating the draft genome, all three species of nematode used in this comparison had longer average gene sequence lengths than the *N. americanus* SCP/TAPS sequences. In the previous annotation, the average *N. americanus* SCP/TAPS sequence length was 244 predicted amino acids, compared with 355 residues in the updated genome. This was likely due to some of the sequences being truncated, yielding similar results to the published *H. polygyrus* proteome [248].

The Circos plot representing *N. americanus* proteases and homologous proteins from the three comparator species (Figure 4B) provides insight into the high degree of sequence similarity. Unlike the SCP/TAPS Circos analysis, all the *N. americanus* ES proteases had homologs with high similarity in the compared species (average 50% identity between *N. americanus* and comparator species protease). This finding supports the notion of using any one of these three parasites to study *N. americanus* proteases in general. We identified 8 aspartyl, 6 cysteine, 9 metallo and 3 serine proteases in the ES products of *N. americanus*. As the adult stage of *N. americanus* feeds on blood, the high abundance of aspartyl proteases was expected [219,269]. Yet when compared to other species, the human hookworm ES products included more of these proteases. Aspartyl proteases play a fundamental role in the digestion of host haemoglobin and have also been implicated in skin penetration, feeding, and host tissue degradation [270,271]. The finding that *N. americanus* has more of this mechanistic class of proteases than the other nematodes assessed here is likely due to split gene models and/or may be a true gene family expansion. Due to the vital role that aspartyl proteases play in parasite feeding, *Na*-APR-1 an *N. americanus* aspartyl protease, was targeted as a vaccine candidate [211]. While we were unable to detect *Na*-APR-1 in the current ES proteome - probably because it is anchored to the gut epithelium [272] - 9 other aspartyl proteases were detected which could potentially be targeted as novel vaccine candidates.

Cysteine proteases, particularly the group belonging to the papain superfamily, are common in nematodes [273]. They have been specifically described for their proteolytic activity against

haemoglobin, antibodies and fibrinogen in the *N. americanus* lifecycle [274]. Similarly, another study highlighted the importance of four cysteine proteases that were upregulated in the *N. americanus* transition from free-living larvae to blood-feeding adult worm, indicating that these proteins are likely to be important for nutrient acquisition [275]. Metalloproteases - particularly the astacins - identified in this study are most likely important for larval and adult migration and invasion through human host tissue [276]. In support of this, astacin metalloproteases were found to be upregulated in *N. brasiliensis* larvae when compared with adult stage parasites [194]. Interestingly, *N. americanus* metalloproteases were reported to inhibit eosinophil recruitment through the cleavage of eotaxin, a potent eosinophil chemoattractant [217]. The least abundant family of proteases in the adult *N. americanus* ES products was the serine proteases. Relatively little is known about these proteases from *N. americanus* specifically; however, a serine protease from the whipworm *T. muris* is involved in degradation of the mucus barrier to facilitate feeding [277]. Due to the importance of these various proteases in parasite feeding, infection, migration and defense, they represent potential targets for chemotherapies and vaccines to limit infection [278].

In this chapter, I have presented the first proteogenomic analysis of a helminth parasite, resulting in a more accurate genome annotation. While the above annotations are valuable additions to the current genome, further improvement requires access to substantially more parasite material including adult stage parasites which are difficult to obtain. Furthermore, I have carried out the first proteomic analysis of the ES products of the human hookworm *N. americanus*. The results presented herein offer significant insight into the validity of these model species while also highlighting differences between these important parasites.

3 Chapter 3 – Recombinant hookworm protein production and their serodiagnostic potential

3.1 Introduction

The proteomic results generated in chapter 2 includes the secretome of the human hookworm. Having access to this database helps to inform the direction and practice of numerous studies into the function of specific molecules from this parasite. To begin to explore these functions, the proteins must first be produced in recombinant form. This is due to the difficulty in obtaining sufficient quantities of purified parasite-derived proteins. Using the example of the human hookworm, ES product yields directly from the worm are very low, require a lot of parasites, and can be costly. The introduction of molecular cloning in the 1970s has afforded researchers the option of producing foreign proteins in a range of prokaryotic and eukaryotic cell lines. Among the many microbial systems, yeast combines the benefits of a unicellular organism (i.e. ease of growth and genetic exploitation) with the advantage of eukaryotic post-translational modifications and protein folding.

P. pastoris is a methylotrophic yeast meaning it uses methanol to grow as a sole source of energy and carbon [279]. The alcohol oxidase gene *AOX1*, induced by methanol, was used in *P. pastoris* to develop an efficient organism for DNA transformations and subsequent protein production [280]. The main advantage of this system is a high secretory capacity in combination with the strongly inducible methanol-dependent promotor [281]. Additionally, the system has a distinct lack of endotoxin by-products or viral DNA [282]. The disadvantageous come with industrial scale production including: high heat production, considerable oxygen demand, and the use of a flammable substrate [279]. As these disadvantageous are linked with very large scale production, they are not necessary to consider for smaller studies.

This system has been used successfully to express a number of helminth secreted proteins. Three glutathione S-transferases from *N. americanus* were expressed in *P. pastoris* with one taken forward as a potential vaccine candidate [283,284]. Notably, SCP/TAPS proteins including Na-ASP-2 and Ac-ASP-7 have been expressed using this system and used for functional and crystal structure determination [285-287]. Given these studies with SCP/TAPS specifically, as well as the previously described advantageous of using *P. pastoris*, this method was ideal for use in this thesis.

Based on the results of the proteomics analysis from chapter 2, eighteen proteins of interest were selected to be produced in yeast. Proteins were selected based on their abundance (emPAI), whether they were SCP/TAPS proteins and chosen based on dissimilarity to other SCP/TAPS proteins from the phylogenetic analyses. To improve the odds of selecting one of the key immunoregulatory molecules, the main aim in chapter 3 was to express SCP/TAPS proteins. I chose SCP/TAPS as they were abundantly represented in the ES products, but also because they are thought to play important roles in the modulation of host immune responses [197,288,289]. This chapter finishes with the testing the recombinant proteins as serodiagnostic markers of infection. This work was carried out to test whether ES antigens from the human hookworm could potentially be used to improve the currently available diagnostic tools.

3.2 Methods

3.2.1 Cloning

Gene of interest

An initiator methionine and signal peptide were used to identify full-length ORFs. The signal peptides were removed and the ORF was designed to be fused in-frame with the encoded signal peptide in the pPICZα vector (Figure 3.1). Sequences were checked for restriction enzyme sites using the online tool

NEBcutter v2.0 available at http://nc2.neb.com/NEBcutter2/. If the sites *Eco*R1, *Xba*1 or *Sac*1 were present, a point mutation was carried out ensuring that the encoded amino acid sequence was not affected. The *Eco*R1 restriction site sequence (GAATTC) was placed at the N-terminus of the ORF while *Xba*1 (TCTAGA) was placed at the C-terminus. In order to retain the ORF, two guanines were inserted upstream of the *Xba*1 sequence, resulting in an additional glycine on the end of each recombinant protein (Supplementary data 7.1). Genes were synthesized by Integrated DNA Technologies (IDT) by following the instructions at www.idtdna.com. Sequences were automatically codon-optimised for *P. pastoris* using the optimization tool through the IDT ordering process.

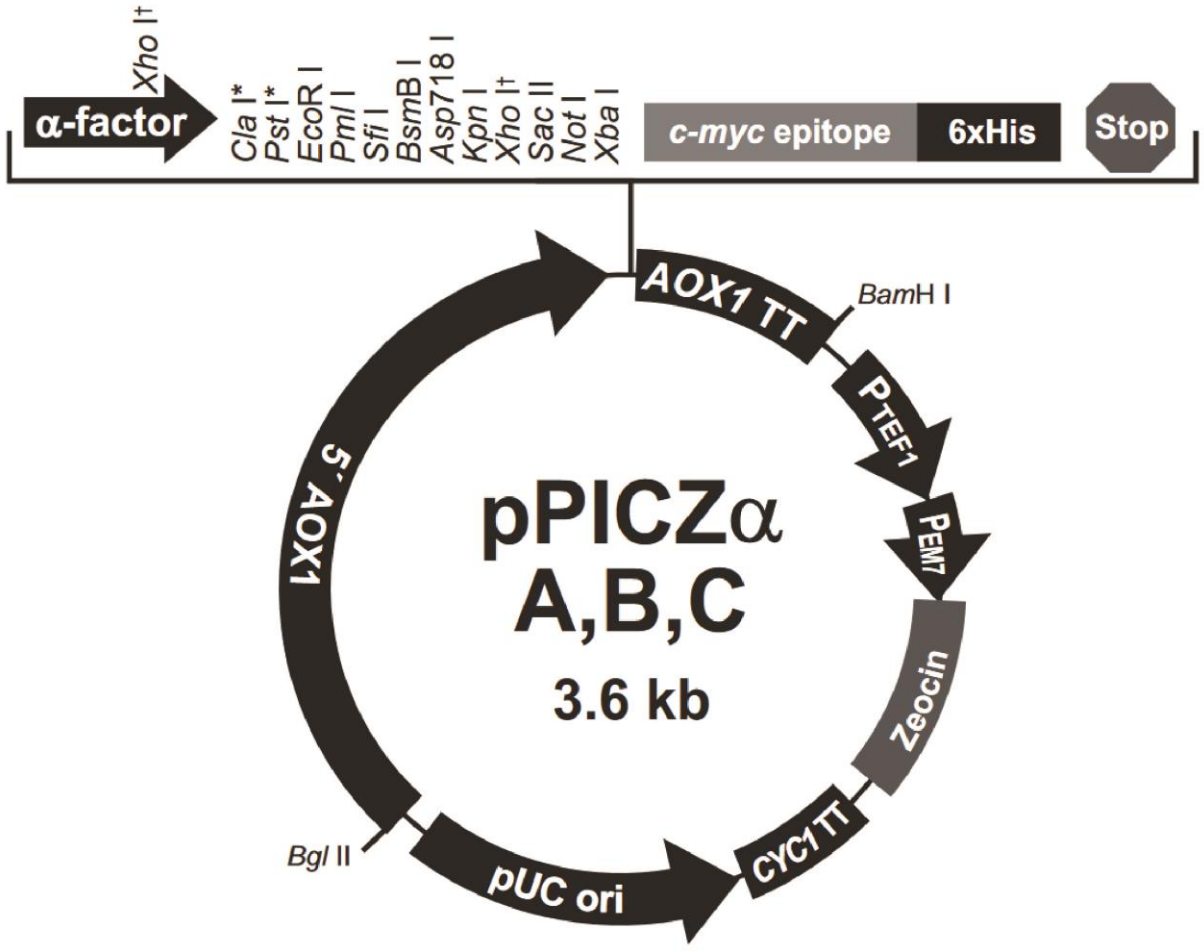


Figure 3.1: Summary of the features of the pPICZα vectors.

Transformation

Aseptic technique using either a Bunsen burner or a fume hood was observed throughout the entirety of this protocol. Four micrograms of gene of interest (GOI) DNA in IDT generic cloning vector was obtained from IDT. Upon arrival, plasmid DNA was hydrated with dH_2O, vortexed and centrifuged briefly. From this plasmid DNA, 0.2 µg was added to 100 µl Ca millilitre Cl_2 competent *E. coli* cells (prepared as described below) and incubated for 20 minutes on ice. The cells were then heat shocked at 42°C for 45 seconds and placed back on ice for a further 2 minutes. After incubation on ice, 500 µl of LB low-salt media was added, the cells were incubated for 1 hour at 37°C, shaking at 220 rpm, before being plated onto LB low-salt plates with 100 µg/ml ampicillin. After overnight incubation at 37°C, single colonies were picked from the plate and cultured in 5 ml LB low-salt media (100 µg/ml ampicillin) in 50 ml tubes. These tubes were incubated O/N at 37°C, shaking at 220 rpm.

Cloudy media indicated good bacterial growth. From this O/N culture 500 µl was added to a solution of 50% glycerol, mixed gently and stored immediately at -80°C. The remaining 4.5 ml of culture was miniprepped using the Nucleospin Plasmid Easypure kit (Macherey-Nagel). Miniprepping was followed according to the kit manual. Flow through (FT) was stored on ice, analysed for DNA content with a Nanodrop 2000 and stored at -30°C for later use.

Using approximately 2 µg of plasmid DNA obtained from the miniprep, a digestion was carried out as follows. A master-mix (MM) was made by combining: 15 µl dH_2O, 3 µl 10X CutSmart Buffer (B7204S, New England BioLabs), 1 µl *Eco*R1 HF (New England BioLabs), 1 µl *Xba*I (New England BioLabs). Twenty microlitres of MM was added to 10 µl of plasmid DNA (2 µg total) to give a final volume of 30 µl. The same format was used for digesting the pPICZα vector (ThermoFisher). Digestion reactions were heated for 2 hours at 37°C on a heat block. An agarose gel (1%) was prepared. Six microlitres of Gel Loading Dye, Purple (New England BioLabs) was added to each of the 30 µl digestion reactions, mixed, and each loaded to a separate well of the 1% agarose gel. HyperLadder 1kb (BioLine) was included as a size ladder. The gel was run for 45-60 minutes at 100 V. Cut plasmids and DNA inserts were viewed using a

Molecular Imager VersaDoc (BioRad). Using a scalpel blade, DNA inserts were cut from the agarose gel. An Isolate II PCR and Gel Kit (Bioline) was used to isolate the DNA inserts from the agarose gel. In brief, gel was solubilized by adding 200 µl of Binding Buffer CB per 100 mg gel, incubating at 50°C for 5-10 minutes and vortexing. Solubilized gel was loaded into a kit column, centrifuged for 30 seconds at 11,000 *g*, and FT discarded. Two washes were carried out by adding 600 µl Wash Buffer CW, centrifuging for 30 seconds at 11,000 *g*, and FT discarded. The membrane was dried by carrying out another 1 min centrifugation at 11,000 *g*. The collection tube was replaced, 15 µl Elution Buffer C added and the tube was incubated at room temperature for 1 min. The tube was then centrifuged for 1 min at 11,000 *g* to elute the DNA in a final volume of 15 µl. DNA was quantified with a Nanodrop 2000 (ThermoFisher) before being stored at -30°C.

Ligation

DNA inserts were ligated into cut pPICZα plasmids (ThermoFisher) following the subsequent procedure. A MM was made by combining: 8.5 µl dH_2O, 1.5 µl 10X T4 DNA Ligase Reaction Buffer (New England BioLabs), 1 µl T4 DNA Ligase (New England BioLabs) and 1 µl cut pPICZα (~300 ng/µl, ThermoFisher). Three microliters of DNA insert was added to 12 µl of MM and left at room temperature for 2 h to ligate. Ligated pPICZα was transformed into *E. coli* and subsequently miniprepped as previously described to a total amount of 15 µg plasmid with GOI. A restriction digest was carried out, as previously described, to confirm the presence of the DNA insert.

Digest and clean-up

Plasmid with GOI was digested to a linear form and cleaned up in preparation for electroporation using the following procedure. MM was made by combining: 7 µl dH_2O, 6 µl 10X CutSmart Buffer (New England BioLabs), 2 µl *Sac*I HF (New England BioLabs). Subsequently, 45 µl of pPICZα with GOI (~350 ng/µl) was added to 15 µl of MM and left for 4 h, at 37°C. Following linearization, plasmid DNA was cleaned up using an Isolate II PCR and Gel Kit (Bioline). The method followed was the same as previously

described with the exception of the initial step. One hundred and twenty microlitres of Binding Buffer CB was added to the 60 µl linearized plasmid with GOI, mixed and loaded into the column.

Electroporation

Competent *P. pastoris* was prepared following the subsequent procedure. *P. pastoris* was streaked on a YPDS plate and allowed to grow until colonies formed. A single colony was picked and grown in YPD media rotating at 200 rpm, 30°C, O/N. From the O/N culture, 500 µl was inoculated into 100 ml of fresh YPD and grown to an optical density (OD) between 1.3-1.5. Throughout the remainder of this protocol the cells were kept on ice as much as possible. The cells were centrifuged for 5 minutes at 1,500 *g*, 4°C, the supernatant (S/N) discarded, and the pellet resuspended in 50 ml ice-cold dH_2O. The cells were centrifuged for 5 minutes at 1,500 *g*, 4°C, the S/N discarded, and pellet resuspended in 25 ml ice-cold dH_2O. The cells were centrifuged for 5 minutes at 1,500 *g*, 4°C, the S/N discarded, and the pellet resuspended in 10 ml ice-cold 1 M Sorbitol. The cells were centrifuged for 5 minutes, 1,500 *g*, 4°C, the S/N discarded, and the pellet resuspended in 1 ml ice-cold 1 M Sorbitol.

Eight microlitres of cells were mixed with 10 µl linearized DNA with GOI (from Digest and Cleanup stage) and transferred to an ice-cold 2 mm electroporation cuvette (Bio-Rad). The cell-DNA mixture was incubated on ice for 5 minutes before being pulsed on a Bio-Rad Gene Pulser Xcell (settings: 2,000 V, 25 µF, and 200 Ω). One millilitre of 1 M sorbitol was added to the sample before it was transferred to a 15 ml falcon tube and incubated for 1.5 h at 30°C. Incubated cells were pelleted gently at 300 *g* for 2 minutes. Nine hundred microlitres of S/N was removed, the cells were resuspended in the remaining 100 µl, and streaked on a YPDS plate with 100 µg/ml Zeocin. Plates were incubated for a minimum of 3 days at 30°C in the dark. Single colonies were picked from plates and patched on to a YPDS plate with 1 mg/ml Zeocin. Patched colonies were allowed to grow for 2 days at 30°C in the dark.

Small scale expression

A single colony was stabbed from the YPDS plate with 1 mg/ml Zeocin and put into 5 ml YPD media (100 µg/ml Zeocin) in a 50 ml tube. The colony was cultured O/N at 30°C, rotating at 200 rpm. One hundred microlitres of O/N culture was aliquoted into 5 ml fresh BMGY media and cultured O/N at 30°C, rotating at 200 rpm. The remaining YPD culture was centrifuged briefly for 5 minutes at 2,000 rpm before the S/N was aspirated. The cell pellet was resuspended in 15% glycerol, transferred to a cryovial and stored immediately at -80°C. The O/N BMGY culture was centrifuged briefly for 5 minutes at 500 *g* and resuspended to an OD of ~1 with BMMY. Five millilitres of OD1 BMMY culture was incubated at 30°C for 72 h, rotating at 200 rpm. At 24 h and 48 h of the culture period, 25 µl of 100% methanol (Sigma-Aldrich) was added to give a final concentration of 0.5% v/v. One hundred microlitres of BMMY culture was collected at 24 h, 48 h and 72 h for use in a dot blot and western blot. The 100 µl culture was centrifuged briefly at 500 *g* for 5 minutes, and the S/N stored at -30°C until use.

Dot blot

Nitrocellulose membrane (GE Healthcare) was cut and boxes were drawn lightly with a pencil. Two microlitres of clarified BMMY S/N from each time point (24 h, 48 h and 72 h) was blotted separately onto the membrane. Each dot was allowed to dry for 10 minutes before another 2 µl of the same S/N was added on top of the original dot. This process was repeated 4 times to a final volume of 8 µl. Five percent skim milk in PBST was prepared and the dried membrane was left to block O/N, rocking gently at 4°C. Blocking buffer was decanted and 20 ml anti-c-myc mAb (Sigma-Aldrich) in PBST diluted 1: 3,000 was added and incubated for 2 hours at room temperature with gentle rocking. The membrane was washed 3 times in PBST for 5 minutes each with gentle rocking. Goat anti-mouse IgG peroxidase (ThermoFisher) diluted 1:2,000 was added and incubated for 1 hour at room temperature with gentle rocking. The membrane was washed 3 times in PBST for 5 minutes each with gentle rocking. The final wash was removed and the membrane was left in PBS until detection. Enhanced

Chemiluminescence (ECL) Western Blotting Substrate (ThermoFisher) was prepared according to the manufacturer's instructions by mixing equal volumes of reagent A and B and incubating the membrane in the solution for 1 minute. The blot was viewed using a Molecular Imager VersaDoc (BioRad)

Western blot

Fifteen microlitres of S/N from the small scale expression was mixed with 5 µl of 4x SDS Protein loading buffer, heated for 5 minutes at 95°C and added to a 12% SDS-PAGE gel locked in an XCell SureLock Mini Cell electrophoresis chamber (ThermoFisher). SDS Running Buffer was used to fill the middle chamber and half the outside chambers. Five microlitres of unheated protein ladder (ThermoFisher) was added to a single lane. The loaded gel was run at 80 V for 20 minutes and then 120 V until the dye front reached the bottom of the gel.

A weigh boat was filled with 4°C western transfer buffer. Three pads were soaked in the buffer before being placed into the transfer cassette. Three Whatman Grade 3MM Chr Blotting papers (GE Healthcare) were soaked and added to the cassette. The completed gel was removed from the holder, trimmed and added to the cassette. Nitrocellulose membrane (GE Healthcare) was soaked and added on top of the gel. Three sheets of Whatman paper and pads were added before the lid was attached and the cassette was placed into the electrophoresis chamber. The middle and outer chambers were filled with western transfer buffer before placing the whole system into ice and running at 20 Amps for 2 h.

Five percent skim milk in PBST was prepared and the membrane was left to block O/N, rocking gently at 4°C. Blocking buffer was poured off and 20 ml anti-c-myc mAb (Sigma-Aldrich) in PBST diluted 1:3,000 was added. The primary antibody was allowed 2 hour to incubate at room temperature with gentle rocking. The membrane was washed 3 times in PBST for 5 minutes each with gentle rocking. Goat anti-mouse IgG peroxidase (ThermoFisher) diluted 1:2,000 was added and incubated for 1 hour at room temperature with gentle rocking. The membrane was washed 3 times in PBST for 5 minutes each with gentle rocking. The final wash was removed and the membrane was left in PBS until detection. ECL

Western Blotting Substrate was prepared as described above. The blot was viewed using a Molecular Imager VersaDoc.

Large scale expression

The best expressing clones from the small scale expression were selected for large scale expression. The glycerol stock was stabbed and placed into culture with 25 ml BMGY in a 250 ml baffled flask (Edwardsco), incubated O/N at 30°C rotating at 200 rpm. The 25 ml O/N culture was added to 225 ml fresh BMGY in a 2 L baffled flask (Edwardsco), incubated O/N at 30°C rotating at 200 rpm. The O/N BMGY culture was centrifuged briefly for 15 minutes at 1,500 *g* and resuspended to an OD of ~1 with 500 ml BMMY. The BMMY culture was incubated at 30°C for 72 h, rotating at 200 rpm. At 24 h and 48 h time points, 2.5 ml of 100% methanol (Sigma-Aldrich) was added to give a final concentration of 0.5% v/v. At 72 h the S/N was harvested by centrifuging the culture at 3,000 *g* for 15 min and filter-sterilised with a 0.22 µm filter (Millipore). The S/N was stored at 4°C until purification.

Protein purification

An ÄKTA Start (GE Healthcare) fast protein liquid chromatography (FPLC) system including lines A, B and sample were washed with 1 M Sodium hydroxide for 30 minutes. The NaOH was rinsed out for 10 minutes with dH_2O. A 5 ml HisTrap excel column (GE Healthcare) was attached and flushed with dH_2O for 10 minutes. Line A was placed into 4°C Buffer A, line B into 4°C Buffer B and the sample line into the 4°C sterile-filtered S/N from the large scale expression. The following settings were used for the protein purification: Flow rate – 5 ml/min, Equilibration volume – 10 ml, Sample volume – 500 ml, Wash column – 50 ml, Elution starting B concentration – 0%, Elution fixed fractionation volume – 3 ml, Elution target B concentration – 100%. Fifteen millilitre tubes (Greiner) were used for fraction collection. Three kDa Amicon Ultra-15 Centrifugal Filter Units (Merck Millipore) were used to concentrate protein-containing fractions from the large scale expression.

Endotoxin quantification and removal

Endotoxin concentration was quantified using a Pierce LAL Chromogenic Endotoxin Quantification Kit (88282, Thermo Fisher). The assay was carried out according to the manufacturer description. Absorbance was measured at 405-410 nm on a FLUOstar Omega plate reader (BMG Labtech).

If the sample contained detectable levels of endotoxin in the above assay (>0.1EU/ml), the following procedure was carried out to remove it. Triton X-114 (Sigma-Aldrich) was added to the protein sample to a final concentration of 1%. Samples were left at 4°C O/N on a tube rotator. Samples were incubated on a heat block at 37°C for 10 minutes before being centrifuged at 16,000 *g* for 5 minutes. The top layer was collected into a fresh microcentrifuge tubes. Samples were incubated on a heat block at 37°C again for 10 minutes before being centrifuged again at 16,000 *g* for 5 minutes. Samples were re-quantified to ensure the concentration of endotoxin was below the limit.

3.2.2 Enzyme-linked immunosorbent assays using human sera

The diagnostic potential of each SCP/TAPS protein was assessed by indirect IgG ELISA using a cohort of hookworm-positive (grouped according to World Health Organisation stratification: heavy infection - ≥4,000 eggs per gram of faeces (epg), n=10; moderate infection - 2,000-3,999 epg, n=11; or light infection - ≤1,999 epg, n=44), hookworm-negative (n=7) and hookworm-negative/*Ascaris*-positive (n=12) samples from a hookworm-endemic region in Minas Gerais, Brazil. Microlon high-binding 96-well microtiter ELISA plates (Greiner) were incubated overnight at 4°C with each antigen (2 µg/ml in 0.1 M Na_2CO_3/$NaHCO_3$, pH 9.6). Somatic extract of *N. americanus* L3 was similarly used as a positive control antigen. After washing with PBST the plates were blocked for 2 hours at room temperature with 100 µl of PBST/5% BSA. One hundred microlitres of sera (1:200 in PBST/1% BSA) were added to the wells and incubated overnight at 4°C, then the plates were washed with PBST and 100 µl of goat anti-human IgG-HRP (Sigma) (1:5,000 in PBST) was added to the plates. Plates were incubated for 2 hours at room temperature and then washed with PBST. Plates were developed by adding 100 µl of

3,3',5,5'-tetramethylbenzidine and incubating for 30 minutes at room temperature in the dark. One hundred microliters of 1 M HCl was added to stop the colorimetric reaction, which was read at a wavelength of 450 nm on a Polarstar Omega (BMG Labtech) microplate reader. The assay was performed in triplicate with two technical replicates and blank-corrected values were plotted using Graphpad Prism 7. The same software was also used to generate single and multi-antigen Receiver-Operating Characteristic (ROC) curves. Reactivity cutoffs were determined as the average plus 1 SD of the values of the hookworm-negative group and the frequency of recognition (FoR) of each antigen was determined using this cutoff. An unpaired Student's t test was used to compare differences in responses between groups *$P \leq 0.05$, **$P \leq 0.01$.

3.3 Results

3.3.1 Cloning

Selected cDNAs of interest

Eighteen cDNAs were selected for protein expression and were uploaded to IDT for synthesising (Table 3.1). The nucleotide sequences for each of these as well as their modifications, and translations are detailed in supplementary data 7.1.

Label	Protein name	SCP/TAPS domains	Length (nt)	Predicted size (Da)	emPAI
NP1	NECAME_01191	Single	579	21462.38	6.07
NP2	NECAME_06164	Double	1074	39572.86	1.11
NP3	NECAME_01366	Double	1332	48821.8	1.6
NP4	NECAME_01369	Double	1383	49434.93	2.05
NP5	NECAME_07493	Single	363	13224.36	15.78
NP6	NECAME_09334	Double	1389	50936.8	1.71
NP7	NECAME_07497	Single	612	22695.06	27.05
NP8	NECAME_15063	Double	1182	42595.39	16.92
NP9	NECAME_15467	Single	444	15919.3	1.2
NP10	NECAME_15644	Double	1143	41768.52	128.67
NP11	NECAME_16121	Double	1143	41745.98	0.5
NP12	NECAME_02290	Single	480	17203.11	0
NP13	NECAME_16708	Single	858	31436.05	5.55
NP14	NECAME_17155	Single	975	34766.46	0.84
NP15	NECAME_17854	Single	516	18953.28	0.55
NP16	NECAME_18970	Single	303	10413.93	50.11
NP17	NECAME_07690	Single	630	23129.11	4.13
NP18	NECAME_16753	Double	1107	41214.65	0.23

Table 3.1: Summary of the genes selected to clone and express.

Transformation into Top10 E. coli

All colonies transformed with IDT ampicillin resistant plasmid with respective gene of interests were successfully propagated on LB ampicillin (100 µg/ml) plates evident by numerous colonies the day after transformation and plating. An example is shown in Figure 3.2. The single colonies chosen from these plates for upscaling all grew well in 5 ml LB broth ampicillin (100 µg/ml) confirmed by visually inspecting the turbidity of the media.

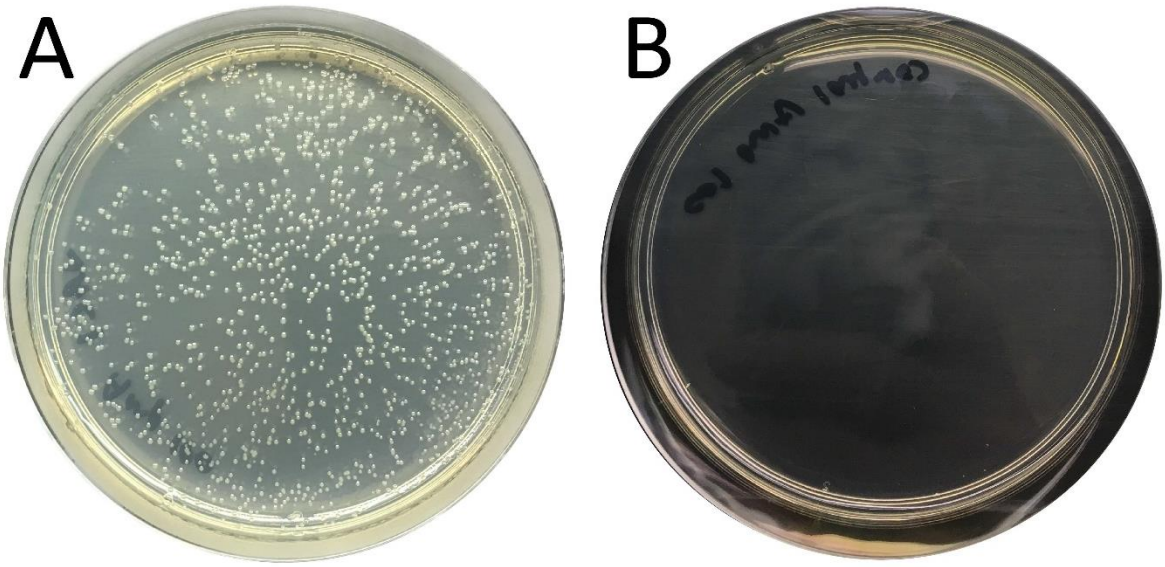


Figure 3.2: Top10 competent cells transformed with gene of interest in plasmid with ampicillin resistance. Transformed cells were *Escherichia coli* grown on a Lysogeny broth plate with ampicillin (100 µg/ml) (A). This is contrasted with a negative control (B) which includes the same cells without transformed resistance to ampicillin.

Digest for GOI

Concentrated DNA was run on an agarose gel to confirm the presence of and isolate each GOI insert for ligation into the expression vector. The uncut vs cut DNA is shown below in Figures 3.3 and 3.4. From the gel it can be clearly seen that all of the DNA preps were successfully digested with the exception of NP5. The uncut lanes of each different prep shows three distinct bands (supercoiled, coiled and nicked plasmid): one at approximately 3,000 bp, one at approximately 6,000 bp and the last above the highest molecular size standard. The cut lanes show three bands at different locations: one at the position reflective of the size of the GOI, another at approximately 2,500 bp (cut plasmid) and the last above the scale. A GOI band was not detected for NP5 in the digest lane (lane K) of Figure 3.3. In addition, the uncut banding pattern visible in lane J was unchanged, indicating that restriction digestion was unsuccessful. Digested samples separated using an agarose gel had inserts excised and purified using a Bioline Isolate II PCR and Gel Kit. These inserts were quantified using a Nanodrop 2000.

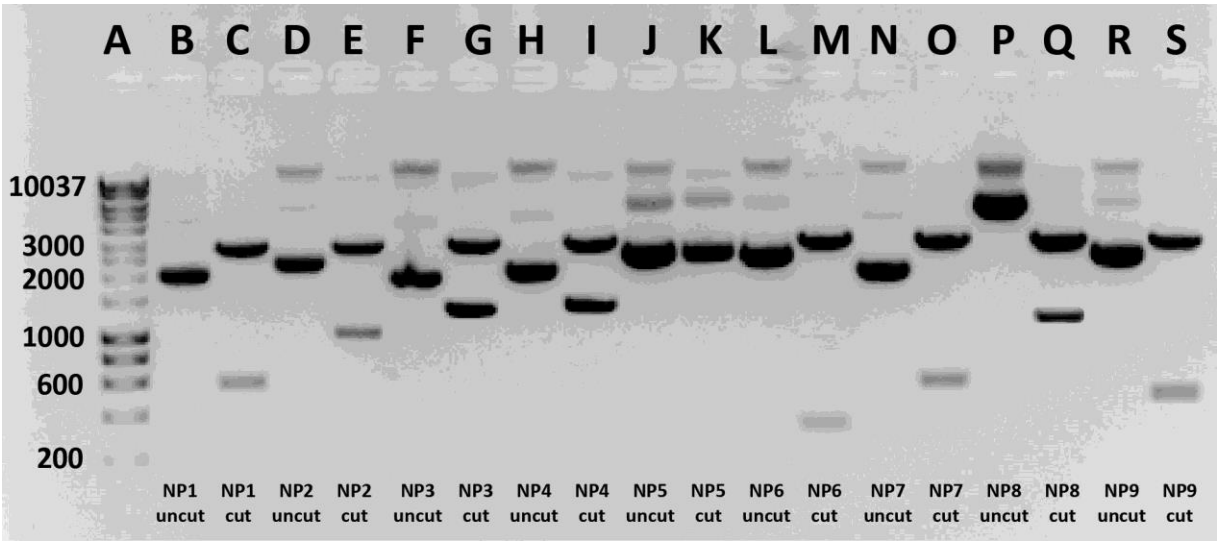


Figure 3.3: Comparison of uncut vs cut plasmid (1-9) with gene of interest. The 1% poured agarose DNA gel was visualised with ethidium bromide and a Molecular Imager VersaDoc (BioRad). Nine individual plasmids (1-9) were run across lanes B-S and the second of each pair was cut with restriction enzymes *Eco*R1 and *Xba*1. One hundred nanograms total DNA was loaded into each well.

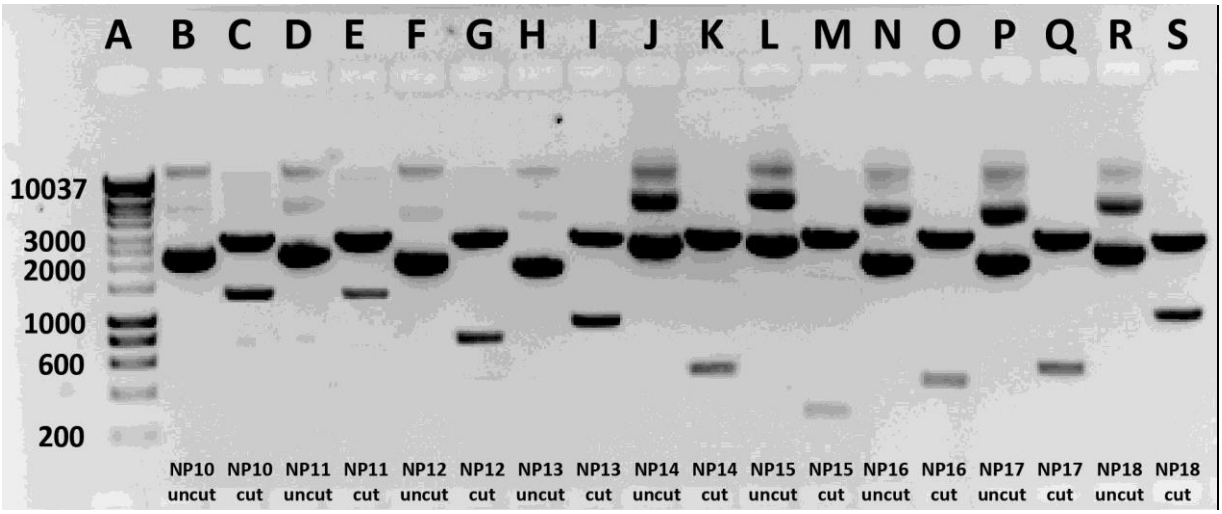


Figure 3.4: Comparison of uncut vs cut plasmid (10-18) with gene of interest. The 1% agarose DNA gel was visualised with ethidium bromide and a Molecular Imager VersaDoc (BioRad). Nine individual plasmids (10-18) were run across lanes B-S and the second of each pair was cut with restriction enzymes *Eco*R1 and *Xba*1. One hundred nanograms total DNA was loaded into each well.

Ligation into pPICZα

All colonies transformed with GOI + pPICZα expression vector were successfully propagated on LB Zeocin (100 µg/ml) plates evident by numerous colonies the day after transformation and plating. An example is shown in Figure 3.5. The single colonies chosen from these plates for upscaling all grew well in 5 ml LB broth Zeocin (100 µg/ml) confirmed by visually inspecting the turbidity of the media. The DNA was concentrated using a Nucleospin Plasmid Easypure kit.

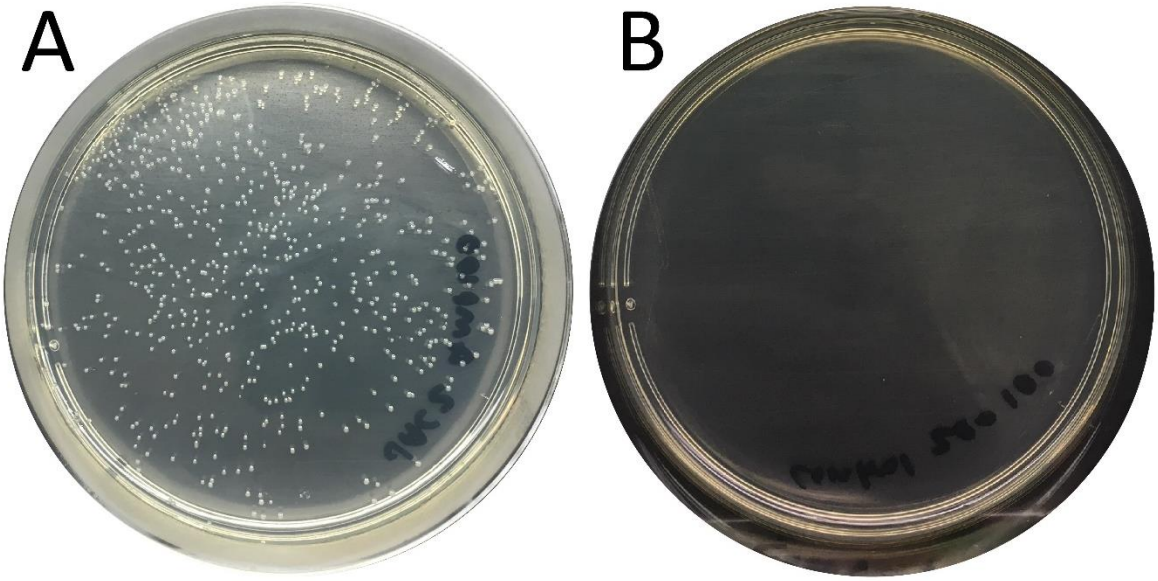


Figure 3.5: Top10 competent cells transformed with gene of interest in plasmid with Zeocin resistance. Transformed *Escherichia coli* were grown on a Lysogeny broth plate with Zeocin (100 µg/ml) (A). This is contrasted with a negative control (B) which includes the same cells without transformed resistance to Zeocin.

Linearisation for electroporation

Concentrated GOI in pPICZα was linearised with *Sac*1 and run on an agarose gel to confirm adequate linearisation. The circular vs linearised DNA is shown below in Figures 3.6 and 3.7. From the gel it can be clearly seen that all of the DNA preps were successfully linearised.

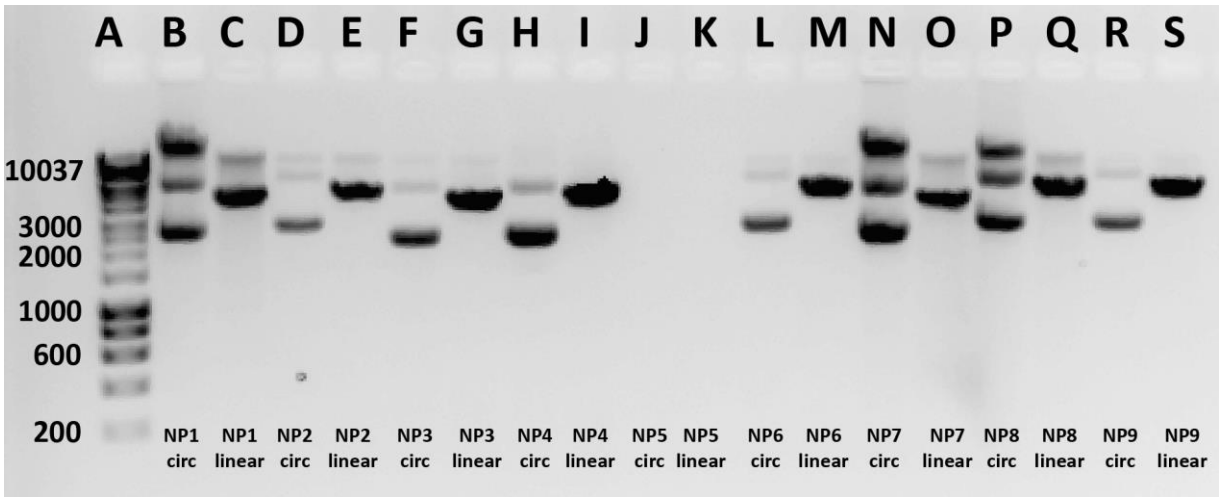


Figure 3.6: Comparison of circular vs linearised plasmid (NP1-9 excluding NP5) with gene of interest. The 1% poured agarose DNA gel was visualised with ethidium bromide and a Molecular Imager VersaDoc (BioRad). Eight individual plasmids were run across lanes B-S and the second of each pair was cut with restriction enzyme *Sac*1. One hundred nanograms total DNA was loaded into each well. Lanes J and K were intentionally left blank.

