

Research review paper

Biotechnology Advances

DRAFT 241108 (~7100 words)

A portrait of the “SCP/TAPS” proteins of eukaryotes – developing a framework for fundamental research and biotechnological outcomes

C. Cantacessi ^a, B.E. Campbell ^a, A. Visser ^b, P. Geldhof ^b, M.J. Nolan ^a, A.J. Nisbet ^c, J.B. Matthews ^c, A. Loukas ^d, A. Hofmann ^e, D. Otranto ^f, R.B. Gasser ^{a,*}

^a *Department of Veterinary Science, The University of Melbourne, Australia*

^b *Laboratory for Parasitology, The University of Gent, Belgium*

^c *Moredun Research Institute, Scotland*

^d *Division of Infectious Diseases, Queensland Institute of Medical Research, Australia*

^e *Eskitis Institute for Cell and Molecular Therapies, Griffith University, Australia*

^f *Department of Veterinary Public Health, Faculty of Veterinary Medicine, Bari, Italy*

* Corresponding author. Tel.: +61-3-97312000; fax: +61-3-97312366.

E-mail address: robinbg@unimelb.edu.au (R.B. Gasser).

Contents

1. Introduction
 2. SCP/TAPS proteins in plants: defence-related proteins
 3. SCP/TAPS proteins in animals: from snake venom to cancer cells
 4. SCP/TAPS proteins in parasitic helminths: ‘activation-associated secreted proteins’
 - 4.1. *Nematodes within clade III (including the Orders Spirurida and Ascaridida)*
 - 4.2. *Nematodes within clade IV (including the Order Tylenchida)*
 - 4.3. *Nematodes within clade V (including the Orders Strongylida and Rhabditida)*
 - 4.3.1. *Superfamily Ancylostomatoidea (Order Strongylida)*
 - 4.3.2. *Superfamily Trichostrongyloidea (Order Strongylida)*
 - 4.3.3. *The free-living nematode C. elegans (Order Rhabditida)*
 - 4.4. *Trematodes (i.e. Subclass Digenea)*
 5. Phylogenetic analysis of eukaryote SCP/TAPS proteins: a proposed classification system
 6. Concluding remarks and biotechnological implications
- Acknowledgements
- References

Abstract

A wide range of proteins belonging to the SCP/TAPS “family” has been described for various eukaryotic organisms, including plants and animals (vertebrates and invertebrates, such as helminths). Although SCP/TAPS proteins have been proposed to play key roles in a number of fundamental biological processes, such as host-pathogen interactions and defence mechanisms, there is a paucity of information on their genetic relationships, structures and functions and there is no standardised nomenclature for these proteins. A detailed analysis of the relationships of members of the SCP/TAPS family of proteins, based on key protein signatures, could provide a foundation for investigating these areas. In this article, we review the current state of knowledge of key SCP/TAPS proteins of eukaryotes, with an emphasis on those from parasitic helminths, undertake a comprehensive, systematic phylogenetic analysis of currently available full-length protein sequence data (considering characteristic protein signatures or motifs) to infer relationships and provide a framework (based on statistical support) for the naming of these proteins. This framework is intended to guide genomic and molecular biological explorations of key SCP/TAPS molecules associated with infectious diseases of plants and animals. In particular, fundamental investigations of these molecules in parasites and the integration of structural and functional data could lead to new and innovative approaches for the control of parasitic diseases, with important biotechnological outcomes.

Keywords: sperm coating protein-extracellular domain; SCP/TAPS; activation associated proteins; hookworms; classification; nomenclature; Bayesian inference; venom allergen; pathogenesis-related proteins

1. Introduction

Pathogens have evolved a range of strategies to invade their hosts, while hosts have developed immune and other defence mechanisms against such pathogens (reviewed by [Schmid-Hempel, 2008](#)). Numerous host-pathogen interactions involve molecular processes, in which proteins and other molecules from a pathogen target host proteins to initiate or maintain the infection, stimulate or evade the host's immune response and/or cause disease (reviewed by [Schmid-Hempel, 2008](#)). Various groups of proteins have been proposed to play major biological roles in the host-pathogen interplay (see e.g., Vermeire et al., 2008). Among these molecules are the sperm-coating protein (SCP)-like extracellular proteins, also called SCP/Tpx-1/Ag5/PR-1/Sc7 (SCP/TAPS; Pfam accession number no. PF00188). SCP/TAPS family members have been identified in various eukaryotes and belong to the cysteine-rich secretory protein (CRISP) "superfamily" (reviewed by Chalmers et al., 2008). SCP/TAPS proteins include rodent sperm-coating glycoproteins (or acidic epididymal glycoproteins, proposed to be involved in sperm maturation during its passage through the epididymis) (Jalkanen et al., 2005), mammalian testis-specific protein (Tpx-1) (Kasahara et al., 1989), glioma pathogenesis-related protein (Murphy et al., 1995; Yamakawa et al., 1998; Rosenzweig et al., 2006), venom allergen 5 from vespid wasps and the venom allergen 3 from fire ants, which mediate allergic reactions to the bites by some insects of the order Hymenoptera (Lu et al., 1993) as well as plant pathogenesis proteins (PRPs) of the PR-1 "subfamily" which are synthesized in response to infections with pathogens or other stress-inducing factors (reviewed by van Loon et al., 2006). Data for a range of eukaryotes suggest that all SCP/TAPS molecules share a common primary structure; the signal peptide is followed by the SCP-extracellular domain (InterProScan: IPR014044) which is proposed to act as a Ca^{2+} -chelator in various signalling processes (Fernandez et al., 1997). In the SCP/TAPS proteins of yeasts, the SCP-domain is variable in length, is polymorphic and is flanked by a threonine-rich region. The C-terminus of family members representing mammals and reptile venoms harbours the CRISP-domain, typified by ten conserved cysteine residues. In invertebrates, such as parasitic worms (= helminths, including flatworms and roundworms), this terminus is smaller and contains only four to six cysteine residues (Yatsuda et al., 2002).

SCP/TAPS homologues have been identified in a range of invertebrates, particularly arthropods, flatworms (trematodes) and roundworms (nematodes), and plants. Despite the roles that SCP/TAPS proteins are proposed to play in fundamental biological processes in various eukaryotes (e.g., van Loon et al., 2006), attempts to define criteria for the unequivocal classification of these molecules are scant and limited to SCP/TAPS proteins of the vinegar fly, *Drosophila melanogaster* (see Kovalick et al., 2005), the blood fluke *Schistosoma mansoni* (see Chalmers et al., 2008) and plants (van Loon, 1990). To date, 17 different "subfamilies" of plant PRPs have been proposed and named in the order in which homologues with defined biological and biochemical properties were discovered (reviewed by van Loon et al., 2006). In parasitic nematodes, *Ancylostoma* secreted proteins or activation-associated proteins (ASPs) were first characterized from hookworms and subsequently from related strongylid nematodes (reviewed by Visser et al., 2008). As ASPs are abundant in the excretory/secretory (ES) products of the infective third-stage larvae (L3), they are thought to play an important role in the transition from the free-living to the parasitic stage during the invasion of the host (Hawdon et al., 1996, 1999; Moser et al., 2005; Datu et al., 2008). Because of their immunogenic properties, one ASP (namely *Na*-ASP-2) is under investigation as a vaccine candidate against the disease caused by the hookworm *Necator americanus* (= necatoriasis) in humans (Bethony et al., 2005; Loukas et al., 2006; Mendez et al., 2008; Xiao et al., 2008). Attempts to classify ASPs have been based on the number and features of the SCP-domain of these molecules. Thus far, three types of ASPs have been described: (1) double domain ASPs, composed of two distinct but related SCP-domains, (2) C-type single domain ASPs and (3) N-type single domain ASPs which have the highest homology to the C- and N-terminus of the SCP-double domain ASPs, respectively (Geldhof et al., 2003). C-type single domain and double-domain ASPs have been identified in a range of nematodes parasitic in animals and plants (reviewed by Visser et al., 2008). To date, N-type single domain ASPs have been characterized only for *Cooperia punctata*, *Ostertagia ostertagi* and *Teladorsagia circumcincta*, three trichostrongylid nematodes of ruminants (Yatsuda et al., 2002; Geldhof et al., 2003; Smith et al., 2008). Usually, new SCP/TAPS gene orthologues or protein homologues identified in parasitic nematodes (based on nucleotide and/or amino acid sequence identities or homologies) are named without consideration of the SCP/TAPS protein group as a whole. Also, names are sometimes assigned to SCP/TAPS gene orthologues/protein homologues for which only partial (nucleotide and/or amino acid) sequence data are provided or available, which has resulted in an unclear relationship among gene(s), transcript(s) and protein(s). The confusion and inconsistencies in classification are exacerbated, as different authors use different names for the same groups of molecules.

Besides their usual designation as ASPs, SCP/TAPS molecules of parasitic nematodes have also been named ‘ASP-like’ (AL), based on their homology to ASPs, identified first in *Ancylostoma* (Saverwyns et al., 2008), and ‘venom allergen-like’ (VAL; Frank et al., 1996; Murray et al., 2001; Chalmers et al., 2008), ‘venom allergen-homologues’ (VAH, Tetteh et al., 1999; Anand et al., 2007) and/or ‘venom allergen-proteins’ (VAP; Gao et al., 2001), based on the homology with the major allergen in the venom of the yellow jacket wasp (Henriksen et al., 2001). In the free-living nematode *Caenorhabditis elegans*, at least 17 different ASP-related genes have been identified (reviewed by Zhan et al., 2003) and given various names (e.g., *vap* or *scl* = ‘SCP-like’). In *C. elegans*, SCP/TAPS molecules are believed to be involved in biological aspects, such as anti-microbial activity (O’Rourke et al., 2006), normal body formation (Wang and Kim, 2003) and lifespan (Ookuma et al., 2003).

Given that SCP/TAPS protein are considered by most authors to be of major biological importance in a wide range of eukaryotes (plants, vertebrates and invertebrates), the inconsistencies in classification as well as the lack of structural and functional information for most of them, we herein reviewed the current state of knowledge of these molecules, with an emphasis on those from parasitic helminths, undertook a comprehensive, systematic phylogenetic analysis of current protein sequence data (considering characteristic protein signatures/motifs) to infer relationships and propose a framework (based on statistical support) for the naming of these proteins. This framework is intended to guide genomic and molecular biological explorations of key SCP/TAPS molecules associated with infectious diseases in plants and animals. In particular, fundamental investigations of these molecules in parasites using novel genomic, proteomic and bioinformatic technologies and the integration of structural and functional data could underpin new approaches for the control of infectious diseases, with biotechnological outcomes.