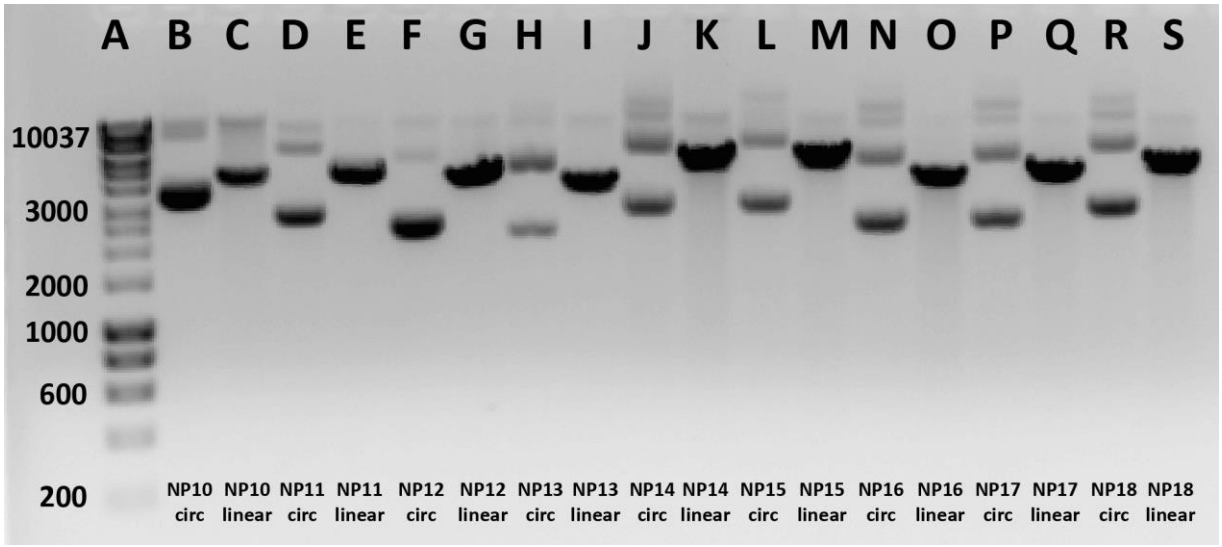


Figure 3.7: Comparison of circular vs linearised plasmid (NP10-18) with gene of interest. The 1% poured agarose DNA gel was visualised with ethidium bromide and a Molecular Imager VersaDoc (BioRad). Nine individual plasmids were run across lanes B-S and the second of each pair was cut with restriction enzyme *Sac*1. One hundred nanograms total DNA was loaded into each well.

Electroporation into P. pastoris

Approximately 5-10 *P. pastoris* colonies were observed for each of the GOIs on each respective plate 72 hours post electroporation and plating. An example of this growth is shown in Figure 3.8 with some individual colonies highlighted with black circles. The single colonies chosen from these plates for small scale expression all grew well in 5 ml YPD broth Zeocin (100 µg/ml) confirmed by visually inspecting the turbidity of the media.

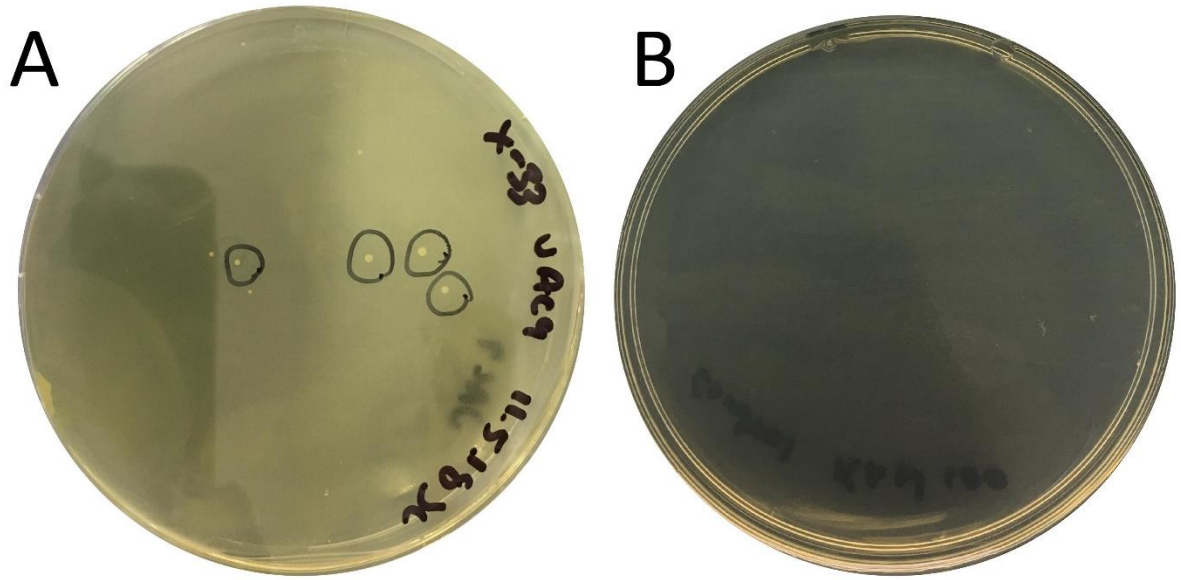


Figure 3.8: Electroporated yeast cells with gene of interest in pPICZα with Zeocin resistance. Transformed *Pichia pastoris* cells were grown on a Yeast Extract Peptone Dextrose Agar with Sorbitol (YPDS) plate with Zeocin (100 µg/ ml) (A). This is contrasted with a negative control (B) which includes the same cells without transformed resistance to Zeocin.

3.3.2 Protein expression

Small scale expression

Three colonies were selected from the electroporation plates for small scale expression, grown in 5 ml YPD Zeocin (100 µg/ml), and the supernatant tested using a dot blot with anti-c-myc mAb to detect

protein expression. The results for each protein are shown in Figures 3.8 and 3.9. The negative control protein, untagged BSA, was not detectable (I48, Figure 3.9). The positive control, an in-house produced and sequenced *N. americanus* c-myc tagged protein, elicited a strong signal (I72, Figure 3.9). NP1 (Row A, Figure 3.9) was the only protein for which expression was undetectable from the selected colonies. The remaining proteins NP2-9 (Figure 3.9) and NP10-18 (Figure 3.10) displayed varying levels of protein expression detected by dot blot. Across all proteins a strong trend was apparent for greater protein expression over time, with 72 h yielding the most intense signal for each of the tested colonies and proteins.

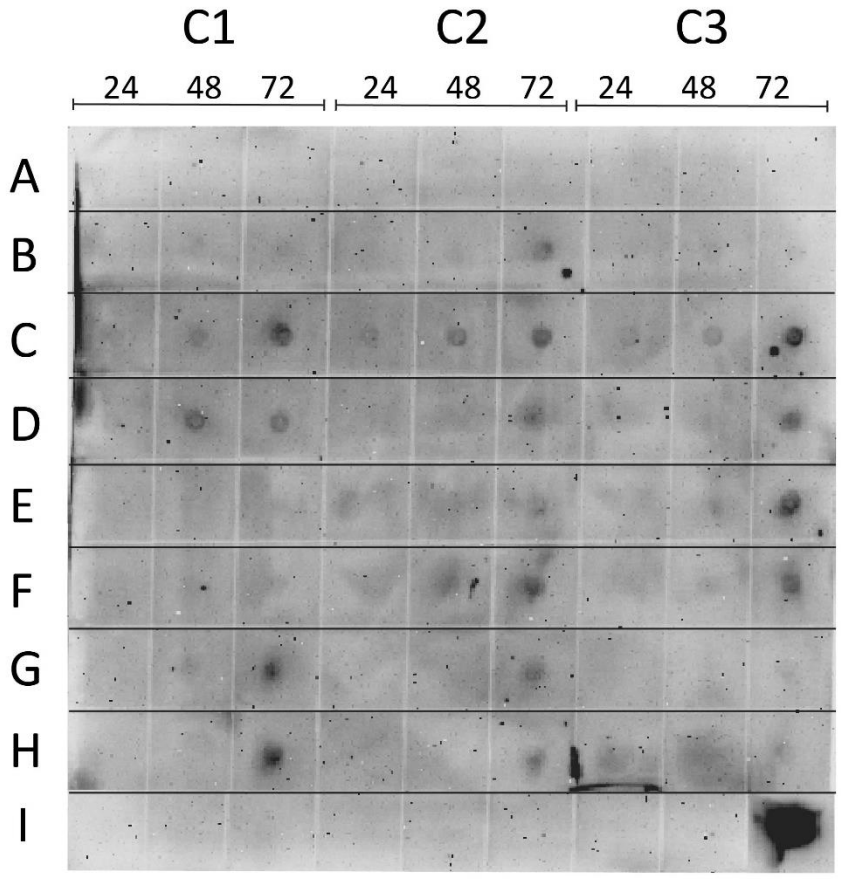


Figure 3.9: Dot blot of small scale expression of hookworm recombinant proteins (NP1-9, excluding 5) expressed in yeast. Image shows three separate colonies (C1-C3) at three time points (24 h, 48 h, 72 h) of eight different proteins expressed in *Pichia pastoris* (A-H). A-H represent NP1-9 respectively, excluding NP5. I72 is the positive control, an established c-myc tagged protein while I48 is a negative control – untagged BSA. Intensity of dot reflects relative protein concentration, probed using anti-c-myc mAb conjugated to horseradish peroxidase and visualised using a Molecular Imager VersaDoc.

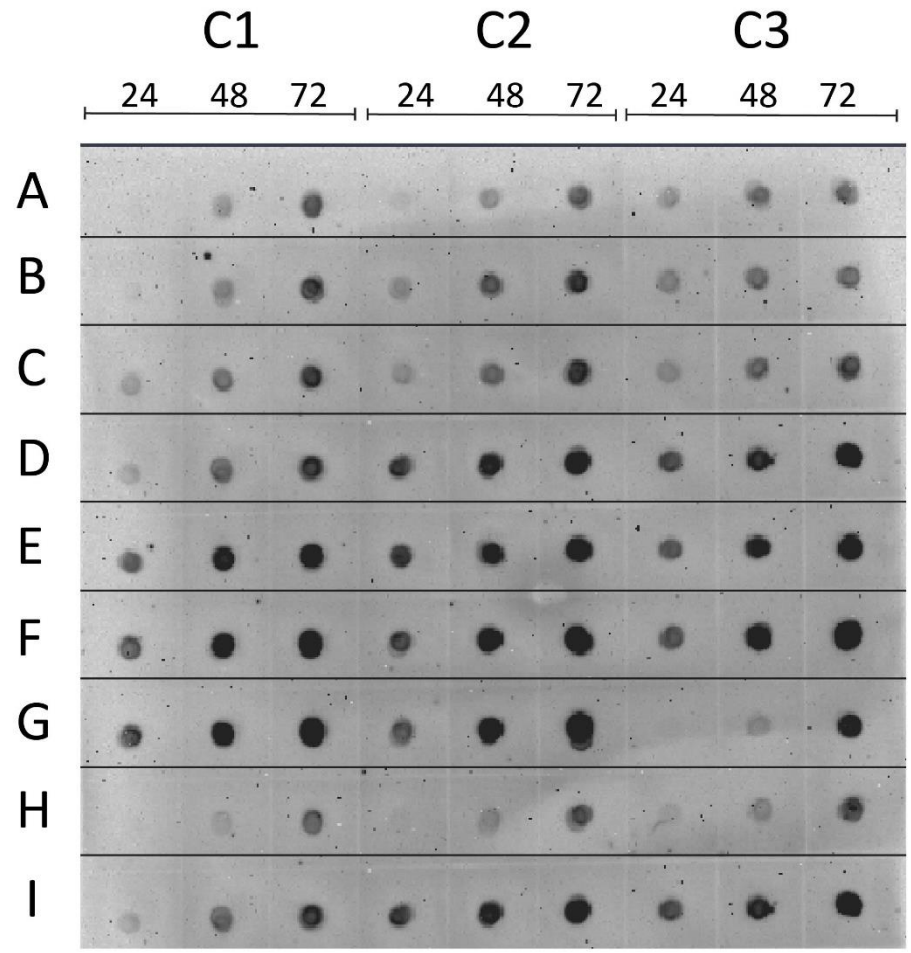


Figure 3.10: Dot blot of small scale expression of hookworm recombinant proteins (NP10-18) expressed in yeast. Image shows three separate colonies (C1-C3) at three time points (24 h, 48 h, 72 h) of nine different protein expressed in *P. pastoris* (A-I). A-I represent NP10-18 respectively. Intensity of dot reflects relative protein concentration, probed using anti-c-myc mAb conjugated to horseradish peroxidase and visualised using a Molecular Imager VersaDoc.

	Protein	Selected		Protein	Selected
Figure 3.9	A - NP1	N/A	Figure 3.10	A - NP10	C1 – 72 h
	B - NP2	C2 – 72 h		B - NP11	C2 – 72 h
	C - NP3	C3 – 72 h		C - NP12	C2 – 72 h
	D - NP4	C3 – 72 h		D - NP13	C3 – 72 h
	E - NP6	C3 – 72 h		E - NP14	C2 – 72 h
	F - NP7	C3 – 72 h		F - NP15	C3 – 72 h
	G - NP8	C1 – 72 h		G - NP16	C3 – 72 h
	H - NP9	C1 – 72 h		H - NP17	C3 – 72 h
	I - controls	N/A		I - NP18	C3 – 72 h

Table 3.2: The best expressing colonies and time points selected for large scale expression based on the dot blots from small scale expression.

Large scale expression

Successfully produced proteins from the small scale expression were taken forward to large scale expression to produce workable quantities. Proteins were purified over a HisTrap excel column and fractions were run on a SDS-PAGE gel to assess purity. Figure 3.11 shows an example of the elution profile of recombinant NP17.

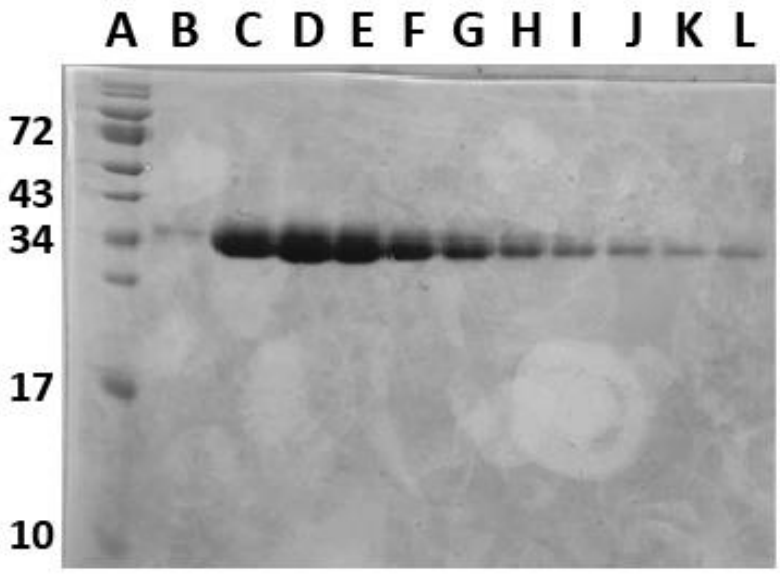


Figure 3.11: Coomassie stained SDS-PAGE gel showing individual fractions of concentrated protein (NP17). The protein was purified using an ÄKTA Start fast protein liquid chromatography system and HisTrap excel column. The concentration of protein, evident from the gel, decreases with each progressive fraction from the ÄKTA. Lane A is the first elution fraction and each consecutive lane is the successive fraction with increasing levels of imidazole.

Each protein was confirmed for correct size by running concentrated fractions on a SDS-PAGE gel. Figure 3.12 is an SDS-PAGE Coomassie stained gel showing NP1-NP9 (C-L). Only proteins NP3 and NP4 were able to be detected via SDS-PAGE. Lane B was a positive control (BSA) present at the correct size/location of approximately 66 kDa.

Figure 3.13 shows the results for NP10-18 (C-L). From the SDS-PAGE gel it can be seen that NP12 and NP17 produced detectable protein, while the remaining proteins were undetectable. These protein bands, were confirmed by Western blot (Figure 3.14).

To test the validity of these results a number of factors were explored as detailed below:

1. All colonies from the original electroporation were trialled for large scale expression
2. Electroporation was repeated and the resultant colonies also tested
3. Large scale expression conditions were explored including varied MeOH concentrations, expression volume, different supernatant collection time points
4. Large scale supernatants were checked pre- and post- purification to determine if there was a column binding issue
5. Cells were lysed to check for a protein secretion issue
6. Both c-myc and histidine tags were explored through Western blotting

Despite this troubleshooting (data not shown) none of the modifications resulted in purification of sufficient protein.

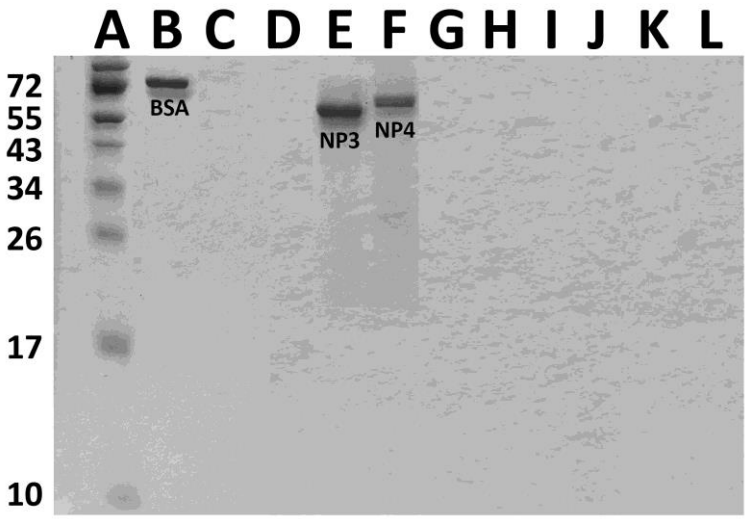


Figure 3.12: Coomassie stained SDS-PAGE gel showing individual concentrated proteins NP1-9 (C-K). The protein was purified using an ÄKTA FPLC workstation and subsequently concentrated before running on the gel. Only proteins NP3 (lane E) and NP4 (lane F) returned detectable protein quantities. Lane B is a positive control, BSA (1 µg), at the correct size of approximately 66 kDa.

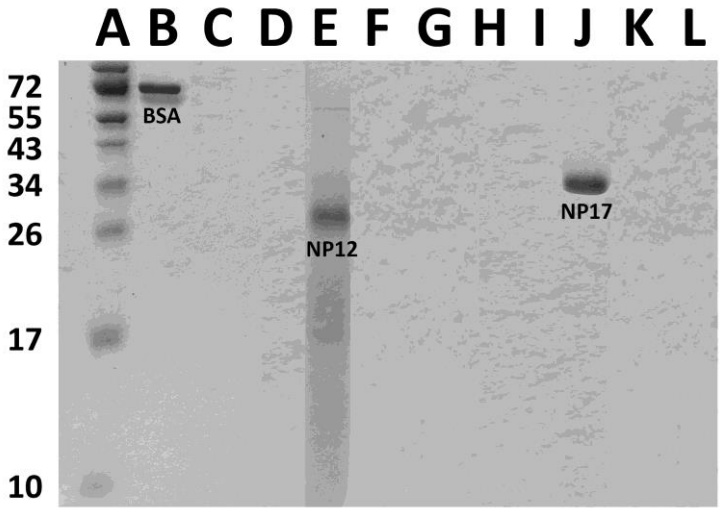


Figure 3.13: Coomassie stained SDS-PAGE gel showing individual concentrated proteins NP10-18 (C-K). The protein was purified using an ÄKTA FPLC workstation and subsequently concentrated before running on the gel. Only proteins NP12 (lane E) and NP17 (lane J) returned detectable protein quantities. Lane B is a positive control, BSA (1 µg), at the correct size of approximately 66 kDa.

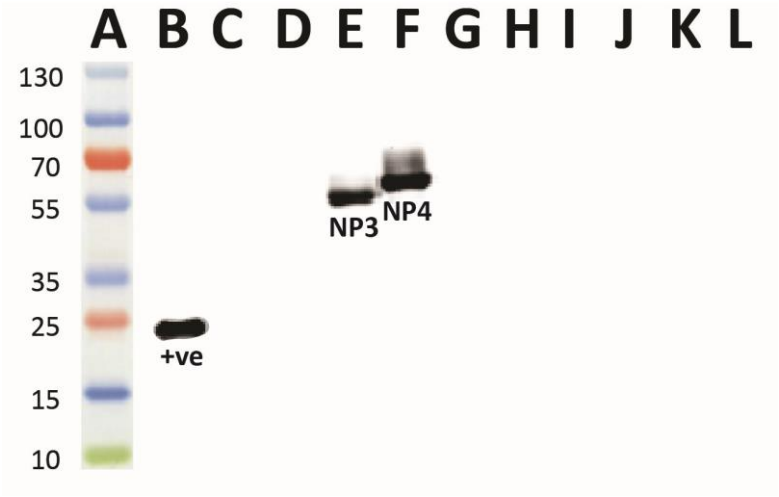


Figure 3.14: Western blots showing proteins NP1-9 (lanes B-K). The protein was purified using an ÄKTA FPLC workstation and subsequently concentrated before running on the gel and carrying out the transfer. Only proteins NP3 (lane E) and NP4 (lane F) returned detectable protein quantities. Lane B is a positive control, a protein with a known c-myc tag, at the correct size ~24 kDa. Blot was visualised using a Molecular Imager VersaDoc.

The protein was purified using an ÄKTA and subsequently concentrated before running on the gel and carrying out the transfer. Lane A is a positive control, a protein with a known c-myc tag, at the correct size ~24 kDa. Blot was visualised using a Molecular Imager VersaDoc.

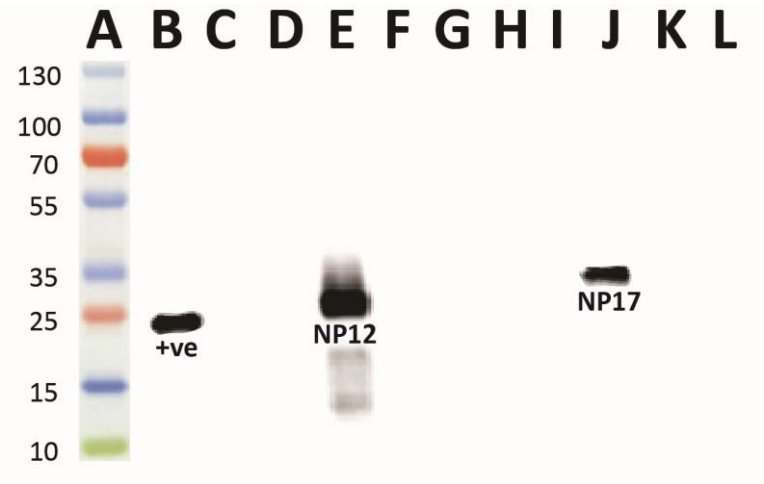


Figure 3.15: Western blots showing proteins NP10-18 (lanes B-K). The protein was purified using an ÄKTA FPLC workstation and subsequently concentrated before running on the gel and carrying out the transfer. Lane B is a positive control, a protein with a known c-myc tag, at the correct size ~24 kDa. Blot was visualised using a Molecular Imager VersaDoc.

3.3.3 Immunogenicity and diagnostic potential of SCP/TAPS proteins

Using the recombinant *N. americanus* SCP/TAPS proteins expressed in this study (and *N. americanus* L3 extract as a comparator), antibody responses in the serum of *N. americanus*-infected individuals were profiled against each recombinant protein (Figure 3.16A-E) and the AUC generated from the ROC curves were used to determine the sensitivity and specificity of each antibody response, and the predictive value of infection (Table 3.3). Antibodies to all antigens were significantly reactive in subjects with a heavy hookworm infection (≥4,000 epg) compared to hookworm-negative controls. Further anti-NP3 and anti-NP4 IgG responses were also significantly elevated in subjects with a light hookworm load (≤1,999 epg) compared to uninfected individuals. For the L3 extract and three of the recombinant antigens, IgG reactivity in the hookworm-negative/*Ascaris*-positive group was higher (not significantly so) compared to the hookworm-negative group with no *Ascaris* infection, which is likely explained by cross-reactivity between *N. americanus* and *Ascaris* antigen. Of the recombinant antigens tested, the highest positive predictive value of infection (PPV) in hookworm-positive vs hookworm-negative subjects was generated by the antibody response to NP4 (AUC = 0.77). Hookworm-positive subjects with antibody levels against a particular antigen above the reactivity cut-off were recorded as positive for recognition of that antigen and, because the recognition pattern of each antigen among the infected populations was different, a combination of antigens was assessed to determine the frequency of recognition (FoR) and/or PPV among hookworm-positive subjects compared to any one antigen alone. A minimum combination of three SCP/TAPS proteins (NP4 + NP3 + NP17) increased the FoR to 82% of hookworm-positive subjects, compared to an FoR of 60% and 66% for NP4 and L3 extract,

respectively (Table 3.3). Further, this antigen combination gave a PPV of 0.83, which was comparable to 0.86 obtained for the L3 extract.

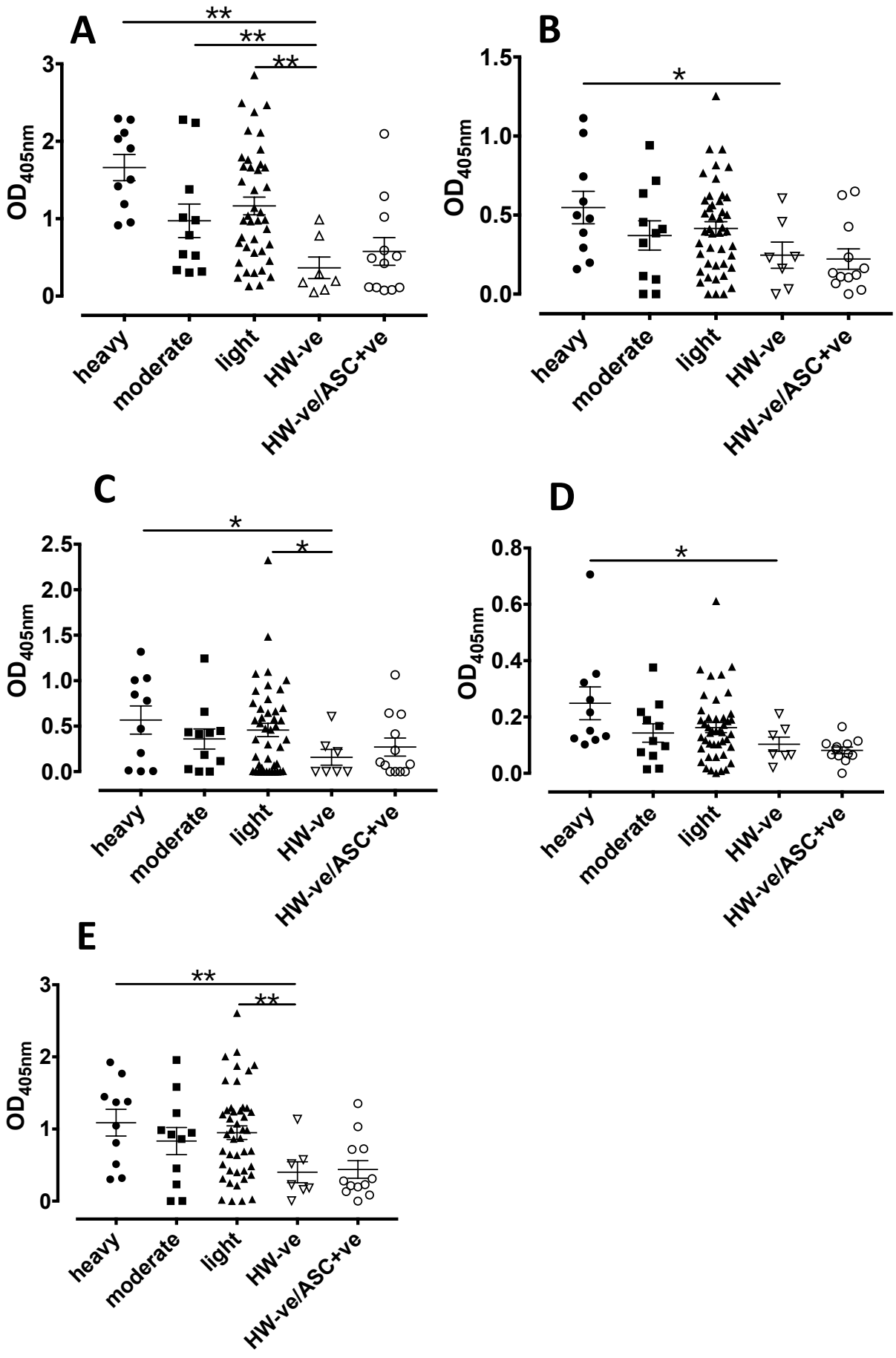

A
B
C
D
E
OD405nm
heavy
moderate
light
HW-ve
HW-ve/ASC+ve
*
**

Figure 3.16: Serological diagnosis of hookworm infection by hookworm extract and recombinant SCP/TAPS proteins. Scatter plots showing the (A) anti-*N. americanus* L3 extract IgG response, (B) anti-NP12 IgG response, (C) anti-NP3 IgG response, (D) anti-NP17 IgG response (E) anti-NP4 IgG response. Hookworm-positive subjects were characterized (WHO stratification) as either having a heavy (≥4,000 epg), moderate (2,000-3,999 epg) or light (≤1,999 epg) infection. "HW -ve" = hookworm egg negative subjects from an endemic area, "HW -ve/ASC +ve" = hookworm egg negative, *Ascaris* egg positive subjects from an endemic area. Differences in responses between groups were analysed by Student's *t* test. **P*≤0.05, ***P*≤0.01. The recognition cutoff was defined as the average response of the HW -ve group + 1SD. # represents the minimal antigen combination (NP4, NP12 and NP17) which gives the highest frequency of recognition. AUC = area under receiver operator curves generated for each antigen and the antigen combination, which determines the positive predictive value of infection.

Antigen	FoR (%)[$]	AUC[&]
N. americanus L3 extract	66	0.86
NP12	43	0.68
NP3	52	0.74
NP17	41	0.64
NP4	60	0.77
NP3 + NP4 + NP17[#]	82	0.83

the minimal antigen combination (NP4, NP12 and NP17) which gives the highest frequency of recognition.

$ the recognition cut-off was defined as the average response of the HW negative group + 1SD.

& area under receiver-operating characteristic curves generated for each antigen and the antigen combination, which determines the positive predictive value of infection.

Table 3.3: Frequency of recognition and area under receiver-operator characteristic curves determined for the human hookworm L3 extract and recombinant SCP/TAPS proteins in the diagnosis of hookworm infection.

3.4 Discussion

In this chapter, *P. pastoris* was utilised for recombinant *N. americanus* protein expression. This system was chosen for a number of factors including: growth rate, high cell density, capacity to produce large quantities of secreted protein, as well the fact that it was a well-established method within our research group [203]. Eighteen proteins were originally chosen for a number of reasons: firstly because I expected that some would not express in this system; second because it was an ambitious but manageable target within the scope of this project; third because these eighteen gave a good representation based on phylogeny and abundance for the greater number of SCP/TAPS proteins. Despite the widespread use of this method of protein production, many proteins of interest are unable to be secreted by or produced in *P. pastoris* [290]. In total, 4/18 selected proteins of interest from the *N. americanus* secretome were successfully produced in this study. The main factors explored to improve this rate were culture conditions such as methanol concentration and collection time point as well as colony selection for varied gene copy number. Increasing methanol concentration has been shown to improve protein production rate and final concentration [291], but despite testing methanol total concentrations of up to 3% (data not shown), usable quantities of protein could not be produced for 14/18 proteins assessed. Many colonies were also screened for each individual protein of interest. In many cases it has been demonstrated that higher gene copy number results in increased product yield [292,293]. However, this trend is not always the rule and in fact it may be protein dependent. High copy number dosage of recombinant secretory proteins was demonstrated to overload host cell physiology and limit secretion of desired proteins [294-296]. In this study, the amount of protein secreted by each colony seemed to vary based on the individual colony as seen from the small scale expression dot blot. For example, colony 1 (C1) from NP9 (H) showed a substantial amount of c-myc tagged protein at 72 h when compared to colony 3 at the same time point. This disparity is likely to be due to differences in the uptake of GOI plasmid during electroporation.

The time point of supernatant collection was shown to be important in maximising the amount of protein produced. From the dot blot it can be clearly seen that most protein was present after 72 h of culture time. It may be beneficial in future studies to extend culture time beyond the 72 h time point, as protein concentration in the supernatant may continue to increase over time. Another aspect that may be explored is the temperature of induction. Lower-temperatures have been shown to reduce proteolytic degradation while limiting oxygen stress [297].

Looking back at the results of chapter 2 reveals a number of points about the four successfully expressed SCP/TAPS. Firstly, I ended up expressing two single and two double domain proteins. This grouping provides a point of difference for the proteins in the chapters to follow and may highlight a functional difference specific to domain number. Secondly, each of the proteins was found to have significant blast scores from an SCP/TAPS protein from all three compared species in section 2.3.3 in chapter 2. This emphasises that these proteins are conserved among similar species and are likely to play related roles.

In a demonstration of the potential translatability of the "omics" data generated herein, four of the SCP/TAPS proteins identified in this study were expressed in recombinant form and their potential to diagnose hookworm infection was assessed using sera from human subjects resident in a hookworm-endemic area of Brazil. Serum IgG levels to each antigen were measured in individuals who were positive for the presence of parasite eggs in the faeces, the current gold standard for diagnosis. When the diagnostic performance of the antigens was assessed individually, each molecule was found to be an accurate predictor of infection as each cognate antibody response generated an AUC value above 0.5, where an AUC of 1.0 is a perfectly accurate test [298]. The L3 extract was used as a control because obtaining adult ES and/or adult somatic extracts is very difficult. Regarding individual recombinant antigens, NP4 displayed the highest FoR (60%) among the infected population but this value was increased by over 20% when the FoR of a combination of antigens (NP4, NP3 and NP17) was calculated, due to the varying recognition patterns of each antigen among hookworm-positive individuals. Further,

this antigen combination had an FoR that exceeded, and an AUC that was comparable to, that of the L3 extract, which could be considered the gold standard for serodiagnosis of hookworm disease, given the widespread use of parasite extracts as markers of helminth infection. The negative correlation of antibody titer and the infection intensity makes these antigens potentially useful vaccine targets.

While the PPV of infection of the L3 extract slightly exceeded that of the antigen combination (understandable given the complexity of the preparation, and so it's ability to capture the breadth of the anti-hookworm immune response) the advantages of a defined, recombinant antigen preparation lies in its reproducibility and sustainably as a resource, compared to, especially in the case of human hookworm, the limitations of generating sufficient reagent for use.

4 Chapter 4 – Assessing the therapeutic properties of hookworm SCP/TAPS proteins in mouse models of inflammatory diseases

4.1 Introduction

IBD and asthma are immune-mediated diseases which may share similar pathology as well as common environmental and genetic risk factors [299]. Neither condition can be cured and currently available treatment options are not effective for all patients. Both of these conditions have a large burden of disease in Australia and worldwide, and therefore represent a significant unmet need in the lives of many people. The 'old friends' hypothesis contributes to the rise of inflammatory diseases in increasingly urban settings by the decreased exposure of people to infectious organisms. The human hookworm is one such organism which has been negatively correlated with a number of allergic and auto-immune diseases including IBD and asthma [146]. Given that established mouse models of both of these diseases were operating in our lab, it was logical to employ them in exploring the potential therapeutic effects of ES molecules from the human hookworm. Mouse models of inflammatory diseases are valuable tools for exploring the efficacy of potentially novel therapeutics. In the case of both IBD and asthma, many models of each disease exist, each with various advantages and disadvantages that should be considered [300,301].

Animal models of intestinal inflammation have been used for decades to better understand disease pathogenesis as well as test novel therapeutics. The complexity of human IBD dictates that no single model is adequate to replicate all aspects of the disease. Therefore a number of models are commonly used which include knockout (KO) mice, chemically-induced models such as dextran sulfate sodium (DSS) colitis and 2,4,6-trinitrobenzene sulfonic acid (TNBS) colitis, as well as adoptive transfer colitis. Each of these models tends to be employed to study a particular mucosal immune function/therapy,

and range from acute models where disease is induced in just days to chronic models that take months [302].

For TNBS colitis an intrarectal (i.r.) injection of the chemical TNBS is given with ethanol to one of a number of susceptible mouse strains [303]. The ethanol plays an essential role in permeating the intestinal barrier, allowing the TNBS to infiltrate further [303]. TNBS is thought to haptenise colonic and microbiota proteins, rendering them immunogenic. This initiates a mucosal immune response from the host that is primarily driven by Th1 cells and is characterized by infiltration of neutrophils, CD4+ T cells, and macrophages [302,303]. The cytokine response that drives the inflammation in this model has been found to depend on the microbiota and genetic strain of mouse used [301,304]. The effects of TNBS on mice include diarrhoea, weight loss, and inflammation of the colon with ulceration and oedema [303]. These clinical manifestations are a key strength of the TNBS colitis model as they overlap significantly with the human disease [302]. While there are advantages to using other, more chronic models, the acute TNBS model means many proteins can be screened for potential therapeutic effects in a short time frame.

In the case of asthma the models differ with respect to type, duration and dose of allergens used to stimulate allergic responses in the airways. The most commonly used allergens are ovalbumin (OVA), cockroach antigen, moulds and house dust mite (HDM) [305-307]. Utilising a variety of asthma-inducing stimuli is useful for studying different asthma subtypes as well as varying aspects of the disease. These factors, along with the timeline chosen mean that some models more closely mimic human sensitisation and disease pathology, while others are more cost effective and higher throughput.

OVA is the most widely used allergen for the study of asthma. It is produced from chicken eggs in large quantities and is effective at inducing intense allergic airway inflammation in mice [300,308]. Despite its widespread use, OVA does not induce pulmonary inflammation in humans [308]. This fact alone has meant it has been scrutinised in terms of relevance to human disease. HDM has become an increasingly

clinically-relevant allergen as exposures are common in the home and up to 85% of asthmatics are allergic to HDM [309].

Although different allergens are often employed, the same basic model of a sensitization and a challenge phase tends to stay consistent. The timeframe of these phases dictates the clinical manifestations. For example, acute models of allergic asthma are often induced in less than 3 weeks. This type of model has been shown to successfully reproduce many features of clinical asthma including airway inflammation, increased IgE, epithelial hypertrophy, goblet cell hyperplasia and airway hyperresponsiveness (AHR) [300,310]. However, when compared to chronic models there are also some drawbacks. Many of the airway lesions and remodeling associated with chronic asthma are missing. In addition, it has been noted that the key features of the acute model such as AHR and inflammation tend to resolve within a few weeks of the final challenge [311]. Acute models are beneficial to use in preliminary tests such as this study where novel therapeutics are being tested for the first time. Chronic models involve sensitisation windows of up to 12 weeks, and are therefore much less cost effective but better represent human disease pathology for the aforementioned reasons [300]. One of the key criticisms of animal models of asthma is that they do not mimic the way that humans develop allergic responses [308]. Allergic asthma in humans involves repeated exposure to aeroallergens through ventilation, which leads to airway inflammation over time. Many models of asthma have traditionally involved intraperitoneal or subcutaneous administration of allergens such as OVA to sensitise animals. To combat this disparity the use of intranasal and intratracheal instillations have become increasingly common [307,312].

Dosage and route of administration for this study were selected based on similar experiments from the literature. A single intraperitoneal dose (1 mg/kg) of the hookworm recombinant protein *Ac*-AIP-1 significantly limited pathology in a TNBS mouse model of colitis [313]. Similarly, multiple doses of *Ac*-AIP-2 (1 mg/kg) during the challenge phase, suppressed airway inflammation and induced regulatory responses in a mouse model of asthma [203].

It is important to highlight that no animal model perfectly reflects the complex interaction of genetic and environmental factors that manifests in human disease. Given this, determining the appropriate model needs to be based on the desired outcome. For studies such as this thesis where the goal was to screen many proteins, the timeframe of models was of primary concern. As both the acute HDM model and the acute TNBS model include many of the fundamental features of their respective diseases, they were selected for this study.

4.2 Methods

4.2.1 Ethics

Ethical approval was obtained from the James Cook University Animal Ethics Committee. Two separate applications were approved: (1) TNBS chemically induced colitis model (ID#A2305) and (2) HDM acute asthma model (ID#A2307).

4.2.2 TNBS colitis

An overview of the procedure employed for TNBS colitis is presented in Figure 4.1.

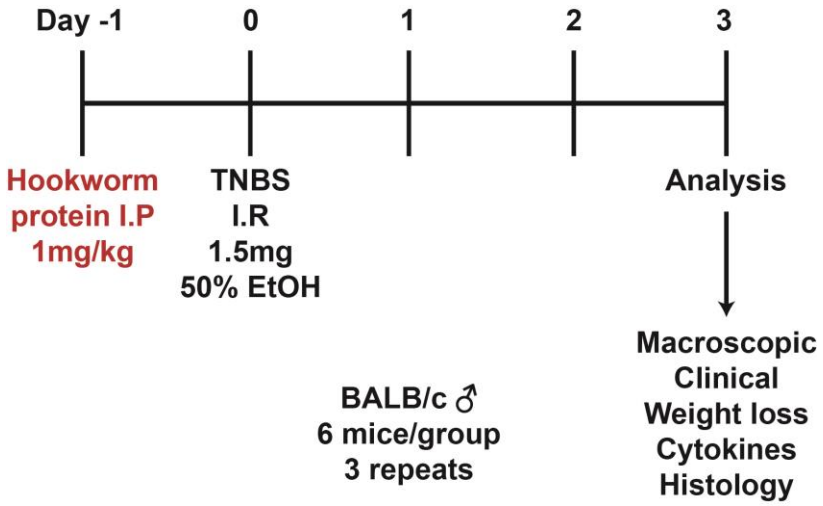


Figure 4.1: An overview of the experimental design for TNBS colitis including mouse strain, treatment concentrations and outcomes. i.p. – Intraperitoneal; i.r. – Intrarectal

Mice

6-week-old BALB/c mice were ordered from the Animal Resources Centre, weight matched (18-19 g), and randomised into groups of six animals per cage. All experiments were repeated three times with six mice in each group. Mice were housed in the JCU animal facility with a 12 hour light/dark cycle, maintained at 22°C with unrestricted access to food pellets and water as per the Australian animal rights and regulation standards.

Treatments

Twenty-four hours prior to TNBS administration, trial compounds in sterile PBS were delivered intraperitoneally (i.p). at a dose of 1 mg/kg to six mice per group. Eight experimental groups were included in the study: (1) healthy naïve; (2) PBS vehicle (colitis control); (3) sulfasalazine (drug control); (4) recombinant human albumin (irrelevant recombinant protein control); (5) NP3; (6) NP4; (7) NP12; and (8) NP17. A 500 mg sulfasalazine tablet was crushed with a mortar and pestle to a fine powder before reconstitution in dH_2O to a final concentration of 20 mg/ml. Two milligrams in 100 µl was administered to mice using an oral gavage. Recombinant NP proteins (produced from chapter 3) that had been frozen at -80°C were electrophoresed on a SDS-PAGE gel prior to administration to check for protein degradation. Endotoxin was confirmed to be below the level of detection using a Pierce LAL Chromogenic Endotoxin Quantification Kit.

Induction of colitis

Mice were anaesthetised using xylazine (5 mg/kg, Bayer, Germany) and ketamine (50 mg/kg, Pfizer Inc., USA). 2,4,6-Trinitrobenzenesulfonic acid (TNBS; Sigma-Aldrich, USA) was prepared by dissolving 1.5 mg in 50% ethanol. Unconscious, unresponsive mice were given an enema with 125 mg/kg TNBS with a lubricated 20-G soft catheter (Terumo, Japan). TNBS was delivered slowly, in 10 µl increments to ensure minimal leakage. Upon delivery of the TNBS mice were inverted for 10 minutes to allow for the TNBS to be adequately distributed.

Monitoring and scoring

Animals were examined daily for clinical signs including weight loss, mobility, faecal consistency/bleeding and piloerection. Clinical scores were carried out as per below and performed by the same person at time points 0, 24 h, 48 h and 72 h (before sacrificing).

Clinical scoring measures:

Mobility:

Normal = 0; Lethargy = 1; Motionless/sickly = 2

Fur:

Normal = 0; Mild piloerection = 1; Severe piloerection = 2

Faeces:

Normal = 0; Mild diarrhoea/soft stool = 1; Bloody/liquid stool = 2

Mice were euthanised (CO_2 asphyxiation) 72 h post administration of TNBS, and the colons removed from the caecum to the rectum. Blinded macroscopic scoring to assess the extent of colitis was carried out on the colon as per below, performed by the same person for every colon in each group to ensure consistency. Colons were imaged and length recorded.

Macroscopic scoring of colon:

Adhesions:

Normal = 0; Minimal adhesions = 1; Adhesions = 2; Severe Adhesions = 3

Ulceration:

None = 0; Minimal ulceration (1 ulcer) = 1; Multiple Ulcers = 2; Severe ulceration/necrosis = 3

Bowel wall thickening:

Normal = 0; Slight thickening in places = 1; Moderate thickening = 2; Severe thickening throughout = 3

Mucosal oedema:

Normal = 0; Slight oedema =1; Moderate oedema = 2; Severe oedema = 3

Histology and cytokine quantification

Pieces of colon were removed in 0.5 cm segments for histological evaluation of inflammation as well as cytokine production. Tissue samples were placed in 4% formalin for fixation before being transferred to 70% ethanol for storage. Paraffin was used to embed the tissue before it was sectioned longitudinally and mounted on a microscope slide. Slides were stained with hematoxylin and eosin (H&E). Scoring of pathology was performed from H&E stained slides by the same individual blinded to the experimental group. Tissue sections were scored on a scale of 0-5 for each of the following parameters; (1) epithelial pathology (crypt elongation, hyperplasia, and erosion), (2) mural inflammation and (3) oedema for an overall maximum total histology score of 15.

Tissue homogenate for measuring cytokine levels was placed into a 2 ml tube containing 500 µl sterile PBS and a stainless steel homogenisation bead. The weight of the tube was measured pre- and post-addition of the tissue to obtain the net tissue weight. The tube containing the tissue, PBS and bead was lysed using a TissueLyser II (Qiagen) at a frequency of 25 for 5 minutes. Samples were centrifuged at 16,000 *g* in a pre-cooled microfuge for 10 minutes before supernatant was collected and stored at -80°C. Supernatant was assayed for IL-6, IL-10, MCP-1, IFN-γ, TNF, and IL-12p70 proteins using a Cytometric Bead Array (CBA) Mouse Inflammation Kit (BD, 552364) as per the manufacturer's instructions.

4.2.3 HDM asthma

An overview of the procedure employed for HDM asthma is presented in Figure 4.2.

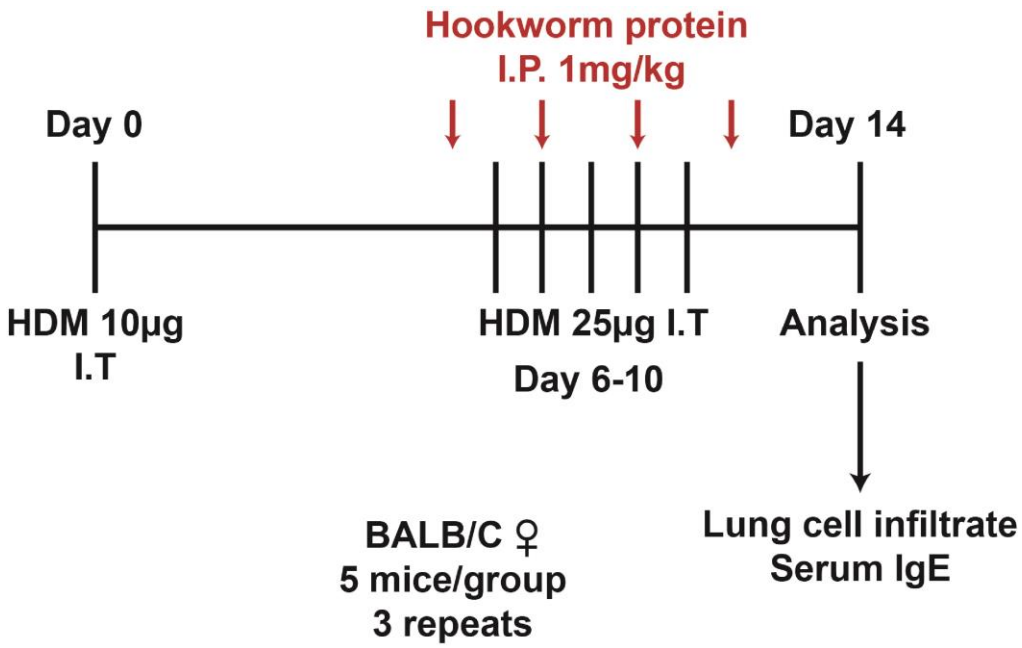


Figure 4.2: An overview of the experimental design for the house dust mite (HDM) asthma model including mouse strain, treatment concentrations and outcomes.

Mice

Mice used for asthma experiments were the same in all regards to the TNBS trials above except they were not weight matched and there were five animals per group. This number of mice was deemed adequate to generate enough statistical power based on previous experiments. Age, handling and housing conditions were all identical as described.

Intranasal vs intratracheal injections

Three mice were assigned to one of three groups: (1) naïve; (2) i.n. food colouring; (3) i.p. food colouring. Mice were anaesthetised with 5% isofluorane for 5 minutes using an isoflurane vaporizer (VetEquip). Group 1 mice received no injection, group 2 received 30 µl i.n. of 1:5 PBS diluted blue food colouring (Queen), group 3 received 30 µl i.t. of 1:5 PBS diluted blue food colouring. All mice were

immediately euthanised (CO_2 asphyxiation) and lungs were removed. Lungs from mice in groups 2 and 3 were compared for the presence and spread of blue dye throughout the lungs.

Treatments

On days 5, 7, 9 and 11 trial compounds in sterile PBS were delivered i.p. at a dose of 1 mg/kg to five mice per group. Eight experimental groups were included in the study: (1) healthy naïve; (2) PBS vehicle (asthma control); (3) dexamethasone (drug control); (4) recombinant human albumin; (5) NP3; (6) NP4; (7) NP12; and (8) NP17. Dexamethasone (1 mg/kg) was administered using an oral gavage in a total volume of 100 µl. Recombinant NP proteins (produced from chapter 3) were electrophoresed on a SDS-PAGE gel prior to administration to check for protein degradation. Endotoxin was confirmed to be below the level of detection using a Pierce LAL Chromogenic Endotoxin Quantification Kit.

HDM-induced asthma

On day 0 mice were anaesthetised with 5% isofluorane for 5 minutes using an isoflurane vaporizer. Groups 2-8 were sensitised with 10 µg house dust mite (HDM) *Dermatophagoides pteronyssinus* extract (Citeq Biologics) i.t. in 30 µl PBS. Beginning on day 6, mice were challenged with 25 µg HDM i.t. once per day for 5 consecutive days (days 6-10). On day 14 mice were euthanised (CO_2 asphyxiation) and blood was immediately collected from the heart using a 23G needle. A bronchoalveolar lavage (BAL) was performed by opening up the trachea, inserting a blunt 18G needle and carrying out 3 x 1 ml washes with PBS. The total BAL wash (3 ml) was stored immediately on ice for subsequent flow cytometry analysis. Upon completion of the dissection blood samples were centrifuged at 10,000 *g* for 10 minutes at room temperature before being stored at -80°C. BAL samples were processed immediately for flow cytometry.

Serum IgE

IgE was measured using the Mouse Ready-SET-Go! ELISA kit (LS885046077, EBioscience) as recommended by the manufacturer. In brief, plates were coated with coating buffer and incubated

overnight at 4°C. Plates were washed 3 times with wash buffer before blocking buffer was added and incubated overnight at 4°C. The plates were washed 3 times before samples (diluted 1:10) and standards were added in duplicate and incubated for 2 hours at room temperature. The plates were washed 3 times before addition of detection antibody. A final 3 washes were carried out and substrate solution was added to each well. The reaction was stopped after 15 minutes with 1 M HCl and absorbance read immediately at 450 nm.

Flow cytometry

BAL samples were maintained on ice throughout the procedure. Red blood cells were lysed using RBC lysis buffer (420301, Biolegend) as per the manufacturer's recommendation. Cells were washed and stained in FACS buffer (PBS containing 1% FCS and 0.01% sodium azide) for 30 minutes in the dark. Antibodies used to stain the cells were PE Siglec F (552126, BD), PerCP-Cy5.5 CD11c (560584, BD), Alexa Fluor700 CD11b (557960, BD), APC-eFluor 780 CD3 (BD), PE-CF594 CD19 (562291, BD), V450 Gr1 (560453, BD) and Alexa Fluor647 F480 (564226, BD). Data was collected using an LSR Fortessa (BD) and analysed using FlowJo. AccuCheck counting beads (PCB100, Life Technologies) were added to determine cell counts.

4.2.4 Statistical analyses

All data from this chapter were analysed using GraphPad Prism (Version 7; Prism). When two groups were compared, a nonparametric Mann-Whitney U test was used. A one-way ANOVA with a Bonferroni post-test was utilized for comparisons of three or more groups. In the experiments looking for an effect over time, between multiple groups, a two-way ANOVA with Bonferroni post-test was selected. 0.05 was the P value chosen for statistical significance. Graphed results represent the mean ± standard error of the mean. Both pooled results from repeated experiments and representative figures are presented throughout, the specifics of which are detailed in figure legends.

4.3 Results

4.3.1 TNBS colitis

Weight-loss and clinical scoring

Mice were pre-treated with one of the four hookworm recombinant proteins before colitis was induced with an i.r. injection of TNBS. Weight changes were measured daily and the results are detailed in Figure 4.3. Naïve mice gained 1-3% of their body weight over the course of the experiment. In comparison, TNBS mice which received either PBS vehicle or recombinant human albumin (negative controls) lost 10-12% of their starting body weight by day 1. This weight loss was maintained so that by day 3, mice were still 10-12% lighter than at the beginning of the experiment. The weights of mice treated with sulfasalazine (positive control) were not statistically different to those of negative control mice. Hookworm protein treated groups NP4, NP12 and NP17 all had significantly greater weight loss than the rHA negative control with approximately 15-18% loss over the course of the trial. NP3-treated mice were not significantly different to the negative control mice.

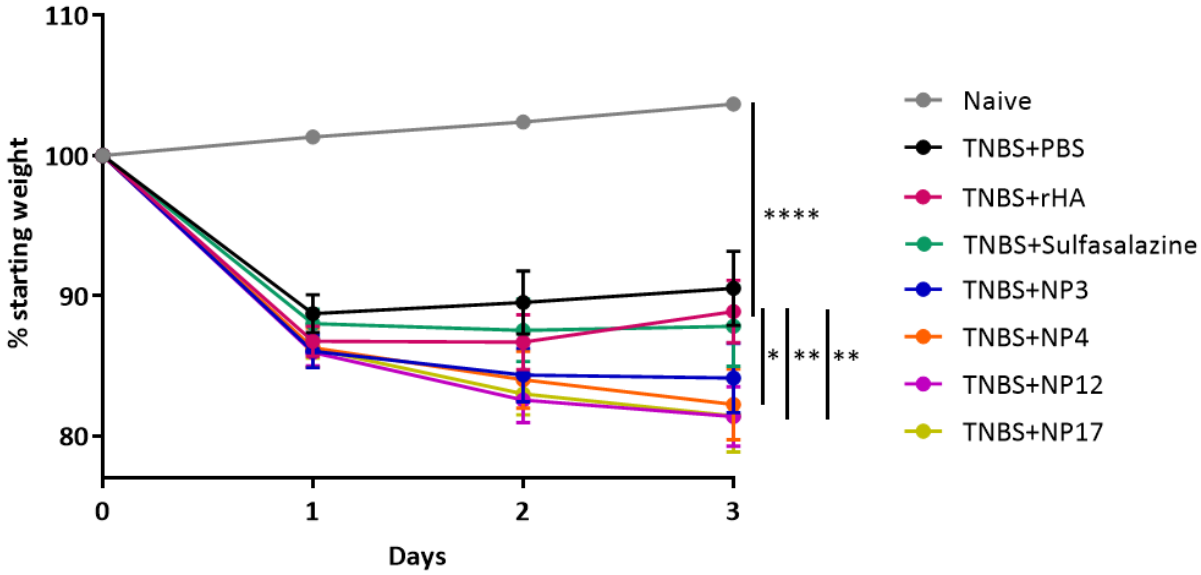


Figure 4.3: Three of the four hookworm recombinant excretory/secretory proteins accentuate TNBS-induced weight loss in mice. Six week old BALB/c mice were pre-treated with or without *Necator americanus* recombinant proteins before colitis was induced with an i.r. injection of TNBS. Weight-loss

was monitored daily for 3 days. Results show two pooled experiments, n = 6 mice for each group in each experiment (total 12 mice). * = p<0.05, ** = p<0.01, **** = p<0.0001.

Mice were examined daily and scored based on mobility, fur appearance and stool consistency for a combined clinical score (Figure 4.4). Compared with naïve mice, all TNBS colitis groups had significantly higher clinical scores across all assessed days post-TNBS-injection. No significant differences in clinical features were observed between the time points (days 1, 2, or 3). Of mice that received TNBS, there were no significant differences in clinical score between protein-treated groups, the negative controls or the sulfasalazine group.

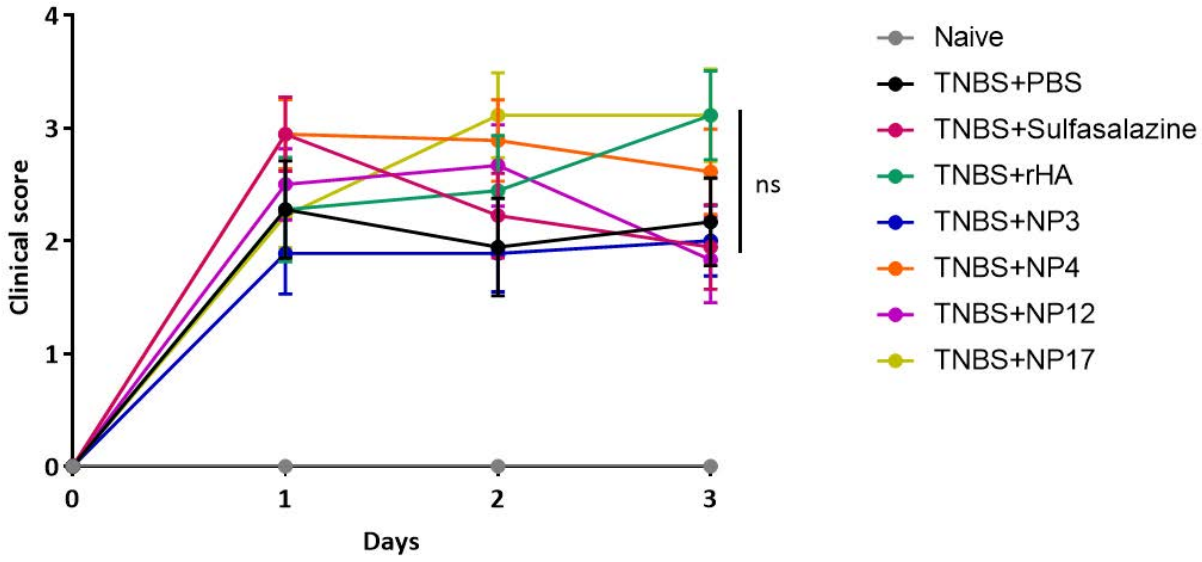


Figure 4.4: Hookworm recombinant excretory/secretory proteins have no effect on clinical manifestations in TNBS colitis. Six week old BALB/c mice were pre-treated with or without *Necator americanus* recombinant proteins before colitis was induced with an intrarectal injection of TNBS. Mice were examined daily and scored based on mobility, fur appearance and stool consistency for a combined clinical score. Results show three pooled experiments, n = 6 mice for each group in each experiment (total 18 mice per group).

Colon length, macroscopic scoring and histopathology

Upon sacrificing mice on day 3, colons were excised and measured from the caecum to the anus. All of the groups that received TNBS had significantly shorter colons than the naïve controls (Figure 4.5). This manifested as 1.5-2 cm reductions in colon lengths. While the sulfasalazine group had slightly longer colons on average than the rHA control, statistical significance was not achieved. For the hookworm recombinant proteins, only NP12 showed a significant difference compared with the rHA control ($p<0.05$). The colons of NP3-, NP4- and NP17-treated mice showed no changes in terms of length compared with the negative controls.

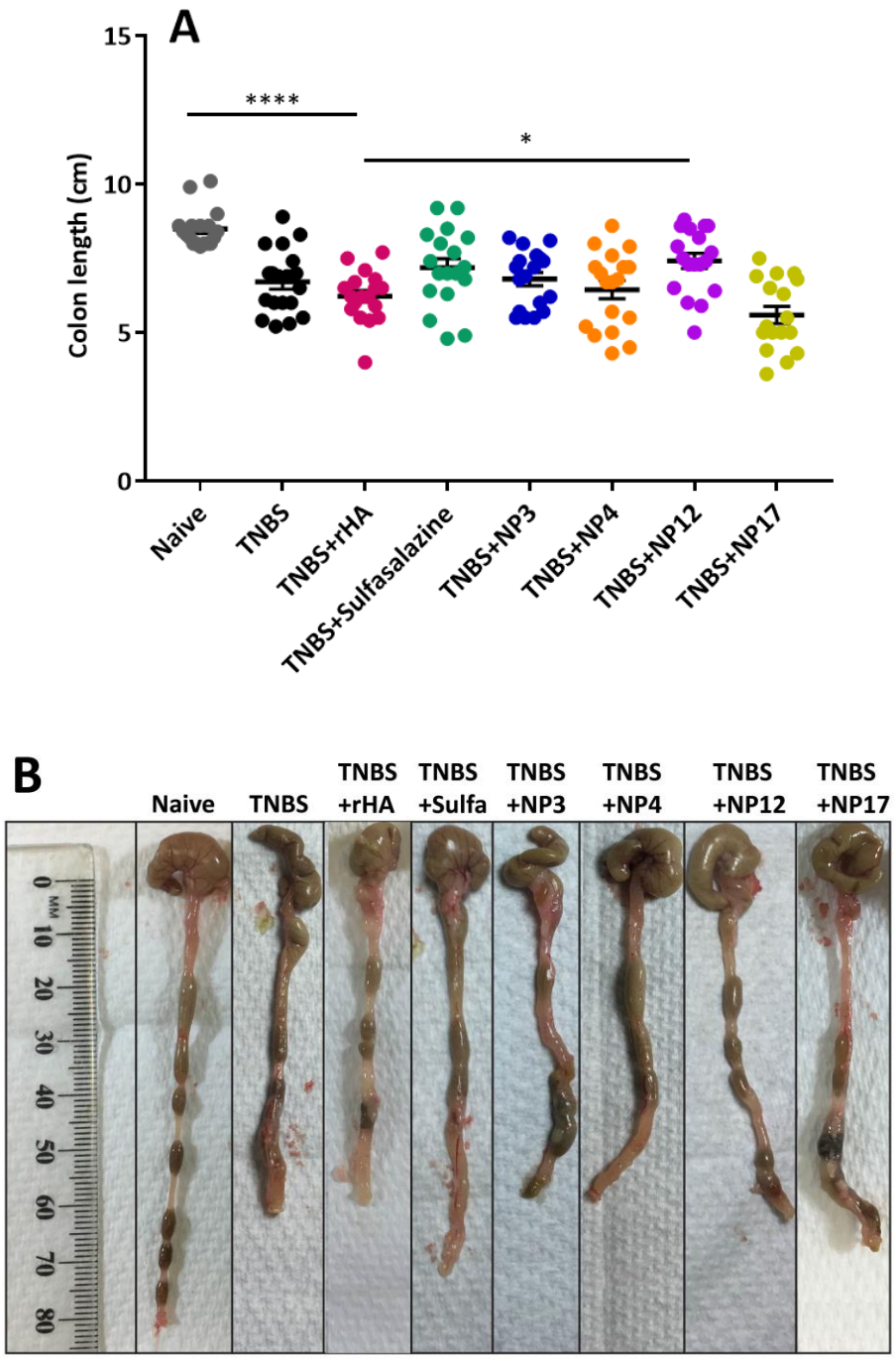


Figure 4.5: Hookworm recombinant excretory/secretory proteins do not protect against colon shortening in TNBS mouse model of colitis. Six week old BALB/c mice were pre-treated with or without recombinant *Necator americanus* proteins before colitis was induced with an intrarectal injection of TNBS. Dissection/colon measurements took place on day 3 (A) and images were taken following measurements (B). Single representative images of the colons are shown. Results show three pooled experiments, n = 6 mice for each group in each experiment (total 18 mice). * = $p<0.05$, **** = $p<0.0001$.

Macroscopic scoring was carried out on the colons based on the presence of adhesions, ulceration, bowel wall thickening and mucosal oedema. Naïve mice did not have any visible pathology whereas all groups of TNBS treated mice had some degree of these features. Due to variation within groups, differences between TNBS groups in Figure 4.6 were not found to be statistically significant. None of the hookworm recombinant protein groups were significantly different from each other, the negative controls or the sulfasalazine group.

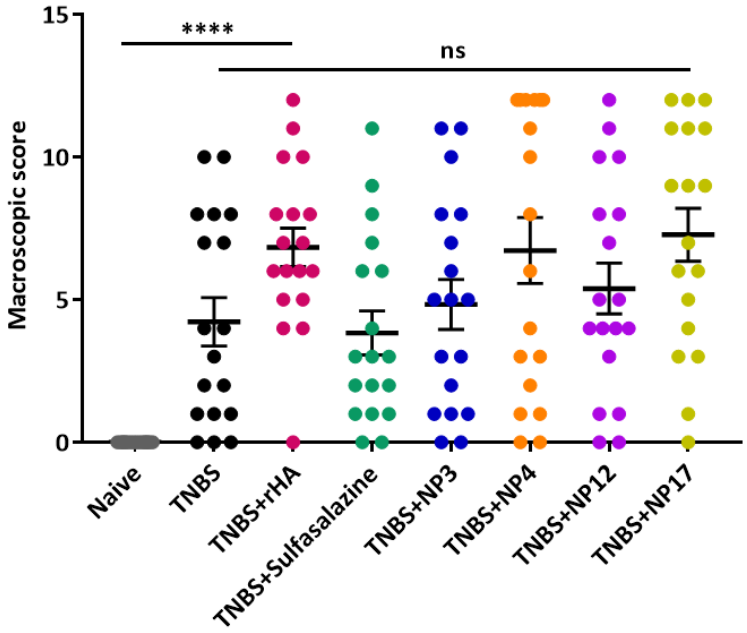


Figure 4.6: Hookworm recombinant excretory/secretory proteins have no effect on colon macroscopic pathology in TNBS colitis. Six week old BALB/c mice were pre-treated with or without recombinant *Necator americanus* proteins before colitis was induced with an intrarectal injection of TNBS. Dissection/colon scoring took place on day 3. Results are reflective of a pooled scoring based on the presence of adhesions, ulceration, bowel wall thickening and mucosal oedema. Results show three pooled experiments, n = 6 mice for each group in each experiment (total 18 mice). **** = $p<0.0001$; ns = non-significant.

Removed colons were either fixed and stained with H&E or homogenised and assessed for inflammatory cytokines. Representative images for each group are shown in Figure 4.7A. Histology images were scored based on inflammation, oedema and hyperplasia individually and these scores were pooled for a combined 'histology score' (Figure 4.7B). From the images and reflected in the scoring, the only significant differences in terms of inflammation, oedema and hyperplasia are present between the naïve controls and all TNBS groups.

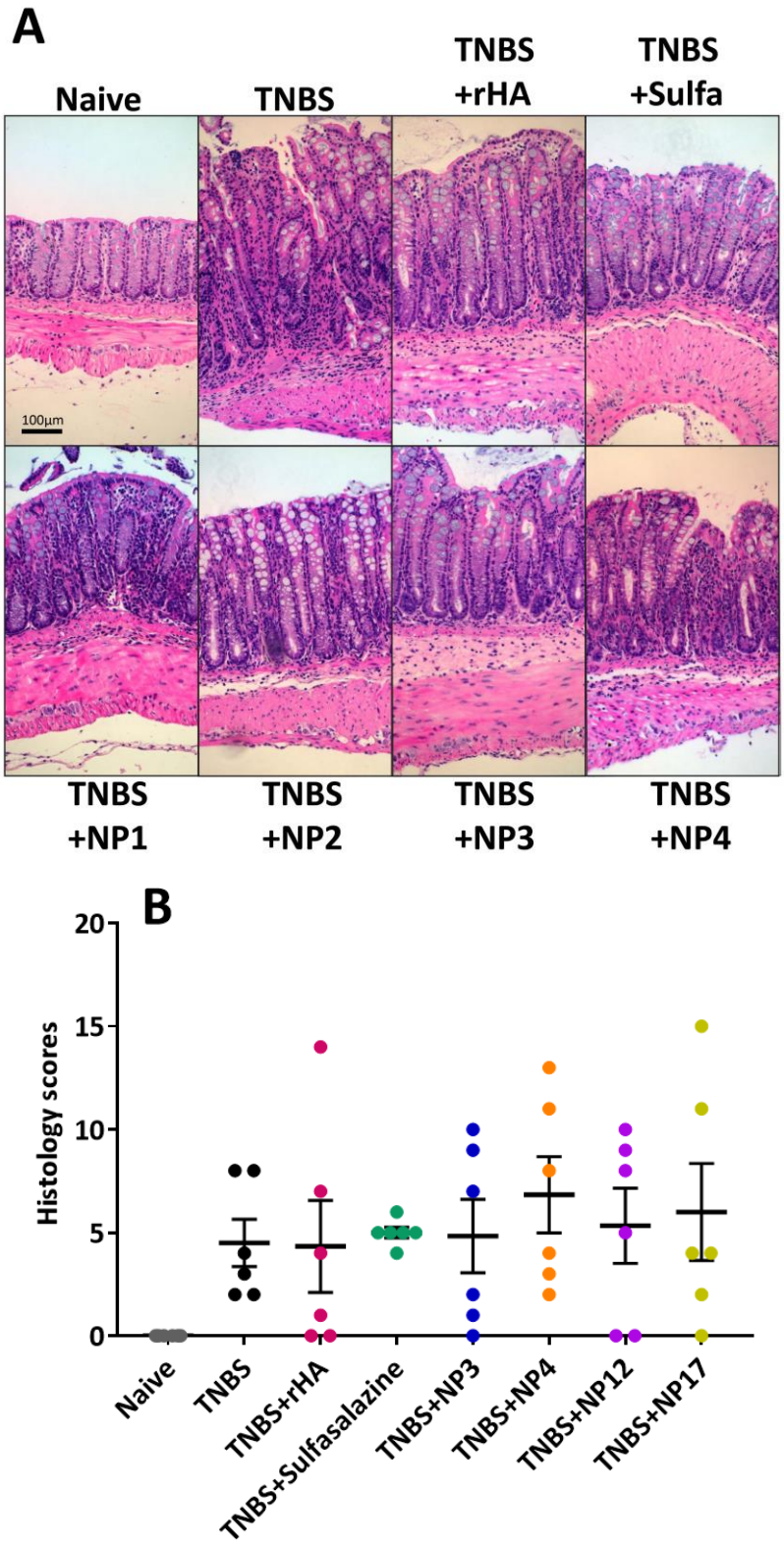


Figure 4.7: Hookworm recombinant excretory/secretory proteins have no effect on microscopic colon pathology in TNBS colitis. Six week old BALB/c mice were pre-treated with or without recombinant *Necator americanus* proteins before colitis was induced with an intrarectal injection of TNBS. Dissection took place on day 3 and colons were fixed and stained with H&E. Images (A; 20x

magnification) were scored on inflammation, oedema and hyperplasia individually and these scores were pooled for a combined 'histology score' (B). Stained images show selected single representatives of three experiments, n = 6 mice for each group in each experiment (total 18 mice).

Cytokine quantification

Homogenised pieces of colon were assayed using a cytometric bead array for IFN-γ, IL-6, IL-10, MCP-1 and TNF-α. IFN-γ and IL-10 were below the level of detection of the assay and therefore these results have not been presented. Marked IL-6, MCP-1 and TNF-α were detectable from the CBA (Figure 4.8A-C respectively). A marked IL-6 response was observed from the colitis groups compared with the naïve controls, $p<0.01$. Between treated groups, no statistically significant differences were observed. A similarly distinct increase was seen in MCP-1 levels from colitis groups compared with the naïve controls. NP12 treated mice were observed on average to have higher levels of MCP-1 in their colons however the variability between mice in this group meant this result did not reach significance when compared with the rHA control ($p>0.05$). TNF-α appeared to be higher in colitis groups in general compared with naïve controls however this induction did not reach statistical significance. Variability in NP3 and NP12 individual results meant that these results were not significantly different compared with the rHA control group.

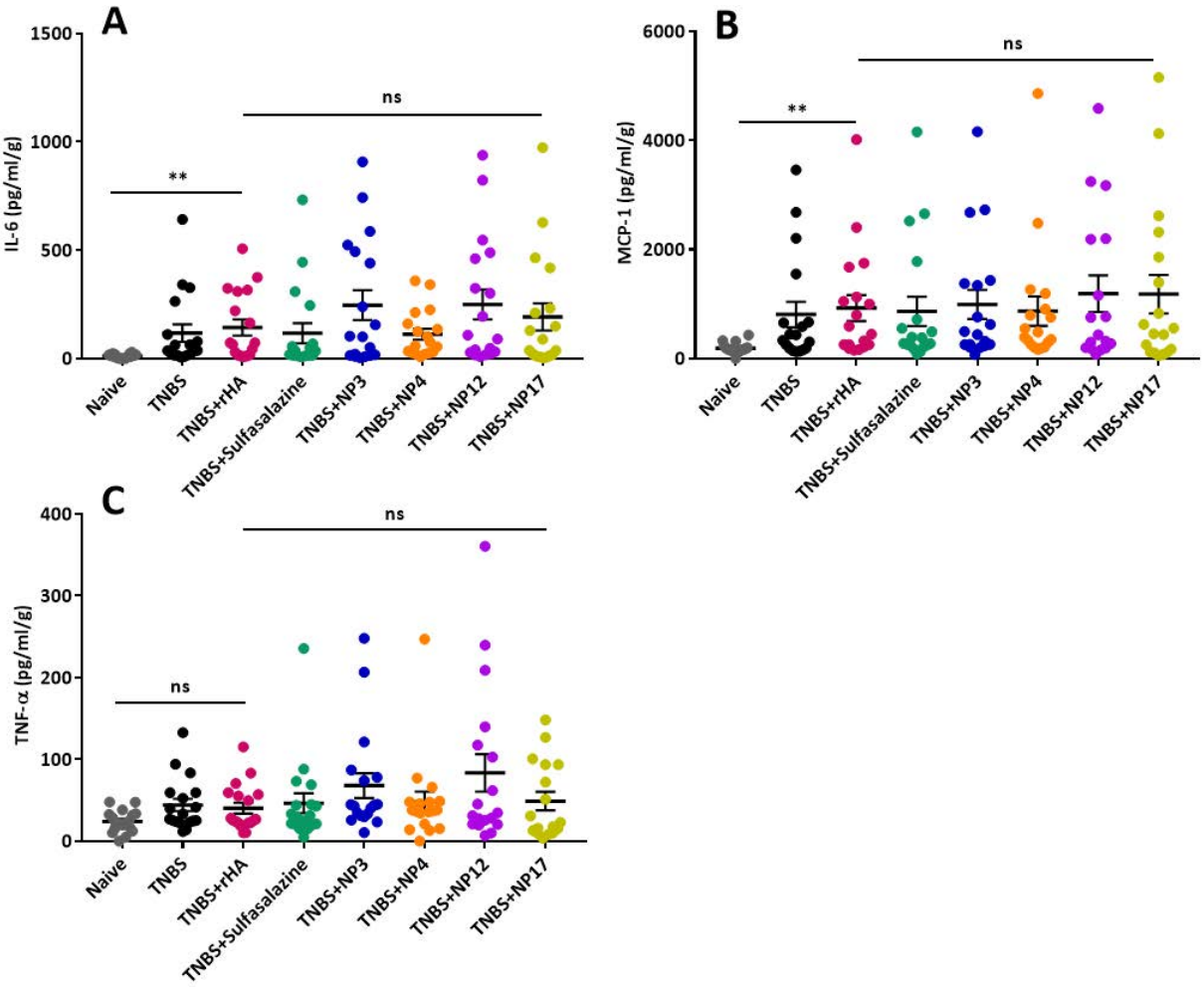


Figure 4.8: Hookworm recombinant excretory/secretory proteins have no effect on intestinal cytokine expression in TNBS colitis. 6 week old BALB/c mice were pre-treated with or without recombinant *Necator americanus* proteins before colitis was induced with an intrarectal injection of TNBS. Dissection took place on day 3 and colons were homogenised and assessed for IL-6, MCP-1 and TNF-α. IL-6, MCP-1 and TNF-α (A-C respectively) were quantified with a cytometric bead array. Data reflect three pooled experiments, n = 6 mice for each group in each experiment (total 18 mice). ** = $p<0.01$; ns = non-significant.

4.3.1 HDM asthma

Comparison of i.n. vs i.t. for HDM delivery

Route of administration of HDM was compared between i.n. and i.t. methods. Figure 4.9 shows an example set of lungs from each group of mice from the experiment. Image A reflects a healthy set of lungs with no blue food colouring administered while B and C are lungs which received blue food colouring i.n. and i.t. respectively. From Figure 4.9C it can be clearly seen that mice that received i.t. administration of the colouring have a much greater quantity of colouration in the lungs. This colouration is visible in the bronchi and throughout both the left and right lobes progressing towards the edge of the lung. By comparison, Figure 4.9B shows some dark colouration close to the base of the bronchi without extensive spreading. This effect was reproduced in all 3 mice included in each experiment with the i.t. route consistently showing greater delivery to the lungs than i.n.

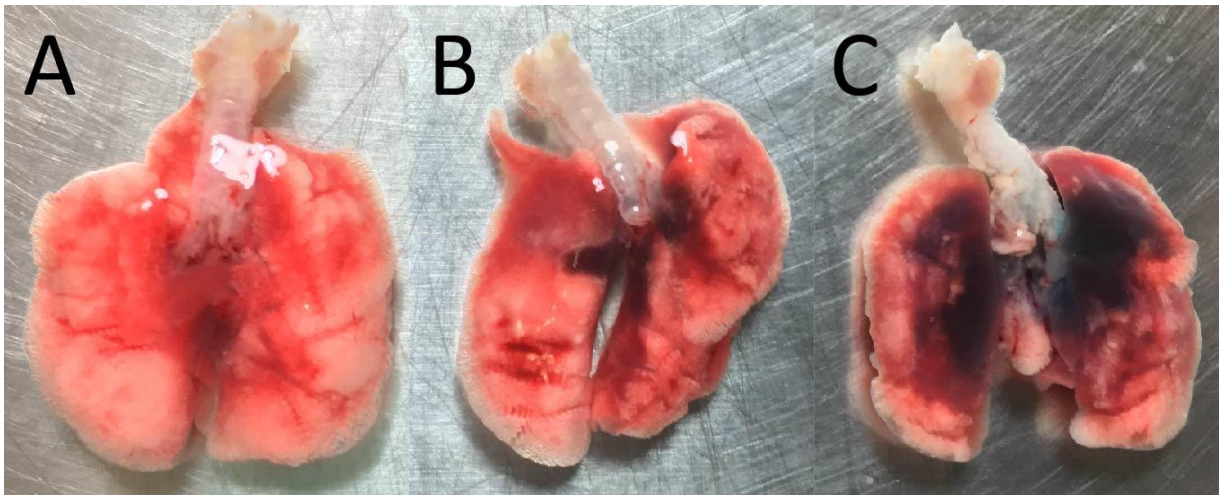


Figure 4.9: Intratracheal (i.t.) administration results in markedly better delivery of solution to the lungs compared with the intranasal (i.n.) route. Naïve lungs (A) are shown compared with lungs administered with blue food colouring i.n. (B) or i.t. (C).

Mice were sensitised and challenged with HDM as described earlier with or without concurrent treatment with hookworm recombinant proteins or dexamethasone. In these experiments the rHA treated group was included as the only negative control. This was reasoned from the earlier TNBS experiments where PBS vehicle treated mice had the same outcomes as recombinant expression-matched rHA mice (no protection). On day 14 mice were sacrificed and a BAL performed to collect lung cells. Cells were stained and immediately subjected to flow cytometry to analyse cell population frequencies. Gating strategy for each of the relevant cell types is shown in Figure 4.10. Figure 4.11A-E shows the breakdown of the number of eosinophils, neutrophils, T cells, B cells, and macrophages. HDM sensitised and challenged mice had significant increases in the number of eosinophils, T cells and B cells. These increases were almost completely ablated with the use of dexamethasone, the positive control. The hookworm recombinant proteins tested had no significant effect on any of the responses when compared with the rHA treated control. These results were reflected in the serum IgE levels (Figure 4.12) whereby only dexamethasone had any effect on HDM-augmented production of IgE.

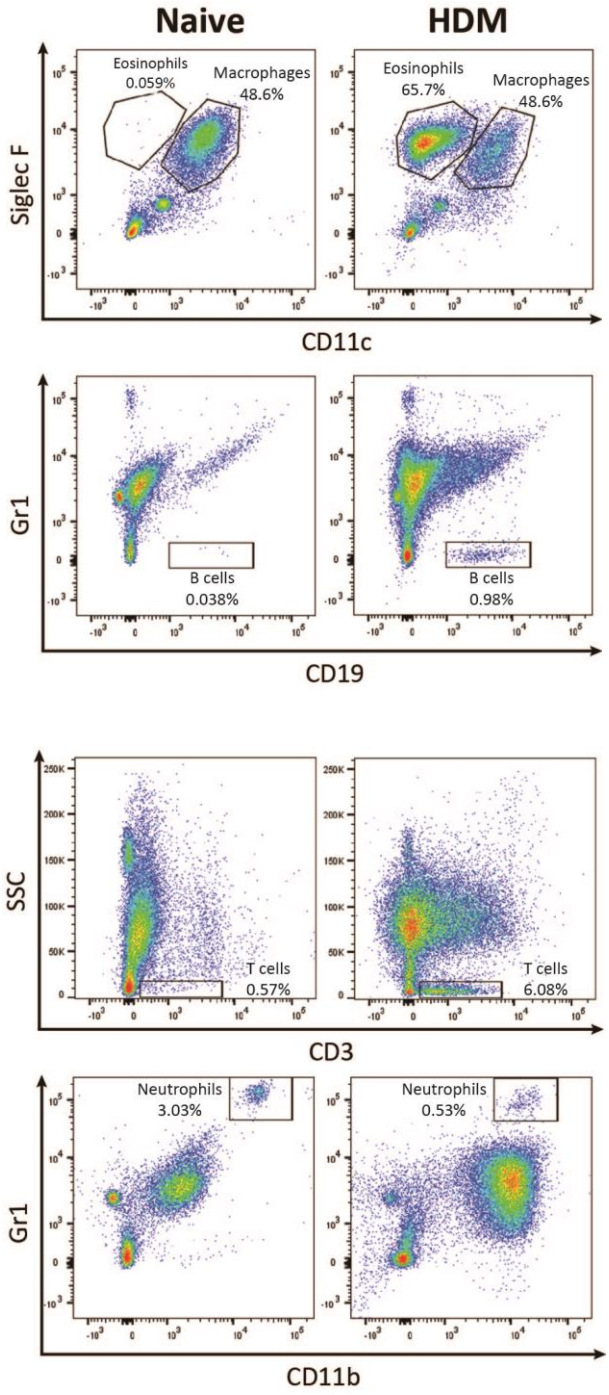


Figure 4.10: Gating strategy for cells collected from bronchoalveolar lavage of healthy vs house dust mite (HDM)-treated mice. Six week old BALB/c mice were sensitised and challenged with HDM intratracheally while being treated with hookworm recombinant proteins intraperitoneally. Mice were sacrificed on day 14 and a bronchoalveolar lavage was performed to collect cells present in the lungs. Eosinophils (Siglec F+ CD11c-), macrophages (Siglec F+, CD11c+), B cells (Gr1- CD19+), T cells (SSClow, CD3+) neutrophils (Gr1+ CD11b+) were stained and gated as shown above.

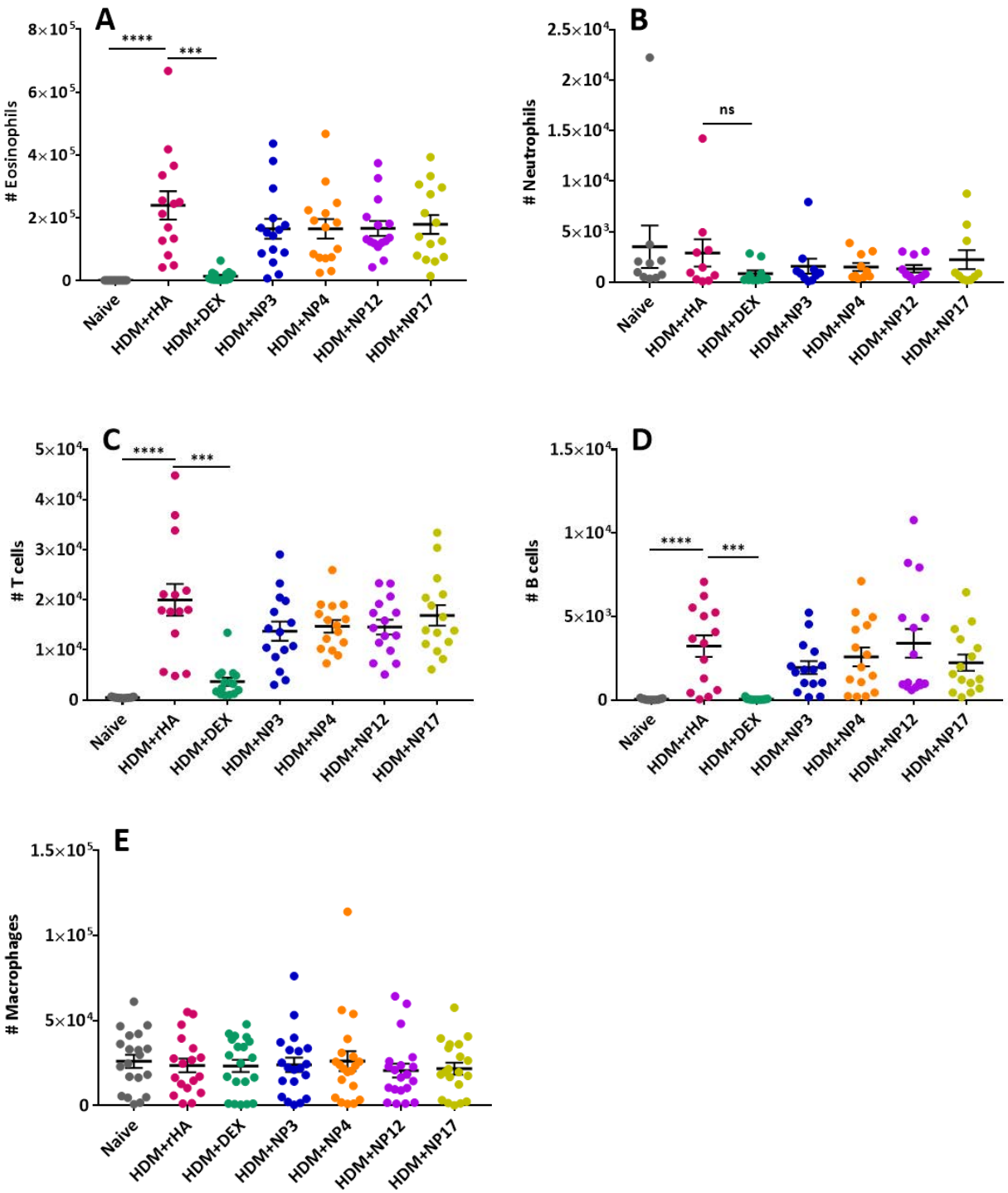


Figure 4.11: Hookworm recombinant excretory/secretory proteins do not protect mice in an acute house dust mite (HDM) asthma model. Six week old BALB/c mice were sensitised and challenged with HDM intratracheally while being treated with recombinant *Necator americanus* proteins intraperitoneally. Mice were sacrificed on day 14 and a bronchoalveolar lavage was performed to collect cells present in the lungs. Eosinophils (Siglec F+ CD11c-), macrophages (Siglec F+, CD11c+), B cells (Gr1- CD19+), T cells (SSClow, CD3+) neutrophils (Gr1+ CD11b+) were stained and gated (B).