2. SCP/TAPS proteins in plants: defence-related proteins

In response to infections by pathogens and/or exposure to other stress-inducing factors, many plants accumulate a variety of pathogenesis-related proteins (PRPs), which are involved in antimicrobial activity and other defence-related mechanisms but are not expressed in healthy plant tissues. PRPs are proposed to be present in all monocotyledonous and dicotyledonous plants and have been classified into 17 “subfamilies” (van Loon et al., 2006). The significance of inducible defence-related proteins in pathogen-infected plants has been reviewed in detail elsewhere (van Loon et al., 2006) and is thus not covered in the present section. The main focus here is on selected members of the PR-1 subfamily, the dominant group of PRPs induced by pathogens (van Loon and van Strien, 1999), whose function is the least understood. The PR-1 proteins show significant sequence similarity to SCP/TAPS from mammals, arthropods and parasitic helminths; although their functions are mostly unknown, their occurrence in this diverse range of organisms is suggestive of roles extending beyond those proposed in relation to plant defence (Datta, 1999).

PR-1 proteins were first discovered in 1970 in tobacco mosaic virus (TMV)-infected tobacco (*Nicotiana tabacum*; family Solanaceae) (Gianinazzi et al., 1970). Three acidic (PR-1a, -1b and -1c, pI ~4.5) and one basic (PR-1g, pI ~10.0) PR-1 proteins with differing biological properties were identified and named in order of decreasing electrophoretic mobility in non-denaturing polyacrylamide gels. Similarly, van Loon et al. (1994) proposed that any protein whose amino acid sequence resembled a characterized PR-1 from tobacco should be designated by assigning a letter to any new members within a family, according to the order in which they were identified. Using biochemical and immunological tools, PR-1a, -1b and -1c were first detected in the extracellular spaces and xylem elements of TMV-infected tobacco leaves (Parent and Asselin, 1984). Subsequently, Carr et al. (1987) described the accumulation of PR-1a, -1b and -1c following experimental TMV-infection of *N. tabacum* in central vacuoles known as crystal idioblasts) of specialized leaf cells. Since tobacco PR-1 genes do not encode a known peptide sequence for vacuolar targeting, the synthesis and accumulation of acidic PR-1 proteins within crystal idioblasts in cells appear to be unique (Carr et al., 1985). In other Solanaceae, extracellular proteins corresponding to the acidic PR-1 from tobacco are basic. For example, in tomato (*Lycopersicon esculentum*), basic protein homologues of acidic PR-1a, -1b and -1c were detected in the apoplastic fluid as well as in leaf homogenates infected with mycetes of *Cladosporium fulvum*, the causative agent of the tomato leaf mould (De Wit, 1977). In tomato, the conformation of PR-1b was determined by nuclear magnetic resonance; its unique molecular architecture has four α -helices and four β -strands arranged in an antiparallel fashion between helices. The tight packing of the α -helices on both sides of the central β -sheet (α - β - α “sandwich” structure) results in a compact, bipartite molecular core, which is stabilized by hydrophobic interactions and multiple hydrogen bonds (Fernández et al., 1997). In potato (*Solanum tuberosum*), a direct inhibitory effect of basic PR-1 family members on fungal

pathogens (i.e. *Phytophthora infestans* and *Uromyces fabae*) has been demonstrated *in vitro* and *in vivo* (Niderman et al., 1995; Rauscher et al., 1999), but the precise mode of action and the cellular and molecular targets of PR-1 proteins remain unknown.

Because of difficulties in studying complex functional aspects in higher plants (such as tobacco and tomato), *Arabidopsis thaliana* (family Brassicaceae) has been used as a model to explore the mechanisms of the hypersensitivity response against pathogens (Metzler et al., 1991) and transcriptional events linked to PR-1 in response to *Botrytis cinerea* infection (Gonzalez et al., 2006). In rice (*Oryza sativa*; family Poaceae), homologues of the acidic tobacco PR-1 are encoded by a multigene family whose transcription has been demonstrated to be enhanced in response to cutting, the administration of jasmonic acid, other phytohormones or protein phosphatase inhibitors in combination with the protein kinase inhibitor staurosporine (Agrawal et al., 2000). Although the mode of action of these molecules is unclear, it has been proposed that the expression of one of the genes encoding PR-1 proteins in rice (i.e., OsPR1a) is regulated by an interaction(s) of phytohormones with light signals, which is most likely mediated *via* one or more protein phosphorylation/dephosphorylation events (Agrawal et al., 2000). Although the roles of the PR-1 proteins are enigmatic in many cases, current sequence data indicate some conservation across invertebrates and vertebrates. Therefore, the occurrence of PR-1 protein homologues in different groups of organisms suggests that these proteins share an evolutionary origin and play crucial roles in the biology, function and survival of organisms (cf. van Loon et al., 2006).

3. SCP/TAPS proteins of animals: from insect allergens to cancer cells

In animals, SCP/TAPS molecules represent major components of the venom of a range of biting insects of the orders Hymenoptera and Diptera (see Charlab et al., 1999; Henriksen et al., 2001; Li et al., 2001). Insect venom allergy is a harmful reaction caused by stinging insects, such as bees, yellow jackets wasps, hornets and fire ants (reviewed by Finegold, 2008). The first exposure to insect venom does not usually result in allergic reactions, but subsequent exposures can cause sensitization and disease. Clinical signs can vary from itchiness, affecting any part of the skin, to anaphylactic shock, which can occur within minutes and result in death (reviewed by Finegold, 2008). Individuals allergic to insect stings often show similar symptoms against a range of different insect species, as a result of sensitization to individual antigens from different insect venoms or cross-reactivity with antigens with similar structural and biochemical features (reviewed by Bilò and Bonifazi, 2008). Among the major components of vespid venom with allergenic properties, Ves v 5 is a hyaluronidase, which is related, based on amino acid sequence identity, to SCP/TAPS proteins in other arthropods, mammals, reptiles, fungi and plants (Henriksen et al., 2001). Ves v 5 was first isolated from the venom of different species of yellow jackets, hornets and paper wasps, and its crystal structure shows hydrogen bonding patterns and hydrophobic interactions which define an α - β - α core structure (Henriksen et al., 2001). Homologues of Ves v 5 are predicted to have the same protein folding due to a high degree of sequence similarity. Although it has been proposed that Ves v 5 proteins in the yellow jacket might act as trypsin inhibitors, *via* the blocking of calcium-dependent ryanodine receptors, their biochemical role in inducing venom insect allergy remains enigmatic (Henriksen et al., 2001).

SCP/TAPS proteins are also components of the venom from fire ants (*Solenopsis* spp.) and the saliva of a range of blood-feeding insects, including mosquitoes (*Anopheles gambiae* and *Aedes aegypti*), sandfly (*Lutzomyia longipalpis*) and tsetse fly (*Glossina morsitans morsitans*) (see Charlab et al., 1999; Li et al., 2001; Valenzuela et al., 2004; Calvo et al., 2006; Hoffman, 2006). Given their abundance in the sialo-transcriptome of adult female *A. gambiae* and *L. longipalpis*, homologues of Ves v 5 have been proposed to be associated with the physiology of the secretory process (e.g., antihemostatic activity, “stabilizers” of the other secretory products, or assist in lubricating the insect mouth parts), thus explaining its ubiquity and relative sequence conservation (Charlab et al., 1999; Calvo et al., 2006). In *G. morsitans*, SCP/TAPS homologues have been shown to be expressed primarily in salivary glands but could also be detected in midgut and proventriculus; therefore, the authors hypothesized a role for these proteins in the mediation of immune reactions in the gut of tsetse flies (Li et al., 2001). In *Drosophila melanogaster*, the SCP/TAPS gene family includes 26 molecules belonging to two distinct groups based on conserved positioning of introns and sequence similarities in the SCP-domain in the genes (Kovalick and Griffin, 2005). Seventy per cent of the *D. melanogaster* SCP/TAPS genes have been shown to be specifically expressed in the testis (seminal vesicles and/or primary spermatocytes), thus suggesting a central role in male reproduction, either by mediating interactions between germ cells and somatic cells within the male reproductive tract or between sperm and egg in the female fly (Haynes et al., 1997; Kovalick and Griffin, 2005).

SCP/TAPS proteins are significant components of snake venom. They belong to a group of toxins affecting smooth muscle contraction and another group affecting cyclic nucleotide-gated (CNG) ion channels (Yamazaki and Morita, 2004). Within the first group, tigrin (isolated from the venom of the Yamakagashi snake, *Rhabdopsis tigrinus tigrinus*), ablomin (from the Japanese mamushi snake, *Agkistrodon blomhoffi*), triflin (from the habu snake, *Trimeresurus flavoviridis*) and latisemin (from the erabu sea snake, *Laticauda semifasciata*) share ~50% sequence homology with mammalian SCP/TAPS proteins, with all cysteine residues being conserved (Yamazaki and Morita, 2004). Tigrin is a helotermine with ion channel-blocking activity (Nobile et al., 1996), whereas ablomin, latisemin and triflin, whose crystal structure was recently determined (Shikamoto et al., 2005), block the depolarization of the smooth muscle membrane (Yamazaki and Morita, 2004). Within the second group of the CNG ion channel blockers, pseudechetoxin (PsTx) and pseudecin have been identified in the venom of the Australian king brown snake (*Pseudechis australis*) and the red-bellied black snake (*Pseudechis porphyriacus*), respectively (Brown et al., 1999; Yamazaki and Morita, 2002). Although these latter molecules are SCP/TAPS proteins, their biological properties and activities differ substantially from other homologues. Indeed, PsTx and pseudecin are basic proteins (pI ~10.0) that block olfactory and retinal α -subunit homotetrameric channels when applied to the exterior surface of membrane patches (Yamazaki and Morita, 2004). Studies have shown that SCP/TAPS proteins are widespread in the venoms from snakes (including those of the families Viperidae, Elapidae and Colubridae) from all five continents (Yamazaki et al., 2003); although the mechanisms of action of some of these proteins have been elucidated (reviewed by Yamazaki and Morita, 2004), the functions of most homologues are unknown. Another SCP/TAPS protein belonging to the second group has been isolated from the venom of the cone snail *Conus textile* and shown to act a substrate-specific protease in the processing of proconotoxins (Milne et al., 2003).