Results are representative of three pooled experiments, n = 5 mice for each group in each experiment (total 15 mice). *** = p<0.001, **** = p<0.0001.

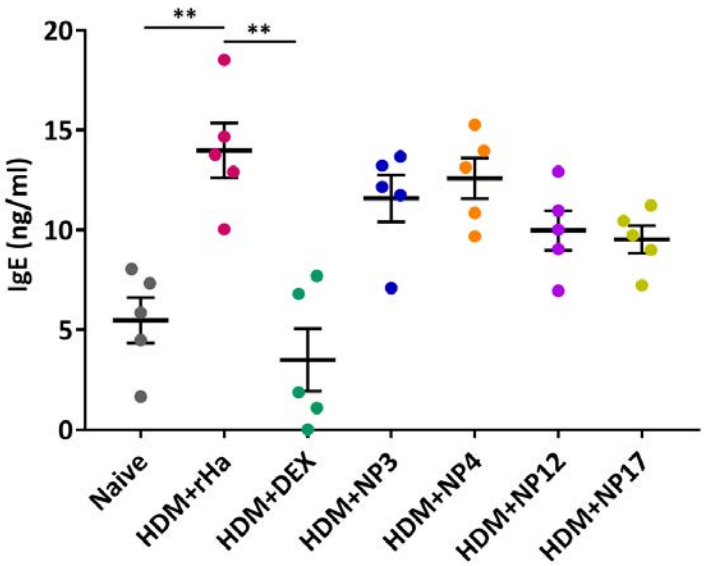


Figure 4.12: Hookworm recombinant excretory/secretory proteins do not protect mice against increased serum IgE in an acute house dust mite (HDM) asthma model. 6 week old BALB/c mice were sensitised and challenged with HDM intratracheally while being treated with recombinant *Necator americanus* proteins intraperitoneally. Mice were sacrificed on day 14 and serum was collected. Results are a single representative of three experiments, n = 5 mice for each group in each. ** = p<0.01.

4.4 Discussion

In this chapter four hookworm recombinant ES proteins were tested for potential therapeutic properties in an IBD mouse model of human disease. In these experiments a single i.r. dose of TNBS was sufficient to induce significant pathology in these mice. The outcomes measured included weight-loss and clinical features, as well as colon pathology, cytokine responses and histology. While vehicle-treated mice administered TNBS were significantly unwell across these categories, in general, the tested outcomes were unaffected by pre-treatment with hookworm proteins. I included a sulfasalazine-treated group as a positive control. Despite sulfasalazine being an anti-rheumatic drug used to treat patients with CD, in this model of colitis it was ineffective at limiting any disease indicators in this study. This is in line with other studies which have reported mixed efficacies for sulfasalazine in murine models of colitis [314,315]. In the future, a more appropriate positive control to employ would be AIP-1 or AIP-2, two hookworm ES proteins which have been shown to protect in inflammatory mouse models [203,313].

The ES products of hookworms contain hundreds of different proteins [195]. Whole mixtures of these proteins, from *A. caninum* and *N. brasiliensis* as well as the soluble proteins from *S. mansoni,* were shown to effectively limit disease induction in the TNBS mouse model [193,255]. The key outcomes of these studies included decreased inflammatory scores and a reduction in the extent of inflammation. This effect was attributed to the suppression of pro-inflammatory cytokines in the colon and was accompanied by upregulation of regulatory responses. Indeed, *A. caninum* ES products also suppressed pathology in the DSS model of chemically induced colitis [192]. While it would have been ideal to include crude ES products from adult *N. americanus* in the current study, this was not a feasible option due to limited accessibility to sufficient material. Given that *A. caninum* and *N. americanus* ES products were shown in chapter 2 to be highly similar in terms of sequence homology, it is reasonable to postulate that they may also protect in colitis models. As these mixtures proved effective across multiple models of colitis it was logical to seek individual molecules responsible for protective

outcomes. Moreover, a single i.p. dose of recombinant *Ac*-AIP-1 significantly suppressed colonic inflammation in the TNBS colitis mouse model [313]. Recombinant *Ac*-NIF, another protein from *A. caninum* ES products, was shown to potently bind neutrophils, inhibiting chemotaxis, adhesion and transmigration of inflammatory cells across the epithelium [316].

The previously mentioned *Ac*-NIF belongs to the SCP/TAPS family, attesting to the role for these proteins in immune regulation. In addition, the vaccine antigen and abundant *N. americanus* L3 secreted protein, *Na*-ASP-2, is a SCP/TAPS family member and was shown to bind to CD79A on human B cells and suppress signaling [317]. Unfortunately only four proteins were successfully produced in recombinant forms in the current study, reducing the size of the pool for identifying therapeutic candidates. Although a therapeutic benefit was not detected for any of the four SCP/TAPS proteins screened in this study, inherent limitations in the models employed, as well as dosing and/or routes of administration of the proteins might not have been optimal.

In the HDM asthma model, I showed stark differences in tissue distribution when comparing i.n. and i.t. delivery routes. While the i.n. model is widely used, it was prudent to compare it with the i.t. model and select the one that worked best. As these differences were so consistent, I opted not to compare responses in a true HDM sensitisation and challenge model to spare animals. When a two-week HDM mouse model was employed using much greater doses of i.n. instillations than those used in the current study, significantly milder inflammatory responses were observed [318]. This comparison attests to the improved delivery of allergen to the lung. Sensitisation and subsequent challenge with HDM was effective in eliciting a significant increase in lung inflammatory cell populations as well as increased serum IgE. The positive control, dexamethasone, potently suppressed the allergic phenotype as seen from the significant reduction in eosinophils, as well as T and B lymphocytes. Dexamethasone was selected for this experiment as it is a currently recommended systemic steroid for moderate to severe asthma exacerbations [319]. Similarly to the colitis experiments, the four tested recombinant proteins had no effect at the dose and delivery frequency used.

Although I did not show any improvements in disease outcomes in the models tested, there are many factors to consider before concluding that the tested proteins have no immunoregulatory capacity. A number of variations may prove useful to consider for future experiments including the route of administration of test proteins as well as dose and timing. For example, in the HDM asthma experiments, it may be beneficial to trial delivery of the proteins specifically to the lung using an i.t. injection. This may allow for a more direct interaction between the proteins and the localised inflammation. Moreover, oral delivery of protein to the gut might be required for obtaining a therapeutic effect in colitis. Other cell populations might also be the target of these SCP/TAPS proteins, as seen with the B cell-specific targeting of CD79A by *Na*-ASP-2 [317].

Another factor to consider across both models is whether the proteins were folded correctly and subsequently active. When recombinant proteins are overproduced it can lead to errors in endoplasmic reticulum folding. When these mechanisms and secretion capacity is impacted the result can be misfolded or unfolded proteins [281]. Given that the function of most proteins depends on correct folding, it would be prudent to confirm this in future experiments. Even partial incorrect folding in any of the proteins tested in this chapter may have played a role in the results obtained. For this reason future experiments should clarify that these proteins are indeed correctly folded and active. Circular dichroism and nuclear magnetic resonance are two methods for determining the folding properties of recombinant proteins. These methods are useful to highlight if any potential mutations that might affect a protein's conformation or stability and can also be used to study protein interactions [320].

While there is a plethora of factors which could be explored in the mouse models, I was interested to see whether the proteins may have a human-specific effect. This was a logical next step as *N. americanus* naturally infects humans and not mice. To test this theory, in the next chapter I co-culture the recombinant proteins with human peripheral blood mononuclear cells (PBMCs).

5 Chapter 5 – Effect of hookworm recombinant SCP/TAPS proteins on human peripheral blood mononuclear cells

5.1 Introduction

Limited clinical efficacy and side effects are key problems associated with immunosuppressive drugs such as corticosteroids in the treatment of allergic and autoimmune diseases. Subsequently there is a need for simple, cheap and rapid screening tools to test the anti-inflammatory capability of novel molecules. One such system utilises purified human peripheral blood mononuclear cells (PBMCs) *ex vivo* from the blood of healthy donors. PBMCs are blood cells which have a round nucleus including lymphocytes, monocytes and macrophages [321]. These cells make up a critical component of the immune system and play roles in the pathogenesis of allergic and autoimmune diseases. When stimulated, innate immune cells begin to release pro-inflammatory cytokines such as IL-12 and IFN-γ [322]. IFN-γ is produced by both innate cells NK and NKT cells) and adaptive cells (CD4 and CD8 T cells) as part of the innate immune response [323]. IFN-γ plays a key role in host-defence mechanisms through the regulation of more than 30 genes which lead to a variety of cellular and physiological responses, such as reactive oxygen species production by macrophages [323]. This immune cascade involves activation of NFκB resulting in, among others, production of TNF-α [323], a key cytokine that drives pathogenesis of many inflammatory diseases including IBD.

Such screening platforms often utilise a stimulus to activate certain cell types. For example, in this chapter PBMCs are stimulated with either lipopolysaccharide (LPS) or a mixture of phorbol myristate acetate (PMA) and ionomycin. LPS is found on the outer layer of gram-negative bacteria and serves as a pathogen-associated molecular pattern (PAMP) binding to TLR4 [324]. TLR4 activation plays a significant role in immune responses to bacteria through recognition of LPS [324]. In contrast, PMA activates protein kinase C, a signal transduction enzyme which along with ionomycin stimulates

production of IL-2 and IL-4 [325]. In this way, LPS primarily activates monocytes and macrophages while PMA and ionomycin stimulate T cells [325]. As mentioned in chapter 1, dysregulation of these varying cell types and the corresponding cytokine profiles are key contributors to IBD. Indeed, the levels of cytokines such as IL-6, IL-8, IL-12, and TNF-α have been shown to be markedly increased in people with IBD. Moreover, TNF-α concentrations are positively correlated with IBD disease severity [92,93]. Any drug that displays a suppressive effect on these pro-inflammatory cytokines may have a beneficial effect in terms of IBD clinical pathology. By treating PBMCs with the recombinant proteins in this chapter I was able to gain insight into their direct effects on immune cells *ex vivo* without the complexity of an animal model.

Later in this chapter a human protein microarray was used to identify interactions between the hookworm SCP/TAPS proteins of interest and putative host cell targets. Microarrays are useful tools for identifying proteins that are the targets of a protein of interest. Typically, protein arrays are probed with serum antibodies, but they can also be used to identify host-pathogen protein-protein interactions [317]. Microarrays have been used to reveal new interactions between humans and a number of pathogens. The binding partners of a protein kinase from *T. gondii* were explored using a human proteome array which lead to further validation of specific candidates [326]. This approach has been employed for non-worm pathogens such as *Legionella pneumophila* whereby a human protein microarray was used to find novel interaction candidates [327]. This method is particularly useful when there are known targets of a tested protein which act as controls when looking for new targets.

These arrays are an ideal platform for which to identify protein-protein or ligand-receptor interactions, particularly for proteins of unknown function [328]. Using these two approaches (the PBMC culturing and arrays) the aim of this chapter was to elucidate putative functions of these *N. americanus* SCP/TAPS proteins.

5.2 Methods

5.2.1 PBMC isolation

Blood was collected from three healthy, hookworm-negative, male donors aged 34-39. This work was covered by human ethics application H7010. PBMCs were isolated from whole blood using Lymphoprep (07801, Stemcell) as per the manufacturer's instructions. In brief, blood was collected from donors into EDTA tubes (BD, 367861) before diluting 1:1 with dPBS + 2% Fetal Bovine Serum (Stemcell). Up to 30 ml diluted blood was carefully layered on 15 ml Lymphoprep. The tube was centrifuged at 800 *g* for 20 minutes at room temperature with no brake before the PBMC layer was collected. PBMCs were washed once with washing media (RPMI Media 1640 (ThermoFisher), 10% FBS, 1% Penicillin-Streptomycin (ThermoFisher); and GlutaMAX (ThermoFisher)). Cells were used immediately or frozen in washing media with 10% dimethyl sulfoxide (DMSO) under liquid nitrogen.

5.2.2 PBMC stimulation and cytokine quantification

Cells were used immediately post isolation or thawed in a 37°C water bath. Cells were diluted with washing media, mixed gently and centrifuged at 500 *g* for 5 minutes at 4°C. The supernatant was aspirated and the cell pellet was resuspended in fresh wash buffer for counting. One hundred thousand cells were added to each well of a 96-well U bottom plate such that each condition could be tested in triplicate. The experiment was carried out a total of three times, each with a different donor's PBMCs.

Cells were split into three main groups: (1) untreated, (2) stimulated with PMA and ionomycin (Cell Stimulation Cocktail (500x) diluted to 1x, 00497093, ThermoFisher), and (3) stimulated with LPS from *Escherichia coli* (10 ng/ml; L6529, Sigma-Aldrich). Before stimulation, cells were treated with either PBS (vehicle control), cyclosporine A (CsA) (10 µg/ml; positive control for group 2; 30024, Sigma-Aldrich), dexamethasone (25 mg/ml; positive control for group 3; D4902, Sigma-Aldrich), or one of the four

hookworm recombinant ES proteins from chapter 3. Three hours post-treatment, cells were stimulated as described above.

Cells were incubated overnight at 37°C with 5% CO_2. Following incubation plates were centrifuged at 500 *g* for 5 minutes and the culture supernatants were collected for cytokine analysis. Cytokines were quantified using a human inflammatory cytometric bead array (551811, BD) as per the manufacturer's instructions. This array measured detectable levels of interleukin-6 (IL-6), interleukin-8 (IL-8), interleukin-10 (IL-10), interleukin-1β (IL-1β), tumor necrosis factor (TNF), and interleukin-12p70.

Statistical analyses

All data were analysed using GraphPad Prism (Version 7; Prism). When two groups were compared, a nonparametric Mann-Whitney U test was used. A one-way ANOVA with a Bonferroni post-test was utilized for comparisons of three or more groups. 0.05 was the P value chosen for statistical significance. Graphed results represent the mean ± standard error of the mean. Both pooled results from repeated experiments and representative figures are presented throughout, the specifics of which are detailed in the respective figure legends.

5.2.3 Human protein microarray

Recipes

Blocking buffer

Five millilitres of 1 M HEPES, 4 ml 5 M NaCl, 800 µl 10% Triton X-100, 50 ml 50% glycerol, 610 mg glutathione powder, and 10 ml 10X Synthetic block were combined before adjusting the pH to 7.5 with NaOH. One hundred microlitres 1 M DTT and dH_2O were added to obtain a total volume of 100 ml. Blocking buffer was stored on ice until use and discarded after 24 hours.

Washing buffer

One hundred millilitres 10X PBS pH 7.4, 100 ml 10X Synthetic block and 10 ml 10% Tween-20 were mixed and made up to 1 L with dH_2O

Biotinylation of recombinant proteins

Biotinylation was carried out using a Biotin-7-NHS Protein Labelling Kit (11418165001, Roche) as per the manufacturer's instructions without modification.

Array

The ProtoArray selected included approximately 8,000 ORFs expressed as N-terminal GST fusion proteins, purified, and printed on a nitrocellulose-coated glass slide in duplicate. Probing of ProtoArray human protein microarrays (PAH0524011, ThermoFisher) was carried out as per the manufacturer's instructions with minor modifications. In brief, microarrays were removed from storage and allowed to equilibrate to 4°C for 30 minutes. Arrays were then blocked with 5 ml blocking buffer (detailed above) in a 4-chamber incubation tray for 1 hour with gentle shaking. Blocking buffer was aspirated before 5 washes of 5 minutes each were carried out at 4°C with gentle shaking. Five millilitres of 3 µM biotinylated hookworm recombinant protein was added to each array and incubated for 90 minutes at 4°C with gentle shaking. Proteins were removed and the arrays were washed 5x as described above. Five millilitres ml of Alexa Fluor 647-streptavidin-conjugated antibody (1 µg/ml) in washing buffer and incubated for 90 minutes at 4°C with gentle shaking. Arrays were washed again 5x as above before they were dried using centrifugation as per the manufacturer's instructions. Each array was scanned using a GeneArray 4000B scanner (Molecular Devices) at 625 nm. The generated multi-TIFF files were analysed with Genepix Propsector software v7.

5.3 Results

5.3.1 Human PBMC cytokine responses

Human PBMCs were treated with hookworm recombinant proteins before being stimulated with endotoxin-free PBS (negative control), LPS or PMA/ionomycin. In the absence of cell stimulation, NP4 elicited a significantly greater IL-6, IL-8, and TNF-α response when compared with the untreated controls ($p<0.001$; Figure 5.1). NP12 and NP3 both induced a slight but nonetheless significant increase in IL-8 responses. NP-4 was the only protein which did not induce any cytokine responses that were statistically different from the untreated control. The cytometric bead array used also included IL-1β, IL-10, and IL-12, none of which were detectable in unstimulated cells.

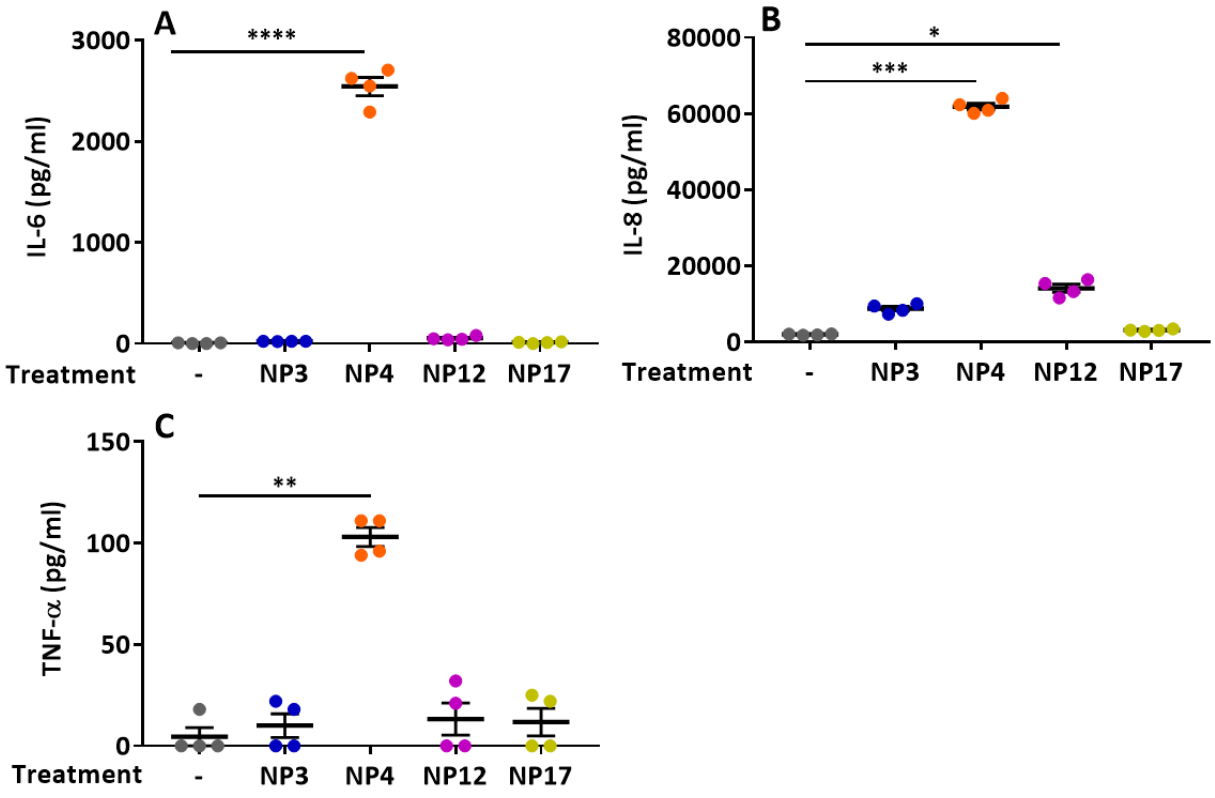


Figure 5.1: Hookworm recombinant excretory/secretory protein NP4 induces an inflammatory response in peripheral blood mononuclear cells (PBMCs). PBMCs were isolated from the whole blood of healthy donors before being treated with or without recombinant *Necator americanus* proteins at a concentration of 20 µg/ml. Supernatant was collected 24 hours later and assayed for inflammatory

cytokines. Graphs show a selected single representative experiment (from a total of three experiments), n=3-4 for each group in each experiment. Error bars are indicative of four technical replicates from the blood sample. **** = $p<0.0001$;** = $p<0.01$; * = $p<0.05$

When PBMCs were stimulated with PMA/ionomycin, strong IL-6, IL-8 and TNF-α responses were detected in the supernatant compared with unstimulated controls. Figure 5.2A and 5.2C show that CsA was highly effective at limiting the production of IL-6 and TNF-α respectively ($p<0.001$). NP-2 significantly increased IL-6 production ($p<0.001$) while the other hookworm proteins had no effect compared with the PMA/ionomycin only group. IL-8 production was reduced by 30% in cells treated with NP17 ($p<0.0001$) while none of the other treatments had any effect on IL-8 production. CsA reduced the production of TNF-α below the level of detection ($p<0.0001$) (Figure 5.2C). NP-2 increased TNF-α production while NP-4 treatment reduced production by approximately 60% compared with the PMA/ionomycin group ($p<0.0001$). Cells treated with PMA/ionomycin did not produce detectable levels of IL-1β, IL-10, and IL-12 and so this data is not presented.

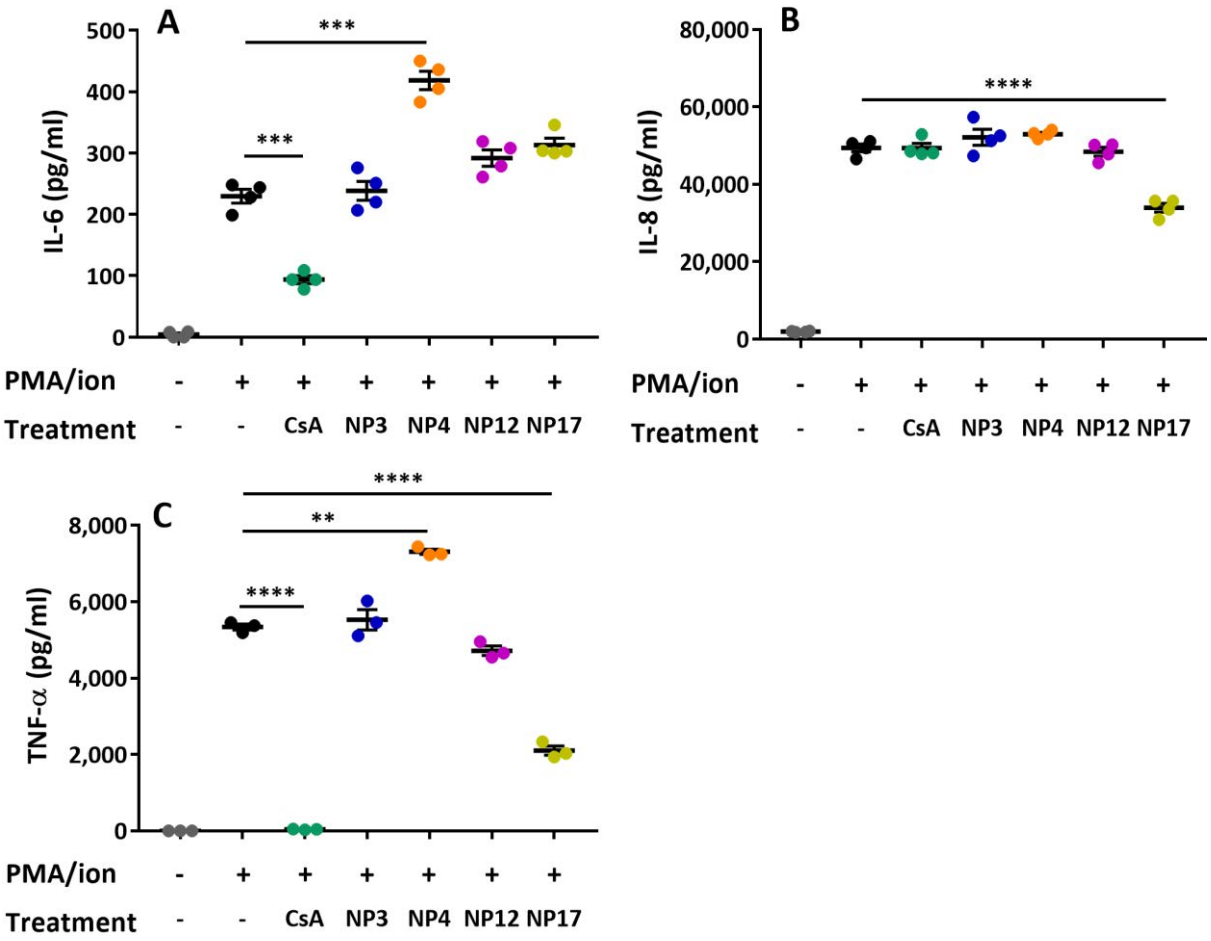


Figure 5.2: Hookworm recombinant excretory/secretory protein NP17 significantly reduces production of TNF-α from PMA/ionomycin stimulated peripheral blood mononuclear cells (PBMCs). PBMCs were isolated from the whole blood of healthy donors, treated with or without recombinant *Necator americanus* proteins (20 µg/ml) and stimulated with PMA/ionomycin 3 hours later. Supernatant was collected 24 hours later and assayed for inflammatory cytokines Graphs show selected single representatives of experiments carried out three times (each experiment was a different individual), n=3-4 for each group in each experiment. Error bars are indicative of four technical replicates from the sample blood sample. **** = $p<0.0001$;** = $p<0.01$; * = $p<0.05$

When PBMCs were stimulated with LPS, strong IL-6, IL-8, IL-10, IL-1β and TNF-α responses were elicited. All of these responses were diminished to varying degrees by dexamethasone treatment ($p<0.001$; Figure 5.3A-E). None of the hookworm protein treatments had an effect on IL-6 or IL-10 production. Similarly to the unstimulated group, NP3, NP4 and NP12 all augmented IL-8 release into the supernatant (Figure 5.3B). NP3 and NP4 both increased IL-1β production (Figure 5.3E). NP17-treated cells had cytokine levels similar to the LPS-only stimulated group with the exception of TNF-α which was reduced to the same level as dexamethasone-treated cells ($p<0.0001$). NP4 and NP12 also reduced TNF-α production ($p<0.001$), although the suppressive effect was less than that achieved for NP17. Cells treated with LPS did not produce detectable levels of IL-12 and so this data is not presented.

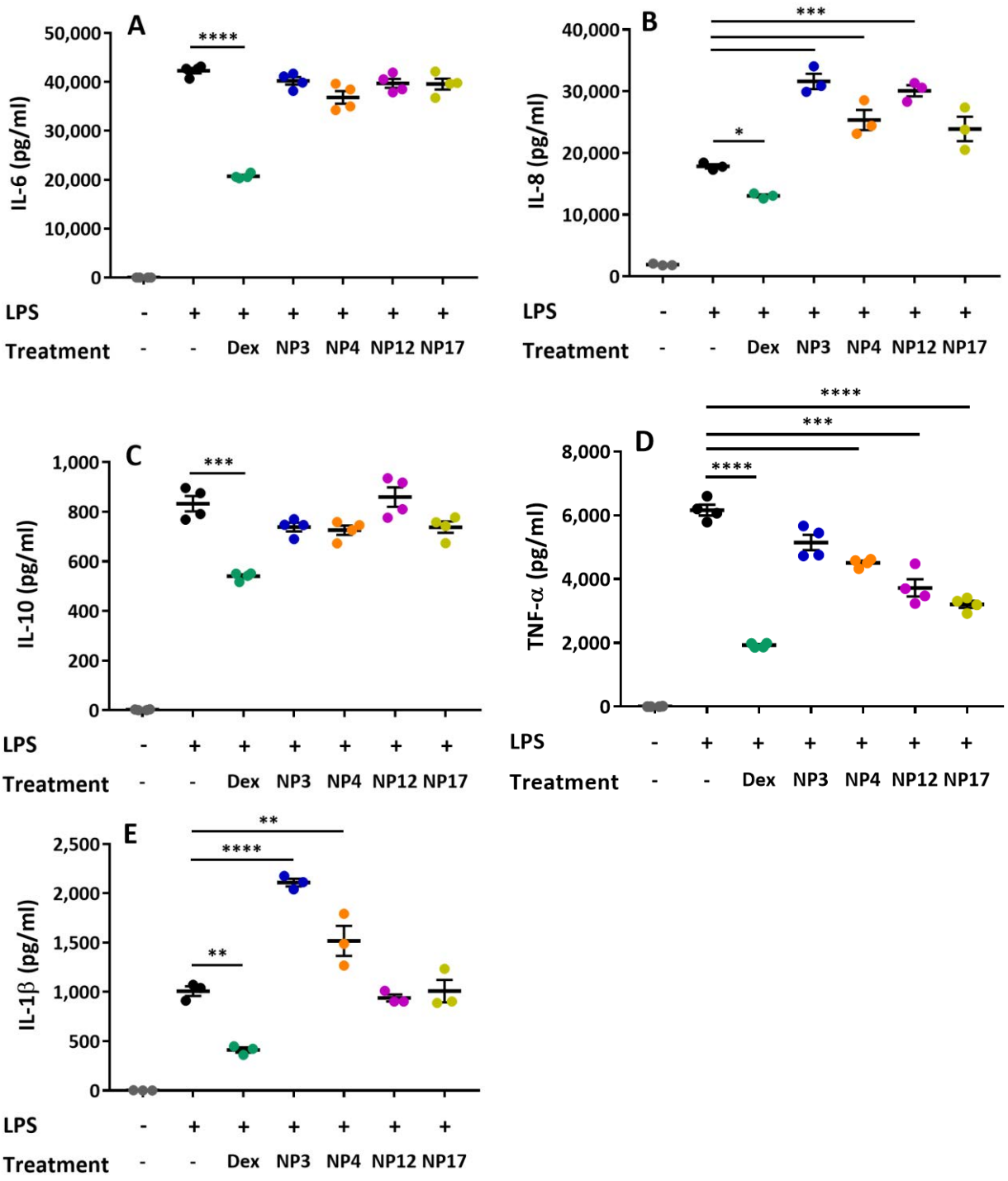


Figure 5.3: Hookworm recombinant excretory/secretory protein NP17 significantly reduces production of TNF-α from LPS stimulated PBMCs. PBMCs were isolated from the whole blood of healthy donors, treated with or without recombinant *Necator americanus* proteins (20 µg/ml) and stimulated with LPS 3 hours later. Supernatant was collected 24 hours later and assayed for inflammatory cytokines. Graphs show single representatives of experiments, each data point is a technical replicate. Experiments were repeated three times. **** = p<0.0001;** = p<0.01; * = p<0.05

5.3.2 ProtoArray human protein microarray

Four ProtoArray human protein microarrays were probed individually with a single biotinylated hookworm recombinant protein (NP3, NP4, NP12, NP17). Each of the proteins elicited a different profile of hits; the top 10 hits for each protein are summarised in Table 5.1. The alpha subunit of propionyl-CoA carboxylase (PCCA), a biotin-binding protein control was a significant hit for all of the proteins (data not shown), however was only a top 10 hit for NP17.

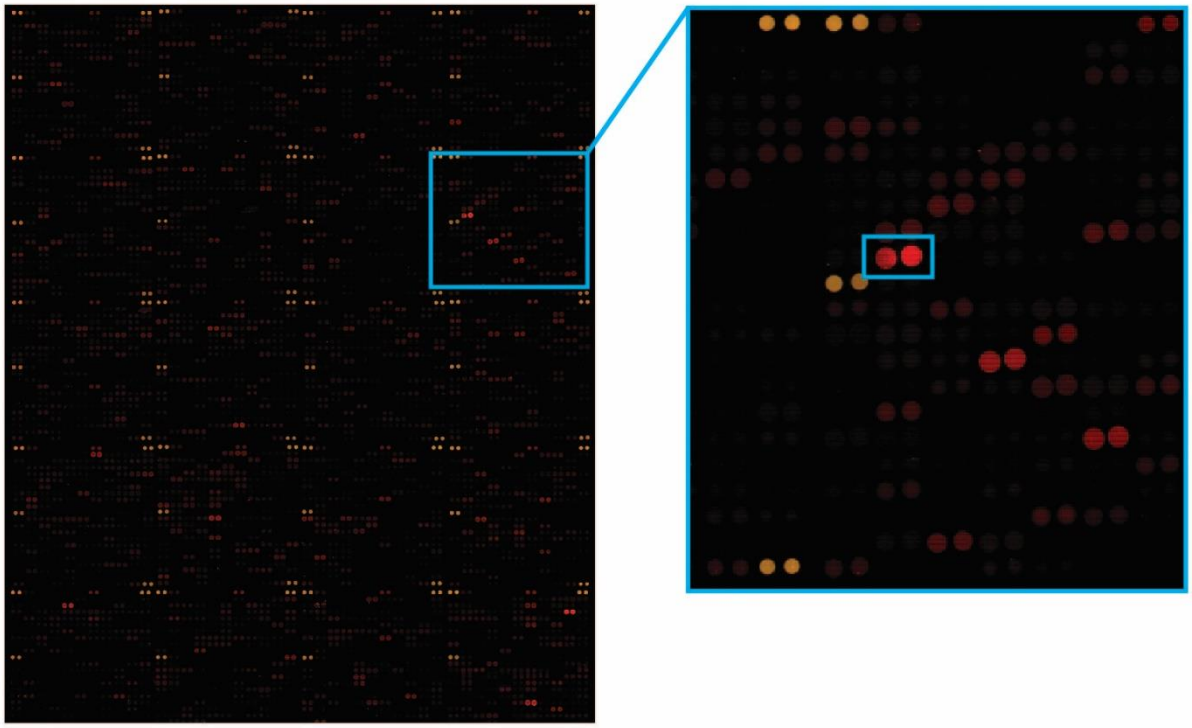


Figure 5.4: Example image of one array from the four tested recombinant proteins, highlighting a significant hit. All proteins on the array are printed in duplicate. Blue boxes denote hits. Yellow spots in the magnified image on the right denote orientation controls that are replicated across the array.

A variety of proteins involved in RNA processing, signaling pathways, mitosis, and immunity were identified in the top 10 binding partners for each of the proteins (Table 5.1). NP4 most significantly bound to TRAF3 interacting protein 1, a protein which binds to and forms part of the TNF receptor complex. NP3 and NP4 both bound to SH2 Domain Containing 2A (SH2D2A), a T cell adapter protein. NP17's second most significant hit was to leukocyte receptor cluster member 1.

	Rank	Protein name	Gene	Brief description	Database ID	Z-Score
NP3	1	Heterogeneous Nuclear Ribonucleoprotein K	HNRNPK	Pre-mRNA processing in nucleus	NM_002140.2	15.5651
	2	Dematin Actin Binding Protein	DMTN	Actin binding - structural role in erythrocytes	BC052805.1	13.6899
	3	Wiskott-Aldrich Syndrome Protein	WAS	Transduction of signals from cell surface receptors to actin cytoskeleton	BC002914.1	13.2461
	4	Alpha Tubulin Acetyltransferase 1	ATAT1	Acetylates alpha tubulin on lysine 40	BC006105.1	11.8336
	5	SH2 Domain Containing 2A	SH2D2A	T-cell signal transduction	NM_003975.1	11.0692
	6	Splicing Factor 1	SF1	Spliceosome assembly	NM_201998.1	10.5161
	7	Proline And Serine Rich Coiled-Coil 1	PSRC1	Mitosis	NM_032636.2	10.3568
	8	hypothetical protein MGC2731	-	-	NM_024068.1	10.2271
	9	G-Patch Domain And KOW Motifs	GPKOW	RNA-binding protein - plays role in pre-mRNA splicing	NM_015698.2	9.19348
	10	Thyroid Hormone Receptor Interactor 6	TRIP6	Signalling from cell surface to nucleus	BC002680.1	9.10054
NP4	1	TRAF3 Interacting Protein 1	TRAF3IP1	Tethers TNF receptor-associated factor 3 to cytoskeletal microtubules	BC059174.1	8.38278
	2	Cortactin	CTTN	Actin branching	NM_138565.1	6.96568
	3	Ubiquitin Like Modifier Activating Enzyme 6	UBA6	Alters proteins with ubiquitin	BC031637.1	5.88249
	4	Polyglutamine Binding Protein 1	PQBP1	Scaffold protein for pre-mRNA splicing, transcription regulation, innate immunity	NM_005710.1	5.77709
	5	SERPINE1 mRNA Binding Protein 1	SERBP1	mRNA stability	BC020555.1	5.71155
	6	Small Nuclear Ribonucleoprotein Polypeptide C	SNRPC	Spliceosome assembly	NM_003093.1	5.62476
	7	Musculoskeletal, Embryonic Nuclear Protein 1	MUSTN1	Development and regeneration of the musculoskeletal system	NM_205853.1	5.60881
	8	SH2 Domain Containing 2A	SH2D2A	T-cell signal transduction	NM_003975.1	5.58579
	9	Sciellin	SCEL	Regulation or assembly of cornified envelope	BC047536.1	5.28643
	10	Proline Rich 16	PRR16	Regulates cell size and mitochondrial respiration	BC038838.1	5.27314
NP12	1	Alpha Tubulin Acetyltransferase 1	ATAT1	Acetylates alpha tubulin on lysine 40	BC006105.1	10.3024
	2	Thyroid Hormone Receptor Interactor 6	TRIP6	Signalling from cell surface to nucleus	BC002680.1	8.55641
	3	Proline Rich 16	PRR16	Regulates cell size and mitochondrial respiration	BC038838.1	7.39087

	4	Polyglutamine Binding Protein 1	PQBP1	Scaffold protein for pre-mRNA splicing, transcription regulation, innate immunity	NM_005710.1	6.52483
	5	Cell Division Cycle Associated 3	CDCA3	Mitosis	NM_031299.2	6.13117
	6	Erythrocyte Membrane Protein Band 4.1 Like 4A	EPB41L4A	Regulates interactions between plasma membrane and cytoskeleton	NM_022140.2	5.8641
	7	Arginine And Glutamate Rich 1	ARGLU1	Required for ESR1 target gene expression	BC050434.1	5.71436
	8	Small Nuclear Ribonucleoprotein Polypeptide C	SNRPC	Spliceosome assembly	NM_003093.1	5.71127
	9	Single Stranded DNA Binding Protein 3	SSBP3	Transcriptional regulation of alpha 2 collagen gene	NM_018070.2	5.56153
	10	CUE Domain Containing 1	CUEDC1	-	NM_017949.1	5.47662
NP17	1	Propionyl-CoA carboxylase (PCC)	PCCA	Biotin binding	NM_000282.1	20.3827
	2	Leukocyte Receptor Cluster Member 1	LENG1	Leukocyte expressed receptors of the immunoglobulin superfamily	NM_024316.1	18.7357
	3	Dead-box helicase 42	DDX42	RNA helicase	BC015505.1	13.5748
	4	RD RNA-binding protein	NELFE	Regulates RNA polymerase II transcription	NM_002904.4	11.307
	5	Alpha Tubulin Acetyltransferase 1	ATAT1	Acetylates alpha tubulin on lysine 40	BC006105.1	11.2321
	6	Actin Binding LIM Protein 1	ABLIM1	Binds actin filaments mediating cytoplasmic and actin interactions	BC002448.2	10.8754
	7	Nuclear speckle splicing regulatory protein 1	NSRP1	Mediates alternative splicing regulation of pre-mRNAs	NM_032141.1	9.9507
	8	Polyglutamine Binding Protein 1	PQBP1	Scaffold protein for pre-mRNA splicing, transcription regulation, innate immunity	NM_005710.1	9.91107
	9	Proline Rich 30	PRR30	Variable function	NM_178553.2	9.40026
	10	Erythrocyte Membrane Protein Band 4.1 Like 4A	EPB41L4	Regulates interactions between plasma membrane and cytoskeleton	NM_022140.2	8.98633

Table 5.1: Top 10 hits from Protoarray human protein microarrays probed with individual human hookworm recombinant proteins. Results are listed based on Z-score.

5.4 Discussion

In this chapter I exposed PBMCs from healthy donors to the four hookworm recombinant proteins. A marked increase in IL-6, IL-8, and TNF-α production was observed from cells exposed to NP4, as well as an increase in IL-8 from NP-1- and NP-3-treated cells. IL-8 is most prominently produced by monocytes and has been shown to be triggered by numerous ligands, metabolic stimulants, bacterial products, and cytokines [329,330]. Given that it is broadly implicated in many inflammatory responses, it is difficult to conclude why most of the proteins elicited some IL-8 release. PBMCs exposed to NP3, NP12 and NP17 had IL-6, IL-8 and TNF-α levels similar to untreated controls. These findings highlight the differences between individual SCP/TAPS, notably that one of them was pro-inflammatory while the others were less so, or even capable of suppressing production of some inflammatory cytokines.

The key finding from this chapter was that NP17 consistently, across multiple donors, inhibited TNF-α release from PBMCs stimulated with either LPS or PMA/ionomycin. The primary TNF-α-producing cells from the PBMC mixture are monocytes and M1 macrophages [331]. This means that the TNF-α-limiting effect of NP17 must involve these cells in some fashion. Under normal circumstances TNF-α is rapidly released following exposure to bacteria-derived LPS or trauma [331]. It plays a pivotal role in the regulation of many pro-inflammatory cytokines and is a key orchestrator of the immune response to a multitude of infections [331]. Anti-TNF-α mAb drugs such as infliximab and adalimumab are frequently used to treat IBD and other immune dysregulated diseases [332]. These drugs are highly effective in treating both UC and CD due to their inhibitory effect of TNF-α [332]. Given that IBD is not curable, the goal is to send patients into remission by restricting the inflammatory responses that drive disease pathology. The finding that NP17 limits TNF-α production is promising and warrants further investigation as a potential therapeutic for treatment of IBD.