Members of the SCP/TAPS protein family have also been identified in vertebrate organisms, such as rat (*Rattus norvegicus*) (Cameo and Blaquier, 1976), mouse (*Mus musculus*) (Mizuki and Kasahara, 1992), rhesus monkey (*Macaca mulatta*) (Sivashanmugam et al., 1999) and horse (*Equus caballus*) (Giese et al., 2002). Due to their high level of expression in the epididymal lumen, they have been proposed to be associated with the binding of sperm to the oocyte and the fertilization process (reviewed by Jalkanen et al., 2005). In humans, the SCP/TAPS family also includes proteins which are expressed specifically in the testes and likely to be involved in sperm maturation (Kasahara et al., 1989). However, a human SCP/TAPS member (i.e. 'glioma pathogenesis-related protein'- GliPR, also designated as 'trypsin inhibitor'- P25TI and 'related to testes-specific, vespid and pathogenesis protein 1'- RTVP-1) has also been isolated from malignant glioma cells. Its high level of expression in glioblastoma multiforme/astrocytoma has been linked to the degree of malignancy of astrocytic tumours and to the regulation of the growth, survival and malignancy of the cancer (Murphy et al., 1995; Yamakawa et al., 1998; Rosenzweig et al., 2006). Another member (i.e. 'glioma pathogenesis related 2'- GLIPR-2, also known as 'Golgi associated PR-1 protein'- GAPR-1 and '*Homo sapiens* chromosome 9 open reading frame 19'- C9orf19) has been shown to be highly expressed in leukocytes, monocytes, lung, spleen and embryonic tissue (Eberle et al., 2002). The highest level of expression of GliPR-1 is in fibrotic kidney cells (Eberle et al., 2002; Eisenberg et al., 2002; Baxter et al., 2007); this latter molecule has been proposed to play a role in the development of fibrosis, due to its ability to stimulate (*via* the recruitment of activated fibroblasts) the transformation/transition of renal epithelial to mesenchymal cells *in vitro* (Baxter et al., 2007).

4. SCP/TAPS from helminths: 'activation- associated secreted proteins'

SCP/TAPS gene orthologues and/or protein homologues are common in parasitic helminths of the phylum Nematoda and class Trematoda. Within the Nematoda, SCP/TAPS have been described in the Orders Spirurida, Ascaridida, Tylenchida, Strongylida and Rhabditida (Skrjabin et al., 1991), which, based on an analysis of sequence data from the small subunit (SSU) of nuclear ribosomal DNA (rDNA) (Blaxter et al., 1998), have been proposed to correspond to clades III, IV and V. In most cases, SCP/TAPS molecules of parasitic helminths have been referred to commonly as ASPs.

4.1. Nematodes within clade III (including the Orders Spirurida and Ascaridida)

In nematodes of clade III, three C-type single-domain ASPs of *Onchocerca volvulus* (Spirurida; a filarioid parasite of humans) have been identified and named Ov-ASP-1, Ov-ASP-2 and Ov-ASP-3 (Lizotte-Waniewski et al., 2000; Tawe et al., 2000). Each of the three proteins shows a distinct transcription pattern

during the development of the parasite; specifically, the transcript *Ov-asp-2* was detected in all developmental stages, *Ov-asp-3* was specific to the third-stage larvae (L3s), whereas *Ov-asp-1* was highly transcribed in second-stage larvae (L2s) and the L3 and present (but at lower levels) in moulting L3s and females of adult *O. volvulus* (see Tawe et al., 2000). Subsequently, the protein *Ov-ASP-1* was studied further in L3s and shown to be localized exclusively to the granules of the glandular oesophagus (MacDonald et al., 2004). This protein is secreted *via* degranulation during development, following the invasion of the host. Due to its immunogenicity, recombinant *Ov-ASP-1* has been demonstrated to induce a partial but significant protection in mice against challenge infection with infective L3s of *O. volvulus* (see MacDonald et al., 2004). However, the presence of *Ov-ASP-1* could not be demonstrated in the supernatants from cultures of L3s, possibly because of a lack of a host-derived signal (MacDonald et al., 2004). Recombinant *Ov-ASP-1* and *Ov-ASP-2* have each been shown to induce an angiogenic response following their injection into cornea of naïve mice, suggesting that these proteins might contribute to corneal neovascularisation in onchocercal keratitis (Tawe et al., 2000). In another parasitic nematode of humans, *Brugia malayi*, only one SCP/TAPS molecule, designated as *Bm-VAL-1*, has been characterised to date (Murray et al., 2001). The transcription of *Bm-val-1* was assessed at several time points during the parasitic phase of the life cycle, revealing a uniform pattern. In contrast, the presence of the protein was shown to be restricted to the larval stages, thus raising the question as to whether *Bm-VAL-1* plays a role in the invasion of the host (Murray et al., 2001). Furthermore, incubation of mosquito-derived L3s (under a variety of conditions) did not yield sufficient *Bm-VAL-1* to be detected by immunochemical means; thus, it was proposed that the exposure to host-derived factors *in vivo* provides the signal required to stimulate the release of the ASP-like protein (Murray et al., 2001). More recently, a cDNA-oligonucleotide array analysis of transcription in adult males and females of *B. malayi* revealed that four EST clusters representing ASPs were “up-regulated” in the male (Li et al., 2005), suggesting a role in the reproductive physiology of the nematode. Also relating to nematodes within clade III, ASP orthologues were identified in an EST dataset representing the arrested larvae of the ascaridoid nematode *Toxocara canis*, the common roundworm of canids (Tetteh et al., 1999). The large insert sizes suggesting to the authors the presence of two double-domain SCP molecules (Tetteh et al., 1999).

4.2. Nematodes within clade IV (including the Order Tylenchida)

Within clade IV, two ASPs, *Meloidogyne* secretory protein (*Mi-MSP-1*) and a venom allergen-like protein (*Mi-VAP-2*), have been recorded for the root knot nematode of plants, *Meloidogyne incognita* (see Ding et al., 2000; Wang et al., 2007). Genes encoding both proteins showed similarly high transcription levels in pre-parasitic and early parasitic stages and low or no transcription in older juveniles and adult females. More specifically, the transcription of *Mi-msp-1* was high in the parasitic, second-stage juveniles (J2) (Ding et al., 2000). The authors hypothesized that this molecule might play a role in the establishment of the parasitic relationship between plant and nematode (Ding et al., 2000; Wang et al., 2007). In a Southern blot analysis, a *Mi-msp-1* probe hybridized to genomic DNA of two more *Meloidogyne* species, *M. arenaria* and *M. javanica*, but neither for *M. hapla* nor for *C. elegans* or the soy bean cyst nematode, *Heterodera glycines* (see Ding et al., 2000). In the latter species, two *vap* genes (namely *Hg-vap-1* and *Hg-vap-2*) were isolated and the corresponding mRNAs localized by *in situ* hybridization specifically to the subventral oesophageal gland cells of unhatched J2s (indicating the synthesis of *Hg-VAP-1* prior to hatching) as well as preparasitic and parasitic J2s. The hybridization signal for both *Hg-vap-1* and *Hg-vap-2* was weak within the gland cells in late parasitic stages, thus reflecting the degeneration of the subventral gland cells in *H. glycines* during parasitism and suggesting an involvement of these molecules in parasite-plant interactions (Gao et al., 2001).

4.3. Nematodes within clade V (including the Orders Strongylida and Rhabditida)

4.3.1. Superfamily Ancylostomatoidea (Order Strongylida)

In clade V, the ASPs have been studied relatively well in the dog hookworm, *A. caninum*. A double SCP-domain ASP, called *Ac-ASP-1*, was firstly identified as the major component of the ES products from serum-activated, infective L3s. The release of the molecule was proposed to be a response to a host-specific signal during the infection process (Hawdon et al., 1996). Subsequently, a C-type single SCP-domain ASP, called *Ac-ASP-2*, was identified as a second, major component of ES from serum-activated L3s (Hawdon et al., 1999) and localized predominantly to granules in the glandular oesophagus (Bethony et al., 2005),

suggesting that it was secreted from *A. caninum* as it invades the host (Bethony et al., 2005). A recent transcriptomic analysis of serum-activated L3s of *A. caninum* demonstrated that the most abundantly represented group of mRNAs associated with serum-activation was that of the *asps* (Datu et al., 2008). In adult *A. caninum*, a SCP/TAPS molecule, designated ‘neutrophil inhibitor factor’ (i.e. ‘NIF’) was first isolated from a cDNA library and shown to play an immunomodulatory role by blocking the adhesion of activated neutrophils to vascular endothelial cells and the release of H₂O₂ from activated neutrophils (Moyle et al., 1994). More recently, two other SCP/TAPS molecules with significant sequence similarity to NIF and Ac-ASP-2 were chromatographically purified and shown to inhibit the aggregation and adhesion of platelets by blocking the function of integrin receptors located on the cell surface (Del Valle et al., 2003). Subsequently, both NIF and platelet inhibitors were shown to be transcribed abundantly in the intestines of both *A. caninum* and *N. americanus* (Ranjit et al., 2006). Four additional ASPs were then identified in adult *A. caninum* and named Ac-ASP-3, Ac-ASP-4, Ac-ASP-5 and Ac-ASP-6 (Zhan et al., 2003). Ac-ASP-3 was identified as a C-type SCP-single domain in ES products, whereas Ac-ASP-4, Ac-ASP-5 and Ac-ASP-6 are double SCP-domain ASPs, originally identified by immunoscreening an *A. caninum* cDNA library using antibodies raised specifically against ES from the adult stage (Zhan et al., 2003). Ac-ASP-3 has been localized to the glands of the oesophagus, Ac-ASP-4 to the cuticular surface, Ac-ASP-5 to the brush border membrane of the posterior end of the gut, and Ac-ASP-6 to the cephalic and excretory glands of the adult stage of *A. caninum*. A recent proteomic analysis of the ES product of adult *A. caninum* (see Mulvenna et al., 2008) has revealed that 28% of ES products comprise ASP proteins. To date, none of the ASPs present in the adult stage of *A. caninum* have been detected by the immunochemical analysis of ES products from serum-activated L3s, although transcripts of *Ac-asp-3* and *Ac-asp-4* mRNAs have been detected in this latter stage (Zhan et al., 2003). While the functions of these molecules in the adult worm are largely unknown, a key involvement in the initiation, establishment and/or maintenance of the host-parasite relationship has been hypothesized (Zhan et al., 2003; Datu et al., 2008; Mulvenna et al., 2008).

Ac-ASP-1 or/and Ac-ASP-2 homologues have been isolated from *Ancylostoma ceylanicum* (Ay-ASP-1 and Ay-ASP-2, respectively), *Ancylostoma duodenale* (Ad-ASP-1) and *Necator americanus* (Na-ASP-1 and Na-ASP-2, respectively) via the screening of (homologous) L3 cDNA libraries using radiolabelled *Ac-asp* probes (Bin et al., 1999; Goud et al., 2004, 2005). Although the ASP-1 homologues are >85% identical at the amino acid level, the sequence of Ac-ASP-1 from *A. caninum* was more similar to the Na-ASP-1 from *N. americanus* than to that of the congeneric species *A. duodenale* (cf. Bin et al., 1999). In addition, the presence of ASPs has also been shown by hybridization in the dog and cat hookworms (i.e. *Ancylostoma tubaeforme*, *Ancylostoma braziliense* and *Uncinaria stenocephala*) (see Hawdon et al., 1999). In *N. americanus*, results from crystallography (Asojo et al., 2005), combined with the observation that Na-ASP-2 induces neutrophil and monocyte migration (Bower et al., 2008), suggest that this molecule has a role as an antagonistic ligand of complement receptor 3 (CR3) and alters the immune cascade by preventing the binding of chemotaxin (see Asojo et al., 2005).