When PBMCs were isolated from *N. americanus* infected individuals and stimulated with adult *A. caninum* ES, a reduced capacity to produce IL-6 and TNF-α was observed [333]. Given that this study used mixtures of hookworm antigens, it is unsurprising that a multitude of effects were observed. In

contrast, I have shown a single human hookworm ES protein which specifically suppresses TNF-α rather than affecting the other cytokines tested. The benefits of this specific targeting of TNF will be discussed later in chapter 6.

I then went on to use human protein microarrays to determine potential binding partners for the hookworm proteins. This was a useful tool for garnering insight into the human host targets to which the proteins might bind. A large number of hits were identified for each of the proteins with varying binding efficacy as measured by a Z score. As each of the proteins were biotinylated, the biotin-binding enzyme, PCCA, was an important control to consider. Each of the proteins significantly bound PCCA (full data not included), while NP3, NP4, and NP12 had numerous more significant binding partners than this protein interaction. Some of the top hits involved immune-related pathways, including TRAF3IP1 (for NP4), a protein which interacts with the TNF-α receptor. NP3 and NP4 also bound SH2D2A which has been implicated in the control of T cell activation. One of the most significant hits for NP17 was binding to the LENG1, a complex involved in a range of immune-related pathways. A majority of the top 10 binding partners for each protein were intracellular targets. This result may be explained by protein endocytosis, with the key hits being present in the cytosol and cell nucleus [334]. Indeed, SCP/TAPS have been shown to be present on *N. brasiliensis* exosomes which were internalised by host cells [255]. The implications for these results in interpreting the PBMC data will be discussed in chapter 6. While array technology is useful, a key limitation should be considered when interpreting the results here. Normally membrane-bound proteins may not adopt correct conformation in the absence of a cell environment or cell membrane [335]. This may mean some of the extracellular targets for the tested proteins may have been affected. To overcome this, future experiments should explore using RNA-Sequencing (RNA-Seq) which would look at expression levels influenced by real interactions between the ligand and their cellular target – this too will be discussed in chapter 6.

The key conclusions from probing the human proteome arrays are that these hookworm proteins bind (at least *in vitro*) to a number of human proteins, some of them immune related, highlighting their potential mechanisms of action and providing leads for targeted validation of these defined pathways.

6 Chapter 6 - General discussion

Humans and other animals across the globe exist in environments heavily populated by pathogenic and non-pathogenic microorganisms, as well as toxic and allergenic substances. These include not only infectious, disease-causing microbes, but also beneficial commensals. The latter must be kept in check and tolerated, while the former is targeted for elimination to ensure optimal health and homeostasis are maintained. Interactions between the host immune system and microbes are both unavoidable and essential for normal development of the immune system. The observed rise in the prevalence of allergic and autoimmune diseases in the developed world, both of which are associated with inappropriate immune responses to harmless stimuli, supports the notion that something is out of balance with immune development in people from these regions. Societies in general have moved towards improved sanitation, cleaner water and superior healthcare, including rapid development of screening tools and anti-infective drugs. These practices have had a dramatic impact on the prevalence of soil-transmitted helminths and other parasites, to the point where many have now been eliminated in most developed nations. In line with increasingly sterile living conditions and decreased exposure to infectious diseases, the incidence of inflammatory conditions has exploded. This effect has led many researchers to propose that these two phenomena are linked, and research efforts are now focused on mechanisms by which to re-establish immune homeostasis without the need for infectious exposures.

Parasites have coexisted with humans throughout our evolution. The eggs from intestinal helminths have been found in both human faeces dating thousands of years and in fossilised dinosaur coprolites [336,337]. Our bodies and our immune systems have been shaped to manage infestations with helminths, while simultaneously some helminths have evolved to manage the pressures imposed on them by the immune system, to ensure their survival. Indeed, hookworms in particular are able to live for long periods of time within their host without provoking protective immunity or serious pathology when present in small numbers. In line with this and through research on related hookworm species,

there is strong evidence that hookworms can modify host immune responses. This is thought to occur primarily through the release of immunomodulatory molecules in the worm's ES products. Thus, these products derived from natural human hookworm pathogens represent a substantial untapped resource of molecules that could be turned into novel therapeutics to treat human inflammatory diseases.

Before beginning to elucidate any protein functions, I first needed to characterise the *N. americanus* ES proteome. As detailed in chapter 2, reannotation of the genome improved the quality of the ES proteome description and was therefore an important first step in this process. The majority of published genomes are first-drafts which have frequent missing data, gaps, and errors in reported sequences [338]. We found 3,425 fewer genes than the original model with increases in gene lengths and the percentage of the genome covered by genes. This new annotation represented a more accurate and refined hookworm draft genome. Not only did this improve the quality of the ES proteome but it also serves as an upgraded reference tool to study numerous aspects of *N. americanus* biology such as genome evolution and pathogenesis. Using the re-annotated genome, nearly 200 ES proteins were identified by LC-MS/MS. Upon analysis I found that SCP/TAPS proteins were the predominant protein family. This finding reflects the published data for closely related species including *N. brasiliensis, A. caninum,* and *H. polygyrus* [194,195,248]. SCP/TAPS proteins have been described from various eukaryotic organisms including many helminths [339].

To date, SCP/TAPS proteins have been reported to have substantial variation in structure and sequence which accordingly reflects their diversity of function. These proteins have been implicated in host-pathogen interactions, defence mechanisms and immunoregulatory roles to name a few. The dominant SCP/TAPS subfamily reported in plants is the PR-1 protein family which falls under the category of pathogenesis-related proteins. PR-1 proteins, identified in every investigated plant species to date, are synthesized in response to stress-inducing factors such as infection with pathogens and plant cutting [340]. *In vitro*, tomato PR-1c impaired growth of *Phytophthora infestans*, a water mould that causes disease in these plants [340]. Similarly, when PR-1 genes in tobacco were selectively expressed,

transformed plants showed greater resistance to certain moulds, but not to other fungus, bacteria or viruses [340]. In animals, SCP/TAPS make up major components of the venom allergens from biting insects such as fire ants and yellow jacket wasps, as well as in the venom of snakes [197]. In the latter, two different SCP/TAPS with ~50% homology to mammalian SCP/TAPS were able to block ion channels and the depolarisation of smooth muscle membranes [341,342]. Indeed, SCP/TAPS have been reported from snakes across five continents [343]. While the mechanism of a few of these proteins have been described, most are yet to be elucidated. In vertebrates, SCP/TAPS have been identified in rats, mice, rhesus monkeys, and horses. Due to their high expression in the epididymal lumen, they have been hypothesised to be involved in sperm-oocyte binding [344]. This aligns with the early finding that particular SCP/TAPS are expressed specifically in the human testes and have been implicated in sperm maturation [345].

SCP/TAPS gene orthologues and protein homologues have been frequently reported in helminths. Not just in the aforementioned species related to the human hookworm, but also in various other nematodes. Three SCP/TAPS proteins - *Ov*-ASP-1, *Ov*-ASP-2, and *Ov*-ASP-3 - have been described from *Onchocera volvulus*, a filarial nematode. Each of these ASPs has been shown to have distinct transcription patterns at various stages of parasite development [346,347]. While little is available on their function, *Ov*-ASP-1 was shown to afford some anti-larval immunity in mice [348]. A single SCP/TAPS protein from *Brugia malayi* called *Bm*-VAL-1 has been identified and found to be specifically restricted to larval stages, highlighting the probability for a role in the invasion of the host [349].

Comparing the expression of SCP/TAPS between free living and parasitic hookworms, the importance of these proteins in parasitism becomes clear. SCP/TAPS were the most abundantly represented group of mRNAs in activated L3 *A. caninum* [266]. *Ac*-ASP-1 and *Ac*-ASP-2 were found to be a key components of the ES proteins of serum-activated L3 larvae [198,199]. Further study showed these proteins to be localised to the granules in the glandular oesophagus, indicating a role in host invasion [200]. This was expanded on in adult *A. caninum* worms when four adult-specific ASPs (*Ac*-ASP-3-6) were described.

Ac-ASP-3, like *Ac*-ASP-1 and *Ac*-ASP-2, was found to localise to the oesophageal glands of adult worms. *Ac*-ASP-4, 5 and 6 were found to localize to the cuticular surface, intestinal brush border membrane, and excretory glands respectively [350]. All four proteins were found in the ES products, supporting the notion that they are released by the parasite [350]. *Ac*-NIF, belonging to the SCP/TAPS family and upregulated in adult *A. caninum*, was reported to block neutrophil recruitment and limit ROS release [351]. SCP/TAPS have also been studied from trematodes that infect humans including *S. mansoni*. Of the thirteen identified SCP/TAPS isolated from *S. mansoni*, transcriptional analysis showed higher expression in the invasive stage of the parasite [196]. From this, roles of these proteins in protease modulation of the immune response have been suggested [196]. SCP/TAPS have been the focus of a large body of research (some of this just described) into understanding host-parasite interactions. The vast number of these proteins, the diversity in stage-specific production and localisation, and the small number which have had their function explored highlights the importance and potential in choosing these proteins to investigate. Of particular importance is the unprecedented expansion of SCP/TAPS found in the *N. americanus [169]*. More than half of the 137 SCP/TAPS proteins were overexpressed in adult worms, with only 6 having orthologues in *C. elegans* [169]. This attests to the notion that nematode SCP/TAPS are likely to have developed before parasitism.

The proteomics data generated in this chapter has implications beyond the scope of this thesis. Analyses of other helminth proteomes have shown that ES products are not only a rich source of potential immunoregulators but also contain promising vaccine candidates and diagnostic markers. For example, SCP/TAPS from both *A. caninum* and *N. americanus* have been used in vaccine studies in mice and humans respectively [200,201,207]. More recently, a cysteine protease from *A. caninum* was shown to induce high levels of antigen-specific antibodies and confer protection against the parasite in a hamster model [352]. I identified 6 cysteine proteases in the ES proteome which could be explored for their vaccine potential in rodents and humans. In places where they are endemic, hookworms contribute significantly to poverty through their devastating effects on childhood nutrition, mental and physical development, productivity, and pregnancy outcomes [353,354]. Given these effects, a vaccine

would have a huge impact on the health and economic prosperity of these communities. While no vaccines currently exist for any human helminths, research into antigen discovery (including this thesis) has undergone exponential growth in the age of systems biology and 'omics' [355].

Given the lack of success in producing an effective vaccine for human helminths thus far, it is unlikely that a single parasite antigen will offer sufficient protection as a protein vaccine. Many parasites have multi-stage lifecycles that utilize a plethora of strategies to evade immune system detection. To add to the complexity, not all individuals elicit immune responses to specific antigens in natural infections [169]. With this in mind, the right combination of antigens in the form of a multivalent vaccine may be key to a successful vaccine for hookworm and other helminths. Indeed, a vaccine against the blood-feeding nematode *Teladorsagia circumcincta* in sheep found that a cocktail of recombinant parasite antigens could be used as an effective vaccine [356]. While vaccine-related studies were outside the scope of this dissertation, it should be noted that the catalogue of ES products generated may be explored for this purpose. The ES molecules are of particular vaccine-relevance due to their direct, frequent interaction with the host.

I also assessed the individual and combined diagnostic potential of each of the four recombinant proteins produced in this study. This was carried out by assaying serum IgG levels from individuals in a hookworm endemic area of Brazil. For individual proteins the highest score was 60% FoR which improved to 82% when 3 of the proteins were combined. This combined score greatly exceeded the FoR of the L3 extract, highlighting the diagnostic potential of ES proteins. Indeed, helminth ES products from a number of species have been shown to hold immunodiagnostic potential [357,358]. Conventional microscopy is the most frequently used method of detecting hookworms. It is however limited by the fact that many nematode eggs require specific expertise to diagnose accurately, it can be time consuming and may be less accurate in cases of low intensity infections. Given that the four recombinant proteins produced in this study were recognized by serum antibodies, the remaining ES

proteins offer a database of potentially new immunogenic targets to overcome the drawbacks of microscopy-based diagnoses.

In addition to the soluble proteins identified in the *N. americanus* ES products, there are other components of the ES milieu that should be explored in future work. Extracellular vesicles (EVs) and metabolites sourced from ES products have been shown to suppress TNBS colitis and impact on cytokine secretion by mammalian host cells [255,359]. Moreover, *H. polygyrus* EVs were shown to suppress eosinophilia and Type 2 innate responses in mice [360]. Recent work demonstrated that a single i.p. injection of EVs from *N. brasiliensis* afforded protection against colitic inflammation in a mouse model of disease [255]. These works collectively support exploration of human hookworm EVs. Indeed, some of the proteins identified in the ES proteome herein are known to be found in EVs, and the ES products were not subjected to ultracentrifugation to remove EVs from the crude prep prior to preparation for LC-MS/MS.

In chapter 3 I successfully expressed 4/18 SCP/TAPS proteins. The remaining 14, despite troubleshooting, were unable to be produced using the *P. pastoris* expression system. This poor rate of expression may be overcome in the future using any number of other systems including *E. coli*, *Saccharomyces cerevisiae,* insect or mammalian cells, or even cell-free expression. Each of these has advantages and drawbacks including cost, complexity, scalability, protein yields, endotoxin content and more [361-363]. Sometimes certain proteins are expressed better in particular systems, however the scope of this project precluded exploring other expression hosts due primarily to time constraints.

Given that I was still able to produce workable quantities of four proteins, I tested their therapeutic efficacy in acute mouse models of inflammatory diseases. I commenced with a chemically-induced colitis model which utilised a single i.r. injection of TNBS with ethanol and was able to successfully induce many of the clinical features of IBD in mice, including weight-loss and colonic inflammation. The TNBS model of colitis was ideal for the purpose of this thesis as it is well characterised in the literature and allows for high throughput screening of multiple proteins.

The outcomes measured from the TNBS experiments included: weight-loss and clinical scores, colon lengths and macroscopic scoring, histology, and colon cytokine profiles. In these experiments none of the four hookworm proteins tested were effective at preventing the induction of TNBS colitis. The TNBS model has been shown to be a valuable tool for identifying molecules from parasites with immune regulatory effects [255,313]. To further explore potential regulatory effects of the recombinant SCP/TAPs proteins identified herein, additional studies need to be undertaken, including dosage and frequency and route of administration. In support of this, a dose-dependent decrease in inflammation was observed in mice treated with *A. caninum* ES products. Similarly, daily i.p. administration of *A. caninum* ES products also produced a dose-dependent decrease in colon inflammatory cytokines in the DSS model of colitis [192]. It may be as simple as adjusting the dose or increasing the frequency of dosage of the recombinant proteins.

No single mouse model encompasses the complexity of human IBD [302]. For this reason, a multitude of mouse models have been established, each providing insight into a particular aspect of the disease [301]. Other models which could be considered for testing the recombinant proteins include: (1) dextran sulfate sodium (DSS), (2) adoptive transfer colitis, and (3) IL-10 KO mice to name a few. In comparing the mechanisms underlying each of these models, it is possible to hypothesise varying ways in which hookworm ES products may be used to treat colitis, despite not being effective in the TNBS model used herein with the current dosing and route of administration schedules. For example, DSS and TNBS colitis both employ chemicals to disrupt the colon surface epithelium. Bacteria and other large molecules are rendered immunogenic and can translocate and provoke inflammation and necrosis [303,364]. In contrast, the IL-10 KO or adoptive transfer colitis models use mice with compromised capacity to produce regulatory responses in the form of Tregs and alternatively activated macrophages [365,366]. If any of the tested proteins exerted their functions by augmenting adaptive immune responses, particularly regulatory cell responses, then the acute nature of the TNBS model may not have been sufficient to capture this effect. Indeed, helminth ES products from a number of species have been shown to promote Treg cell expansion and limit pathology across multiple models

of disease [367]. Moving forward, if future studies were able to produce greater numbers of *N. americanus* recombinant ES proteins and conduct similar studies, it is likely that new immunomodulatory molecules would be found. These molecule/s could be characterised by examining various immune populations – e.g. effector T cells, DCs, macrophages, Tregs – both at the site of inflammation and in lymph nodes. If certain cell types were specifically regulated, this might indicate which pathways to explore further. For instance, if regulatory cells were upregulated in mice administered with a particular ES protein, it would be prudent to use KO mice to test which receptors or pathways might be involved. The alternative to this theory is that these particular proteins have a distinct mechanism of action that may not be captured by the model employed. These proteins might also need to act in concert with other ES proteins to exert an anti-inflammatory effect.

To maximise delivery of HDM to the lungs for the asthma model, route of administration was compared between i.n. and i.t. A mixture of PBS with blue food colouring was prepared and delivered to mice under anaesthetic. Following dissection and using visual inspection the i.t. instillation showed markedly greater delivery to the lungs. This was important to confirm as I was utilising one of the most acute asthma models, with the smallest number of HDM administrations [307,310]. Following sensitisation and challenge with i.t. injections of HDM, I successfully observed HDM-induced airway inflammation and a corresponding increase in serum IgE. In particular, eosinophils as well as T and B lymphocytes were significantly augmented in HDM mice. This compared well with other studies of OVA and HDM mouse models which found increases in these cell populations [203,318]. None of the four hookworm recombinant proteins had any effect on the inflammatory response when administered throughout the challenge phase of the experiment. This model was selected because it allowed for testing molecules in the shortest experimental timeframe as well as requiring a relatively lesser quantity of hookworm protein than if employing a chronic model. While I did not see differences with the tested indices, others have reported a multitude of factors which are differentially expressed in similar asthma models. One study of HDM-challenged mice found 31 mRNAs which were significantly upregulated in lung tissue [310]. This included genes linked with human asthma (e.g. *ear11, il-10, il-21, clca3*) and with

leukocyte recruitment in the lungs (e.g. *ccl11, 12* and *24*). The same study also identified a panel of cytokines including IL-4, IL-5, IL-6, IFN-γ, IL-1β, TNF-α, IL-13 and IL-33 which were elevated in the BAL and lung tissue. I observed in this study that key markers of inflammation (immune cells and IgE) were upregulated in allergen-treated mice. Although there were no differences between protein-treated and untreated groups in terms of these inflammatory markers, this does not necessarily mean that other factors weren't affected. Future studies with these proteins would benefit from a more comprehensive analysis of disease indicators including gene regulation and lung cytokine levels.

Earlier in this thesis the sequence similarities were compared between the human hookworm secretome and those of published similar species. A key outcome from this comparison was that the majority of ES proteins from *N. americanus* had a significant match to a compared species. Despite the relatively high sequence conservation, these matches still contained large tracts of sequence variation. It is difficult to draw convincing conclusions as to whether the variation in sequence – and therefore folding and function – could be a factor in why the proteins from the human hookworm weren't effective in mouse models of disease. The vast majority of published roundworm-derived immune regulatory molecules have come from species which infect non-human animals which have been tested in their relevant hosts. Furthermore, the ES products of *N. americanus* and *A. caninum* have been reported to have varying protein patterns and bind to distinct leukocyte populations [368]. This may mean that the human hookworm ES molecules may have distinct effects on the human immune system [369]. To rule out host-specific differences it may be useful for future work to utilise more human applicable models, building on the *ex vivo* findings herein using human PBMCs or even intestinal biopsy tissue co-cultured with hookworm recombinant proteins.

As there were no successful therapeutic outcomes from mouse inflammatory models described in chapter 4, I moved on to *ex vivo* screening of the hookworm proteins with human PBMCs. This assay is frequently used in the selection of biotherapeutics for its potential to predict immunomodulatory compounds as well as test molecule toxicity [321,370]. Initially I sought to determine whether any of

the proteins had an inherent pro-inflammatory effect by culturing them for 24 hours with unstimulated PBMCs. NP4 induced potent IL-6, IL-8, and TNF-α release into the supernatant when compared with untreated cells. Given that all of the proteins were screened for endotoxin prior to the assay, the presence of an endotoxin contaminant can be ruled out in interpreting this result. Furthermore, the absence of an IL-1β response supports the notion that LPS is unlikely to be present [371]. Further purification to remove other contaminants, such as yeast derived zymosan, would be useful in better attributing the cytokine responses to the hookworm protein. However, given that these yeast derived contaminants were likely present in all samples, the cytokine responses are most likely caused by the hookworm proteins. It is likely, therefore, that NP4 through some undetermined mechanism causes the release of IL-6, IL-8 and TNF-α. IL-6 is produced by multiple cell types and is involved in a number of anti- and pro-inflammatory processes [372]. These processes include the activation and differentiation of macrophages and lymphocytes [373,374]. A study of IL-6 KO mice infected with *H. polygyrus* showed that IL-6 limits the Th2 response, modifies regulatory responses and increases host susceptibility to infection [374]. While this doesn't reveal a mechanism for NP4, it does provide a possible rationale for an ES protein from *N. americanus* that causes increased IL-6 production.

When PBMCs from healthy donors were stimulated with PMA/ionomycin, significantly elevated amounts of IL-6, IL-8 and TNF-α were released into the supernatant compared to buffer-treated cells. This result was expected, as PMA/ionomycin-stimulated cytokine responses have been well characterised [370,375]. As with unstimulated cultures, IL-6 and TNF-α were further augmented for cells treated with NP4. Interestingly, both IL-8 and TNF-α were significantly diminished in cells treated with NP17 before being stimulated with PMA/ion. TNF-α is a particularly important pro-inflammatory cytokine which has been implicated in both Th1 and Th2 mediated immune responses [376]. As mentioned earlier, TNF-α upregulation is significantly linked with the development of IBD, but it also plays a role in allergic airway inflammation. This has led to the development and use of anti-TNF molecules for the treatment of these and other immunological conditions [97]. Furthermore, blocking TNF-α in IL-4 KO mice delayed *Trichuris muris* expulsion, implicating this cytokine in defence against

parasitic infections [377]. Given the central role TNF-α plays in immunity against parasitic infections, as well as in immune diseases, it makes sense that a hookworm ES molecule would selectively inhibit its release. The follow up experiments to determine the molecular mechanism by which NP17 inhibits TNF-α release will be discussed later in this chapter.

When PBMCs were stimulated with LPS, I observed an increase in a number of cytokines. Significant levels of IL-1β, IL-6, IL-8, IL-10, and TNF-α were observed from PBMCs 24 hours post-stimulation with LPS. This profile of cytokines has been observed in other studies [378,379]. None of the hookworm recombinant proteins had an effect on LPS-induced IL-6 or IL-10 production. NP3 significantly increased IL-1β, a pro-inflammatory cytokine predominantly released by monocytes and macrophages with a protective role against viral and bacterial infections [380,381]. This IL-1β release is dependent on nod-like receptor (NLR) family, Pyrin Domain Containing 3 (NLRP3) inflammasome activation which can be triggered by various PAMPs [380]. In the case of helminth infections, *H. polygyrus* and *T. muris* infections increased activation of NLRP3 and IL-1β production [382,383]. In contrast to the usually protective role, *T. muris* NLRP3 activation inhibited Th2 immunity [383]. When NLRP3 was ablated, a reduction in IL-1β followed, with increases in Th2 responses as well as accelerated helminth expulsion [383]. Interestingly, helminth ES products were sufficient *in vitro* to induce NLRP3 and the corresponding IL-1β release [383,384]. This result aligns with my finding that NP3 selectively augmented IL-1β release from PBMCs. A rationale for this outcome is that helminths have evolved to manipulate these immune pathways, dampen host responses, and prolong their longevity.

As with the PMA/ionomycin-stimulation experiments, LPS-induced TNF-α release by PBMCs was significantly reduced by NP17. Similarly, the TNF-α response from LPS stimulated cells was also significantly reduced by NP4 and NP12 but to a lesser degree than NP17. While NP17 decreased TNF-α from PMA/ionomycin stimulated cells, NP4 and NP12 did not have this same effect. This result draws attention to the differences in cell activation and the corresponding immune signalling pathways activated by PMA/ionomycin vs. LPS, and their eventual impact on TNF-α secretion. PMA activates

protein kinase C while ionomycin regulates calcium influx [325]. Together these molecules lead to activation of intracellular signalling from T cells and increased TNF-α [325]. In contrast, LPS binds to TLR4 on monocytes, activates NFκB signalling and consequently augments TNF-α gene expression among other pro-inflammatory cytokines [385]. NP4 and NP12 selectively reduced TNF-α production from LPS-stimulated cells, indicating that these proteins are affecting monocytes specifically in some way. I cannot however draw meaningful conclusions as to the subsequent cell-cell interactions which may also be having an effect on cytokine production. For example, TNF-α produced by LPS-activated monocytes can cause T cell proliferation and increased cytokine production by those proliferating cells [386,387].

It is important to consider the next steps, as well as limitations of the current study, so that future experiments can elucidate the mechanisms of action by which these *N. americanus* SCP/TAPS proteins regulate human immune responses. One of the main limitations of this experiment is that I used whole PBMCs rather than specific cell subsets. PBMCs are a heterogeneous cell population made up of many immune cell types including, but not exclusively, lymphocytes (70%), B cells (15%), NK cells (10%) and monocytes (5%) [388]. This breakdown is highly variable and as such the interpretation of results from mixed populations can be challenging. This is due to the fact that certain cell populations communicate with and affect other cell responses. For example, LPS-activated macrophages release cytokines that can act on lymphocytes to induce more inflammatory cytokines. In this way, the use of stimulants (LPS or PMA/ionomycin) in this chapter did not exclusively target specific cell responses. Another point to consider is that in a PBMC mixture, as noted above, the ratios of different cell populations can change substantially. For this reason, lymphocyte responses (70% of PBMC total) may mask monocyte responses (5% total). These complications can be overcome through the isolation of individual populations of immune cells prior to experimentation [388,389]. For example, monocyte populations can be isolated with positive and negative magnetic sorting techniques using markers such as CD14 [390]. The same sorting systems could be employed to selectively remove the monocytes and culture the lymphocytes specifically. Another option would be using whole PBMCs, to capture cell-cell

interactions, and using intracellular staining to look at cytokine production from defined cell subsets. It would also be prudent to compare immune cells isolated from humans with those from mice. This is particularly important due to the frequent use of murine-infecting helminths as a model for the human hookworm. It may also add to the overall understanding for the lack of efficacy of these proteins in mouse models of disease. Indeed, there are dissimilarities between gene expression profiles of human and mouse monocyte subsets [391].

Isolation of cell populations would be particularly beneficial for future experiments into the mechanism of action of these proteins, particularly for exploring gene expression. RNA sequencing (RNA-seq) is the gold standard for characterisation of gene expression changes and would be instrumental in further characterisation of the proteins and informing future work [392]. By comparing the expression levels of specific cell populations treated with and without the ES proteins, the immune pathways they affect may be elucidated. All of these future experiments are essential for better understanding the function of individual ES proteins from the human hookworm. The next steps include manufacturing recombinant versions of many more of these proteins and carrying out similar screening as in this thesis. The vast majority of the information available for these proteins is theoretical, based on protein family analyses and sequence similarities. Future work should build on the experiments as discussed here using functional assays to determine potential effects *in situ*. By looking more thoroughly at the underlying potential gene regulation of these proteins on immune cells the mechanisms of action can begin to be unravelled.

In chapter 5 I used human protein microarrays to begin to explore some of possible binding partners for each of the recombinant proteins. Given the lack of functional characterisation for these SCP/TAPS proteins, this was a useful way to start to explain the immune regulatory effects seen from the *in vitro* work with PBMCs. NP4 notably increased production of IL-6, IL-8 and TNF-α from otherwise unstimulated whole PBMCs. Probing the microarray with NP4 revealed that the most significant binding partner was TNF receptor-associated factor 3 (TRAF3) interacting protein 1, a protein involved in

tethering TRAF3 to cytoskeletal microtubules [393]. TRAF3 is involved in the mediation of CD40 signal transduction which plays a role in TNF receptor immune responses [394]. Through its binding with TRAF2, this protein has been shown to regulate activation of NFκB and MAPK [393]. While this result does not indicate the specific mechanism by which NP4 functions, it validates the *in vitro* result and provides support for this particular SCP/TAPS protein's regulation of host immunity. Interestingly, NP3 and NP4 also bound to SH2D2A on the array; SH2D2A is a T cell-specific adapter protein (TSAd) protein involved in T cell activation [395]. This protein has been shown to be induced in activated NK cells as well as T cells exposed to activated macrophages [395]. This provides a potential mechanism by which NP4 upregulated inflammatory responses in whole PBMC mixtures. Interestingly, a protein from *N. americanus* was shown to bind mouse and human NK cells, inducing IFN-γ production [368]. Future work with NP3 and NP4 should include purifying NK cells and checking whether they induce IFN-γ. NP17 conversely downregulated inflammatory responses, specifically TNF-α. When the human proteome array was probed with NP17, the second most significant binding partner identified was LENG1, a leukocyte expressed immunoglobulin receptor. The leukocyte receptor complex encompasses a wide array of different ligands expressed on NK, cytotoxic T, and myeloid cells [396]. This gene family has been broadly implicated in many immune recognition pathways [396]. While little information is available on the LENG1 receptor in the literature, this does allude to an immune regulatory role for NP17.

This project began with a reannotation of the human hookworm genome as well as a proteomic description of the ES products. This is the first report of the *N. americanus* secretome and will increase capacity for future research in the areas of hookworm derived vaccines, diagnostics and therapeutics. I used this data to inform my approach in the following chapters, selecting numerous SCP/TAPS to express in a yeast expression system. The aim was to begin to characterise some of these proteins through the use of mouse inflammatory models as well as *in vitro* cultures, and microarray screening tools. I have identified one candidate protein which robustly downregulates specific cytokine responses *in vitro* with human PBMCs, and begun to identify putative binding partners using human proteome

microarrays. This research broadly advances the knowledge of the variable functions of helminth ES proteins. It also adds specific insight into the relatively unknown function of SCP/TAPS proteins, and some of the hits from the human proteome array studies now need to be validated. Further work should aim at exploring more of these proteins given their abundance and seeming importance across many parasitic helminths. Given the increasing burden of chronic inflammatory diseases worldwide, there is a greater need than ever to better understand, treat and manage these conditions. This project offers a novel resource for the biomedical community and ultimately might contribute to improving the outcomes for people with hookworm infection as well as those with chronic immune disorders.

7 Chapter 7 – Supplementary data

7.1 Supplementary material from Chapter 3

Sequences of selected *Necator americanus* proteins sent to IDT.

NECAME_01191 (NP1)

Nucleotide (NT) sequence optimised without stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCATGTGCTCAGGAAATATTGGCATGTCAAACGATGATGAGCGTAACGCATTTTTATCAAAACACAATGA
GCTTAGGTCTAACCTGGCCAGGGGCTTAGTCAAGAACGGCAAAACCAATTCTTATGCCAGGCGTGCCTCAAAA
ATGAACAAGCTAAAGTACGATTGCAGTGCTGAAAAAATAGCATACAATTGGGCCTCTAAGTGTATTGATAAGC
AGAGTCCTAAAGAACAGAGACCCGGTCTTGTCGAGAATAGGTATGTTTTTAACAACATGAACATAAATAACCCC
CCCCTAAGAGCCTCAAACAGGTGGTTCAACGAAATCACTACAGTTGGTGTTGTCCAAGATAATGGTAGTAATTA
TTTCACGTTCAGTCTAGGCATAACCCATTTTGCAAAGATGGCCTGGGACAGTGCAACGAAGATGGGTTGCGCT
GTGTACAAGTGCAGTACTTATTTGCATGCAGTATGTGAATATGACTCTCCTGTTAAGGATAACGCACAGATATA
CAAAATGGGTCCACCCTGTCACAGATGCCCATCAGGAGTCTCATGTTCCAGTGGCCTTTGCGTAGTCGGTCTAG
A

Translated protein sequence

MCSGNIGMSNDDERNAFLSKHNELRSNLARGLVKNGKTNSYARRASKMNKLKYDCSAEKIAYNWASKCIDKQSPK
EQRPGLVENRYVFNNMNINNPPLRASNRWFNEITTVGVVQDNGSNYFTFSLGITHFAKMAWDSATKMGCAVYKC
STYLHAVCEYDSPVKDNAQIYKMGPPCHRCPSGVSCSSGLCVV

NECAME_06164 (NP2)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCACCGCAACTACTAAAATGACGGAAGTGCTTGCTCAAGCCGAGGCTAAAAAATGCGGCCCTATCACCA
CACCCAAAAATTACGGTTCTACTCAAAATATGATCTCTGTCAAAAAGACGTGTGACGCCACCGAAACAGTTCGT
AAGGAGATTTCAACGTGGTGGAAAGACGGTGCTACACAACAGACGTCTTCCACTAAAGTAGACAAAAATGATA
AGTTCTCACAGATGGCTTACTTTGAAACTAGTGCCGTAGCCTGCTCCTATTACCCATGTACTGCAACGCAACTAT
CCCTAGTTTGTTTATATAACAAAAATGGTGCTGCAAAAACGGTAAAAGAACTTTACAATAACGCTACGGGAAAG
TGTAATTGCACAGGTTGCACGGCTTATCTTTGTCCATCTGAGTTTACACCTGCCCAGCTTCCTTCCTCTGTCTGTT
CTTCCTGCGCACCTTCTTCTGAGCTATTTAGGGAGGTCGCCCTTTACTCACACAACTATTATAGACATGACAGAC
TAACGAGATTGGTGGCAACAGGTTGGGCTGAAGATAAGAAGTCAAAATACGCTAAACCTGCAAAGGCTATGA
TGGAGCTTCAATATAATAAGACGCTGGAGGACAAGGCAAAAAAGCACGTTAACAAGGCTACCTGCCCAACAAC
ACCTCAAAATGCAGAATACGCAGGCGAAAGTTTCTGGACATCAGAGAAGTACTCACTATCCGAAGAAGAGGCA

GTTAAGAAAGCAATAGAGAAATGGTTCTCCTACTTAAAGAACTCCGGACTAGGTGACAATATGGATTACACATC
ATTGAAAAAGGCCTCCGGTAAACAACTAGCAAATGTCATACATGATAAGACTACATCCGTGGGTTGCTTTGTTA
AGCGTTGTGAAAAGCAAGGCGCAATAGTTGTGGACTGTAGATACGATCCTTATGTCATCACCCACGAGACCAA
CAATACCGCAACTTTTTTAGGAGTGAATGTAGAGGATGCTAGATACGGTCTAGA

Translated protein sequence

MQGTRYLSSKAAIIVVFTVPAIDATATTKMTEVLAQAEAKKCGPITTPKNYGSTQNMISVKKTCDATETVRKEISTW
WKDGATQQTSSTKVDKNDKFSQMAYFETSAVACSYYPCTATQLSLVCLYNKNGAAKTVKELYNNATGKCNCTGCT
AYLCPSEFTPAQLPSSVCSSCAPSSELFREVALYSHNYYRHDRLTRLVATGWAEDKKSKYAKPAKAMMELQYNKTLE
DKAKKHVNKATCPTTPQNAEYAGESFWTSEKYSLSEEEAVKKAIEKWFSYLKNSGLGDNMDYTSLKKASGKQLANVI
HDKTTSVGCFVKRCEKQGAIVVDCRYDPYVITHETNNTATFLGVNVEDARY

NECAME_01366 (NP3)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCGTGCCACCTATGGCCGCCAAGCTAGAGGCATGCGACGAAAATAACCTCGGCGCTTATCGGGAACTGT
ACTTGAAAACCCACAACGATTACAGACTCTCACTTGCAAAAGGAGAAGTAGAAGCAGAATCCGGGAAACTGCC
AGGATCGGAGAGCTTGTTCAAATTTAAGTATAACTGTCAACTCGAAGGGCTCGCTATGGCTGTTGCTATGACAT
GTGAACGTAAACCTTTGTTAAACATGAAACAACTCGGATATGGTCAAAACTATTACTGGACTAGAAAAGTCAAC
AATGATATAGTAGACGCCGTAGACAAGGTTCAAGATCGTTTCAATTTTGCTGTTGAAGAATGGTGGTATACGGT
AGAGGATGCGAAAATCGGTCCAGATGTCATATATGACGATGAAGCGGTGGAGCCATTTGCAAACATAGCTTAC
AACAAATCGATATCCTTTGGGTGTTATAACAATCGATGTGGATCCAATCCTAGCAGTCTCGCTACGGCCTATGC
ACCGGCTGCAGGCACAACAGAAAACACCACACCAGTAACATCGGCTACAGACGAAACTACGGTGGACTCGACT
ACAGCTGAAACTACGGTGGACTCGACGACAACCACAGTCACATCGACTACAGTCGAAACTACAGTGAAATCGA
CTACTGACGAGACTACGGTAGAATCCAATACACAACAAACGACTGTTGCACCAACAACAGCTGCTAGTGAAAT
GGATGAAACAATACGAAAAGAAATACTTCGCACGCACAACGATTACAGATCACTGCTTTCTCAAGGGTATGTTA
GAAATGGGAAGGTAGGAAAGCTTAATCTGCCAACAGCTACCAATATGCTCACAATGAAATACGATGAGTCTAT
GGAAACGGAGGCTCAAGCGTATGCGGATCAGTATTGTCTCGTTGGATCTGGTCCTAATTCGGGCCAAAACACT
CATGTCATTAAATCAACATCGGTTACTGCGCTCGATGCATTTAAGCAGTCTATGCGAACATGGTGGGAACAGAT
CTCTTCAACCTCCGTGGGATCAAATTTGAAATACACTTCCACTTTAGAAGAGAAGAAAAACGCACCCACGAGTT
TCACTCAGATGGCGTGGGCTGAAACCTACAAAGTCGGATGCGGGATAAAACGCTGCTCCTCCAAAACCTTCGT
AGTCTGCCGCTACGATCCACCTGGAAACATCCTCGACAAAAACATCTACGTACAAGGCGCTATCTGTGGAATTT
GCCGAAGCAACTGCATTGCAGCACTCTGTCCAACACCTGAAAACGGTCTAGA

Translate protein sequence

MLPILLVILILYHAVPPMAAKLEACDENNLGAYRELYLKTHNDYRLSLAKGEVEAESGKLPGSESLFKFKYNCQLEGLA
MAVAMTCERKPLLNMKQLGYGQNYYWTRKVNNDIVDAVDKVQDRFNFAVEEWWYTVEDAKIGPDVIYDDEAV
EPFANIAYNKSISFGCYNNRCGSNPSSLATAYAPAAGTTENTTPVTSATDETTVDSTTAETTVDSTTTTVTSTTVETTV
KSTTDETTVESNTQQTTVAPTTAASEMDETIRKEILRTHNDYRSLLSQGYVRNGKVGKLNLPTATNMLTMKYDESM
ETEAQAYADQYCLVGSGPNSGQNTHVIKSTSVTALDAFKQSMRTWWEQISSTSVGSNLKYTSTLEEKKNAPTSFTQ
MAWAETYKVGCGIKRCSSKTFVVCRYDPPGNILDKNIYVQGAICGICRSNCIAALCPTPEN

NECAME_01369 (NP4)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCAACTCACCATCAGTCCGGGTTGTGTGTTCAAATCCACCGAGACCGGATGATAAAGTGCGTCTGGAGTT
CCAAACTCAGCACAACAACTACAGAACGAATCTTGCGACTGGAACCGTAGACGATGCCGGGGGCACTCTACCA
GGATCGACAAGCTTGTTTATGATGTCATACAACTGCGAACTGGAAATACAGGCTCTAACGGTTACCAGCGCATG
TGATACTCAGAGCAACACACCAACTGTCATAGGGGGAAATTCTATAAACTATGGAATAATTCACAGTCACCCTG
GCGCTCCTTCTGATGATGATGCTGTAAAGGACATTCAAGCGCTGATGACTAAATGGATGGAAACGAGTTACGA
ATATAATTTCGACAAAAAGACCGTTATCTACGGTAGCGAAGATGCTGGACCGTTTGCAAGGATTGTTTACAACA
AGTCGATTTCGCTGGGATGTGCGATAACTCAGTGTGCAGCGAGAAAAACTGCATACGCTTGTGCATACAGCTC
AAGTCCAGTCGTTGGAGAACCCCTTTATTGGCCGTCAAAAAATCCTACCGGTTGCACGAGACCCGGTCAGTGTA
GAAAGGCAATACCTGGTACAGCTACCACGTGCGATACTAATCTAAACCTCTGCTCAACGAATTTGTTAACATTA
GATGATAGCACTTCAGCAGCACCAACAACTCAATCCCAAACCACCACACCAGCTCCTGCAGAATCAACTACGCC
TGCGGCTGTGCCTGGTGCCAGTGGTGTGATGACAGATGCGCTTCGTCAGATGGTAGTCGATACACATAATAGG
GACAGATCGCTGCTTGCTCAGGGAACCATCAGAAATGGAAAACCAGGAAAGCCGAACCTTGGAACAGCCACA
AATATGTTAAGGATGAGGTACGACATGGCGATGGAGACGGAAGCACAAGCCTACGCTGACACTTGTATTCTCA
ACAAATCTGCTGTAGCTATTCGCCCGAACTCTGGTGAAAACGTTTATGTTATTCAGTCATCGACAATTACAGTAT
CCGACGCATTGACACAATCAATGCAATCATGGTGGAACCAAATCTTTATCAATGGATTGAACGCCAGACTGATG
TACACACTGTATCTGGAAACAAAGTCCAATGCTCCTCTTGATTTTACTCAGATGGGATGGGCTGCCACCTACAA
AGTTGGATGTGGAGTTAGCCGTTGTGGCATGTCCACTTTCGTCGTTTGCCGCTACAATCCACGTGGAAATATTA
TCGATCAGTTCATCTACAATGTTGGCTCGATTTGTGCAACATGTATGAGTTCATGCACGGATGCACTTTGTGCAA
CACCCGCGTACGGTCTAGA

Translated protein sequence

MFTIVLVLLCAACATANSPSVRVVCSNPPRPDDKVRLEFQTQHNNYRTNLATGTVDDAGGTLPGSTSLFMMSYNC
ELEIQALTVTSACDTQSNTPTVIGGNSINYGIIHSHPGAPSDDDAVKDIQALMTKWMETSYEYNFDKKTVIYGSEDA
GPFARIVYNKSISLGCAITQCAARKTAYACAYSSSPVVGEPLYWPSKNPTGCTRPGQCRKAIPGTATTCDTNLNLCST
NLLTLDDSTSAAPTTQSQTTTPAPAESTTPAAVPGASGVMTDALRQMVVDTHNRDRSLLAQGTIRNGKPGKPNLG
TATNMLRMRYDMAMETEAQAYADTCILNKSAVAIRPNSGENVYVIQSSTITVSDALTQSMQSWWNQIFINGLNA
RLMYTLYLETKSNAPLDFTQMGWAATYKVGCGVSRCGMSTFVVCRYNPRGNIIDQFIYNVGSICATCMSSCTDALC
ATPAY

NECAME_07493 (NP5)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCGTCTATGAGTGTGATGCCGGCGGAAAAGTCCCCAATTCAATGAAGAAGCAGATTGTAGATTTACACA
ACCAACTTCGTCAGCAGTTGGCTAATGGAGAAGTACAAGGCGCAACTGGAAAACTGCCTAAGGCAAAGAATAT
GCCAATGTTGCAGTGGGATTGTGACATGGAAGCCCTGGCAGCCGAGCGTCTGCGTAACGGTAGGCTAGATGT
GAATTTTCTAGGTATGGCCGGCTACAATGACGACGCCTTCACAGGTAGAATTAGAGCCAATGCCGTCATATACG
AGAAGGCAGAAGCAAAAGCCGTCGGTCTAGA