4.3.2 Superfamily Trichostrongyloidea (Order Strongylida)

In adult *Haemonchus contortus* (the ‘barber’s pole worm’) of small ruminants, two ASPs have been identified to date, namely Hc24 (a C-type single SCP-domain) and Hc40 (double SCP-domain) (Schallig et al., 1997a; Rehman and Jasmer, 1998). Hc24 was first described by Takats et al. (1995) as one of the two low molecular weight oesophageal antigens recognized by sera from sheep hyperimmunized (by infection) with L3s of *H. contortus*. Subsequently, the same protein was identified in ES products from *H. contortus* adults (Schallig et al., 1997a). In contrast to hookworms, the transcript representing Hc24 was shown to be present in L4s and adult worms but not in eggs or L3s (Schallig et al., 1997b). Hc40 was detected via the immunoscreening of an adult worm cDNA library using polyclonal antiserum specifically against multiple membrane- and/or secreted proteins from the gut of adult *H. contortus*. However, interestingly, the mRNA encoding Hc40 was preferentially detected in tissues external to the gut (Rehman and Jasmer, 1998). Although only these two molecules have been isolated and characterized to date, an analysis of *Haemonchus* ES proteins revealed 9 distinct spots representing Hc24 or related molecules, which showed variation in relative immunogenicity in sheep (Yatsuda et al., 2003). A subsequent appraisal of the EST database available for *H. contortus* revealed that a range of expressed sequence tags (ESTs), encoding *asps* (n = 36; cf. www.nematode.net) and differing in their predicted amino acid sequence by 5-13%, could be assigned to 22 distinct clusters (Yatsuda et al., 2003).

In adult *Ostertagia ostertagi* (the ‘brown stomach worm’), two N-type single SCP-domain ASP homologues, designated *Oo*-ASP-1 and *Oo*-ASP-2, were characterized as major antigens in a protective ES-thiol fraction of a partially purified ES products. The highest levels of transcription and expression of these molecules were detected in L4s and adults of *O. ostertagi*, particularly in ES products from adults. These proteins could be localized to the reproductive tract of the adult female stage (Geldhof et al., 2003). In the closely related nematode *Teladorsagia circumcincta*, Nisbet et al. (2008) recently characterized 93 ASP-encoding ESTs which were uniquely represented in L4s but not in exsheathed L3s. These ESTs from the L4 stage formed 40 contigs and encoded three different types of ASPs. In the EST dataset representing the L3 stage, only two *asp* contigs (comprising six ESTs) were identified. A proteomic analysis of ES products harvested from *T. circumcincta* L4s (five days after inoculation of sheep with L3s) also identified a number of ASPs (Smith et al., 2008). Small numbers of ESTs (n = 2-3) encoding ASPs have also been identified in *Trichostrongylus vitrinus* (black scour worm of small ruminants) and *Oesophagostomum dentatum* (the nodule worm of the large intestines of pigs) (Nisbet and Gasser, 2004; Cottee et al., 2006), and were demonstrated to be male-enriched. Also, an analysis of PCR products from genomic DNA revealed substantial sequence variation within an individual of *Cooperia punctata* (see Yatsuda et al., 2002). Indeed, two double SCP-domain ASPs were identified and named *Cp*-ASP-1a and *Cp*-ASP-1b, in addition to seven different single SCP-domain ASPs, of which three showed higher homology in the N-terminus and four in the C-terminus of the identified double domain molecules, respectively (Yatsuda et al., 2002).

4.3.3. The free-living nematode *C. elegans* (Order Rhabditida)

At least 17 genes encoding SCP/TAPS homologues have been characterized in *C. elegans* (see Zhan et al., 2003). Of these, *vap-1* encodes a unique double SCP-domain protein which is predicted to be secreted by the adult stage of this nematode. However, information on its localization prior to the proposed secretion/excretion and its function is currently lacking. Other orthologues include *scl* genes encoding single SCP-domain proteins. Three *scl* genes (i.e. *scl-2*, *scl-20* and *scl-27*) show high transcription levels in *C. elegans* in response to individual infections with multiple bacterial pathogens (O’Rourke et al., 2006; Troemel et al., 2006; Wong et al., 2007). However, a specific function has been assessed only for *scl-2*, which is likely to act as a putative anti-microbial effector in response to infection by *Microbacterium nematophilum* (see O’Rourke et al., 2006). Transcription for *scl-2*, *scl-20* and *scl-27* has been observed to be increased in response to infection by more than one pathogen (O’Rourke et al., 2006; Troemel et al., 2006; Wong et al., 2007). Other *scl* genes have been proposed to be involved in the regulation of longevity and stress resistance (i.e. *scl-1*; Ookuma et al., 2003), ageing or stress resistance (i.e. *scl-1* to *scl-15*; Patterson, 2003) and normal fat storage (i.e. *scl-9*; Ashrafi et al., 2003). In a study of *C. elegans*, aimed at investigating the possible role of *gld-1* (for ‘defective in germline development’) in the cell proliferation in a germline cancer, the genes *scl-19* (renamed as *dct-2*, DAF-16/FOXO-controlled tumor gene) and *scl-20* (renamed as *dct-4*) were inferred to suppress the tumor growth promoted by *gld-1* (Pinkston-Gosse and Kenyon, 2007). The authors suggested a possible biological function of these molecules, based on the sequence similarity with the human GliPR proteins (see section 3); *scl-19* promoted apoptosis, and *scl-20* was inferred to play a role in signalling and in the innate immune response (Pinkston-Gosse and Kenyon, 2007). Another SCP/TAPS gene orthologue (gene code F48E8.1), designated *lon-1* (for *long* body), was characterized in long-bodied *C. elegans* and demonstrated to regulate body morphogenesis (Maduzia et al., 2002). Selected mutations in the *lon-1* gene resulted in worms that were 25% longer than those of the wild-type (N2) strain, whereas an over-expression resulted in a small-bodied (‘Lon’) phenotype (Maduzia et al., 2002).

4.4. Trematodes (i.e. Subclass Digenea)

SCP/TAPS proteins have been studied comprehensively in *S. mansoni*, a dioecious trematode infecting humans (Chalmers et al., 2008). Thirteen SCP/TAPS proteins (SmVAL1 to SmVAL13) have been isolated and characterized from this parasite, of which 12 display a single SCP-domain and one a double SCP-domain. However, an analysis of genomic sequence data for *S. mansoni* has predicted the existence of 15 other SmVAL proteins (<http://www.genedb.org/genedb/smansoni/>), and transcription has been assessed for 11 of them (Chalmers et al., 2008). Transcriptional analysis of SmVAL1 to SmVAL13 in different developmental stages of *S. mansoni* revealed significant variation in profiles, with some proteins exhibiting higher transcription in the invasive stage of the parasite. While roles in regulating protease activity, protease

inhibition and/or a modulation of the immune response during the invasion of the vertebrate host has been postulated (Chalmers et al., 2008), they remain to be investigated in detail.

5. Phylogenetic analysis of eukaryote SCP/TAPS: a proposed classification system

To investigate the phylogenetic relationships of SCP/TAPS proteins, 72 full-length amino acid sequences representing 35 species of eukaryote were obtained from the GenBank (<http://www.ncbi.nlm.nih.gov/>) and EMBL-EBI (<http://www.ebi.ac.uk/>) databases (cf. Tables 1 and 2) and subjected to InterProScan analysis (<http://www.ebi.ac.uk/interpro>) to identify single (i.e. $n = 60$, cf. Table 1) and double SCP-domains (i.e. $n = 12$, cf. Table 2). For each domain-type, sequences were aligned separately using BioEdit (Hall, 1999) and adjusted manually. Sequence alignments (560 and 561 amino acids for single and double-SCP domain sequences respectively) are available from primary author upon request. Amino acid sequence variation determined based on pairwise comparisons (p -distance, Chilton et al., 1995) using MEGA v4.0 software (Tamura et al., 2007) ranged from 0.034 to 0.909 (mean: 0.649) for single- and from 0.025 to 0.759 (mean: 0.606) for double-SCP domain sequences, respectively.

The data for each of the two domains were subjected to analysis by Bayesian inference (BI) using the program MrBayes v.3.1.2 (Ronquist and Huelsenbeck, 2003). The latter program incorporates a “codon-position model”, which is considered the most appropriate substitution model for the phylogenetic analysis of protein sequence datasets (Shapiro et al., 2006). Each analysis was conducted for 1,000,000 generations (ngen = 1,000,000), with every 100-th tree being saved. Each analysis used the following parameters: rates = gamma, aamodelpr = mixed, and the prior parameters left at the default settings. Tree and branch lengths were summarized using the parameter ‘sumt burnin = 1000’; a consensus tree was constructed, with ‘contype = halfcompat’ nodal support determined using consensus posterior probabilities and displayed employing the program TreeView v.1.6.6 (Page, 2002). A nodal support of ≥ 0.70 was considered significant. This procedure was applied to the set of holotype sequences (see Tables 1 and 2) to produce the phylogenetic trees shown in Figs. 1 and 2.

For single SCP-domain SCP/TAPS proteins, the most likely tree was generated using the protein PR-5 of *Or. sativa* as the outgroup (Fig. 1). The analysis identified one main clade comprising SCP/TAPS proteins from animals (nodal support: 0.96) to the exclusion of clades formed by homologues from plants (nodal supports: 0.68 and 0.91, cf. Fig. 1). Within the former clade, 3 sub-clades representing SCP/TAPS sequences from nematodes, snake venom, *S. mansoni* and arthropods were formed. Nodal support for each of these clades ranged from 0.73 to 0.88 (see Fig. 1). The proteins from nematodes (C-type and N-type) grouped to the exclusion of snake venom, supporting the existence of a monophyletic group of SCP/TAPS for parasitic nematodes. Conversely, a distinct group of SCP/TAPS proteins from trematodes and humans was not supported as monophyletic, based on the analysis. Indeed, two SCP/TAPS proteins from *S. mansoni* (i.e. SmVAL6 and SmVAL13) formed a separate cluster with GAPR1 from *H. sapiens* and G2 from *D. melanogaster*. A recent phylogenetic analysis of the SCP/TAPS protein family for *S. mansoni* identified two groups, with both SmVAL6 and SmVAL13 grouping to the exclusion of the remaining family members; this division could be attributed to a lack of predicted signal peptides in both of these molecules (Chalmers et al., 2008). Similarly, two separate groups of genes encoded the SCP/TAPS proteins in *D. melanogaster*, with proteins G2 and G3 representing each of the two clades (Kovalick and Griffin, 2005). In the present analysis, the clade of proteins from nematodes included two distinct sub-clades of SCP/TAPS proteins representing strongylids of animals (nodal support: 0.73) and those from filarioids and plant-parasitic nematodes (nodal support: 1.00), respectively (see Fig. 1). The paraphyletic relationship of the SCP/TAPS proteins from animals with PR-1 proteins from plants suggests a single-gene inheritance of SCP/TAPS within the animal and plant kingdoms, in which a mutation in a single ancestor gene has followed a predictable pattern of inheritance in each kingdom. This mode of inheritance has also been proposed for PR-5 proteins from plants (Shatters et al., 2006).

For double SCP-domain SCP/TAPS proteins, the most likely tree was generated using sequence data available for organisms other than nematodes as the outgroup (i.e. SmVAL11, cf. Table 2 and Fig. 2). The analysis identified a monophyletic clade (nodal support: 0.94) that included all but one sequence (i.e. Ac-ASP-4). No distinct separation between SCP/TAPS from hookworms and those from other strongylid nematodes was supported.

Phylogenetic approaches, such as that used herein, are commonly utilized to construct frameworks for the classification of gene and protein families for a range of prokaryote and eukaryote organisms (cf. Nelson et al., 1996; Crickmore et al., 1998; Jehle et al., 2006). Although there is no consensus opinion as to

which is the best algorithm for phylogenetic analysis/reconstruction, the present approach (considering clades with strong nodal support) provides a foundation for an improved classification of nematode SCP/TAPS proteins. Together with the positioning of newly described nematode SCP/TAPS proteins within the current tree, it would also be possible to provide a standardized nomenclature. Proteins could be named with a code, such as ‘Nem’ (for nematode), followed by an abbreviation of the species name (the first letter of the genus name followed by the first two letters of the species name - *in italics*), followed by the wording ‘SCP/TAPS’ indicating the protein family and ‘1’ or ‘2’ for single- and double SCP-domain proteins, respectively. A lower case letter might be added to indicate the order of isolation. Following these criteria, as an example, the presently known secretory protein *Mi-MSP-1*, this single domain SCP/TAPS member, that was first isolated from the nematode *M. incognita* (see Ding et al., 2000), would be named Nem-*Min*-SCP/TAPS-1a. Using this method, each SCP/TAPS protein could be positioned in the phylogenetic tree (with statistical support for its position) and then assigned a name containing all of the relevant information required.

6. Concluding remarks and biotechnological implications

SCP/TAPS proteins have been the focus of numerous studies aimed at understanding the mechanisms determining the interplay between hosts and pathogens. The vast number of SCP/TAPS proteins that have been identified to date, and described from a range of plant and animal species, exhibit variation in sequence, structure and/or biochemistry, which might reflect a diversification of function. However, the presence of SCP/TAPS proteins across a broad range of taxa suggests key roles. Unless a clear classification based on phenetic and genetic characters for these molecules is provided, the opportunity to explore the function of those molecules in the network of the host-pathogen interaction mechanisms is somewhat compromised. In the present review, we examined in detail the relationships of the SCP/TAPS proteins of parasitic nematodes (by comparing amino acid sequence data for homologues from other eukaryotic organisms), using a phylogenetic approach based on an objective mathematical/statistical model/algorithm. The sequence data and molecular phylogenetic analysis can be transformed into quantifiable parameters, such as degrees of sequence identity, genetic distances and nodal support for groupings, which might allow the definition of criteria for the demarcation of groups. The present classification system does not take into account the localization of a molecule, participation in biochemical pathways or functional role/s assigned to a small number of SCP/TAPS proteins. Indeed, it is well accepted that functional classification should be distinct and independent from the classification based on structure, as the first should allow structurally and/or phylogenetically unrelated molecules to be included in the same functional category based on similarities in biological roles and other features (reviewed by Ouzounis et al., 2003). Moreover, a classification based on phylogenetic relationships can be generated automatically using software designed specifically to facilitate the grouping of molecules, unlike a functional classification which requires extensive experimentation in a defined biological system and can provide detailed insights into the cellular and molecular aspects of genes and their products.

Given the fundamental biological roles that SCP/TAPS proteins have been proposed to play in a range of eukaryotic systems (e.g., defence against pathogens, allergens and induction of cell proliferation), a clear understanding of the function of such molecules in each biological system has become a priority. The major advances in proteomic, genomic and bioinformatic technologies now provide unique opportunities for a ‘systems biological’ exploration of the structure, function and molecular interactions of these proteins. Such an integrated approach will be advantageous in elucidating the fundamental molecular biology and functional roles of these proteins. Such an approach should lead to important biotechnological outcomes, including the development of drugs or vaccines designed to disrupt or interrupt key biological pathways or processes associated with SCP/TAPS proteins. For example, current evidence indicates that a range of SCP/TAPS molecules are transcribed/expressed by hookworms (as larvae) during their invasion of the host (i.e. transition from free-living to parasitic stage) and as adults in the intestine of the host (see section 4.3.1). The likely involvement of these molecules at the parasite-host interface might make some of them attractive as vaccine candidates. Currently, the protein *Na*-ASP-2 from the human hookworm *N. americanus* is being evaluated in clinical trials aimed at assessing its safety and tolerability as a vaccine against necatoriasis in humans (Bethony et al., 2008). Elucidating the molecular function of this molecule (as well as of other SCP/TAPS members of other parasitic nematodes), together with additional structural information on the proteins, could assist in the identification of key ‘structural antigens’; such small protein domains could then be synthesized chemically and to be used as a “conformational epitope” vaccine (cf. Corradin et al., 2007).

The identification of structural antigenic or immunogenic epitopes, which are conserved among multiple hookworm species, has the potential to lead to the development of a “multivalent” vaccine against hookworm disease in humans.

Acknowledgements

The research was supported by the Australian Research Council (ARC) (DP0665230; RBG and AL), the Australian Academy of Science and the Australian-American Fulbright Commission (RBG). JBM and AJN are funded by the Scottish Government RERAD, CC is the grateful recipient of an International Postgraduate Research Scholarship (IPRS) from the Australian Government *via* The University of Melbourne.

References

- Agrawal GK, Jwa NS, Rakwal R. A novel rice (*Oryza sativa* L.) acidic PR1 gene highly responsive to cut, phytohormones, and protein phosphatase inhibitors. *Biochem Biophys Res Commun* 2000;274:157-65.
- Anand SB, Gnanasekar M, Thangadurai M, Prabhu PR, Kaliraj P, Ramaswamy K. Immune response studies with *Wuchereria bancrofti* vespinal allergen homologue (WbVAH) in human lymphatic filariasis. *Parasitol Res* 2007;101:981-8.
- Ashrafi K, Chang FY, Watts JL, Fraser AG, Kamath RS, Ahringer J, Ruvkun G. Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature* 2003;421:268-72.
- 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;61:391-4.
- [Baxter RM](#), [Crowell TP](#), [George JA](#), [Getman ME](#), [Gardner H](#). The plant pathogenesis related protein GLIPR-2 is highly expressed in fibrotic kidney and promotes epithelial to mesenchymal transition in vitro. [Matrix Biol](#) 2007;26:20-9.
- Bethony J, Loukas A, Smout M, Brooker S, Mendez S, Plieskatt J, Goud G, Bottazzi ME, Zhan B, Wang Y, Williamson A, Lustigman S, Correa-Oliveira R, Xiao S, Hotez PJ. Antibodies against a secreted protein from hookworm larvae reduce the intensity of hookworm infection in humans and vaccinated laboratory animals. *FASEB J* 2005;19:1743-5.
- 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;26:2408-17.
- Bilò BM, Bonifazi F. Epidemiology of insect-venom anaphylaxis. *Curr Opin Allergy Clin Immunol* 2008;8:330-7.
- [Blaxter ML](#), [De Ley P](#), [Garey JR](#), [Liu LX](#), [Scheldeman P](#), [Vierstraete A](#), [Vanfleteren JR](#), [Mackey LY](#), [Dorris M](#), [Frisse LM](#), [Vida JT](#), [Thomas WK](#). A molecular evolutionary framework for the *phylum* Nematoda. *Nature* 1998;392:71-5.
- 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;118:569-75.
- Brown MP, Grundy WN, Lin D, Cristianini N, Sugnet CW, Furey TS, Ares M Jr, Haussler D. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc Natl Acad Sci U S A* 2000;97:262-7.
- Brown RL, Haley TL, West KA, Crabb JW. Pseudochetoxin: a peptide blocker of cyclic nucleotide-gated ion channels. *Proc Natl Acad Sci U S A* 1999;96:754-9.
- Calvo E, Pham VM, Lombardo F, Arcà B, Ribeiro JM. The sialotranscriptome of adult male *Anopheles gambiae* mosquitoes. *Insect Biochem Mol Biol* 2006;36:570-5.
- Cameo MS, Blaquier JA. Androgen-controlled specific proteins in rat epididymis. *J Endocrinol* 1976;69:47-55.
- Carr JP, Dixon DC, Klessig DF. Synthesis of pathogenesis-related proteins in tobacco is regulated at the level of mRNA accumulation and occurs on membrane-bound polysomes. *Proc Natl Acad Sci U S A* 1985;82:7999-8003.
- Carr JP, Dixon DC, Nikolau BJ, Voelkerding KV, Klessig DF. Synthesis and localization of pathogenesis-related proteins in tobacco. *Mol Cell Biol* 1987;7:1580-3.
- 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.
- Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. *Proc Natl Acad Sci U S A* 1999;96:15155-60.
- Chilton NB, Gasser RB, Beveridge I. Differences in a ribosomal DNA sequence of morphologically indistinguishable species within the *Hypodontus macropi* complex (Nematoda: Strongyloidea). *Int J Parasitol* 1995;25:647-51.