Translated protein sequence

MMFAIYILILFVSVASAVYECDAGGKVPNSMKKQIVDLHNQLRQQLANGEVQGATGKLPKAKNMPMLQWDCDM
EALAAERLRNGRLDVNFLGMAGYNDDAFTGRIRANAVIYEKAEAKAV-

NECAME_09334 (NP6)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCGCTATATCCCGTCAAAAGAGGGCCACTTTTGGATGTAGTTCTGATATCATGTCTGATAATACCCGTCAA
GTTTTTCTGAACTACCACAATGAAGCCAGATTGCGTGTAGCAAAAGGTATAGAACCCAATAAAAAAGGTTTTTT
AAACCCCGCCAAAAACATGTATAAACTGGAATGGGACTGTAACATGGAAAAACAAGCACAGGAAGCCATCGCC
AGTTGCCCTAATAGTTTCTCTCCATGGCCAAACATGGGTCAGAATCTTATGAAATTTACATCTGGAAGTGGATTT
TCTAATCCCCCTGGACAGATAAAGTTTGCACTTGATAACTGGTGGTCTAGAGCTAAACAATATGGTGTTACAGA
TGCCGAAAACAGGTACACCAATGGCTACCTGTATACATTCGCAAACATGGTATTTGCAGAAACCACGAAGTTG
GGATGCGCCTATGCTATATGTGGCAAGGAATTAGTCATCACTTGCTTATATAACGCTATAGCTTATTATACCAAC
AACCCTATGTGGGAAACGGGAAAGGCTTGCGAAACGGCAGAGAACTGCACTACCTATGCCAACTCAAGGTGTT
CAAACGGTCTATGTACGAAAGGAGCTGACCTACCAGAGACAAATAACCAGTGTTCCTCAAACAATGGCATGAC
AGATTCTGTCAGGCAGACTTTCCTAAGTATGCATAATAAATTCAGAAGTTCTGTGGCCAGGGGTCTAGAACCTG
ATGCATTGGGTGGTAACGCTCCCAAAGCTGCTAAAATGTTAAAGATGGTTTACGATTGTTCAATAGAAGCCTCT
GCATTAAAACATGCCCAAAAGTGCATAAGTGAGCATAGTAATAAAAAGGATAGGCCTGGACTTGGAGAGAAC
ATTTATACGACAACTGCACTTAATTTCGACAAGGTGAAGGCTGCTTCACAGGCTTCACAGCTTTGGTGGGGTGA
ACTAAAGAAATTTGGTGTGGGACCCTCAAACATACTTACAACAGAGTTATGGAATAGGGAGAACACTCAAATA
GGCCACTACAGTCAAATGGCTTGGGAGACCTCTTATAGGTTAGGATGTGCAATTTCACATTGTCCAAAGTTTAC
GCTTGGAGTCTGTCAGTACGGCCCAGGTGGAAACATTCTTGACCACTTAATCTATACCATCGGAAACCCCTGTA
CGAGTGATGCCGGTTGTCCTGGATCTTACACATGTTCAGTTGAAGAAGGATTATGCAATGTTGTAGGTCTAGA

Translated protein sequence

MTQAAAYICITGYCGYKYKGFLSSRHLSAWMIPTSLLPILLFAVSTSNAAISRQKRATFGCSSDIMSDNTRQVFLNYH
NEARLRVAKGIEPNKKGFLNPAKNMYKLEWDCNMEKQAQEAIASCPNSFSPWPNMGQNLMKFTSGSGFSNPPG
QIKFALDNWWSRAKQYGVTDAENRYTNGYLYTFANMVFAETTKLGCAYAICGKELVITCLYNAIAYYTNNPMWET
GKACETAENCTTYANSRCSNGLCTKGADLPETNNQCSSNNGMTDSVRQTFLSMHNKFRSSVARGLEPDALGGNA
PKAAKMLKMVYDCSIEASALKHAQKCISEHSNKKDRPGLGENIYTTTALNFDKVKAASQASQLWWGELKKFGVGPS
NILTTELWNRENTQIGHYSQMAWETSYRLGCAISHCPKFTLGVCQYGPGGNILDHLIYTIGNPCTSDAGCPGSYTCS
VEEGLCNVV-

NECAME_07497 (NP7)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1 with point mutation

GAATTCCAAGACGTTGGAAACACCCCTGTTAAGCCACTGTGCAAGGGTGGAACATACGACGAGGAGTATATAC
ACGATTATGTAGTTCAATCTATCAATGGACACAGATACAAGTTATTGAGGGGCACTGAGCTGAACGGCCCATG
GAATGGAAAAGGTTATGGTAAAAAGTTTCCCCTAGCTAAGACTATGAACAAACTTGAATACAACTGCAATTTAG
AAAGGAAAGCCAGTGCTCTGTTAGATAAATCATGCTCCACCAATACCCCTACAGGCCCAAACGGTACCGCAGC
ACTATTCTATCACATTGACATAGACTTTGACGACCCAGAACTGTTTTCCGCCATATGGGAGTGGTCAGAAAAAA

TAAGGGAGTTCGCCGTATCTGATGATGCCATAACCAAGAAACAGGTGGTGTATAAGGATAAAGATAGAAATCT
TTATGAGTACCTGAACCTTATGCGTACGGCAATCTCAAAAATCGGATGCGCCGATATAACGTGTAGAAATGGC
GACCTGAGGAAATATAGGGCAGTGTGTCTTATCAATAAAGAGGGTCTAGA

Translated protein sequence

MLIIFSIAVGALVLPTTLPQDVGNTPVKPLCKGGTYDEEYIHDYVVQSINGHRYKLLRGTELNGPWNGKGYGKKFPLA
KTMNKLEYNCNLERKASALLDKSCSTNTPTGPNGTAALFYHIDIDFDDPELFSAIWEWSEKIREFAVSDDAITKKQVV
YKDKDRNLYEYLNLMRTAISKIGCADITCRNGDLRKYRAVCLINK

NECAME_15063 (NP8)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1 with point mutation

GAATTCGCTGGCTTTAAGTGCCAAAACAGTCTTATCTCTGACCAGTGGAGAAAGGGAATACTTGAACTGATAAA
CGGTTTCAGGCGTACGGTAGCCAAAGGTGAGCAACTAGGCAAAGACGACAAAAAATTACCAACGGCCAAGAA
AATGAACCAGCTTTCATGGGATTGTAATATAGAAGCCGAGGCACAAGCTGCTGCAAAAACCTGTCCTACTTCCC
TTACCACGGCCCAACTGCCAGCCTTTCAGGCTACCGCATACAAGTTGGGTTACATCACTGCCACTGGAACAATC
AAGGCTAACTCATGTGACGCCACTAAAACGGCCAAAGACCTGCTAACACAGCAATGGAATACAGCCGTAGCAA
ATCAAGAAGATGAAAAGGTCGCCAGAGATGCTAAAGCTTTTGCTCAACTGGCTCATGATGAGGTTTCCGCCTTC
GCTTGCACGTATCAAATATGCCAACCAACGAGTTACTATCTACTTTGCTTCTTCAATCAAAAGGCCCCTACCAAA
GACGATGTTTATACTGCCGACGCCAAGGGCAAGTATTGTTCTAACGCCACCTGCACCTGTGTAGATGGTCTATG
CGAGCGTACCCCAAAGCCCGGTTGGGCCAAGTCTGGAGACAACTACGCTAAACCAGCTACTAAGATGATAGAG
CTGCAGTATGATAAGACGCTGGAAGATGCTGCTATTAAGGAAATAAAGGACTGTAAGGCCAACTTGAATAATG
CTGTTGAAAAGAACTTCTGGTCTCTTGACGTCACAACGAGTAAGAATATAGGTGAAGATGCTATTAAAAAGGCT
GTAGCTAAGTGGTGGAAACCCAGTGGTAACCAAGCCTTCGGCGAGAATAGAAAGAATACTAACGGTTTGATGT
CAGCAGCAAATATCGCCTACGATGAAGCAACTAAGGTCGGCTGCGCTATAGACGCCGGCGCTGCTTGTTTGAA
ACTAGGCAAACTATACATAATGTGTAAGTATGATAATAGTCCTAATGTTGGCGATGTAATATACGAAGCTGGAA
ATAGAGCCTGCTCTAAATGTGGCACGACCGGAACAAACCCTTCTTGCTCACCATTGGGTGGACTTTGCGTCAAA
AGTTCCGGTCTAGA

Translated protein sequence

MEEDIIIKTLFLASLQLYNNAVTSAAGFKCQNSLISDQWRKGILELINGFRRTVAKGEQLGKDDKKLPTAKKMNQLS
WDCNIEAEAQAAAKTCPTSLTTAQLPAFQATAYKLGYITATGTIKANSCDATKTAKDLLTQQWNTAVANQEDEKVA
RDAKAFAQLAHDEVSAFACTYQICQPTSYYLLCFFNQKAPTKDDVYTADAKGKYCSNATCTCVDGLCERTPKPGWA
KSGDNYAKPATKMIELQYDKTLEDAAIKEIKDCKANLNNAVEKNFWSLDVTTSKNIGEDAIKKAVAKWWKPSGNQA
FGENRKNTNGLMSAANIAYDEATKVGCAIDAGAACLKLGKLYIMCKYDNSPNVGDVIYEAGNRACSKCGTTGTNPS
CSPLGGLCVKSS-

NECAME_15467 (NP9)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCATGCGTTCACAGATTGCTCTGGGCACGGCACCCAATAAGAACGGAACCTGTCCAAGAGGACAAAATG
TGTACAAACTTAGTTGGGACTGCTTTTTGGAGATGCAAGCACAAAACGCAGCCGATCAATGCTCTGAGAACGT
GAAGGGCCCAACCGGTTACTCTCAACTTGTACAAAAAGTGCGTATAACCACTTGTAATCTAGCCCCAATCCCCA
AGTCCACAGTCGATGGTTGGTGGAGTGAGGTGAAGTCATTGGCTAATGGCAAGGCCACTAAAATCGGTTGTGC
CCAGCGTAATTGCGGTGCTGATTTGTACGTCGTATGCGTGGTTTACGATAGGGTTTTCACAACAGGCGGTCAAA
TATACAAGATGGGAGAGCCCTGTAAAAGGTGTAGTGCAGTAGGCCAGGCAGTATGTAAAGATAACTTGTGTG
CTCTAAACGGTCTAGA

Translated protein sequence

MRSQIALGTAPNKNGTCPRGQNVYKLSWDCFLEMQAQNAADQCSENVKGPTGYSQLVQKVRITTCNLAPIPKSTV
DGWWSEVKSLANGKATKIGCAQRNCGADLYVVCVVYDRVFTTGGQIYKMGEPCKRCSAVGQAVCKDNLCALN-

NECAME_15644 (NP10)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1 with point mutation

GAATTCATGTTCTCTGACTTTCAATGCTGGAACTTCGACTCTACGGACAACATTAGGGGTGCATATCGTACTAAT
ATCAACAACCTACGTTCTTCCCTAGCTAAAGGCGACGCTACGGGCAAAAATGGAAAGCTGCCCCAAGGCAAGA
ACATTTATAGGCTGGATTGGGATTGTATGCTTGAGGTGGAGGCCCAAAAAGTAGTAGATAAATGCGATGCCAC
GGCCAGTGCTCCTTCCGAATTATCTATGGTGATTGCTTCAGTTCCCATTACCACATGCGATCCAATAACTATGAT
AAAGCCCCAAGTGAACAAGTGGTGGAATACCGTCAAGAACGCAGCCGTTGACGGTACTTCCCCAACTTATAGT
AACAAAAAGCTGCATGACTTTGCAATGTTAGCCCACGGTTCCGCCGTCAAGCTGGGCTGCGCCCAGCGTAATT
GTAACGGTAAATTAGTGATGGCATGTATGCTGTATCCAAAGGGTCCCGACAATGGTCAGCCTATATACGAAAA
AGGCGGTGGATGTACGAAGTTAGATAACGGAGATGTGTCAACCGGAACCATGTGTCCTCAGAACTCCCAGATG
AGTACTGACGACGTGAGAAGAGAGTTCTTGAATGCACATAATGGTTACAGATCCACAATAGCTCAAGGTAAGG
GAGTTATGTCAGGAGGAATACGTGCTAGGCCTGCATCCCAGATGAGAAAAATGGAGTATGACTGCGAGGCCG
AAAAGCTTGCCTATAACGCTGTGTGCGGCACGTCAACTAATGTCGAAAACAGAATGACATTTGCCAATGCCAAC
AAGGATCCAGCCGAGGCTGCAAGACTTGTCCATTTGTTTGCCGTAAGGAGGTGGTATAATGAGATACAAACGG
GTACCATGGCTCAACAAGCAGGCGACAAGAACTTGTATAGTGTTAATTTAGCTATACCAAGATTTGCAAAAGTG
GTGTGGGAAACTAATGTCAAGCTAGGATGTGCTGTCCGTCAGTGTAATGGAAAGTTGAGAGTACTATGCATAT
ATTCAAATTCTACGTTCGGTGTAAACCAGCAGATATACAAAATGGGTGGTACTTGCAATAGGTGTAAAAACGTG
TGCGACAACAATACCGGCCTGTGCCCTTCCAACTGGACTCCATCAGGTCTAGA

Translated protein sequence

MFSDFQCWNFDSTDNIRGAYRTNINNLRSSLAKGDATGKNGKLPQGKNIYRLDWDCMLEVEAQKVVDKCDATAS
APSELSMVIASVPITTCDPITMIKPQVNKWWNTVKNAAVDGTSPTYSNKKLHDFAMLAHGSAVKLGCAQRNCNGK
LVMACMLYPKGPDNGQPIYEKGGGCTKLDNGDVSTGTMCPQNSQMSTDDVRREFLNAHNGYRSTIAQGKGVM
SGGIRARPASQMRKMEYDCEAEKLAYNAVCGTSTNVENRMTFANANKDPAEAARLVHLFAVRRWYNEIQTGTM
AQQAGDKNLYSVNLAIPRFAKVVWETNVKLGCAVRQCNGKLRVLCIYSNSTFGVNQQIYKMGGTCNRCKNVCDN
NTGLCPSNWTPS-

NECAME_16121 (NP11)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCATGGTTGAACCATGCAGTGACGACATAAAAGACACGCCTCCTAAAAGTAGTAAAGTTACCTGCCTATC
ACACCCACTTGGAAGACAAATAATCGAATTTGCTAGTTTAATTGGTGACTCTCACTCAGACAGAGCTTTCGGCT
GTCGTAACTCTCTTATATCCGATGAGTGGAGAGCAATGGTCCTTACGTTCCATAATGACAAGAGGAAGACACTG
GCTAGAGGCGATCAACCTGGCAAAGACGGTAATAAACTAAGTCCTGCCAAGCAAATGAACGAACTGACATGG
GACTGTAATCTTGAGAGATTCGCTTCAGATGCCGCAGCCCAGTGCAAAGATTACGATGATCCCCAGAGGGGAG
TAAACCAAAAAACACAGGTCTTAATTGTACCAGACTTCTGTAGTCTTCCTATGAAGAAAACAGAGACTGTTCTTA
ATGAGTGGTGGAATGAAGTAACCGTCGTGGGACTAGGTGCCGATTTGAAGTATACCGACAATTCCGGCCTTGA
GGAGTTTGCCAACTGCTTAAATGTCTCTGCAGCAAACGACGACCAGTGTGCCGATGACGGTATGACCAACGAT
ATGACAAACAGTGCCCTGTATATGCACAACTACTACAGAGGTCTTTTAGGAACAGGTTGGGCTGAGGATAAAA
TATCCACTTATGCTCCTTTAGCTAAAGCCATGCCACAATTAAATTACGAATGTGGCACAATCGCCAAGGAAGCA
AAAGATAGGGCCCAAGCATGCAAGGAAACTCCAGACGACCCTACTGCTCCCAGATCACAAAACTACCTTGTAA
TCAAAGATTTATCCATGGATAAGGAGGATGCCTTGAAGCAGGCAATATCTACGTGGTGGAAAGGCCTGGCCGA
CGTCGGATTGGGAGAGGACACAACTTACAAAGCCGCCATGAAGAATACACTAGGCCAATTTGCTAATATGGCC
TATGATCAAACTAAGCAAGTGGGCTGCTCAGTTGAGACTTGCACAAAACAGGGCTTTACTTTGGTCGTATGCCA
ATACGACAAGTTAATCTCAGACGGAAAGAAAATCTACGAGACGGGTAAGTCAGGCGACAAAGCATGCGCTGG
TTGCGCTACTACCAAGTGCGGAGACCTTCAGGGCCTTTGCGCACCTGGTCTAGA

Translated protein sequence

MVEPCSDDIKDTPPKSSKVTCLSHPLGRQIIEFASLIGDSHSDRAFGCRNSLISDEWRAMVLTFHNDKRKTLARGDQP
GKDGNKLSPAKQMNELTWDCNLERFASDAAAQCKDYDDPQRGVNQKTQVLIVPDFCSLPMKKTETVLNEWWNE
VTVVGLGADLKYTDNSGLEEFANCLNVSAANDDQCADDGMTNDMTNSALYMHNYYRGLLGTGWAEDKISTYAP
LAKAMPQLNYECGTIAKEAKDRAQACKETPDDPTAPRSQNYLVIKDLSMDKEDALKQAISTWWKGLADVGLGEDT
TYKAAMKNTLGQFANMAYDQTKQVGCSVETCTKQGFTLVVCQYDKLISDGKKIYETGKSGDKACAGCATTKCGDL
QGLCAP-

NECAME_02290 (NP12)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCATGCTAATCACCATCGGTGCTTACTGGCCTCCTTTCTTCGAACAGAACTGGGACTACAATGCGGAAGA
ACTGGCTCAAATAGAGGCTGCAAAATGTAACGCAGCCATCACAACACCCGCTACTTACGGATCTATTCAGAGCA
TGATTCCTGTGAAGAAATCATGTGATGCCATTGCTTCAACGCAGGCGGAAATAAAGAAATGGTGGGACGATGG
TGCAAAAGAGCAAACTGACCAGGCCAAAGTTGCTGGCAACGATAAGTTTTCTCAAATGGCAAACGCTGAAACC
AGTGCTGTTGCTTGTTCATACTCCACATGCTCTTCTAATACACAGTTGAATTTGGTCTGCTTCTACAACAAAGAC
GGAACTCAGTCACAAAATCTCTACACTGCAAATGCTGCCGGCTGTACCTGTGCGGGTTGCACCGCTTATCTTTG
CCCGTCGACATTCGATCCTGGTGAGCTTTTTGGGGATTATGGTCTAGA

Translated protein sequence

MLITIGAYWPPFFEQNWDYNAEELAQIEAAKCNAAITTPATYGSIQSMIPVKKSCDAIASTQAEIKKWWDDGAKEQ
TDQAKVAGNDKFSQMANAETSAVACSYSTCSSNTQLNLVCFYNKDGTQSQNLYTANAAGCTCAGCTAYLCPSTFD
PGELFGDY-

NECAME_16708 (NP13)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1 with point mutation

GAATTCGCAGGCTTTAAATGCCAGAACTCTCTGATTTCAGATCAGTGGCGTAAAGGCATACTAGAACTTATCAA
CGGCTTCCGTCGTACGGTGGCTAAAGGCCAACAGCTGGGCAAGGATGGAGCTAAACTGCCATCAGCTAAGAA
AATGAACCAGCTGACCTGGGATTGCAACATTGAAGCTGAAGCTCAGGGAGCAGCTAAAACATGCCCAACATCA
CTTACTTCAGCTCAGTTACCCGTATTTCAGGACACAGCATACAAGCTGGGATACATTACCGCAACTAAAGCAATT
AAGGCCAATTCCTGCGATACCACGAAAACTGCCAAGGACCTATTGACAGAACAGTGGAATGCTGCCGTCGCAA
ATCAACCTGATGAAAATGTGATCAAAGGCGAGAAAGTTTTCTCACAGTTAGCAAATGAAGAAATGGATGCCTT
TGCATGTACCTACCAAATTTGCCAAACGACATCCTATTACCTTCTATGCTTCTTCAATAAGAAGCCTGCAGCTGC
AGGAGCACCTGTATATGAGAAAGAAACGGGAACAAAGTACTGCACGGATCCCGATTGCAAGTATGCAGTCGC
TAGTTGGTGGAAATCTTTTGAAAACAAGGGATTCGGCGAAGATAGAGAGAACAAAAAATCTTTAATTAGTGCT
GCTAACATTGCTTATGACGAGACAACTAAAGTTGGATGCGCCATAGACGTTGGTGATGCCTGCTTGAAGCTAG
GTAAGCTATACATTATGTGCATGTACGACAAATCTACGCATGGTCTAGA

Translated protein sequence

MVLLEILIPVFKTLFLASLQLYNNVVTSAAGFKCQNSLISDQWRKGILELINGFRRTVAKGQQLGKDGAKLPSAKKMN
QLTWDCNIEAEAQGAAKTCPTSLTSAQLPVFQDTAYKLGYITATKAIKANSCDTTKTAKDLLTEQWNAAVANQPDE
NVIKGEKVFSQLANEEMDAFACTYQICQTTSYYLLCFFNKKPAAAGAPVYEKETGTKYCTDPDCKYAVASWWKSFE
NKGFGEDRENKKSLISAANIAYDETTKVGCAIDVGDACLKLGKLYIMCMYDKSTH-

NECAME_17155 (NP14)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCATGTGTCCAGGTAATGACCTGGACGACAATGATCGTACCACACTGAGGAAAGCTCATAACGATTTGA
GGAGGAAGATAGTCAAAGGATCAGCCAAAAACTACGCCGGAAATTTAAACGCAGGAAAGAATATGTATAGTT
TGTTGGCTTATGACAGAAATGTGGCCCTTGGATGTCAGTATAAGAAATGTACCGATAAAGTCGTCATTACTTGC
ATGTATTCAAATATAGTCCCTGATAACGCAGTCCTATATGAGAGAGGTACCCCTTGCACTAAGGATACCGACTG
CACTTTATACAGTCCCGCAGTATGCAACAACGGACTATGTAAGGCAAAGGCTGTCATTCCCAATCCTAATCCCC
AGGGCACAACTACTGCTGCTCCTGGTGCAGCAAAGTGTGCCAATGCTGAGGTTACCGATGCAGTTCGTAACGC
TGTTATCGATGTTCACAATTCATATCGTTCCCAGGTTGCCAGGGGCTTGATAATAAACGGCAAAAGTGGACAGA
ATTGTCCCACTGCTGCAAACATGTATAAGATGGAGTACGATTGTTCTTTAGAGTCATCTGCATTGGTTCAAGCCA
AGACTTGTACTCTTACTGCTTCAGCCGCTTCAAGTCGTGGAGGACAGGGAGAGAACTATTATAGTGGAGCCCT
AGTCAACGATCTTGAGCAGGCTGTGCGTACGGCAATGCAACAATGGTTTAACCAGATAACAACCAGTGGTGTA
AACCCTGTGATGCAGTTCCGTGCCAGAGTCAGAGACAAACCAAATGCCCCTGTCGCATTCACGCAAATGGTGT
GGGGAACATCCACAAAAGTGGGATGTGCTGTCATTAAGTGTCCCAGTAACACTTATACGGTATGTAGATATTCA
CCTGCTGGTAACATTGTAGATGGTTACGTCTACAAAAGGGGTAACCTATGCGGAGACTGCCCTGCTATGTGCA
ATAACGGCTTGTGCGGAACTGCTGGTCTAGA

Translated protein sequence

MCPGNDLDDNDRTTLRKAHNDLRRKIVKGSAKNYAGNLNAGKNMYSLLAYDRNVALGCQYKKCTDKVVITCMYS
NIVPDNAVLYERGTPCTKDTDCTLYSPAVCNNGLCKAKAVIPNPNPQGTTTAAPGAAKCANAEVTDAVRNAVIDVH
NSYRSQVARGLIINGKSGQNCPTAANMYKMEYDCSLESSALVQAKTCTLTASAASSRGGQGENYYSGALVNDLEQA
VRTAMQQWFNQITTSGVNPVMQFRARVRDKPNAPVAFTQMVWGTSTKVGCAVIKCPSNTYTVCRYSPAGNIVD
GYVYKRGNLCGDCPAMCNNGLCGTA-

NECAME_17854 (NP15)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCATGAGGGTGTCTCTGAACAACGAAAGGTACCGTTATGATTGCGAACTGGAGCAAGCTGCTATTGACG
CCGTCGGCGATTCTTGCTCAGCCGCCCTTGCCGAGCCAGCCAAATACGGTCAAAACGTCCAGGTTTATGAGACA
CAAGCTGTCTCTGCCGTGGCAACCGACGAATTGCTTAAAGATGCAGTAAAGCAGTGGTACCAGCCTGTTCTATA
TTACGGATTGAGGAATGACGGCAATAAGTATGAAGATAAGAGGCTATACACCTTTGCTAATATCGCCTACGATC
GTAATACTGCAATAGGCTGCCATCATAAAAAATGCGGCGACAAGGTCGTCATAACTTGTATGTATAGTAACATA
GTGCCTGATAATGCTACTCTGTACGAGAGAGGCACCCCCTGCACAAAGGATGACGACTGTACGATATATACGC
CCTCTACGTGTAAATCTAGTCTATGTACGGCTAAGGCCATGATTCCTAATCCTTCTCCTTCCTCTACCACGGGACT
AGGTCTAGA

Translated protein sequence

MRVSLNNERYRYDCELEQAAIDAVGDSCSAALAEPAKYGQNVQVYETQAVSAVATDELLKDAVKQWYQPVLYYGL
RNDGNKYEDKRLYTFANIAYDRNTAIGCHHKKCGDKVVITCMYSNIVPDNATLYERGTPCTKDDDCTIYTPSTCKSSL
CTAKAMIPNPSPSSTTGL-

NECAME_18970 (NP16)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCATGGCTAAATGGTGGAAACCCTCTGGAAACCAAGCATTTGGAGAAAACCGTAAGAACACAAACGGTC
TTATGTCAGCCGCTAATATAGCTTACGACGAGGCTACGAAAGTGGGATGCGCTATCGATGCAGGAGCTGCATG
CCTGAAGCTAGGTAAGTTATATATTTTGTGCAAGTATGACAAATCTCCAGCCATTGGAGAAGAAATCTACGAGG
CAGGTAATAAGGCATGTAGTAAATGCGGCACAACCGGCACGAATCCATCTTGTAGTCCACTTGGCGGACTGTG
TGTTAAGTCCTCAGGTCTAGA

Translated protein sequence

MAKWWKPSGNQAFGENRKNTNGLMSAANIAYDEATKVGCAIDAGAACLKLGKLYILCKYDKSPAIGEEIYEAGNK
ACSKCGTTGTNPSCSPLGGLCVKSS-

NECAME_07690 (NP17)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and Xba1

GAATTCCTTGAATCACAATTCGGCTGTAGTGGAGAAATGGGTGATGACAAAAGAGCAATCTTTCTTGAGGCTC
ACAATGAACTTCGCAGGAGAGTAGCACTAGGTAAAGAGGAAAACAAATTTGGATATCTAGGACCAGCACGGA
ATATGTACCAGCTTAATTGGGACTGTACATTGGAAGACGATGCAGCTAATTCTATTGCCGACTGTAGGCTTGAT
TCAACGGTCGATTCTGCTCGAAACAACATCGCGCTTTCCGGTGGCTATCTTCCTGATGAACTTTTAATTACGCTT
GGAATGTACCAATGGGAAAACGCGGTGAAATTCTTCGGTAAAAGTGAGCCAACCAACGTGAACGATGGAACT
CTCTGCACATTTGCCAACATGGTACATGCAAACACCACGTCGGTAGGATGTGCTTACAAATCTTGTAATTCGAA
TTTCCTTATTACATGTCTCTACGACAAATGTGGCACCAAGGAACCTCTATTGTGGGAGATAGGTACTGCTTGTAA
GGAAGATAAAGACTGCACCACTTACGAAGATTCGAAGTGTGTTGATAGTCTTTGTGTGGTCGACGGTCTAGA

Translated protein sequence

MRPNLAVLVVLFISSAYALESQFGCSGEMGDDKRAIFLEAHNELRRRVALGKEENKFGYLGPARNMYQLNWDCTLE
DDAANSIADCRLDSTVDSARNNIALSGGYLPDELLITLGMYQWENAVKFFGKSEPTNVNDGTLCTFANMVHANTTS
VGCAYKSCNSNFLITCLYDKCGTKEPLLWEIGTACKEDKDCTTYEDSKCVDSLCVVD-

NECAME_16753 (NP18)

NT sequence optimised without signal peptide or stop codon, with restriction enzyme sites EcoR1 and

Xba1

GAATTCATGACCGGCAAAAACTGTGACACAAAGGCAAACACAAAGAAAGCATTGCAAACATGGTGGGAGGAG
GTCAAAGTCGAAGATCTTTCGGCAGATCCAAAATATGACGACAAATATAAAAACTTTGCTCAAATGGTAAACGC
TGAAGTTACTGGATTTGCCTGCTCTTACAATCTGTGCAACGGAGCGGGCAAATTGGTGTGTCTCTACAACAAAA
TGCTCTCCCAACCTGCCGATTTGTTGTATAAAGTGGCTGCTGATGAACCCAATGTCTGTGCAGACTGCACCCCTC
CAACTCTTGCCACCTCTTGTGTAGCACACCTTTGTCAGCCAGAATACAAACTCGGTATGTTCTTTGAATTAGAAC
TGTTACCAGCCGCCGATACGGCTCCACCAGTACAATGTCCGTCTCAGCCCGCTGCAGCCGATGACATGACATAC
GATTTGGAACAGACGGCAATATACATGCACAATTATTACAGAAGGCTTATTGCAACTGGATGGGCTATGGATA
ACAACAACAAAAGTTATGCAAAACCGGCGGCAAAAATGCCTGCATTGAAAATCAATTGTCAAACCGTGGCGGA
CTCAAAAATAAAAGCGGACTGCACAAATCCACAAAATACTCCGCTGGCAGATCACTCCCTGAATACCTACACAA
AGAATTACCCTGCCTCAAGGCAAGAAGCCCTCGAAGAGGCGATAAATTCGTGGTATGGACAACTTCTGGATGT
CGCACTCGACGAAAAGGCTACGTATACGTCAGCAATAAAGAAAACAGCGGAAAGTTTCGCTAATTTAGCTCAA
GATGAGGCCACTGAGATTGGATGTCACGTGCAAGAATGCGCGAAACAAGGTTTTACTATCGCCATCTGCCAGT
ACAACAAGTACGTTCCAATGATTCTTAAGGGAACGGATAGATTTGAGACAATGACGCAAATCTTAGCACAAACT
AAAAATTTTGCTGATCATTACGATCAACAGCTGCTCGGCATTCGGAAAGAAAACAGAAAAGAACAAGTCGAAA
GAACAAGTTCCTACCGAAGATTCTCCTCTGTATGCAGTGGGCAACACGTGCTCGAAATGTTCGACGCTGAACAA
AAAGTGGGTCTAGA

Translated protein sequence

MTGKNCDTKANTKKALQTWWEEVKVEDLSADPKYDDKYKNFAQMVNAEVTGFACSYNLCNGAGKLVCLYNKML
SQPADLLYKVAADEPNVCADCTPPTLATSCVAHLCQPEYKLGMFFELELLPAADTAPPVQCPSQPAAADDMTYDLE
QTAIYMHNYYRRLIATGWAMDNNNKSYAKPAAKMPALKINCQTVADSKIKADCTNPQNTPLADHSLNTYTKNYPA
SRQEALEEAINSWYGQLLDVALDEKATYTSAIKKTAESFANLAQDEATEIGCHVQECAKQGFTIAICQYNKYVPMILK
GTDRFETMTQILAQTKNFADHYDQQLLGIRKENRKEQVERTSSYRRFSSVCSGQHVLEMFDAEQKV-

References

1. Medzhitov R: **Inflammation 2010: new adventures of an old flame**. *Cell* 2010, **140**:771-776.
2. El-Gabalawy H, Guenther LC, Bernstein CN: **Epidemiology of immune-mediated inflammatory diseases: incidence, prevalence, natural history, and comorbidities**. *J Rheumatol Suppl* 2010, **85**:2-10.
3. Galli SJ, Tsai M, Piliponsky AM: **The development of allergic inflammation**. *Nature* 2008, **454**:445-454.
4. Holgate ST, Wenzel S, Postma DS, Weiss ST, Renz H, Sly PD: **Asthma**. *Nat Rev Dis Primers* 2015, **1**:15025.
5. To T, Wang C, Guan J, McLimont S, Gershon AS: **What is the lifetime risk of physician-diagnosed asthma in Ontario, Canada?** *Am J Respir Crit Care Med* 2010, **181**:337-343.
6. Tai A, Tran H, Roberts M, Clarke N, Gibson AM, Vidmar S, Wilson J, Robertson CF: **Outcomes of childhood asthma to the age of 50 years**. *J Allergy Clin Immunol* 2014, **133**:1572-1578.e1573.
7. Kuenzig ME, Bishay K, Leigh R, Kaplan GG, Benchimol EI, Crowdscreen SRRT: **Co-occurrence of Asthma and the Inflammatory Bowel Diseases: A Systematic Review and Meta-analysis**. *Clinical and translational gastroenterology* 2018, **9**:188-188.
8. Sartor RB: **Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis**. *Nat Clin Pract Gastroenterol Hepatol* 2006, **3**:390-407.
9. Xavier RJ, Podolsky DK: **Unravelling the pathogenesis of inflammatory bowel disease**. *Nature* 2007, **448**:427-434.
10. Sands BE: **From symptom to diagnosis: clinical distinctions among various forms of intestinal inflammation**. *Gastroenterology* 2004, **126**:1518-1532.
11. Danese S, Fiocchi C: **Ulcerative colitis**. *N Engl J Med* 2011, **365**:1713-1725.
12. Asher I, Pearce N: **Global burden of asthma among children**. *Int J Tuberc Lung Dis* 2014, **18**:1269-1278.
13. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, et al.: **Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010**. *Lancet* 2012, **380**:2163-2196.
14. Statistics ABo: **National Health Survey: First Results, 2014–15**. Edited by ABS; 2016.
15. Moonie SA, Sterling DA, Figgs L, Castro M: **Asthma status and severity affects missed school days**. *J Sch Health* 2006, **76**:18-24.
16. Papi A, Brightling C, Pedersen SE, Reddel HK: **Asthma**. *Lancet* 2018, **391**:783-800.
17. Cabieses B, Uphoff E, Pinart M, Anto JM, Wright J: **A systematic review on the development of asthma and allergic diseases in relation to international immigration: the leading role of the environment confirmed**. *PLoS One* 2014, **9**:e105347.
18. Burisch J, Jess T, Martinato M, Lakatos PL: **The burden of inflammatory bowel disease in Europe**. *J Crohns Colitis* 2013, **7**:322-337.
19. Ng SC, Tang W, Ching JY, Wong M, Chow CM, Hui AJ, Wong TC, Leung VK, Tsang SW, Yu HH, et al.: **Incidence and phenotype of inflammatory bowel disease based on results from the Asia-pacific Crohn's and colitis epidemiology study**. *Gastroenterology* 2013, **145**:158-165.e152.
20. Betteridge JD, Armbruster SP, Maydonovitch C, Veerappan GR: **Inflammatory bowel disease prevalence by age, gender, race, and geographic location in the U.S. military health care population**. *Inflamm Bowel Dis* 2013, **19**:1421-1427.
21. Prideaux L, Kamm MA, De Cruz PP, Chan FK, Ng SC: **Inflammatory bowel disease in Asia: a systematic review**. *J Gastroenterol Hepatol* 2012, **27**:1266-1280.
22. Benchimol EI, Fortinsky KJ, Gozdyra P, Van den Heuvel M, Van Limbergen J, Griffiths AM: **Epidemiology of pediatric inflammatory bowel disease: a systematic review of international trends**. *Inflamm Bowel Dis* 2011, **17**:423-439.

23. Logan I, Bowlus CL: **The geoepidemiology of autoimmune intestinal diseases**. *Autoimmun Rev* 2010, **9**:A372-378.
24. Burisch J, Pedersen N, Cukovic-Cavka S, Brinar M, Kaimakliotis I, Duricova D, Shonova O, Vind I, Avnstrom S, Thorsgaard N, et al.: **East-West gradient in the incidence of inflammatory bowel disease in Europe: the ECCO-EpiCom inception cohort**. *Gut* 2014, **63**:588-597.
25. Ober C, Yao TC: **The genetics of asthma and allergic disease: a 21st century perspective**. *Immunol Rev* 2011, **242**:10-30.
26. Agrawal DK, Shao Z: **Pathogenesis of allergic airway inflammation**. *Curr Allergy Asthma Rep* 2010, **10**:39-48.
27. Thomsen SF, van der Sluis S, Kyvik KO, Skytthe A, Skadhauge LR, Backer V: **Increase in the heritability of asthma from 1994 to 2003 among adolescent twins**. *Respir Med* 2011, **105**:1147-1152.
28. Subbarao P, Mandhane PJ, Sears MR: **Asthma: epidemiology, etiology and risk factors**. *Cmaj* 2009, **181**:E181-190.
29. Dezateux C, Stocks J, Dundas I, Fletcher ME: **Impaired airway function and wheezing in infancy: the influence of maternal smoking and a genetic predisposition to asthma**. *Am J Respir Crit Care Med* 1999, **159**:403-410.
30. Wright RJ, Finn P, Contreras JP, Cohen S, Wright RO, Staudenmayer J, Wand M, Perkins D, Weiss ST, Gold DR: **Chronic caregiver stress and IgE expression, allergen-induced proliferation, and cytokine profiles in a birth cohort predisposed to atopy**. *J Allergy Clin Immunol* 2004, **113**:1051-1057.
31. Sears MR, Greene JM, Willan AR, Wiecek EM, Taylor DR, Flannery EM, Cowan JO, Herbison GP, Silva PA, Poulton R: **A longitudinal, population-based, cohort study of childhood asthma followed to adulthood**. *N Engl J Med* 2003, **349**:1414-1422.
32. **Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee**. *Lancet* 1998, **351**:1225-1232.
33. von Hertzen L, Laatikainen T, Pitkanen T, Vlasoff T, Makela MJ, Vartiainen E, Haahtela T: **Microbial content of drinking water in Finnish and Russian Karelia - implications for atopy prevalence**. *Allergy* 2007, **62**:288-292.
34. Waser M, Michels KB, Bieli C, Floistrup H, Pershagen G, von Mutius E, Ege M, Riedler J, Schram-Bijkerk D, Brunekreef B, et al.: **Inverse association of farm milk consumption with asthma and allergy in rural and suburban populations across Europe**. *Clin Exp Allergy* 2007, **37**:661-670.
35. Ball TM, Castro-Rodriguez JA, Griffith KA, Holberg CJ, Martinez FD, Wright AL: **Siblings, day-care attendance, and the risk of asthma and wheezing during childhood**. *N Engl J Med* 2000, **343**:538-543.
36. Riedler J, Braun-Fahrlander C, Eder W, Schreuer M, Waser M, Maisch S, Carr D, Schierl R, Nowak D, von Mutius E: **Exposure to farming in early life and development of asthma and allergy: a cross-sectional survey**. *Lancet* 2001, **358**:1129-1133.
37. Ownby DR, Johnson CC, Peterson EL: **Exposure to dogs and cats in the first year of life and risk of allergic sensitization at 6 to 7 years of age**. *Jama* 2002, **288**:963-972.
38. Holgate ST: **Pathogenesis of asthma**. *Clin Exp Allergy* 2008, **38**:872-897.
39. Murdoch JR, Lloyd CM: **Chronic inflammation and asthma**. *Mutat Res* 2010, **690**:24-39.
40. Sporik R, Holgate ST, Platts-Mills TA, Cogswell JJ: **Exposure to house-dust mite allergen (Der p I) and the development of asthma in childhood. A prospective study**. *N Engl J Med* 1990, **323**:502-507.
41. Permaul P, Hoffman E, Fu C, Sheehan W, Baxi S, Gaffin J, Lane J, Bailey A, King E, Chapman M, et al.: **Allergens in urban schools and homes of children with asthma**. *Pediatr Allergy Immunol* 2012, **23**:543-549.

42. Olivieri M, Zock JP, Accordini S, Heinrich J, Jarvis D, Kunzli N, Anto JM, Norback D, Svanes C, Verlato G: **Risk factors for new-onset cat sensitization among adults: a population-based international cohort study**. *J Allergy Clin Immunol* 2012, **129**:420-425.
43. Lambrecht BN, Hammad H: **The immunology of asthma**. *Nat Immunol* 2015, **16**:45-56.
44. Holgate ST, Polosa R: **Treatment strategies for allergy and asthma**. *Nat Rev Immunol* 2008, **8**:218-230.
45. Huber HL, Koessler KK: **The pathology of bronchial asthma**. *JAMA Internal Medicine* 1922, **30**:689-760.
46. Roche WR, Beasley R, Williams JH, Holgate ST: **Subepithelial fibrosis in the bronchi of asthmatics**. *Lancet* 1989, **1**:520-524.
47. Kanemitsu Y, Ito I, Niimi A, Izuhara K, Ohta S, Ono J, Iwata T, Matsumoto H, Mishima M: **Osteopontin and periostin are associated with a 20-year decline of pulmonary function in patients with asthma**. *Am J Respir Crit Care Med* 2014, **190**:472-474.
48. Lopez-Guisa JM, Powers C, File D, Cochrane E, Jimenez N, Debley JS: **Airway epithelial cells from asthmatic children differentially express proremodeling factors**. *J Allergy Clin Immunol* 2012, **129**:990-997.e996.
49. Chanez P, Vignola M, Stenger R, Vic P, Michel FB, Bousquet J: **Platelet-derived growth factor in asthma**. *Allergy* 1995, **50**:878-883.
50. Puddicombe SM, Polosa R, Richter A, Krishna MT, Howarth PH, Holgate ST, Davies DE: **Involvement of the epidermal growth factor receptor in epithelial repair in asthma**. *Faseb j* 2000, **14**:1362-1374.
51. Levy BD, Vachier I, Serhan CN: **Resolution of inflammation in asthma**. *Clin Chest Med* 2012, **33**:559-570.
52. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA, et al.: **Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease**. *Nature* 2012, **491**:119-124.
53. Shaw MH, Kamada N, Warner N, Kim YG, Nunez G: **The ever-expanding function of NOD2: autophagy, viral recognition, and T cell activation**. *Trends Immunol* 2011, **32**:73-79.
54. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barmada MM, Datta LW, et al.: **Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis**. *Nat Genet* 2007, **39**:596-604.
55. McCarroll SA, Huett A, Kuballa P, Chilewski SD, Landry A, Goyette P, Zody MC, Hall JL, Brant SR, Cho JH, et al.: **Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease**. *Nat Genet* 2008, **40**:1107-1112.
56. Zhernakova A, Festen EM, Franke L, Trynka G, van Diemen CC, Monsuur AJ, Bevova M, Nijmeijer RM, van 't Slot R, Heijmans R, et al.: **Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP**. *Am J Hum Genet* 2008, **82**:1202-1210.
57. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Karaliuskas R, Duerr RH, et al.: **A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease**. *Nature* 2001, **411**:603-606.
58. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, Lee JC, Goyette P, Imielinski M, Latiano A, et al.: **Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47**. *Nat Genet* 2011, **43**:246-252.
59. Brand S: **Crohn's disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease**. *Gut* 2009, **58**:1152-1167.
60. Tremelling M, Cummings F, Fisher SA, Mansfield J, Gwilliam R, Keniry A, Nimmo ER, Drummond H, Onnie CM, Prescott NJ, et al.: **IL23R variation determines susceptibility but not disease phenotype in inflammatory bowel disease**. *Gastroenterology* 2007, **132**:1657-1664.