- Corradin G, Villard V, Kajava AV. Protein structure based strategies for antigen discovery and vaccine development against malaria and other pathogens. *Endocr Metab Immune Disord Drug Targets* 2007;7:259-65.
- Cottee PA, Nisbet AJ, Abs El-Osta YG, Webster TL, Gasser RB. Construction of gender-enriched cDNA archives for adult *Oesophagostomum dentatum* by suppressive-subtractive hybridization and a microarray analysis of expressed sequence tags. *Parasitology* 2006;132:691-708.
- Crickmore N, Zeigler DR, Feitelson J, Schnepf E, Van Rie J, Lereclus D, Baum J, Dean DH. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. *Microbiol Mol Biol Rev* 1998;62:807-13.
- Datta S K. Subbaratnam Muthukrishnan. Pathogenesis-related Proteins in Plants. CRC Press, 1999.
- 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.
- Del Valle A, Jones BF, Harrison LM, Chadderdon RC, Cappello M. Isolation and molecular cloning of a secreted hookworm platelet inhibitor from adult *Ancylostoma caninum*. *Mol Biochem Parasitol* 2003;129:167-77.
- De Wit PJGM. A light and scanning-electron microscopic study of infection of tomato plants by virulent and avirulent races of *Cladosporium fulvum*. *Neth J Plant Pathol* 1977;83:109-22.
- Ding X, Shields J, Allen R, Hussey RS. Molecular cloning and characterisation of a venom allergen AG5-like cDNA from *Meloidogyne incognita*. *Int J Parasitol* 2000;30:77-81.
- [Eberle HB](#), [Serrano RL](#), [Füllekrug J](#), [Schlosser A](#), [Lehmann WD](#), [Lottspeich F](#), [Kaloyanova D](#), [Wieland FT](#), [Helms JB](#). Identification and characterization of a novel human plant pathogenesis-related protein that localizes to lipid-enriched microdomains in the Golgi complex. *J Cell Sci* 2002;115:827-38.
- Eberspacher U, Roosterman D, Krätzschmar J, Haendler B, Habenicht UF, Becker A, Quensel C, Petri T, Schleuning WD, Donner P. Mouse androgen-dependent epididymal glycoprotein CRISP-1 (DE/AEG): isolation, biochemical characterization, and expression in recombinant form. *Mol Reprod Dev* 1995;42:157-72.
- Eisenberg I, Barash M, Kahan T, Mitrani-Rosenbaum S. Cloning and characterization of a human novel gene C9orf19 encoding a conserved putative protein with an SCP-like extracellular protein domain. *Gene* 2002;293:141-8.
- Fernández C, Szyperski T, Bruyère T, Ramage P, Mössinger E, Wütrich K. NMR solution structure of the pathogenesis-related protein p14a. *J Mol Biol* 1997;266:576-93.
- Finegold I. Issues in stinging insect allergy immunotherapy: a review. *Curr Opin Allergy Clin Immunol* 2008;8:343-7.
- Frank GR, Tripp CA, Grieve RB. Molecular cloning of a developmentally regulated protein isolated from excretory-secretory products of larval *Dirofilaria immitis*. *Mol Biochem Parasitol* 1996;75:231-40.
- [Gao B](#), [Allen R](#), [Maier T](#), [Davis EL](#), [Baum TJ](#), [Hussey RS](#). Molecular characterisation and expression of two venom allergen-like protein genes in *Heterodera glycines*. *Int J Parasitol* 2001;31:1617-25.
- Geldhof P, Meyvis Y, Vercruysse J, Claerebout E. Vaccine testing of a recombinant activation-associated secreted protein (ASP1) from *Ostertagia ostertagi*. *Parasite Immunol* 2008;30:57-60.
- Geldhof P, Vercauteren I, Gevaert K, Staes A, Knox DP, Vandekerckhove J, Vercruysse J, Claerebout E. Activation-associated secreted proteins are the most abundant antigens in a host protective fraction from *Ostertagia ostertagi*. *Mol Biochem Parasitol* 2003;128:111-4.
- Gianinazzi S, Martin C, Valée JC. Hypersensibilité aux virus, température et protéines solubles chez le *Nicotiana Xanthi* n.c. Apparition de nouvelles macromolécules lors de la répression de la synthèse virale. *C R Acad Sci Paris* 1970;270:2383-6.
- Giese A, Jude R, Kuiper H, Piumi F, Schambony A, Guérin G, Distl O, Töpfer-Petersen E, Leeb T. Molecular characterization of the equine AEG1 locus. *Gene* 2002;292:65-72.
- González J, Reyes F, Salas C, Santiag M, Codriansky Y, Coliheuque N, Silva H. *Arabidopsis thaliana*: a model host plant to study plant-pathogen interaction using Chilean field isolates of *Botrytis cinerea*. *Biol Res* 2006;39:221-8.
- Goud GN, Bottazzi ME, Zhan B, Mendez S, Deumic V, Plieskatt J, Liu S, Wang Y, Bueno L, Fujiwara R, Samuel A, Ahn SY, Solanki M, Asojo OA, Wang J, Bethony JM, Loukas A, Roy M, Hotez PJ. 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;23:4754-64.

- Goud GN, Zhan B, Ghosh K, Loukas A, Hawdon J, Dobardzic A, Deumic V, Liu S, Dobardzic R, Zook BC, Jin Q, Liu Y, Hoffman L, Chung-Debose S, Patel R, Mendez S, Hotez PJ. Cloning, yeast expression, isolation, and vaccine testing of recombinant *Ancylostoma*-secreted protein (ASP)-1 and ASP-2 from *Ancylostoma ceylanicum*. *J Infect Dis* 2004;189:919-29.
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 1999;41:95-8.
- Haynes SR, Cooper MT, Pye S, Stolow DT. Involvement of a tissue-specific RNA recognition motif protein in *Drosophila* spermatogenesis. *Mol Cell Biol* 1997;17:2708-15.
- 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;271:6672-8.
- 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;99:149-65.
- Henriksen A, [King TP](#), [Mirza O](#), [Monsalve RI](#), [Meno K](#), [Ipsen H](#), [Larsen JN](#), [Gajhede M](#), [Spangfort MD](#). Major venom allergen of yellow jackets, Ves v 5: structural characterization of a pathogenesis-related protein superfamily. [Proteins](#) 2001;45:438-48.
- Hoffman DR. Hymenoptera venom allergens. *Clin Rev Allergy Immunol* 2006;30:109-28.
- 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;5:527-34.
- Kauffman S, Legrand M, Geoffroy P, Fritig B. Biological function of pathogenesis-related proteins: Four PRproteins have beta-1, 3-glucanase activity. *EMBO J* 1987;6:3209-12.
- Kovalick GE, Griffin DL. Characterization of the SCP/TAPS gene family in *Drosophila melanogaster*. *Insect Biochem Mol Biol* 2005;35:825-35.
- 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;72:1268-74.
- [Jehle JA](#), [Blissard GW](#), [Bonning BC](#), [Cory JS](#), [Herniou EA](#), [Rohrmann GF](#), [Theilmann DA](#), [Thiem SM](#), [Vlak JM](#). On the classification and nomenclature of baculoviruses: a proposal for revision. [Arch Virol](#) 2006;151:1257-66.
- Legrand M, Kauffmann S, Geoffroy P, Fritig B. Biological function of pathogenesis-related proteins: Four tobacco pathogenesis-related proteins are chitinases. *Proc Natl Acad Sci U S A* 1987;84:6750-4.
- Li BW, Rush AC, Crosby SD, Warren WC, Williams SA, Mitreva M, Weil GJ. Profiling of gender-regulated gene transcripts in the filarial nematode *Brugia malayi* by cDNA oligonucleotide array analysis. *Mol Biochem Parasitol* 2005;143:49-57.
- Li S, Kwon J, Aksoy S. Characterization of genes expressed in the salivary glands of the tsetse fly, *Glossina morsitans morsitans*. *Insect Mol Biol* 2001;10:69-76.
- Linthorst HJM. Pathogenesis related proteins of plants. *Crit Rev Plant Sci* 1991;10:123.
- [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;68:3491-501.
- [Loukas A](#), [Bethony J](#), [Brooker S](#), [Hotez P](#). Hookworm vaccines: past, present, and future. [Lancet Infect Dis](#) 2006;6:733-41.
- Lu G, Villalba M, Coscia MR, Hoffman DR, King TP. Sequence analysis and antigenic cross-reactivity of a venom allergen, antigen 5, from hornets, wasps, and yellow jackets. *J Immunol* 1993;150:2823-30.
- 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.
- Maduzia LL, Gumienny TL, Zimmerman CM, Wang H, Shetgiri P, Krishna S, Roberts AF, Padgett RW. *lon-1* regulates *Caenorhabditis elegans* body size downstream of the *dbl-1* TGF beta signaling pathway. *Dev Biol* 2002;246:418-28.
- Mendez S, D' Samuel A, Antoine AD, Ahn S, Hotez P. Use of the air pouch model to investigate immune responses to a hookworm vaccine containing the Na-ASP-2 protein in rats. *Parasite Immunol* 2008;30:53-6.
- Metzler MC, Cutt JR, Klessig DF. Isolation and Characterization of a Gene Encoding a PR-1-Like Protein from *Arabidopsis thaliana*. *Plant Physiol* 1991;96:346-8.

- [Milne TJ](#), [Abbenante G](#), [Tyndall JD](#), [Halliday J](#), [Lewis RJ](#). Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. [J Biol Chem](#) 2003;278:31105-10.
- [Mizuki N](#), [Kasahara M](#). Mouse submandibular glands express an androgen-regulated transcript encoding an acidic epididymal glycoprotein-like molecule. [Mol Cell Endocrinol](#) 1992;89:25-32.
- Moser JM, Freitas T, Arasu P, Gibson G. Gene expression profiles associated with the transition to parasitism in *Ancylostoma caninum* larvae. [Mol Biochem Parasitol](#) 2005;143:39-48.
- [Moyle M](#), [Foster DL](#), [McGrath DE](#), [Brown SM](#), [Laroche Y](#), [De Meutter J](#), [Stanssens P](#), [Bogowitz CA](#), [Fried VA](#), [Ely JA](#), Soule HR, Vlasuk GP. A hookworm glycoprotein that inhibits neutrophil function is a ligand of the integrin CD11b/CD18. [J Biol Chem](#) 1994;269:10008-15.
- Mulvenna J, Hamilton B, Nagaraj S, Smyth D, Loukas A, Gorman J. Proteomic analysis of the excretory/secretory component of the blood-feeding stage of the hookworm, *Ancylostoma caninum*. [Mol Cell Proteomics](#) 2008; in press.
- Murphy EV, Zhang Y, Zhu W, Biggs J. The human glioma pathogenesis-related protein is structurally related to plant pathogenesis-related proteins and its gene is expressed specifically in brain tumors. [Gene](#) 1995;159:131-5.
- [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.
- Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC, Nebert DW. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. [Pharmacogenetics](#) 1996;6:1-42.
- Niderman T, Genetet I, Bruyère T, Gees R, Stintzi A, Legrand M, Fritig B, Möisinger E. Pathogenesis-related PR-1 proteins are antifungal. Isolation and characterization of three 14-kilodalton proteins of tomato and of a basic PR-1 of tobacco with inhibitory activity against *Phytophthora infestans*. [Plant Physiol](#) 1995;108:17-27.
- Nisbet AJ, Gasser RB. Profiling of gender-specific gene expression for *Trichostrongylus vitrinus* (Nematoda: Strongylida) by microarray analysis of expressed sequence tag libraries constructed by suppressive-subtractive hybridisation. [Int J Parasitol](#) 2004;34:633-43.
- Nisbet AJ, Redmond DL, Matthews JB, Watkins C, Yaga R, Jones JT, Nath M, Knox DP. Stage-specific gene expression in *Teladorsagia circumcincta* (Nematoda: Strongylida) infective larvae and early parasitic stages. [Int J Parasitol](#) 2008;38:829-38.
- 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.
- [Ookuma S](#), [Fukuda M](#), [Nishida E](#). Identification of a DAF-16 transcriptional target gene, *scl-1*, that regulates longevity and stress resistance in *Caenorhabditis elegans*. [Curr Biol](#) 2003;13:427-31.
- O'Rourke D, Baban D, Demidova M, Mott R, Hodgkin J. Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. [Genome Res](#) 2006;16:1005-16.
- Ouzounis CA, Coulson RM, Enright AJ, Kunin V, Pereira-Leal JB. Classification schemes for protein structure and function. [Nat Rev Genet](#) 2003;4:508-19.
- Page RD. Visualizing phylogenetic trees using TreeView. [Curr Protoc Bioinformatics](#) 2002;6: 6.2
- Parent JG, Asselin A. Detection of pathogenesis related proteins (PR or b) and of other proteins in the intercellular fluid of hypersensitive plants infected with tobacco mosaic virus. [Can J Bot](#) 1984;62:564-9.
- Patterson GI. Aging: new targets, new functions. [Curr Biol](#) 2003;13:R279-81.
- [Pinkston-Gosse J](#), [Kenyon C](#). DAF-16/FOXO targets genes that regulate tumor growth in *Caenorhabditis elegans*. DAF-16/FOXO targets genes that regulate tumor growth in *Caenorhabditis elegans*. [Nat Genet](#) 2007;39:1403-9.
- 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;36:701-10.
- Rauscher M, Ádám AL, Wirtz S, Guggenheim R, Mendgen K, Deising HB. PR-1 protein inhibits the differentiation of rust infection hyphae in leaves of acquired resistant broad bean. [Plant J](#) 1999;19: 625-33.