61. Zuk O, Hechter E, Sunyaev SR, Lander ES: **The mystery of missing heritability: Genetic interactions create phantom heritability**. *Proc Natl Acad Sci U S A* 2012, **109**:1193-1198.
62. Zhang YZ, Li YY: **Inflammatory bowel disease: pathogenesis**. *World J Gastroenterol* 2014, **20**:91-99.
63. de Castella H: **Non-smoking: a feature of ulcerative colitis**. *Br Med J (Clin Res Ed)* 1982, **284**:1706.
64. Cosnes J: **What is the link between the use of tobacco and IBD?** *Inflamm Bowel Dis* 2008, **14 Suppl 2**:S14-15.
65. Lakatos PL, Szamosi T, Lakatos L: **Smoking in inflammatory bowel diseases: good, bad or ugly?** *World J Gastroenterol* 2007, **13**:6134-6139.
66. Birrenbach T, Bocker U: **Inflammatory bowel disease and smoking: a review of epidemiology, pathophysiology, and therapeutic implications**. *Inflamm Bowel Dis* 2004, **10**:848-859.
67. Shaw SY, Blanchard JF, Bernstein CN: **Association between the use of antibiotics and new diagnoses of Crohn's disease and ulcerative colitis**. *Am J Gastroenterol* 2011, **106**:2133-2142.
68. Ananthakrishnan AN, Higuchi LM, Huang ES, Khalili H, Richter JM, Fuchs CS, Chan AT: **Aspirin, nonsteroidal anti-inflammatory drug use, and risk for Crohn disease and ulcerative colitis: a cohort study**. *Ann Intern Med* 2012, **156**:350-359.
69. Shaw SY, Blanchard JF, Bernstein CN: **Association between the use of antibiotics in the first year of life and pediatric inflammatory bowel disease**. *Am J Gastroenterol* 2010, **105**:2687-2692.
70. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, et al.: **A human gut microbial gene catalogue established by metagenomic sequencing**. *Nature* 2010, **464**:59-65.
71. Lucas Lopez R, Grande Burgos MJ, Galvez A, Perez Pulido R: **The human gastrointestinal tract and oral microbiota in inflammatory bowel disease: a state of the science review**. *Apmis* 2017, **125**:3-10.
72. Joossens M, Huys G, Cnockaert M, De Preter V, Verbeke K, Rutgeerts P, Vandamme P, Vermeire S: **Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives**. *Gut* 2011, **60**:631-637.
73. Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M, et al.: **The treatment-naive microbiome in new-onset Crohn's disease**. *Cell Host Microbe* 2014, **15**:382-392.
74. Andoh A, Imaeda H, Aomatsu T, Inatomi O, Bamba S, Sasaki M, Saito Y, Tsujikawa T, Fujiyama Y: **Comparison of the fecal microbiota profiles between ulcerative colitis and Crohn's disease using terminal restriction fragment length polymorphism analysis**. *J Gastroenterol* 2011, **46**:479-486.
75. Martinez C, Antolin M, Santos J, Torrejon A, Casellas F, Borruel N, Guarner F, Malagelada JR: **Unstable composition of the fecal microbiota in ulcerative colitis during clinical remission**. *Am J Gastroenterol* 2008, **103**:643-648.
76. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Gratadoux JJ, Blugeon S, Bridonneau C, Furet JP, Corthier G, et al.: **Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients**. *Proc Natl Acad Sci U S A* 2008, **105**:16731-16736.
77. Pascal V, Pozuelo M, Borruel N, Casellas F, Campos D, Santiago A, Martinez X, Varela E, Sarrabayrouse G, Machiels K, et al.: **A microbial signature for Crohn's disease**. *Gut* 2017, **66**:813-822.
78. Hirano A, Umeno J, Okamoto Y, Shibata H, Ogura Y, Moriyama T, Torisu T, Fujioka S, Fuyuno Y, Kawarabayasi Y, et al.: **Comparison of the microbial community structure between inflamed and non-inflamed sites in patients with ulcerative colitis**. *J Gastroenterol Hepatol* 2018.
79. Antoni L, Nuding S, Wehkamp J, Stange EF: **Intestinal barrier in inflammatory bowel disease**. *World J Gastroenterol* 2014, **20**:1165-1179.
80. Bruewer M, Samarin S, Nusrat A: **Inflammatory bowel disease and the apical junctional complex**. *Ann N Y Acad Sci* 2006, **1072**:242-252.

81. Arslan G, Atasever T, Cindoruk M, Yildirim IS: **(51)CrEDTA colonic permeability and therapy response in patients with ulcerative colitis**. *Nucl Med Commun* 2001, **22**:997-1001.
82. Welcker K, Martin A, Kolle P, Siebeck M, Gross M: **Increased intestinal permeability in patients with inflammatory bowel disease**. *Eur J Med Res* 2004, **9**:456-460.
83. Moehle C, Ackermann N, Langmann T, Aslanidis C, Kel A, Kel-Margoulis O, Schmitz-Madry A, Zahn A, Stremmel W, Schmitz G: **Aberrant intestinal expression and allelic variants of mucin genes associated with inflammatory bowel disease**. *J Mol Med (Berl)* 2006, **84**:1055-1066.
84. de Mattos BR, Garcia MP, Nogueira JB, Paiatto LN, Albuquerque CG, Souza CL, Fernandes LG, Tamashiro WM, Simioni PU: **Inflammatory Bowel Disease: An Overview of Immune Mechanisms and Biological Treatments**. *Mediators Inflamm* 2015, **2015**:493012.
85. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, et al.: **Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease**. *Nature* 2001, **411**:599-603.
86. Smith AM, Rahman FZ, Hayee B, Graham SJ, Marks DJ, Sewell GW, Palmer CD, Wilde J, Foxwell BM, Gloger IS, et al.: **Disordered macrophage cytokine secretion underlies impaired acute inflammation and bacterial clearance in Crohn's disease**. *J Exp Med* 2009, **206**:1883-1897.
87. Hart AL, Al-Hassi HO, Rigby RJ, Bell SJ, Emmanuel AV, Knight SC, Kamm MA, Stagg AJ: **Characteristics of intestinal dendritic cells in inflammatory bowel diseases**. *Gastroenterology* 2005, **129**:50-65.
88. Matricon J, Barnich N, Ardid D: **Immunopathogenesis of inflammatory bowel disease**. *Self Nonself* 2010, **1**:299-309.
89. Martin B, Banz A, Bienvenu B, Cordier C, Dautigny N, Becourt C, Lucas B: **Suppression of CD4+ T lymphocyte effector functions by CD4+CD25+ cells in vivo**. *J Immunol* 2004, **172**:3391-3398.
90. Smith PD, Ochsenbauer-Jambor C, Smythies LE: **Intestinal macrophages: unique effector cells of the innate immune system**. *Immunol Rev* 2005, **206**:149-159.
91. Wahl SM, Hunt DA, Wakefield LM, McCartney-Francis N, Wahl LM, Roberts AB, Sporn MB: **Transforming growth factor type beta induces monocyte chemotaxis and growth factor production**. *Proc Natl Acad Sci U S A* 1987, **84**:5788-5792.
92. Dave M, Papadakis KA, Faubion WA, Jr.: **Immunology of inflammatory bowel disease and molecular targets for biologics**. *Gastroenterol Clin North Am* 2014, **43**:405-424.
93. Lee SH, Kwon JE, Cho ML: **Immunological pathogenesis of inflammatory bowel disease**. *Intest Res* 2018, **16**:26-42.
94. Mullin JM, Snock KV: **Effect of tumor necrosis factor on epithelial tight junctions and transepithelial permeability**. *Cancer Res* 1990, **50**:2172-2176.
95. Rampart M, De Smet W, Fiers W, Herman AG: **Inflammatory properties of recombinant tumor necrosis factor in rabbit skin in vivo**. *J Exp Med* 1989, **169**:2227-2232.
96. Amiri P, Locksley RM, Parslow TG, Sadick M, Rector E, Ritter D, McKerrow JH: **Tumour necrosis factor alpha restores granulomas and induces parasite egg-laying in schistosome-infected SCID mice**. *Nature* 1992, **356**:604-607.
97. Adegbola SO, Sahnan K, Warusavitarne J, Hart A, Tozer P: **Anti-TNF Therapy in Crohn's Disease**. *Int J Mol Sci* 2018, **19**.
98. Barnes PJ: **Current therapies for asthma. Promise and limitations**. *Chest* 1997, **111**:17s-26s.
99. Selroos O, Edsbacker S, Hultquist C: **Once-daily inhaled budesonide for the treatment of asthma: clinical evidence and pharmacokinetic explanation**. *J Asthma* 2004, **41**:771-790.
100. Barnes PJ, Adcock IM: **How do corticosteroids work in asthma?** *Ann Intern Med* 2003, **139**:359-370.
101. Rhen T, Cidlowski JA: **Antiinflammatory action of glucocorticoids--new mechanisms for old drugs**. *N Engl J Med* 2005, **353**:1711-1723.
102. Barnes PJ: **Biochemical basis of asthma therapy**. *J Biol Chem* 2011, **286**:32899-32905.
103. Wijesinghe M, Weatherall M, Perrin K, Harwood M, Beasley R: **Risk of mortality associated with formoterol: a systematic review and meta-analysis**. *Eur Respir J* 2009, **34**:803-811.

104. Yates DH, Sussman HS, Shaw MJ, Barnes PJ, Chung KF: **Regular formoterol treatment in mild asthma. Effect on bronchial responsiveness during and after treatment**. *Am J Respir Crit Care Med* 1995, **152**:1170-1174.
105. Newnham DM, McDevitt DG, Lipworth BJ: **Bronchodilator subsensitivity after chronic dosing with eformoterol in patients with asthma**. *Am J Med* 1994, **97**:29-37.
106. McIvor RA, Pizzichini E, Turner MO, Hussack P, Hargreave FE, Sears MR: **Potential masking effects of salmeterol on airway inflammation in asthma**. *Am J Respir Crit Care Med* 1998, **158**:924-930.
107. Humbert M, Beasley R, Ayres J, Slavin R, Hebert J, Bousquet J, Beeh KM, Ramos S, Canonica GW, Hedgecock S, et al.: **Benefits of omalizumab as add-on therapy in patients with severe persistent asthma who are inadequately controlled despite best available therapy (GINA 2002 step 4 treatment): INNOVATE**. *Allergy* 2005, **60**:309-316.
108. Corren J, Lemanske RF, Hanania NA, Korenblat PE, Parsey MV, Arron JR, Harris JM, Scheerens H, Wu LC, Su Z, et al.: **Lebrikizumab treatment in adults with asthma**. *N Engl J Med* 2011, **365**:1088-1098.
109. Sana A, Ben Salem C, Ahmed K, Abdelbeki A, Jihed S, Imene BS, Mohamed B: **Allergen specific immunotherapy induced multi-organ failure**. *Pan Afr Med J* 2013, **14**:155.
110. Lockey RF, Benedict LM, Turkeltaub PC, Bukantz SC: **Fatalities from immunotherapy (IT) and skin testing (ST)**. *J Allergy Clin Immunol* 1987, **79**:660-677.
111. Bernstein DI, Epstein T: **Systemic reactions to subcutaneous allergen immunotherapy**. *Immunol Allergy Clin North Am* 2011, **31**:241-249, viii-ix.
112. Frew AJ: **Allergen immunotherapy**. *J Allergy Clin Immunol* 2010, **125**:S306-313.
113. Rousseaux C, Lefebvre B, Dubuquoy L, Lefebvre P, Romano O, Auwerx J, Metzger D, Wahli W, Desvergne B, Naccari GC, et al.: **Intestinal antiinflammatory effect of 5-aminosalicylic acid is dependent on peroxisome proliferator-activated receptor-gamma**. *J Exp Med* 2005, **201**:1205-1215.
114. Allgayer H: **Review article: mechanisms of action of mesalazine in preventing colorectal carcinoma in inflammatory bowel disease**. *Aliment Pharmacol Ther* 2003, **18 Suppl 2**:10-14.
115. Lim WC, Wang Y, MacDonald JK, Hanauer S: **Aminosalicylates for induction of remission or response in Crohn's disease**. *Cochrane Database Syst Rev* 2016, **7**:Cd008870.
116. Rezaie A, Kuenzig ME, Benchimol EI, Griffiths AM, Otley AR, Steinhart AH, Kaplan GG, Seow CH: **Budesonide for induction of remission in Crohn's disease**. *Cochrane Database Syst Rev* 2015:Cd000296.
117. Steiner S, Daniel C, Fischer A, Atreya I, Hirschmann S, Waldner M, Neumann H, Neurath M, Atreya R, Weigmann B: **Cyclosporine A regulates pro-inflammatory cytokine production in ulcerative colitis**. *Arch Immunol Ther Exp (Warsz)* 2015, **63**:53-63.
118. Nielsen CH, Albertsen L, Bendtzen K, Baslund B: **Methotrexate induces poly(ADP-ribose) polymerase-dependent, caspase 3-independent apoptosis in subsets of proliferating CD4+ T cells**. *Clin Exp Immunol* 2007, **148**:288-295.
119. Glatzer T, Killig M, Meisig J, Ommert I, Luetke-Eversloh M, Babic M, Paclik D, Bluthgen N, Seidl R, Seifarth C, et al.: **RORgammat(+) innate lymphoid cells acquire a proinflammatory program upon engagement of the activating receptor NKp44**. *Immunity* 2013, **38**:1223-1235.
120. Colombel JF, Sandborn WJ, Reinisch W, Mantzaris GJ, Kornbluth A, Rachmilewitz D, Lichtiger S, D'Haens G, Diamond RH, Broussard DL, et al.: **Infliximab, azathioprine, or combination therapy for Crohn's disease**. *N Engl J Med* 2010, **362**:1383-1395.
121. Hanauer SB, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, Colombel JF, Rachmilewitz D, Wolf DC, Olson A, Bao W, et al.: **Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial**. *Lancet* 2002, **359**:1541-1549.
122. Guo Y, Lu N, Bai A: **Clinical use and mechanisms of infliximab treatment on inflammatory bowel disease: a recent update**. *Biomed Res Int* 2013, **2013**:581631.

123. Hanauer SB, Sandborn WJ, Rutgeerts P, Fedorak RN, Lukas M, MacIntosh D, Panaccione R, Wolf D, Pollack P: **Human anti-tumor necrosis factor monoclonal antibody (adalimumab) in Crohn's disease: the CLASSIC-I trial**. *Gastroenterology* 2006, **130**:323-333; quiz 591.
124. Colombel JF, Sandborn WJ, Rutgeerts P, Enns R, Hanauer SB, Panaccione R, Schreiber S, Byczkowski D, Li J, Kent JD, et al.: **Adalimumab for maintenance of clinical response and remission in patients with Crohn's disease: the CHARM trial**. *Gastroenterology* 2007, **132**:52-65.
125. Sandborn WJ, Hanauer SB, Rutgeerts P, Fedorak RN, Lukas M, MacIntosh DG, Panaccione R, Wolf D, Kent JD, Bittle B, et al.: **Adalimumab for maintenance treatment of Crohn's disease: results of the CLASSIC II trial**. *Gut* 2007, **56**:1232-1239.
126. Hueber W, Sands BE, Lewitzky S, Vandemeulebroecke M, Reinisch W, Higgins PD, Wehkamp J, Feagan BG, Yao MD, Karczewski M, et al.: **Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial**. *Gut* 2012, **61**:1693-1700.
127. Maxwell JR, Zhang Y, Brown WA, Smith CL, Byrne FR, Fiorino M, Stevens E, Bigler J, Davis JA, Rottman JB, et al.: **Differential Roles for Interleukin-23 and Interleukin-17 in Intestinal Immunoregulation**. *Immunity* 2015, **43**:739-750.
128. Reinisch W, Hommes DW, Van Assche G, Colombel JF, Gendre JP, Oldenburg B, Teml A, Geboes K, Ding H, Zhang L, et al.: **A dose escalating, placebo controlled, double blind, single dose and multidose, safety and tolerability study of fontolizumab, a humanised anti-interferon gamma antibody, in patients with moderate to severe Crohn's disease**. *Gut* 2006, **55**:1138-1144.
129. Ferrari L, Krane MK, Fichera A: **Inflammatory bowel disease surgery in the biologic era**. *World J Gastrointest Surg* 2016, **8**:363-370.
130. Strachan DP, Taylor EM, Carpenter RG: **Family structure, neonatal infection, and hay fever in adolescence**. *Archives of disease in childhood* 1996, **74**:422-426.
131. Strachan DP: **Hay fever, hygiene, and household size**. *Bmj* 1989, **299**:1259-1260.
132. Braun-Fahrlander C, Gassner M, Grize L, Neu U, Sennhauser FH, Varonier HS, Vuille JC, Wuthrich B: **Prevalence of hay fever and allergic sensitization in farmer's children and their peers living in the same rural community. SCARPOL team. Swiss Study on Childhood Allergy and Respiratory Symptoms with Respect to Air Pollution**. *Clin Exp Allergy* 1999, **29**:28-34.
133. Kilpelainen M, Terho EO, Helenius H, Koskenvuo M: **Farm environment in childhood prevents the development of allergies**. *Clin Exp Allergy* 2000, **30**:201-208.
134. Von Ehrenstein OS, Von Mutius E, Illi S, Baumann L, Bohm O, von Kries R: **Reduced risk of hay fever and asthma among children of farmers**. *Clin Exp Allergy* 2000, **30**:187-193.
135. Bach JF: **The effect of infections on susceptibility to autoimmune and allergic diseases**. *N Engl J Med* 2002, **347**:911-920.
136. Frolkis A, Dieleman LA, Barkema HW, Panaccione R, Ghosh S, Fedorak RN, Madsen K, Kaplan GG: **Environment and the inflammatory bowel diseases**. *Can J Gastroenterol* 2013, **27**:e18-24.
137. Aujnarain A, Mack DR, Benchimol EI: **The role of the environment in the development of pediatric inflammatory bowel disease**. *Curr Gastroenterol Rep* 2013, **15**:326.
138. Dogaru CM, Nyffenegger D, Pescatore AM, Spycher BD, Kuehni CE: **Breastfeeding and childhood asthma: systematic review and meta-analysis**. *Am J Epidemiol* 2014, **179**:1153-1167.
139. Beigelman A, Bacharier LB: **Early-life respiratory infections and asthma development: role in disease pathogenesis and potential targets for disease prevention**. *Curr Opin Allergy Clin Immunol* 2016, **16**:172-178.
140. Rook GA, Martinelli R, Brunet LR: **Innate immune responses to mycobacteria and the downregulation of atopic responses**. *Curr Opin Allergy Clin Immunol* 2003, **3**:337-342.
141. Rook GAW: **99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: darwinian medicine and the 'hygiene' or 'old friends' hypothesis**. *Clinical and experimental immunology* 2010, **160**:70-79.

142. Valdes AM, Walter J, Segal E, Spector TD: **Role of the gut microbiota in nutrition and health**. *Bmj* 2018, **361**:k2179.
143. Fukuda S, Ohno H: **Gut microbiome and metabolic diseases**. *Semin Immunopathol* 2014, **36**:103-114.
144. Vitetta L, Bambling M, Alford H: **The gastrointestinal tract microbiome, probiotics, and mood**. *Inflammopharmacology* 2014, **22**:333-339.
145. Taylor MD, van der Werf N, Harris A, Graham AL, Bain O, Allen JE, Maizels RM: **Early recruitment of natural CD4+ Foxp3+ Treg cells by infective larvae determines the outcome of filarial infection**. *Eur J Immunol* 2009, **39**:192-206.
146. Loukas A, Hotez PJ, Diemert D, Yazdanbakhsh M, McCarthy JS, Correa-Oliveira R, Croese J, Bethony JM: **Hookworm infection**. *Nat Rev Dis Primers* 2016, **2**:16088.
147. Allen JE, Wynn TA: **Evolution of Th2 immunity: a rapid repair response to tissue destructive pathogens**. *PLoS Pathog* 2011, **7**:e1002003.
148. Gieseck RL, 3rd, Wilson MS, Wynn TA: **Type 2 immunity in tissue repair and fibrosis**. *Nat Rev Immunol* 2018, **18**:62-76.
149. Maizels RM, McSorley HJ: **Regulation of the host immune system by helminth parasites**. *J Allergy Clin Immunol* 2016, **138**:666-675.
150. Turner JD, Jackson JA, Faulkner H, Behnke J, Else KJ, Kamgno J, Boussinesq M, Bradley JE: **Intensity of intestinal infection with multiple worm species is related to regulatory cytokine output and immune hyporesponsiveness**. *J Infect Dis* 2008, **197**:1204-1212.
151. Akdis CA, Akdis M: **Advances in allergen immunotherapy: aiming for complete tolerance to allergens**. *Sci Transl Med* 2015, **7**:280ps286.
152. Akdis M, Akdis CA: **Therapeutic manipulation of immune tolerance in allergic disease**. *Nat Rev Drug Discov* 2009, **8**:645-660.
153. Metenou S, Dembele B, Konate S, Dolo H, Coulibaly SY, Coulibaly YI, Diallo AA, Soumaoro L, Coulibaly ME, Sanogo D, et al.: **At homeostasis filarial infections have expanded adaptive T regulatory but not classical Th2 cells**. *J Immunol* 2010, **184**:5375-5382.
154. Nausch N, Midzi N, Mduluza T, Maizels RM, Mutapi F: **Regulatory and activated T cells in human Schistosoma haematobium infections**. *PLoS One* 2011, **6**:e16860.
155. Toulza F, Tsang L, Ottenhoff TH, Brown M, Dockrell HM: **Mycobacterium tuberculosis-specific CD4+ T-cell response is increased, and Treg cells decreased, in anthelmintic-treated patients with latent TB**. *Eur J Immunol* 2016, **46**:752-761.
156. Wilson MS, Taylor MD, O'Gorman MT, Balic A, Barr TA, Filbey K, Anderton SM, Maizels RM: **Helminth-induced CD19+CD23hi B cells modulate experimental allergic and autoimmune inflammation**. *Eur J Immunol* 2010, **40**:1682-1696.
157. Finney CA, Taylor MD, Wilson MS, Maizels RM: **Expansion and activation of CD4(+)CD25(+) regulatory T cells in Heligmosomoides polygyrus infection**. *Eur J Immunol* 2007, **37**:1874-1886.
158. Rausch S, Huehn J, Kirchhoff D, Rzepecka J, Schnoeller C, Pillai S, Loddenkemper C, Scheffold A, Hamann A, Lucius R, et al.: **Functional analysis of effector and regulatory T cells in a parasitic nematode infection**. *Infect Immun* 2008, **76**:1908-1919.
159. Hotez PJ, Brindley PJ, Bethony JM, King CH, Pearce EJ, Jacobson J: **Helminth infections: the great neglected tropical diseases**. *J Clin Invest* 2008, **118**:1311-1321.
160. Hotez PJ, Brooker S, Bethony JM, Bottazzi ME, Loukas A, Xiao S: **Hookworm infection**. *N Engl J Med* 2004, **351**:799-807.
161. Brooker S, Bethony J, Hotez PJ: **Human hookworm infection in the 21st century**. *Adv Parasitol* 2004, **58**:197-288.
162. Christian P, Khatry SK, West KP, Jr.: **Antenatal anthelmintic treatment, birthweight, and infant survival in rural Nepal**. *Lancet* 2004, **364**:981-983.
163. Brooker S, Hotez PJ, Bundy DAP: **Hookworm-related anaemia among pregnant women: a systematic review**. *PLoS neglected tropical diseases* 2008, **2**:e291-e291.

164. Pullan RL, Smith JL, Jasrasaria R, Brooker SJ: **Global numbers of infection and disease burden of soil transmitted helminth infections in 2010**. *Parasit Vectors* 2014, **7**:37.
165. Hotez P: **Hookworm and poverty**. *Ann N Y Acad Sci* 2008, **1136**:38-44.
166. Hotez PJ, Bethony J, Bottazzi ME, Brooker S, Buss P: **Hookworm: "the great infection of mankind"**. *PLoS Med* 2005, **2**:e67.
167. Easton AV, Oliveira RG, O'Connell EM, Kepha S, Mwandawiro CS, Njenga SM, Kihara JH, Mwatele C, Odiere MR, Brooker SJ, et al.: **Multi-parallel qPCR provides increased sensitivity and diagnostic breadth for gastrointestinal parasites of humans: field-based inferences on the impact of mass deworming**. *Parasit Vectors* 2016, **9**:38.
168. Cimino RO, Jeun R, Juarez M, Cajal PS, Vargas P, Echazu A, Bryan PE, Nasser J, Krolewiecki A, Mejia R: **Identification of human intestinal parasites affecting an asymptomatic peri-urban Argentinian population using multi-parallel quantitative real-time polymerase chain reaction**. *Parasit Vectors* 2015, **8**:380.
169. Tang YT, Gao X, Rosa BA, Abubucker S, Hallsworth-Pepin K, Martin J, Tyagi R, Heizer E, Zhang X, Bhonagiri-Palsikar V, et al.: **Genome of the human hookworm Necator americanus**. *Nat Genet* 2014, **46**:261-269.
170. Horn D: **Antigenic variation in African trypanosomes**. *Mol Biochem Parasitol* 2014, **195**:123-129.
171. Scherf A, Lopez-Rubio JJ, Riviere L: **Antigenic variation in Plasmodium falciparum**. *Annu Rev Microbiol* 2008, **62**:445-470.
172. Braschi S, Wilson RA: **Proteins exposed at the adult schistosome surface revealed by biotinylation**. *Mol Cell Proteomics* 2006, **5**:347-356.
173. Mulvenna J, Moertel L, Jones MK, Nawaratna S, Lovas EM, Gobert GN, Colgrave M, Jones A, Loukas A, McManus DP: **Exposed proteins of the Schistosoma japonicum tegument**. *Int J Parasitol* 2010, **40**:543-554.
174. Sotillo J, Pearson M, Becker L, Mulvenna J, Loukas A: **A quantitative proteomic analysis of the tegumental proteins from Schistosoma mansoni schistosomula reveals novel potential therapeutic targets**. *Int J Parasitol* 2015, **45**:505-516.
175. Figueiredo CA, Barreto ML, Rodrigues LC, Cooper PJ, Silva NB, Amorim LD, Alcantara-Neves NM: **Chronic intestinal helminth infections are associated with immune hyporesponsiveness and induction of a regulatory network**. *Infect Immun* 2010, **78**:3160-3167.
176. Gaze S, McSorley HJ, Daveson J, Jones D, Bethony JM, Oliveira LM, Speare R, McCarthy JS, Engwerda CR, Croese J, et al.: **Characterising the mucosal and systemic immune responses to experimental human hookworm infection**. *PLoS Pathog* 2012, **8**:e1002520.
177. Geiger SM, Fujiwara RT, Santiago H, Correa-Oliveira R, Bethony JM: **Early stage-specific immune responses in primary experimental human hookworm infection**. *Microbes Infect* 2008, **10**:1524-1535.
178. Scrivener S, Yemaneberhan H, Zebenigus M, Tilahun D, Girma S, Ali S, McElroy P, Custovic A, Woodcock A, Pritchard D, et al.: **Independent effects of intestinal parasite infection and domestic allergen exposure on risk of wheeze in Ethiopia: a nested case-control study**. *Lancet* 2001, **358**:1493-1499.
179. Obata-Ninomiya K, Ishiwata K, Tsutsui H, Nei Y, Yoshikawa S, Kawano Y, Minegishi Y, Ohta N, Watanabe N, Kanuka H, et al.: **The skin is an important bulwark of acquired immunity against intestinal helminths**. *J Exp Med* 2013, **210**:2583-2595.
180. Gerbe F, Sidot E, Smyth DJ, Ohmoto M, Matsumoto I, Dardalhon V, Cesses P, Garnier L, Pouzolles M, Brulin B, et al.: **Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites**. *Nature* 2016, **529**:226-230.
181. Grove DI, Burston TO, Forbes IJ: **Fall in IgE levels after treatment for hookworm**. *Clin Exp Immunol* 1974, **18**:565-569.
182. Quinnell RJ, Pritchard DI, Raiko A, Brown AP, Shaw MA: **Immune responses in human necatoriasis: association between interleukin-5 responses and resistance to reinfection**. *J Infect Dis* 2004, **190**:430-438.

183. Broadhurst MJ, Leung JM, Kashyap V, McCune JM, Mahadevan U, McKerrow JH, Loke P: **IL-22+ CD4+ T cells are associated with therapeutic trichuris trichiura infection in an ulcerative colitis patient**. *Sci Transl Med* 2010, **2**:60ra88.
184. Summers RW, Elliott DE, Urban JF, Jr., Thompson RA, Weinstock JV: **Trichuris suis therapy for active ulcerative colitis: a randomized controlled trial**. *Gastroenterology* 2005, **128**:825-832.
185. Scholmerich J, Fellermann K, Seibold FW, Rogler G, Langhorst J, Howaldt S, Novacek G, Petersen AM, Bachmann O, Matthes H, et al.: **A Randomised, Double-blind, Placebo-controlled Trial of Trichuris suis ova in Active Crohn's Disease**. *J Crohns Colitis* 2017, **11**:390-399.
186. Huang X, Zeng LR, Chen FS, Zhu JP, Zhu MH: **Trichuris suis ova therapy in inflammatory bowel disease: A meta-analysis**. *Medicine (Baltimore)* 2018, **97**:e12087.
187. Feary JR, Venn AJ, Mortimer K, Brown AP, Hooi D, Falcone FH, Pritchard DI, Britton JR: **Experimental hookworm infection: a randomized placebo-controlled trial in asthma**. *Clin Exp Allergy* 2010, **40**:299-306.
188. Croft AM, Bager P, Kumar S: **Helminth therapy (worms) for allergic rhinitis**. *Cochrane Database Syst Rev* 2012:Cd009238.
189. Croese J, Giacomin P, Navarro S, Clouston A, McCann L, Dougall A, Ferreira I, Susianto A, O'Rourke P, Howlett M, et al.: **Experimental hookworm infection and gluten microchallenge promote tolerance in celiac disease**. *J Allergy Clin Immunol* 2015, **135**:508-516.
190. Correale J, Farez MF: **The impact of parasite infections on the course of multiple sclerosis**. *J Neuroimmunol* 2011, **233**:6-11.
191. Cancado GG, Fiuza JA, de Paiva NC, Lemos Lde C, Ricci ND, Gazzinelli-Guimaraes PH, Martins VG, Bartholomeu DC, Negrao-Correa DA, Carneiro CM, et al.: **Hookworm products ameliorate dextran sodium sulfate-induced colitis in BALB/c mice**. *Inflamm Bowel Dis* 2011, **17**:2275-2286.
192. Ferreira I, Smyth D, Gaze S, Aziz A, Giacomin P, Ruyssers N, Artis D, Laha T, Navarro S, Loukas A, et al.: **Hookworm excretory/secretory products induce interleukin-4 (IL-4)+ IL-10+ CD4+ T cell responses and suppress pathology in a mouse model of colitis**. *Infect Immun* 2013, **81**:2104-2111.
193. Ruyssers NE, De Winter BY, De Man JG, Loukas A, Pearson MS, Weinstock JV, Van den Bossche RM, Martinet W, Pelckmans PA, Moreels TG: **Therapeutic potential of helminth soluble proteins in TNBS-induced colitis in mice**. *Inflamm Bowel Dis* 2009, **15**:491-500.
194. Sotillo J, Sanchez-Flores A, Cantacessi C, Harcus Y, Pickering D, Bouchery T, Camberis M, Tang SC, Giacomin P, Mulvenna J, et al.: **Secreted proteomes of different developmental stages of the gastrointestinal nematode Nippostrongylus brasiliensis**. *Mol Cell Proteomics* 2014, **13**:2736-2751.
195. Morante T, Shepherd C, Constantinoiu C, Loukas A, Sotillo J: **Revisiting the Ancylostoma Caninum Secretome Provides New Information on Hookworm-Host Interactions**. *Proteomics* 2017, **17**.
196. Chalmers IW, McArdle AJ, Coulson RM, Wagner MA, Schmid R, Hirai H, Hoffmann KF: **Developmentally regulated expression, alternative splicing and distinct sub-groupings in members of the Schistosoma mansoni venom allergen-like (SmVAL) gene family**. *BMC Genomics* 2008, **9**:89.
197. Cantacessi C, Campbell BE, Visser A, Geldhof P, Nolan MJ, Nisbet AJ, Matthews JB, Loukas A, Hofmann A, Otranto D, et al.: **A portrait of the "SCP/TAPS" proteins of eukaryotes--developing a framework for fundamental research and biotechnological outcomes**. *Biotechnol Adv* 2009, **27**:376-388.
198. Hawdon JM, Jones BF, Hoffman DR, Hotez PJ: **Cloning and characterization of Ancylostoma-secreted protein. A novel protein associated with the transition to parasitism by infective hookworm larvae**. *J Biol Chem* 1996, **271**:6672-6678.
199. Hawdon JM, Narasimhan S, Hotez PJ: **Ancylostoma secreted protein 2: cloning and characterization of a second member of a family of nematode secreted proteins from Ancylostoma caninum**. *Mol Biochem Parasitol* 1999, **99**:149-165.

200. Bethony J, Loukas A, Smout M, Brooker S, Mendez S, Plieskatt J, Goud G, Bottazzi ME, Zhan B, Wang Y, et al.: **Antibodies against a secreted protein from hookworm larvae reduce the intensity of hookworm infection in humans and vaccinated laboratory animals**. *Faseb j* 2005, **19**:1743-1745.
201. Bethony JM, Simon G, Diemert DJ, Parenti D, Desrosiers A, Schuck S, Fujiwara R, Santiago H, Hotez PJ: **Randomized, placebo-controlled, double-blind trial of the Na-ASP-2 hookworm vaccine in unexposed adults**. *Vaccine* 2008, **26**:2408-2417.
202. Schnyder-Candrian S, Maillet I, Le Bert M, Brault L, Jacobs M, Ryffel B, Schnyder B, Moser R: **Neutrophil Inhibitory Factor Selectively Inhibits the Endothelium-Driven Transmigration of Eosinophils In Vitro and Airway Eosinophilia in OVA-Induced Allergic Lung Inflammation**. *J Allergy (Cairo)* 2012, **2012**:245909.
203. Navarro S, Pickering DA, Ferreira IB, Jones L, Ryan S, Troy S, Leech A, Hotez PJ, Zhan B, Laha T, et al.: **Hookworm recombinant protein promotes regulatory T cell responses that suppress experimental asthma**. *Sci Transl Med* 2016, **8**:362ra143.
204. Johnston CJC, Smyth DJ, Kodali RB, White MPJ, Harcus Y, Filbey KJ, Hewitson JP, Hinck CS, Ivens A, Kemter AM, et al.: **A structurally distinct TGF-beta mimic from an intestinal helminth parasite potently induces regulatory T cells**. *Nat Commun* 2017, **8**:1741.
205. Smyth DJ, Harcus Y, White MPJ, Gregory WF, Nahler J, Stephens I, Toke-Bjolgerud E, Hewitson JP, Ivens A, McSorley HJ, et al.: **TGF-beta mimic proteins form an extended gene family in the murine parasite Heligmosomoides polygyrus**. *Int J Parasitol* 2018, **48**:379-385.
206. Osbourn M, Soares DC, Vacca F, Cohen ES, Scott IC, Gregory WF, Smyth DJ, Toivakka M, Kemter AM, le Bihan T, et al.: **HpARI Protein Secreted by a Helminth Parasite Suppresses Interleukin-33**. *Immunity* 2017, **47**:739-751.e735.
207. Bottazzi ME: **The human hookworm vaccine: recent updates and prospects for success**. *J Helminthol* 2015, **89**:540-544.
208. Schneider B, Jariwala AR, Periago MV, Gazzinelli MF, Bose SN, Hotez PJ, Diemert DJ, Bethony JM: **A history of hookworm vaccine development**. *Hum Vaccin* 2011, **7**:1234-1244.
209. Diemert DJ, Pinto AG, Freire J, Jariwala A, Santiago H, Hamilton RG, Periago MV, Loukas A, Tribolet L, Mulvenna J, et al.: **Generalized urticaria induced by the Na-ASP-2 hookworm vaccine: implications for the development of vaccines against helminths**. *J Allergy Clin Immunol* 2012, **130**:169-176.e166.
210. Knox DP, Redmond DL, Newlands GF, Skuce PJ, Pettit D, Smith WD: **The nature and prospects for gut membrane proteins as vaccine candidates for Haemonchus contortus and other ruminant trichostrongyloids**. *Int J Parasitol* 2003, **33**:1129-1137.
211. Hotez PJ, Diemert D, Bacon KM, Beaumier C, Bethony JM, Bottazzi ME, Brooker S, Couto AR, Freire Mda S, Homma A, et al.: **The Human Hookworm Vaccine**. *Vaccine* 2013, **31 Suppl 2**:B227-232.
212. Diemert DJ, Freire J, Valente V, Fraga CG, Talles F, Grahek S, Campbell D, Jariwala A, Periago MV, Enk M, et al.: **Safety and immunogenicity of the Na-GST-1 hookworm vaccine in Brazilian and American adults**. *PLoS Negl Trop Dis* 2017, **11**:e0005574.
213. Mulvenna J, Hamilton B, Nagaraj SH, Smyth D, Loukas A, Gorman JJ: **Proteomics analysis of the excretory/secretory component of the blood-feeding stage of the hookworm, Ancylostoma caninum**. *Mol Cell Proteomics* 2009, **8**:109-121.
214. Hotez PJ, Bethony JM, Diemert DJ, Pearson M, Loukas A: **Developing vaccines to combat hookworm infection and intestinal schistosomiasis**. *Nat Rev Microbiol* 2010, **8**:814-826.
215. Hotez PJ, Beaumier CM, Gillespie PM, Strych U, Hayward T, Bottazzi ME: **Advancing a vaccine to prevent hookworm disease and anemia**. *Vaccine* 2016, **34**:3001-3005.
216. Hunt VL, Tsai IJ, Coghlan A, Reid AJ, Holroyd N, Foth BJ, Tracey A, Cotton JA, Stanley EJ, Beasley H, et al.: **The genomic basis of parasitism in the Strongyloides clade of nematodes**. *Nat Genet* 2016, **48**:299-307.

217. Culley FJ, Brown A, Conroy DM, Sabroe I, Pritchard DI, Williams TJ: **Eotaxin is specifically cleaved by hookworm metalloproteases preventing its action in vitro and in vivo**. *J Immunol* 2000, **165**:6447-6453.
218. Kumar S, Pritchard DI: **Secretion of metalloproteases by living infective larvae of Necator americanus**. *J Parasitol* 1992, **78**:917-919.
219. Ranjit N, Zhan B, Hamilton B, Stenzel D, Lowther J, Pearson M, Gorman J, Hotez P, Loukas A: **Proteolytic degradation of hemoglobin in the intestine of the human hookworm Necator americanus**. *J Infect Dis* 2009, **199**:904-912.
220. Cantacessi C, Mitreva M, Jex AR, Young ND, Campbell BE, Hall RS, Doyle MA, Ralph SA, Rabelo EM, Ranganathan S, et al.: **Massively parallel sequencing and analysis of the Necator americanus transcriptome**. *PLoS Negl Trop Dis* 2010, **4**:e684.
221. Wang Z, Abubucker S, Martin J, Wilson RK, Hawdon J, Mitreva M: **Characterizing Ancylostoma caninum transcriptome and exploring nematode parasitic adaptation**. *BMC Genomics* 2010, **11**:307.
222. Schwarz EM, Hu Y, Antoshechkin I, Miller MM, Sternberg PW, Aroian RV: **The genome and transcriptome of the zoonotic hookworm Ancylostoma ceylanicum identify infection-specific gene families**. *Nat Genet* 2015, **47**:416-422.
223. Palevich N, Britton C, Kamenetzky L, Mitreva M, de Moraes Mourao M, Bennuru S, Quack T, Scholte LLS, Tyagi R, Slatko BE: **Tackling Hypotheticals in Helminth Genomes**. *Trends Parasitol* 2018, **34**:179-183.
224. Borchert N, Dieterich C, Krug K, Schütz W, Jung S, Nordheim A, Sommer RJ, Macek B: **Proteogenomics of Pristionchus pacificus reveals distinct proteome structure of nematode models**. *Genome Res* 2010, **20**:837-846.
225. Nesvizhskii AI: **Proteogenomics: concepts, applications and computational strategies**. *Nat Methods* 2014, **11**:1114-1125.
226. Jafari M, Primo V, Smejkal GB, Moskovets EV, Kuo WP, Ivanov AR: **Comparison of in-gel protein separation techniques commonly used for fractionation in mass spectrometry-based proteomic profiling**. *Electrophoresis* 2012, **33**:2516-2526.
227. Holt C, Yandell M: **MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects**. *BMC Bioinformatics* 2011, **12**:491.
228. Bao W, Kojima KK, Kohany O: **Repbase Update, a database of repetitive elements in eukaryotic genomes**. *Mob DNA* 2015, **6**:11.
229. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data**. *Bioinformatics* 2014, **30**:2114-2120.
230. Kim D, Langmead B, Salzberg SL: **HISAT: a fast spliced aligner with low memory requirements**. *Nat Methods* 2015, **12**:357-360.
231. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL: **StringTie enables improved reconstruction of a transcriptome from RNA-seq reads**. *Nat Biotechnol* 2015, **33**:290-295.
232. Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M: **BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS**. *Bioinformatics* 2016, **32**:767-769.
233. **UniProt: the universal protein knowledgebase**. *Nucleic Acids Res* 2017, **45**:D158-d169.
234. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M: **WormBase ParaSite - a comprehensive resource for helminth genomics**. *Mol Biochem Parasitol* 2017, **215**:2-10.
235. Campbell MS, Holt C, Moore B, Yandell M: **Genome Annotation and Curation Using MAKER and MAKER-P**. *Curr Protoc Bioinformatics* 2014, **48**:4.11.11-39.
236. Lomsadze A, Burns PD, Borodovsky M: **Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm**. *Nucleic Acids Res* 2014, **42**:e119.
237. Stanke M, Diekhans M, Baertsch R, Haussler D: **Using native and syntenically mapped cDNA alignments to improve de novo gene finding**. *Bioinformatics* 2008, **24**:637-644.

238. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al.: **InterProScan 5: genome-scale protein function classification**. *Bioinformatics* 2014, **30**:1236-1240.
239. Campbell MS, Law M, Holt C, Stein JC, Moghe GD, Hufnagel DE, Lei J, Achawanantakun R, Jiao D, Lawrence CJ, et al.: **MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations**. *Plant Physiol* 2014, **164**:513-524.
240. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM: **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs**. *Bioinformatics* 2015, **31**:3210-3212.
241. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research**. *Bioinformatics* 2005, **21**:3674-3676.
242. Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S: **Feature-based prediction of non-classical and leaderless protein secretion**. *Protein Eng Des Sel* 2004, **17**:349-356.
243. Emanuelsson O, Brunak S, von Heijne G, Nielsen H: **Locating proteins in the cell using TargetP, SignalP and related tools**. *Nat Protoc* 2007, **2**:953-971.
244. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes**. *J Mol Biol* 2001, **305**:567-580.
245. Supek F, Bosnjak M, Skunca N, Smuc T: **REVIGO summarizes and visualizes long lists of gene ontology terms**. *PLoS One* 2011, **6**:e21800.
246. Conway JR, Lex A, Gehlenborg N: **UpSetR: an R package for the visualization of intersecting sets and their properties**. *Bioinformatics* 2017, **33**:2938-2940.
247. Parkinson J, Blaxter M: **SimiTri--visualizing similarity relationships for groups of sequences**. *Bioinformatics* 2003, **19**:390-395.
248. Hewitson JP, Harcus Y, Murray J, van Agtmaal M, Filbey KJ, Grainger JR, Bridgett S, Blaxter ML, Ashton PD, Ashford DA, et al.: **Proteomic analysis of secretory products from the model gastrointestinal nematode Heligmosomoides polygyrus reveals dominance of venom allergen-like (VAL) proteins**. *J Proteomics* 2011, **74**:1573-1594.
249. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA: **Circos: an information aesthetic for comparative genomics**. *Genome Res* 2009, **19**:1639-1645.
250. Letunic I, Bork P: **Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation**. *Bioinformatics* 2007, **23**:127-128.
251. Jaffe JD, Berg HC, Church GM: **Proteogenomic mapping as a complementary method to perform genome annotation**. *Proteomics* 2004, **4**:59-77.
252. Sotillo J, Toledo R, Mulvenna J, Loukas A: **Exploiting Helminth-Host Interactomes through Big Data**. *Trends Parasitol* 2017, **33**:875-888.
253. Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA, De Silva N, Velarde GS, Anderson TJ, Clark RC, et al.: **A systematically improved high quality genome and transcriptome of the human blood fluke Schistosoma mansoni**. *PLoS Negl Trop Dis* 2012, **6**:e1455.
254. Zarowiecki M, Berriman M: **What helminth genomes have taught us about parasite evolution**. *Parasitology* 2015, **142 Suppl 1**:S85-97.
255. Eichenberger RM, Ryan S, Jones L, Buitrago G, Polster R, Montes de Oca M, Zuvelek J, Giacomin PR, Dent LA, Engwerda CR, et al.: **Hookworm Secreted Extracellular Vesicles Interact With Host Cells and Prevent Inducible Colitis in Mice**. *Front Immunol* 2018, **9**:850.
256. Coakley G, Buck AH, Maizels RM: **Host parasite communications-Messages from helminths for the immune system: Parasite communication and cell-cell interactions**. *Mol Biochem Parasitol* 2016, **208**:33-40.
257. Marcilla A, Martin-Jaular L, Trelis M, de Menezes-Neto A, Osuna A, Bernal D, Fernandez-Becerra C, Almeida IC, Del Portillo HA: **Extracellular vesicles in parasitic diseases**. *J Extracell Vesicles* 2014, **3**:25040.

258. Eichenberger RM, Talukder MH, Field MA, Wangchuk P, Giacomin P, Loukas A, Sotillo J: **Characterization of Trichuris muris secreted proteins and extracellular vesicles provides new insights into host-parasite communication**. *J Extracell Vesicles* 2018, **7**:1428004.
259. Hewitson JP, Grainger JR, Maizels RM: **Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity**. *Mol Biochem Parasitol* 2009, **167**:1-11.
260. Jian X, Sen L, Hui-Qin Q, Hai-Nan R, Tie-Hua L, Hai-Chou X, Hotez PJ, Shu-Hua X: **Necator americanus: maintenance through one hundred generations in golden hamsters (Mesocricetus auratus). I. Host sex-associated differences in hookworm burden and fecundity**. *Exp Parasitol* 2003, **104**:62-66.
261. Xue J, Hui-Qing Q, Jun-Ming Y, Fujiwara R, Zhan B, Hotez P, Shu-Hua X: **Necator americanus: optimization of the golden hamster model for testing anthelmintic drugs**. *Exp Parasitol* 2005, **111**:219-223.
262. Jian X, Shu-Hua X, Hui-Qing Q, Sen L, Hotez P, Bing-Gui S, Hai-Chou X, Tie-Hua L, Bin Z: **Necator americanus: maintenance through one hundred generations in golden hamsters (Mesocricetus auratus). II. Morphological development of the adult and its comparison with humans**. *Exp Parasitol* 2003, **105**:192-200.
263. Knox D: **Proteases in blood-feeding nematodes and their potential as vaccine candidates**. *Adv Exp Med Biol* 2011, **712**:155-176.
264. Abdulla MH, Lim KC, Sajid M, McKerrow JH, Caffrey CR: **Schistosomiasis mansoni: novel chemotherapy using a cysteine protease inhibitor**. *PLoS Med* 2007, **4**:e14.
265. Hewitson JP, Ivens AC, Harcus Y, Filbey KJ, McSorley HJ, Murray J, Bridgett S, Ashford D, Dowle AA, Maizels RM: **Secretion of protective antigens by tissue-stage nematode larvae revealed by proteomic analysis and vaccination-induced sterile immunity**. *PLoS Pathog* 2013, **9**:e1003492.
266. Datu BJ, Gasser RB, Nagaraj SH, Ong EK, O'Donoghue P, McInnes R, Ranganathan S, Loukas A: **Transcriptional changes in the hookworm, Ancylostoma caninum, during the transition from a free-living to a parasitic larva**. *PLoS Negl Trop Dis* 2008, **2**:e130.
267. Lo SK, Rahman A, Xu N, Zhou MY, Nagpala P, Jaffe HA, Malik AB: **Neutrophil inhibitory factor abrogates neutrophil adhesion by blockade of CD11a and CD11b beta(2) integrins**. *Mol Pharmacol* 1999, **56**:926-932.
268. Shepherd C, Wangchuk P, Loukas A: **Of dogs and hookworms: man's best friend and his parasites as a model for translational biomedical research**. *Parasit Vectors* 2018, **11**:59.
269. Ranjit N, Jones MK, Stenzel DJ, Gasser RB, Loukas A: **A survey of the intestinal transcriptomes of the hookworms, Necator americanus and Ancylostoma caninum, using tissues isolated by laser microdissection microscopy**. *Int J Parasitol* 2006, **36**:701-710.
270. Williamson AL, Brindley PJ, Abbenante G, Prociv P, Berry C, Girdwood K, Pritchard DI, Fairlie DP, Hotez PJ, Dalton JP, et al.: **Cleavage of hemoglobin by hookworm cathepsin D aspartic proteases and its potential contribution to host specificity**. *Faseb j* 2002, **16**:1458-1460.
271. Lin X, Koelsch G, Wu S, Downs D, Dashti A, Tang J: **Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein**. *Proc Natl Acad Sci U S A* 2000, **97**:1456-1460.
272. Loukas A, Bethony JM, Mendez S, Fujiwara RT, Goud GN, Ranjit N, Zhan B, Jones K, Bottazzi ME, Hotez PJ: **Vaccination with recombinant aspartic hemoglobinase reduces parasite load and blood loss after hookworm infection in dogs**. *PLoS Med* 2005, **2**:e295.
273. Ono Y, Sorimachi H: **Calpains: an elaborate proteolytic system**. *Biochim Biophys Acta* 2012, **1824**:224-236.
274. Baig S, Damian RT, Peterson DS: **A novel cathepsin B active site motif is shared by helminth bloodfeeders**. *Exp Parasitol* 2002, **101**:83-89.
275. Ranjit N, Zhan B, Stenzel DJ, Mulvenna J, Fujiwara R, Hotez PJ, Loukas A: **A family of cathepsin B cysteine proteases expressed in the gut of the human hookworm, Necator americanus**. *Mol Biochem Parasitol* 2008, **160**:90-99.

276. Williamson AL, Lustigman S, Oksov Y, Deumic V, Plieskatt J, Mendez S, Zhan B, Bottazzi ME, Hotez PJ, Loukas A: **Ancylostoma caninum MTP-1, an astacin-like metalloprotease secreted by infective hookworm larvae, is involved in tissue migration**. *Infect Immun* 2006, **74**:961-967.
277. Hasnain SZ, McGuckin MA, Grencis RK, Thornton DJ: **Serine protease(s) secreted by the nematode Trichuris muris degrade the mucus barrier**. *PLoS Negl Trop Dis* 2012, **6**:e1856.
278. Mendez S, Zhan B, Goud G, Ghosh K, Dobardzic A, Wu W, Liu S, Deumic V, Dobardzic R, Liu Y, et al.: **Effect of combining the larval antigens Ancylostoma secreted protein 2 (ASP-2) and metalloprotease 1 (MTP-1) in protecting hamsters against hookworm infection and disease caused by Ancylostoma ceylanicum**. *Vaccine* 2005, **23**:3123-3130.
279. Zahrl RJ, Pena DA, Mattanovich D, Gasser B: **Systems biotechnology for protein production in Pichia pastoris**. *FEMS Yeast Res* 2017, **17**.
280. Cregg JM, Barringer KJ, Hessler AY, Madden KR: **Pichia pastoris as a host system for transformations**. *Mol Cell Biol* 1985, **5**:3376-3385.
281. Puxbaum V, Mattanovich D, Gasser B: **Quo vadis? The challenges of recombinant protein folding and secretion in Pichia pastoris**. *Appl Microbiol Biotechnol* 2015, **99**:2925-2938.
282. Weinacker D, Rabert C, Zepeda AB, Figueroa CA, Pessoa A, Farías JG: **Applications of recombinant Pichia pastoris in the healthcare industry**. *Braz J Microbiol* 2013, **44**:1043-1048.
283. Zhan B, Perally S, Brophy PM, Xue J, Goud G, Liu S, Deumic V, de Oliveira LM, Bethony J, Bottazzi ME, et al.: **Molecular cloning, biochemical characterization, and partial protective immunity of the heme-binding glutathione S-transferases from the human hookworm Necator americanus**. *Infect Immun* 2010, **78**:1552-1563.
284. Curti E, Seid CA, Hudspeth E, Center L, Rezende W, Pollet J, Kwityn C, Hammond M, Matsunami RK, Engler DA, et al.: **Optimization and revision of the production process of the Necator americanus glutathione S-transferase 1 (Na-GST-1), the lead hookworm vaccine recombinant protein candidate**. *Hum Vaccin Immunother* 2014, **10**:1914-1925.
285. Osman A, Wang CK, Winter A, Loukas A, Tribolet L, Gasser RB, Hofmann A: **Hookworm SCP/TAPS protein structure--A key to understanding host-parasite interactions and developing new interventions**. *Biotechnol Adv* 2012, **30**:652-657.
286. Goud GN, Bottazzi ME, Zhan B, Mendez S, Deumic V, Plieskatt J, Liu S, Wang Y, Bueno L, Fujiwara R, et al.: **Expression of the Necator americanus hookworm larval antigen Na-ASP-2 in Pichia pastoris and purification of the recombinant protein for use in human clinical trials**. *Vaccine* 2005, **23**:4754-4764.
287. Asojo OA, Goud G, Dhar K, Loukas A, Zhan B, Deumic V, Liu S, Borgstahl GE, Hotez PJ: **X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, Necator americanus, and a vaccine antigen for human hookworm infection**. *J Mol Biol* 2005, **346**:801-814.
288. Bower MA, Constant SL, Mendez S: **Necator americanus: the Na-ASP-2 protein secreted by the infective larvae induces neutrophil recruitment in vivo and in vitro**. *Exp Parasitol* 2008, **118**:569-575.
289. Asojo OA, Loukas A, Inan M, Barent R, Huang J, Plantz B, Swanson A, Gouthro M, Meagher MM, Hotez PJ: **Crystallization and preliminary X-ray analysis of Na-ASP-1, a multi-domain pathogenesis-related-1 protein from the human hookworm parasite Necator americanus**. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2005, **61**:391-394.
290. Ahmad M, Hirz M, Pichler H, Schwab H: **Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production**. *Appl Microbiol Biotechnol* 2014, **98**:5301-5317.
291. Khatri NK, Hoffmann F: **Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant Pichia pastoris**. *Biotechnol Bioeng* 2006, **93**:871-879.
292. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM: **Heterologous protein production using the Pichia pastoris expression system**. *Yeast* 2005, **22**:249-270.

293. Damasceno LM, Huang CJ, Batt CA: **Protein secretion in Pichia pastoris and advances in protein production**. *Appl Microbiol Biotechnol* 2012, **93**:31-39.
294. Marx H, Mecklenbrauker A, Gasser B, Sauer M, Mattanovich D: **Directed gene copy number amplification in Pichia pastoris by vector integration into the ribosomal DNA locus**. *FEMS Yeast Res* 2009, **9**:1260-1270.
295. Inan M, Aryasomayajula D, Sinha J, Meagher MM: **Enhancement of protein secretion in Pichia pastoris by overexpression of protein disulfide isomerase**. *Biotechnol Bioeng* 2006, **93**:771-778.
296. Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D: **Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant Pichia pastoris**. *Biotechnol Bioeng* 2004, **85**:367-375.
297. Yang Z, Zhang Z: **Engineering strategies for enhanced production of protein and bio-products in Pichia pastoris: A review**. *Biotechnol Adv* 2018, **36**:182-195.
298. Greiner M, Pfeiffer D, Smith RD: **Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests**. *Prev Vet Med* 2000, **45**:23-41.
299. Kuenzig ME, Bishay K, Leigh R, Kaplan GG, Benchimol EI: **Co-occurrence of Asthma and the Inflammatory Bowel Diseases: A Systematic Review and Meta-analysis**. *Clin Transl Gastroenterol* 2018, **9**:188.
300. Nials AT, Uddin S: **Mouse models of allergic asthma: acute and chronic allergen challenge**. *Dis Model Mech* 2008, **1**:213-220.
301. Kiesler P, Fuss IJ, Strober W: **Experimental Models of Inflammatory Bowel Diseases**. *Cell Mol Gastroenterol Hepatol* 2015, **1**:154-170.
302. Wirtz S, Popp V, Kindermann M, Gerlach K, Weigmann B, Fichtner-Feigl S, Neurath MF: **Chemically induced mouse models of acute and chronic intestinal inflammation**. *Nat Protoc* 2017, **12**:1295-1309.
303. Antoniou E, Margonis GA, Angelou A, Pikouli A, Argiri P, Karavokyros I, Papalois A, Pikoulis E: **The TNBS-induced colitis animal model: An overview**. *Ann Med Surg (Lond)* 2016, **11**:9-15.
304. Dohi T, Fujihashi K, Rennert PD, Iwatani K, Kiyono H, McGhee JR: **Hapten-induced colitis is associated with colonic patch hypertrophy and T helper cell 2-type responses**. *J Exp Med* 1999, **189**:1169-1180.
305. Havaux X, Zeine A, Dits A, Denis O: **A new mouse model of lung allergy induced by the spores of Alternaria alternata and Cladosporium herbarum molds**. *Clin Exp Immunol* 2005, **139**:179-188.
306. Kim J, Merry AC, Nemzek JA, Bolgos GL, Siddiqui J, Remick DG: **Eotaxin represents the principal eosinophil chemoattractant in a novel murine asthma model induced by house dust containing cockroach allergens**. *J Immunol* 2001, **167**:2808-2815.
307. Cates EC, Fattouh R, Wattie J, Inman MD, Goncharova S, Coyle AJ, Gutierrez-Ramos JC, Jordana M: **Intranasal exposure of mice to house dust mite elicits allergic airway inflammation via a GM-CSF-mediated mechanism**. *J Immunol* 2004, **173**:6384-6392.
308. Aun MV, Bonamichi-Santos R, Arantes-Costa FM, Kalil J, Giavina-Bianchi P: **Animal models of asthma: utility and limitations**. *J Asthma Allergy* 2017, **10**:293-301.
309. Gregory LG, Lloyd CM: **Orchestrating house dust mite-associated allergy in the lung**. *Trends Immunol* 2011, **32**:402-411.
310. Piyadasa H, Altieri A, Basu S, Schwartz J, Halayko AJ, Mookherjee N: **Biosignature for airway inflammation in a house dust mite-challenged murine model of allergic asthma**. *Biol Open* 2016, **5**:112-121.
311. McMillan SJ, Lloyd CM: **Prolonged allergen challenge in mice leads to persistent airway remodelling**. *Clin Exp Allergy* 2004, **34**:497-507.
312. Johnson JR, Wiley RE, Fattouh R, Swirski FK, Gajewska BU, Coyle AJ, Gutierrez-Ramos JC, Ellis R, Inman MD, Jordana M: **Continuous exposure to house dust mite elicits chronic airway inflammation and structural remodeling**. *Am J Respir Crit Care Med* 2004, **169**:378-385.

313. Ferreira IB, Pickering DA, Troy S, Croese J, Loukas A, Navarro S: **Suppression of inflammation and tissue damage by a hookworm recombinant protein in experimental colitis**. *Clin Transl Immunology* 2017, **6**:e157.
314. Radi ZA, Heuvelman DM, Masferrer JL, Benson EL: **Pharmacologic evaluation of sulfasalazine, FTY720, and anti-IL-12/23p40 in a TNBS-induced Crohn's disease model**. *Dig Dis Sci* 2011, **56**:2283-2291.
315. Witaicenis A, Luchini AC, Hiruma-Lima CA, Felisbino SL, Garrido-Mesa N, Utrilla P, Galvez J, Di Stasi LC: **Suppression of TNBS-induced colitis in rats by 4-methylesculetin, a natural coumarin: comparison with prednisolone and sulphasalazine**. *Chem Biol Interact* 2012, **195**:76-85.
316. Madden K, Janczak J, McEnroe G, Lim D, Hartman T, Liu D, Stanton L: **A peptide derived from neutrophil inhibitory factor (NIF) blocks neutrophil adherence to endothelial cells**. *Inflamm Res* 1997, **46**:216-223.
317. Tribolet L, Cantacessi C, Pickering DA, Navarro S, Doolan DL, Trieu A, Fei H, Chao Y, Hofmann A, Gasser RB, et al.: **Probing of a human proteome microarray with a recombinant pathogen protein reveals a novel mechanism by which hookworms suppress B-cell receptor signaling**. *J Infect Dis* 2015, **211**:416-425.
318. Woo LN, Guo WY, Wang X, Young A, Salehi S, Hin A, Zhang Y, Scott JA, Chow CW: **A 4-Week Model of House Dust Mite (HDM) Induced Allergic Airways Inflammation with Airway Remodeling**. *Sci Rep* 2018, **8**:6925.
319. Shefrin AE, Goldman RD: **Use of dexamethasone and prednisone in acute asthma exacerbations in pediatric patients**. *Can Fam Physician* 2009, **55**:704-706.
320. Greenfield NJ: **Using circular dichroism spectra to estimate protein secondary structure**. *Nat Protoc* 2006, **1**:2876-2890.
321. Jenny M, Klieber M, Zaknun D, Schroecksnadel S, Kurz K, Ledochowski M, Schennach H, Fuchs D: **In vitro testing for anti-inflammatory properties of compounds employing peripheral blood mononuclear cells freshly isolated from healthy donors**. *Inflamm Res* 2011, **60**:127-135.
322. Romagnani S: **Immunologic influences on allergy and the TH1/TH2 balance**. *J Allergy Clin Immunol* 2004, **113**:395-400.
323. Schoenborn JR, Wilson CB: **Regulation of interferon-gamma during innate and adaptive immune responses**. *Adv Immunol* 2007, **96**:41-101.
324. Park BS, Lee JO: **Recognition of lipopolysaccharide pattern by TLR4 complexes**. *Exp Mol Med* 2013, **45**:e66.
325. Ai W, Li H, Song N, Li L, Chen H: **Optimal method to stimulate cytokine production and its use in immunotoxicity assessment**. *Int J Environ Res Public Health* 2013, **10**:3834-3842.
326. Yang Z, Hou Y, Hao T, Rho HS, Wan J, Luan Y, Gao X, Yao J, Pan A, Xie Z, et al.: **A Human Proteome Array Approach to Identifying Key Host Proteins Targeted by Toxoplasma Kinase ROP18**. *Mol Cell Proteomics* 2017, **16**:469-484.
327. Yu X, Decker KB, Barker K, Neunuebel MR, Saul J, Graves M, Westcott N, Hang H, LaBaer J, Qiu J, et al.: **Host-pathogen interaction profiling using self-assembling human protein arrays**. *J Proteome Res* 2015, **14**:1920-1936.
328. Schweitzer B, Predki P, Snyder M: **Microarrays to characterize protein interactions on a whole-proteome scale**. *Proteomics* 2003, **3**:2190-2199.
329. Bacon KB, Oppenheim JJ: **Chemokines in disease models and pathogenesis**. *Cytokine Growth Factor Rev* 1998, **9**:167-173.
330. Smyth MJ, Zachariae CO, Norihisa Y, Ortaldo JR, Hishinuma A, Matsushima K: **IL-8 gene expression and production in human peripheral blood lymphocyte subsets**. *J Immunol* 1991, **146**:3815-3823.
331. Parameswaran N, Patial S: **Tumor necrosis factor-α signaling in macrophages**. *Crit Rev Eukaryot Gene Expr* 2010, **20**:87-103.
332. Rawla P, Sunkara T, Raj JP: **Role of biologics and biosimilars in inflammatory bowel disease: current trends and future perspectives**. *J Inflamm Res* 2018, **11**:215-226.

333. Geiger SM, Fujiwara RT, Freitas PA, Massara CL, Dos Santos Carvalho O, Corrêa-Oliveira R, Bethony JM: **Excretory-secretory products from hookworm l(3) and adult worms suppress proinflammatory cytokines in infected individuals**. *J Parasitol Res* 2011, **2011**:512154.
334. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y, et al.: **The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses**. *Curr Protoc Bioinformatics* 2016, **54**:1.30.31-31.30.33.
335. Kodadek T: **Protein microarrays: prospects and problems**. *Chem Biol* 2001, **8**:105-115.
336. Cox FE: **History of human parasitology**. *Clin Microbiol Rev* 2002, **15**:595-612.
337. Poinar G, Jr., Boucot AJ: **Evidence of intestinal parasites of dinosaurs**. *Parasitology* 2006, **133**:245-249.
338. Denton JF, Lugo-Martinez J, Tucker AE, Schrider DR, Warren WC, Hahn MW: **Extensive error in the number of genes inferred from draft genome assemblies**. *PLoS Comput Biol* 2014, **10**:e1003998.
339. Cantacessi C, Gasser RB: **SCP/TAPS proteins in helminths--where to from now?** *Mol Cell Probes* 2012, **26**:54-59.
340. van Loon LC, Rep M, Pieterse CM: **Significance of inducible defense-related proteins in infected plants**. *Annu Rev Phytopathol* 2006, **44**:135-162.
341. Nobile M, Noceti F, Prestipino G, Possani LD: **Helothermine, a lizard venom toxin, inhibits calcium current in cerebellar granules**. *Exp Brain Res* 1996, **110**:15-20.
342. Shikamoto Y, Suto K, Yamazaki Y, Morita T, Mizuno H: **Crystal structure of a CRISP family Ca2+ -channel blocker derived from snake venom**. *J Mol Biol* 2005, **350**:735-743.
343. Yamazaki Y, Morita T: **Structure and function of snake venom cysteine-rich secretory proteins**. *Toxicon* 2004, **44**:227-231.
344. Jalkanen J, Huhtaniemi I, Poutanen M: **Mouse cysteine-rich secretory protein 4 (CRISP4): a member of the Crisp family exclusively expressed in the epididymis in an androgen-dependent manner**. *Biol Reprod* 2005, **72**:1268-1274.
345. Kasahara M, Gutknecht J, Brew K, Spurr N, Goodfellow PN: **Cloning and mapping of a testis-specific gene with sequence similarity to a sperm-coating glycoprotein gene**. *Genomics* 1989, **5**:527-534.
346. Tawe W, Pearlman E, Unnasch TR, Lustigman S: **Angiogenic activity of Onchocerca volvulus recombinant proteins similar to vespid venom antigen 5**. *Mol Biochem Parasitol* 2000, **109**:91-99.
347. Lizotte-Waniewski M, Tawe W, Guiliano DB, Lu W, Liu J, Williams SA, Lustigman S: **Identification of potential vaccine and drug target candidates by expressed sequence tag analysis and immunoscreening of Onchocerca volvulus larval cDNA libraries**. *Infect Immun* 2000, **68**:3491-3501.
348. MacDonald AJ, Tawe W, Leon O, Cao L, Liu J, Oksov Y, Abraham D, Lustigman S: **Ov-ASP-1, the Onchocerca volvulus homologue of the activation associated secreted protein family is immunostimulatory and can induce protective anti-larval immunity**. *Parasite Immunol* 2004, **26**:53-62.
349. Murray J, Gregory WF, Gomez-Escobar N, Atmadja AK, Maizels RM: **Expression and immune recognition of Brugia malayi VAL-1, a homologue of vespid venom allergens and Ancylostoma secreted proteins**. *Mol Biochem Parasitol* 2001, **118**:89-96.
350. Zhan B, Liu Y, Badamchian M, Williamson A, Feng J, Loukas A, Hawdon JM, Hotez PJ: **Molecular characterisation of the Ancylostoma-secreted protein family from the adult stage of Ancylostoma caninum**. *Int J Parasitol* 2003, **33**:897-907.
351. Moyle M, Foster DL, McGrath DE, Brown SM, Laroche Y, De Meutter J, Stanssens P, Bogowitz CA, Fried VA, Ely JA, et al.: **A hookworm glycoprotein that inhibits neutrophil function is a ligand of the integrin CD11b/CD18**. *J Biol Chem* 1994, **269**:10008-10015.
352. Noon JB, Schwarz EM, Ostroff GR, Aroian RV: **A highly expressed intestinal cysteine protease of Ancylostoma ceylanicum protects vaccinated hamsters from hookworm infection**. *PLoS Negl Trop Dis* 2019, **13**:e0007345.

353. Haagsma JA, Graetz N, Bolliger I, Naghavi M, Higashi H, Mullany EC, Abera SF, Abraham JP, Adofo K, Alsharif U, et al.: **The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study 2013**. *Inj Prev* 2016, **22**:3-18.
354. Bartsch SM, Hotez PJ, Asti L, Zapf KM, Bottazzi ME, Diemert DJ, Lee BY: **The Global Economic and Health Burden of Human Hookworm Infection**. *PLoS Negl Trop Dis* 2016, **10**:e0004922.
355. Diemert DJ, Bottazzi ME, Plieskatt J, Hotez PJ, Bethony JM: **Lessons along the Critical Path: Developing Vaccines against Human Helminths**. *Trends Parasitol* 2018, **34**:747-758.
356. Nisbet AJ, McNeilly TN, Wildblood LA, Morrison AA, Bartley DJ, Bartley Y, Longhi C, McKendrick IJ, Palarea-Albaladejo J, Matthews JB: **Successful immunization against a parasitic nematode by vaccination with recombinant proteins**. *Vaccine* 2013, **31**:4017-4023.
357. Rivera Marrero CA, Santiago N, Hillyer GV: **Evaluation of immunodiagnostic antigens in the excretory-secretory products of Fasciola hepatica**. *J Parasitol* 1988, **74**:646-652.
358. Ju JW, Joo HN, Lee MR, Cho SH, Cheun HI, Kim JY, Lee YH, Lee KJ, Sohn WM, Kim DM, et al.: **Identification of a serodiagnostic antigen, legumain, by immunoproteomic analysis of excretory-secretory products of Clonorchis sinensis adult worms**. *Proteomics* 2009, **9**:3066-3078.
359. Wangchuk P, Shepherd C, Constantinoiu C, Ryan RYM, Kouremenos KA, Becker L, Jones L, Buitrago G, Giacomin P, Wilson D, et al.: **Hookworm-Derived Metabolites Suppress Pathology in a Mouse Model of Colitis and Inhibit Secretion of Key Inflammatory Cytokines in Primary Human Leukocytes**. *Infect Immun* 2019, **87**.
360. Buck AH, Coakley G, Simbari F, McSorley HJ, Quintana JF, Le Bihan T, Kumar S, Abreu-Goodger C, Lear M, Harcus Y, et al.: **Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity**. *Nat Commun* 2014, **5**:5488.
361. Brondyk WH: **Selecting an appropriate method for expressing a recombinant protein**. *Methods Enzymol* 2009, **463**:131-147.
362. Jackson AM, Boutell J, Cooley N, He M: **Cell-free protein synthesis for proteomics**. *Brief Funct Genomic Proteomic* 2004, **2**:308-319.
363. Hunt I: **From gene to protein: a review of new and enabling technologies for multi-parallel protein expression**. *Protein Expr Purif* 2005, **40**:1-22.
364. Chassaing B, Aitken JD, Malleshappa M, Vijay-Kumar M: **Dextran sulfate sodium (DSS)-induced colitis in mice**. *Curr Protoc Immunol* 2014, **104**:Unit 15.25.
365. Keubler LM, Buettner M, Häger C, Bleich A: **A Multihit Model: Colitis Lessons from the Interleukin-10-deficient Mouse**. *Inflamm Bowel Dis* 2015, **21**:1967-1975.
366. Eri R, McGuckin MA, Wadley R: **T cell transfer model of colitis: a great tool to assess the contribution of T cells in chronic intestinal inflammation**. *Methods Mol Biol* 2012, **844**:261-275.
367. Logan J, Navarro S, Loukas A, Giacomin P: **Helminth-induced regulatory T cells and suppression of allergic responses**. *Curr Opin Immunol* 2018, **54**:1-6.
368. Hsieh GC, Loukas A, Wahl AM, Bhatia M, Wang Y, Williamson AL, Kehn KW, Maruyama H, Hotez PJ, Leitenberg D, et al.: **A secreted protein from the human hookworm necator americanus binds selectively to NK cells and induces IFN-gamma production**. *J Immunol* 2004, **173**:2699-2704.
369. McSorley HJ, Loukas A: **The immunology of human hookworm infections**. *Parasite Immunol* 2010, **32**:549-559.
370. Grievink HW, Moerland M: **Sample Aging Profoundly Reduces Monocyte Responses in Human Whole Blood Cultures**. *J Immunol Res* 2018, **2018**:8901485.
371. Eggesbo JB, Hjermann I, Lund PK, Joo GB, Ovstebo R, Kierulf P: **LPS-induced release of IL-1 beta, IL-6, IL-8, TNF-alpha and sCD14 in whole blood and PBMC from persons with high or low levels of HDL-lipoprotein**. *Cytokine* 1994, **6**:521-529.
372. Hunter CA, Jones SA: **IL-6 as a keystone cytokine in health and disease**. *Nat Immunol* 2015, **16**:448-457.

373. Akdis M, Burgler S, Crameri R, Eiwegger T, Fujita H, Gomez E, Klunker S, Meyer N, O'Mahony L, Palomares O, et al.: **Interleukins, from 1 to 37, and interferon-gamma: receptors, functions, and roles in diseases**. *J Allergy Clin Immunol* 2011, **127**:701-721.e701-770.
374. Smith KA, Maizels RM: **IL-6 controls susceptibility to helminth infection by impeding Th2 responsiveness and altering the Treg phenotype in vivo**. *Eur J Immunol* 2014, **44**:150-161.
375. Jeurink PV, Vissers YM, Rappard B, Savelkoul HF: **T cell responses in fresh and cryopreserved peripheral blood mononuclear cells: kinetics of cell viability, cellular subsets, proliferation, and cytokine production**. *Cryobiology* 2008, **57**:91-103.
376. Hernandez-Pando R, Rook GA: **The role of TNF-alpha in T-cell-mediated inflammation depends on the Th1/Th2 cytokine balance**. *Immunology* 1994, **82**:591-595.
377. Artis D, Potten CS, Else KJ, Finkelman FD, Grencis RK: **Trichuris muris: host intestinal epithelial cell hyperproliferation during chronic infection is regulated by interferon-gamma**. *Exp Parasitol* 1999, **92**:144-153.
378. Jansky L, Reymanova P, Kopecky J: **Dynamics of cytokine production in human peripheral blood mononuclear cells stimulated by LPS or infected by Borrelia**. *Physiol Res* 2003, **52**:593-598.
379. De Groote D, Zangerle PF, Gevaert Y, Fassotte MF, Beguin Y, Noizat-Pirenne F, Pirenne J, Gathy R, Lopez M, Dehart I, et al.: **Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation**. *Cytokine* 1992, **4**:239-248.
380. Murakami T, Ockinger J, Yu J, Byles V, McColl A, Hofer AM, Horng T: **Critical role for calcium mobilization in activation of the NLRP3 inflammasome**. *Proc Natl Acad Sci U S A* 2012, **109**:11282-11287.
381. Inclan-Rico JM, Siracusa MC: **First Responders: Innate Immunity to Helminths**. *Trends Parasitol* 2018, **34**:861-880.
382. Zaiss MM, Maslowski KM, Mosconi I, Guenat N, Marsland BJ, Harris NL: **IL-1beta suppresses innate IL-25 and IL-33 production and maintains helminth chronicity**. *PLoS Pathog* 2013, **9**:e1003531.
383. Alhallaf R, Agha Z, Miller CM, Robertson AAB, Sotillo J, Croese J, Cooper MA, Masters SL, Kupz A, Smith NC, et al.: **The NLRP3 Inflammasome Suppresses Protective Immunity to Gastrointestinal Helminth Infection**. *Cell Rep* 2018, **23**:1085-1098.
384. Ferguson BJ, Newland SA, Gibbs SE, Tourlomousis P, Fernandes dos Santos P, Patel MN, Hall SW, Walczak H, Schramm G, Haas H, et al.: **The Schistosoma mansoni T2 ribonuclease omega-1 modulates inflammasome-dependent IL-1beta secretion in macrophages**. *Int J Parasitol* 2015, **45**:809-813.
385. Prêle CM, Woodward EA, Bisley J, Keith-Magee A, Nicholson SE, Hart PH: **SOCS1 regulates the IFN but not NFkappaB pathway in TLR-stimulated human monocytes and macrophages**. *J Immunol* 2008, **181**:8018-8026.
386. Scheurich P, Thoma B, Ucer U, Pfizenmaier K: **Immunoregulatory activity of recombinant human tumor necrosis factor (TNF)-alpha: induction of TNF receptors on human T cells and TNF-alpha-mediated enhancement of T cell responses**. *J Immunol* 1987, **138**:1786-1790.
387. Yokota S, Geppert TD, Lipsky PE: **Enhancement of antigen- and mitogen-induced human T lymphocyte proliferation by tumor necrosis factor-alpha**. *J Immunol* 1988, **140**:531-536.
388. Corkum CP, Ings DP, Burgess C, Karwowska S, Kroll W, Michalak TI: **Immune cell subsets and their gene expression profiles from human PBMC isolated by Vacutainer Cell Preparation Tube (CPT™) and standard density gradient**. *BMC Immunol* 2015, **16**:48.
389. Autissier P, Soulas C, Burdo TH, Williams KC: **Evaluation of a 12-color flow cytometry panel to study lymphocyte, monocyte, and dendritic cell subsets in humans**. *Cytometry A* 2010, **77**:410-419.
390. Bhattacharjee J, Das B, Mishra A, Sahay P, Upadhyay P: **Monocytes isolated by positive and negative magnetic sorting techniques show different molecular characteristics and immunophenotypic behaviour**. *F1000Res* 2017, **6**:2045.

391. Ingersoll MA, Spanbroek R, Lottaz C, Gautier EL, Frankenberger M, Hoffmann R, Lang R, Haniffa M, Collin M, Tacke F, et al.: **Comparison of gene expression profiles between human and mouse monocyte subsets**. *Blood* 2010, **115**:e10-19.
392. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szczesniak MW, Gaffney DJ, Elo LL, Zhang X, et al.: **A survey of best practices for RNA-seq data analysis**. *Genome Biol* 2016, **17**:13.
393. He L, Grammer AC, Wu X, Lipsky PE: **TRAF3 forms heterotrimers with TRAF2 and modulates its ability to mediate NF-{kappa}B activation**. *J Biol Chem* 2004, **279**:55855-55865.
394. Bishop GA, Stunz LL, Hostager BS: **TRAF3 as a Multifaceted Regulator of B Lymphocyte Survival and Activation**. *Front Immunol* 2018, **9**:2161.
395. Kolltveit KM, Granum S, Aasheim HC, Forsbring M, Sundvold-Gjerstad V, Dai KZ, Molberg O, Schjetne KW, Bogen B, Shapiro VS, et al.: **Expression of SH2D2A in T-cells is regulated both at the transcriptional and translational level**. *Mol Immunol* 2008, **45**:2380-2390.
396. Barrow AD, Trowsdale J: **The extended human leukocyte receptor complex: diverse ways of modulating immune responses**. *Immunol Rev* 2008, **224**:98-123.

Appendices

Recipe List

Lysogeny broth (LB) low-salt media

10 g Tryptone (Sigma Aldrich), 5 g NaCl and 5 g Yeast Extract (Sigma Aldrich) were combined, made up to 950 ml with dH_2O, pH adjusted to 7.5 with 1 M NaOH, and topped up to 1 L before autoclaving on a liquid cycle at 121°C for 20 minutes. Solution was stored at room temperature until used.

Lysogeny broth (LB) low-salt agar

10 g Tryptone, 5 g NaCl and 5 g Yeast Extract were combined, made up to 950 ml with dH_2O, pH adjusted to 7.5 with 1 M NaOH, 15 g of agar added, and topped up to 1 L before autoclaving on a liquid cycle at 121°C for 20 minutes. Solution was stored at room temperature until used.

50% Glycerol

50 ml Glycerol (Sigma-Aldrich) was added to 50 ml dH_2O. Solution was autoclaved on a liquid cycle at 121°C for 20 minutes or filter sterilized with a 0.22 µm filter (Millipore) before storing at room temperature.

15% Glycerol

15 ml Glycerol was added to 85 ml dH_2O. Solution was autoclaved on a liquid cycle at 121°C for 20 minutes or filter sterilized with a 0.22 µm filter before storing at room temperature.

1% Agarose gel

1.5 g agarose was added to 150 ml of 1X TAE Buffer and microwaved for 30 second intervals until dissolved.

1X TAE Buffer

4.84 g Trizma Base (Sigma-Aldrich), 0.372 g Ethylenediaminetetraacetic acid (EDTA) disodium salt dihydrate (Sigma-Aldrich) and 1.14 ml Glacial acetic acid (Chem-supply) were combined and made up to 1 L with dH_2O.

Yeast Extract Peptone Dextrose Agar with Sorbitol (YPDS)

10 g Yeast Extract, 20 g Tryptone, 182.2 g D-Sorbitol (Sigma-Aldrich) and 20 g Agar were combined and made up to 900 ml with dH_2O. The solution was autoclaved on a liquid cycle at 121°C for 20 minutes. 20 g Dextrose was dissolved in 100 ml dH_2O, filter-sterilised with a 0.22 µm filter, and added to the solution to make a total volume of 1 L. For YPDS plates with 100 µg/ml Zeocin (Invivogen), agar was allowed to cool below 55°C before adding Zeocin. YPDS agar plates with Zeocin were stored in the dark at 4°C and used within 1 week of preparation.

YPD media

10 g Yeast Extract, 20 g Tryptone were combined and made up to 900 ml with dH_2O. The solution was autoclaved on a liquid cycle at 121°C for 20 minutes. 20 g Dextrose was dissolved in 100 ml dH_2O, filter-sterilised with a 0.22 µm filter, and added to the solution to make a total volume of 1 L.

1 M Sorbitol

18.22 g D-Sorbitol (Sigma-Aldrich) was made up to 100 ml with dH_2O before autoclaving on a liquid cycle at 121°C for 20 minutes.

BMGY

10 g Yeast Extract, 20 g Tryptone were combined and made up to 700 ml with dH_2O. The solution was autoclaved on a liquid cycle at 121°C for 20 minutes. 100 ml of 1 M Potassium Phosphate Buffer pH6.0, 100 ml 10X YNB, 100 ml 10X Glycerol and 2 ml 500X Biotin was added before use. Media that was not used immediately was stored at 4°C.

BMMY

10 g Yeast Extract, 20 g Tryptone were combined and made up to 700 ml with dH_2O. The solution was autoclaved on a liquid cycle at 121°C for 20 minutes. 100 ml of 1 M Potassium Phosphate Buffer pH6.0, 100 ml 10X YNB, 100 ml 10X Methanol and 2 ml 500X Biotin was added before use. Media that was not used immediately was stored at 4°C.

1 M Potassium Phosphate Buffer pH6.0

174 g Potassium Phosphate Dibasic (Sigma-Aldrich) was dissolved in 1 L of dH_2O to make a 1 M solution. 136 g Potassium Phosphate Monobasic (Sigma-Aldrich) was dissolved in 1 L of dH_2O to make a 1 M solution. 868 ml of 1 M Potassium Phosphate Monobasic was added to 132 ml of 1 M Potassium Phosphate Dibasic. The pH of the solution was adjusted to 6.0 using 10M Potassium hydroxide (KOH). The solution was autoclaved on a liquid cycle at 121°C for 20 minutes.

10X Yeast Nitrogen Base (YNB)

67 g of Yeast Nitrogen Base without Amino Acids (Sigma-Aldrich) was made up 500 ml with dH_2O. The solution was filter-sterilised with a 0.22 µm filter. Media that was not used immediately was stored at 4°C.

10X Glycerol

100 ml Glycerol was made up to 1 L of dH_2O. The solution was autoclaved on a liquid cycle at 121°C for 20 minutes.

500X Biotin

10 mg Biotin (Sigma-Aldrich) was dissolved in 50 ml dH_2O. The solution was filter-sterilised with a 0.22 µm filter. Media that was not used immediately was stored at 4°C.

10X Methanol

25 ml Methanol (Sigma-Aldrich) was added to 475 ml dH_2O. The solution was filter-sterilised with a 0.22 µm filter. Media that was not used immediately was stored at 4°C.

Phosphate Buffered Saline 0.05% Tween-20 (PBST)

500 µl Tween 20 (Sigma-Aldrich) was added to 1 L of PBS. The solution was stored at room temperature until used.

5% Skim milk in PBST

2.5 g of Skim milk powder (Coles) was added to 50 ml PBST.

4X SDS Protein loading buffer

3.1 ml dH_2O, 4 ml Glycerol, 0.8g Sodium dodecyl sulfate (Sigma-Aldrich), 4 mg Bromophenol Blue (Sigma-Aldrich), 0.5 ml 2-Mercaptoethanol (Sigma-Aldrich) were combined. 0.378 g of Trizma hydrochloride (Sigma-Aldrich) was made up to 2.4 ml with dH_2O to make a 1 M solution, and the pH was adjusted to 6.8. The 2.4 ml 1 M Trizma HCl was added to the other reagents to complete the loading buffer.

10X SDS Running Buffer

30 g of Trizma Base, 144 g of Glycine (Sigma-Aldrich) and 10 g of Sodium dodecyl sulfate were dissolved in and made up to 1 L with dH_2O.

Western transfer Buffer

200 ml of Methanol was added to 100 ml of 10X SDS Running Buffer and made up to 1 L with dH_2O.

Buffer A

DPBS (-/-) (1X) (Life Technologies)

Buffer B

DPBS (-/-) (1X) with 500 mM Imidazole (Sigma-Aldrich)

1 M Sodium hydroxide

40 g Sodium hydroxide (Sigma-Aldrich) was made up to 1 L with dH_2O.