- Rehman A, Jasmer DP. A tissue specific approach for analysis of membrane and secreted protein antigens from *Haemonchus contortus* gut and its application to diverse nematode species. *Mol Biochem Parasitol* 1998;97:55-68.
- Ronquist F, Huelsenbeck J. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 2003;19:1572-4.
- Rosenzweig T, [Ziv-Av A](#), [Xiang C](#), [Lu W](#), [Cazacu S](#), [Taler D](#), [Miller CG](#), [Reich R](#), [Shoshan Y](#), [Anikster Y](#), [Kazimirsky G](#), [Sarid R](#), [Brodie C](#). Related to testes-specific, vespid, and pathogenesis protein-1 (RTVP-1) is overexpressed in gliomas and regulates the growth, survival, and invasion of glioma cells. *Cancer Res* 2006;66:4139-48.
- Saverwyns H, Visser A, Nisbet AJ, Peelaers I, Gevaert K, Vercruysse J, Claerebout E, Geldhof P. Identification and characterization of a novel specific secreted protein family for selected members of the subfamily Ostertagiinae (Nematoda). *Parasitology* 2008;135:63-70.
- Schallig HD, van Leeuwen MA, Cornelissen AW. Protective immunity induced by vaccination with two *Haemonchus contortus* excretory secretory proteins in sheep. *Parasite Immunol* 1997;19:447-53.
- Schallig HD, van Leeuwen MA, Verstrepen BE, Cornelissen AW. Molecular characterization and expression of two putative protective excretory secretory proteins of *Haemonchus contortus*. *Mol Biochem Parasitol* 1997;88:203-13.
- Shapiro B, Rambaut A, Drummond AJ. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol Biol Evol* 2006;23:7-9.
- [Shatters RG Jr](#), [Boykin LM](#), [Lapointe SL](#), [Hunter WB](#), [Weathersbee AA 3rd](#). Phylogenetic and structural relationships of the PR5 gene family reveal an ancient multigene family conserved in plants and select animal taxa. *J Mol Evol* 2006;63:12-29.
- 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;350:735-43.
- Schmid-Hempel P. Parasite immune evasion: a momentous molecular war. *Trends Ecol Evol* 2008;23:318-26.
- Skrjabin KI, Shikhobalova NP, Schulz RS, Popova TI, Boev SN, Delyamure SL. (1991). Strongylata. In: Skrjabin KI, editor. Keys to Parasitic Nematodes. [English Translation of Opredelitel' Paraziticheskikh Nematod. Izdatel'stvo Akademii Nauk SSSR, Moscow]. Vol. 3, Amerind Publishing Co. Pvt. Ltd., pp. 230-247 (1991).
- Sivashanmugam P, Richardson RT, Hall S, Hamil KG, French FS, O'Rand MG. Cloning and characterization of an androgen-dependent acidic epididymal glycoprotein/CRISP1-like protein from the monkey. *J Androl* 1999;20:384-93.
- [Yamakawa T](#), [Miyata S](#), [Ogawa N](#), [Koshikawa N](#), [Yasumitsu H](#), [Kanamori T](#), [Miyazaki K](#). cDNA cloning of a novel trypsin inhibitor with similarity to pathogenesis-related proteins, and its frequent expression in human brain cancer cells. *Biochim Biophys Acta* 1998;1395:202-8.
- Yamazaki Y, Brown RL, Morita T. Purification and cloning of toxins from elapid venoms that target cyclic nucleotide-gated ion channels. *Biochemistry* 2002;41:11331-7.
- Yamazaki Y, Hyodo F, Morita T. Wide distribution of cysteine-rich secretory proteins in snake venoms: isolation and cloning of novel snake venom cysteine-rich secretory proteins. *Arch Biochem Biophys* 2003;412:133-41.
- Yamazaki Y, Morita T. Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* 2004;44\227-31.
- Yatsuda AP, Eysker M, Vieira-Bressan MC, De Vries E. A family of activation associated secreted protein (ASP) homologues of *Cooperia punctata*. *Res Vet Sci* 2002;73:297-306.
- Yatsuda AP, Krijgsveld J, Cornelissen AW, Heck AJ, de Vries E. Comprehensive analysis of the secreted proteins of the parasite *Haemonchus contortus* reveals extensive sequence variation and differential immune recognition. *J Biol Chem* 2003;278:16941-51.
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007;24:1596-9.
- Tetteh KK, Loukas A, Tripp C, Maizels RM. Identification of abundantly expressed novel and conserved genes from the infective larval stage of *Toxocara canis* by an expressed sequence tag strategy. *Infect Immun* 1999;67:4771-9.
- 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;109:91-9.

- Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, Kim DH. p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. PLoS Genet 2006;2:e183.
- Valenzuela JG, Garfield M, Rowton ED, Pham VM. Identification of the most abundant secreted proteins from the salivary glands of the sand fly *Lutzomyia longipalpis*, vector of *Leishmania chagasi*. J Exp Biol 2004;207:3717-29.
- van Loon LC. The nomenclature of pathogenesis related proteins. Physiol Mol Plant Pathol 1990;37:229.
- van Loon LC, Pierpoint WS, Boller T, Conejero V. Recommendations for naming plant pathogenesis related proteins. Plant Mol Biol Rep 1994;12:245-64.
- van Loon LC, Rep M, Pieterse CM. Significance of inducible defense-related proteins in infected plants. Annu Rev Phytopathol 2006;44:135-62.
- van Loon LC, Van Strien EA The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol Mol Plant Pathol 1999;55:85-97.
- Vermeire JJ, Cho Y, Lolis E, Bucala R, Cappello M. Orthologs of macrophage migration inhibitory factor from parasitic nematodes. Trends Parasitol 2008;24:355-63.
- Visser A, Van Zeveren AM, Meyvis Y, Peelaers I, Van den Broeck W, Gevaert K, Vercruysse J, Claerebout E, Geldhof P. Gender-enriched transcription of activation associated secreted proteins in *Ostertagia ostertagi*. Int J Parasitol 2008;38:455-65.
- Wang J, Kim SK. Global analysis of dauer gene expression in *Caenorhabditis elegans*. Development 2003;130:1621-34.
- Wang X, Li H, Hu Y, Fu P, Xu J. Molecular cloning and analysis of a new venom allergen-like protein gene from the root-knot nematode *Meloidogyne incognita*. Exp Parasitol 2007;117:133-40.
- Wong D, Bazopoulou D, Pujol N, Tavernarakis N, Ewbank JJ. Genome-wide investigation reveals pathogen-specific and shared signatures in the response of *Caenorhabditis elegans* to infection. Genome Biol 2007;8:R194.
- Xiao S, Zhan B, Xue J, Goud GN, Loukas A, Liu Y, Williamson A, Liu S, Deumic V, Hotez P. The evaluation of recombinant hookworm antigens as vaccines in hamsters (*Mesocricetus auratus*) challenged with human hookworm, *Necator americanus*. Exp Parasitol 2008;118:32-40.
- 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.

LEGENDS TO TABLES AND FIGURES:

Table 1

Full-length single SCP-domain SCP/TAPS protein sequences included in the phylogenetic analysis, listed according the main eukaryote taxa (i.e. plants, arthropods, mammals, parasitic and free-living helminths). Current nomenclature, sequence accession number and available information on function are also given.

Table 2

Full-length double SCP-domain SCP/TAPS protein sequences characterized for parasitic and free-living helminths and included in the phylogenetic analysis. Current nomenclature, sequence accession number and information on function (if available) are also given.

Fig. 1. Phylogenetic reconstruction of single SCP-domain SCP/TAPS protein sequence data using Bayesian inference. Accession numbers, linked to [numbers in the tree], are: NM_001056008 [1]; Q20609 [2]; AF020586 [3]; AF334661 [4]; AAD16985 [5]; AF001100 [6]; H39490 [7]; AA917267 [8]; AF013289 [9]; EF370396 [10]; AF374388 [11]; AY033601 [12]; P90958 [13]; AAC35986 [14]; AAP41951 [15]; AAP41953 [16]; AAP41952 [17]; AM747038 [18]; U64793 [19]; AY217004 [20]; AJ310819 [21]; AJ515523 [22]; X91911 [23]; U16307 [24]; AF384218 [25]; AF384219 [26]; AY093955 [27]; AF384220 [28]; AY994061 [29]; AY994062 [30]; AY994063 [31]; EF421456 [32]; DQ269978 [33]; DQ269980 [34]; DQ269979 [35]; DQ060001 [36]; DQ060000 [37]; NP650264 [38]; P35786 [39]; P35783 [40]; M98858 [41]; P35784 [42]; AF012919 [43]; Y17702 [44]; [AF132511](#) [45]; [AF466605](#) [46]; AF259957 [47]; BC126292 [48]; AJ491318 [49]; AY953433 [50]; DQ269977 [51]; AY039756 [52]; NP608663 [53]; AY050221 [54]; A22633 [55]; 815949 (Gene ID) [56]; U82200 [57]; U89895 [58]; X05959 [59]; X03465 [60]; X05454 [61].

Fig. 2. Phylogenetic reconstruction of double SCP-domain SCP/TAPS protein sequence data using Bayesian inference. Accession numbers, linked to [numbers in the tree], are: DQ151891 [1]; AY217005 [2]; Q19348 [3]; Q16937 [4]; AAD13340 [5]; AAD13339 [6]; AAN11402 [7]; AM747039 [8]; AAC03562 [9]; AF352702 [10]; AF352714 [11]; AY217007 [12]; AY217006 [13].

Table 1. Full-length single SCP-domain SCP/TAPS protein sequences included in the phylogenetic analysis, listed according the main eukaryote taxa (i.e. plants, arthropods, mammals, parasitic and free-living helminths). Current nomenclature, sequence accession number and available information on function are also given.

Organism	Protein name	Current nomenclature	Accession number	Inferred function	References
Plants					
<i>Arabidopsis thaliana</i>	Pathogenesis related protein 1	AT2G14610	815949 (Gene ID)	Defence against mycete infections	Gonzalez et al., 2006
<i>Lycopersicon esculentum</i>		SIP14	A22633	Defence against mycete infections	Niderman et al., 1995
<i>Nicotiana tabacum</i>		NtPR-1a	X05959	Defence against tobacco mosaic virus infections	Cornelissen et al., 1987
<i>Oryza sativa</i>		NtPR-1b	X03465		
<i>Solanum tuberosum</i>		NtPR-1c	X05454	Response to cutting, application of jasmonic acid or other phytohormones	Agrawal et al., 2000a Agrawal et al., 2000b
		OsPR-1a	U82200		
		OsPR-1b	U89895		
		StPR-1b	AY050221	Defence against mycete infections	Hoegen et al., 2002
Arthropods					
<i>Drosophila melanogaster</i>	SCP/TAPS	G2 G3	NP608663 NP650264	Male reproduction	Kovalick and Griffin, 2005
<i>Vespula vulgaris</i>	Venom allergen 5	Ves v 5	M98858	Venom allergen	Lu et al., 2001

<i>Vespula squamosa</i>	Venom allergen 5 homologue	VsAG5	P35786	Venom allergen	Hoffman, 1993
<i>Vespula flavopilosa</i>		VfAG5	P35783		
<i>Vespula germanica</i>		VgAG5	P35784		
<i>Anopheles gambiae</i>		gVaG	Y17702	Secretion of saliva	Arca' et al., 1999
<i>Aedes aegypti</i>		AaAG5	AF466605	Unknown	Valenzuela et al., 2002
<i>Lutzomyia longipalpis</i>		LuLoAG5	AF132511	Secretion of saliva	Charlab et al., 1999
<i>Glossina morsitans</i>		TAG5	AF259957	Unknown	Li et al., 2001
<i>Solenopsis invicta</i>	Venom allergen 3	Vesv3	AF012919	Venom allergen	Lu et al., 1993
Snakes					
<i>Agkistrodon blomhoffi</i>	Ablomin		AF384218	Smooth muscle contraction blockers	Yamazaki et al., 2002
<i>Laticauda semifasciata</i>	Latisemin		AF384220		
<i>Rhabdopsis tigrinus</i>	Tigrin		AY093955		
<i>Trimeresurus flavoviridis</i>	Triflin		AF384219		
Cone shells					
<i>Conus textile</i>	Venom allergen homologue	Tex31	AJ491318	Ca2+ channel blocker	Milne et al., 2003
Human					
<i>Homo sapiens</i>	Related to testes-specific, vespid and pathogenesis protein 1	RTVP-1	X91911	Regulation of the growth, survival and invasion of glioma cancer cells	Rosenzweig et al., 2006
	Trypsin inhibitor	P25TI	BC126292		Yamakawa et al., 1998
	Glioma pathogenesis-related protein	GliPR	U16307		Murphy et al., 1995
	Golgi associated PR-1 protein	GAPR1	AY039756	Regulation of development of fibrosis	Eberle et al., 2002 Eisenberg et al., 2002 Baxter et al., 2007
Parasitic helminths					
<i>Ancylostoma caninum</i>	<i>Ancylostoma caninum</i> secreted protein- 2	<i>Ac</i> -ASP-2	AAC35986	Transition from the free-living to the parasitic stage of activated third stage larvae	Hawdon et al., 1999

	<i>Ancylostoma caninum</i> secreted protein- 3	<i>Ac</i> -ASP-3	AY217004	Establishment or maintenance of host-parasite relationships	Zhan et al., 2003
<i>Necator americanus</i>	<i>Necator americanus</i> activation associated secreted protein- 2	<i>Na</i> -ASP-2	AAP41952	Unknown	Bin et al., 1999
<i>Ancylostoma duodenale</i>	<i>Ancylostoma duodenale</i> activation associated secreted protein- 2	<i>Ad</i> -ASP-2	AAP41951	Unknown	
<i>Ancylostoma ceylanicum</i>	<i>Ancylostoma ceylanicum</i> activation associated secreted protein- 2	<i>Ay</i> -ASP-2	AAP41953	Unknown	
<i>Ostertagia ostertagi</i>	<i>Ostertagia</i> secreted protein 1*	<i>Oo</i> -ASP-1	AJ310819	Unknown	Vercauteren et al., 2003
	<i>Ostertagia</i> secreted protein 2*	<i>Oo</i> -ASP-2	AJ515523	Unknown	Geldhof et al., 2003
	<i>Ostertagia</i> secreted protein 3	<i>Oo</i> -ASP-3	AM747038	Unknown	Visser et al., 2008
<i>Haemonchus contortus</i>	<i>Haemonchus</i> 24 kd protein	Hc24	U64793	Unknown	Shallig et al., 1997a
<i>Onchocerca volvulus</i>	<i>Onchocerca</i> secreted protein 1	<i>Ov</i> -ASP-1	AF020586	Role in corneal neovascularisation in onchocercal keratitis?	Tawe et al., 2000
	<i>Onchocerca</i> secreted protein 2	<i>Ov</i> -ASP-2	H39490		
	<i>Onchocerca</i> secreted protein 3	<i>Ov</i> -ASP-3	AA917267	Unknown	
<i>Brugia malayi</i>	<i>Brugia malayi</i> venom allergen homologue <i>asp</i> -like 1	<i>Bm</i> -VAL-1	AF334661	Invasion of the vertebrate host	Murray et al., 2001
<i>Wuchereria bancrofti</i>	<i>Wuchereria bancrofti</i> venom allergen homologue 1	<i>Wb</i> -VAH	AAD16985	Unknown	Balakrishnan et al., 2007
<i>Dirofilaria immitis</i>	<i>Dirofilaria immitis</i> venom allergen-like 1	<i>Di</i> -VAL-1	AF001100	Unknown	Frank et al., 1996

<i>Meloidogyne incognita</i>	<i>Meloidogyne incognita</i> - <i>Meloidogyne</i> secretory protein	Mi-MSP-1	AF013289	Establishment of the relationships with the plant host	Ding et al., 2000
<i>Meloidogyne incognita</i>	<i>Meloidogyne incognita</i> venom allergen protein 2	Mi-VAP-2	EF370396	Unknown	Wang et al., 2007
<i>Heterodera glycines</i>	<i>Heterodera glycines</i> venom allergen protein 1	Hg-VAP-1	AF374388	Establishment of the relationships with the plant host	Gao et al., 2001
	<i>Heterodera glycines</i> venom allergen protein 2	Hg-VAP-2	AY033601		
<i>Schistosoma mansoni</i>	<i>Schistosoma mansoni</i> venom allergen like 1	SmVAL1	AY994061	Protease activity, protease inhibition and/or modulation of the host immune response during the infection process	Chalmers et al., 2008
	<i>Schistosoma mansoni</i> venom allergen like 2	SmVAL2	AY994062		
	<i>Schistosoma mansoni</i> venom allergen like 3	SmVAL3	DQ060000		
	<i>Schistosoma mansoni</i> venom allergen like 4	SmVAL4	AY994063		
	<i>Schistosoma mansoni</i> venom allergen like 5	SmVAL5	DQ269980		
	<i>Schistosoma mansoni</i> venom allergen like 6	SmVAL6	AY953433		
	<i>Schistosoma mansoni</i> venom allergen like 7	SmVAL7	DQ060001		
	<i>Schistosoma mansoni</i> venom allergen like 9	SmVAL9	DQ269979		
	<i>Schistosoma mansoni</i> venom allergen like 10	SmVAL10	EF421456		
	<i>Schistosoma mansoni</i> venom allergen like 12	SmVAL12	DQ269978		
	<i>Schistosoma mansoni</i> venom allergen like 13	SmVAL13	DQ269977		

<i>Free-living helminths</i>					
<i>Caenorhabditis elegans</i>	<i>scp</i> -like 10 (F49E11.5)	SCL-10	Q20609	Ageing and stress-resistance	Patterson, 2003
	<i>scp</i> -like 22 (T05A10.5)	SCL-22	P90958	Resistance to pathogen infection	O'Rourke et al., 2006 Troemel et al., 2006 Wong et al., 2007

*N-type single SCP domain SCP/TAPS proteins.

Table 2. Full-length double SCP-domain SCP/TAPS protein sequences characterized for parasitic and free-living helminths and included in the phylogenetic analysis. Current nomenclature, sequence accession number and information on function (when available) are also given.

Species	Protein name	Current nomenclature	Accession number	Inferred function	References
<i>Ancylostoma caninum</i>	<i>Ancylostoma caninum</i> secreted protein- 2	<i>Ac</i> -ASP-1	Q16937	Response to host-specific signals during the infection process	Hawdon et al., 1996
	<i>Ancylostoma caninum</i> secreted protein- 4	<i>Ac</i> -ASP-4	AY217005	Establishment or maintenance of host-parasite relationships	Zhan et al., 2003
	<i>Ancylostoma caninum</i> secreted protein- 5	<i>Ac</i> -ASP-5	AY217006		
	<i>Ancylostoma caninum</i> secreted protein- 6	<i>Ac</i> -ASP-6	AY217007		
<i>Necator americanus</i>	<i>Necator americanus</i> <i>Ancylostoma</i> -secreted protein 1	<i>Na</i> -ASP-1	AAD13340	Unknown	Bin et al., 1999
<i>Ancylostoma duodenale</i>	<i>Ancylostoma duodenale</i> <i>Ancylostoma</i> -secreted protein 1	<i>Ad</i> -ASP-1	AAD13339	Unknown	
<i>Ancylostoma ceylanicum</i>	<i>Ancylostoma ceylanicum</i> <i>Ancylostoma</i> -secreted protein 1	<i>Ay</i> -ASP-1	AAN11402	Unknown	Goud et al., 2004
	<i>Ostertagia ostertagi</i> <i>Ancylostoma</i> secreted protein 4	<i>Oo</i> -ASP-4	AM747039	Unknown	
<i>Haemonchus contortus</i>	<i>Haemonchus</i> 40 kd protein	Hc40	AAC03562	Unknown	Shallig et al., 1997a
<i>Cooperia punctata</i>	<i>Cooperia punctata</i> <i>Ancylostoma</i> -secreted protein 1a	<i>Cp</i> -ASP-1a	AF352702	Unknown	Yatsuda et al., 2002
	<i>Cooperia punctata</i> <i>Ancylostoma</i> -secreted protein 1b	<i>Cp</i> -ASP-1b	AF352714	Unknown	
<i>Schistosoma mansoni</i>	<i>Schistosoma mansoni</i> venom allergen-like 11	SmVAL11	DQ151891	Protease activity, protease inhibition and/or modulation of the host immune response during the infection process	Chalmers et al., 2008

Caenorhabditis elegans

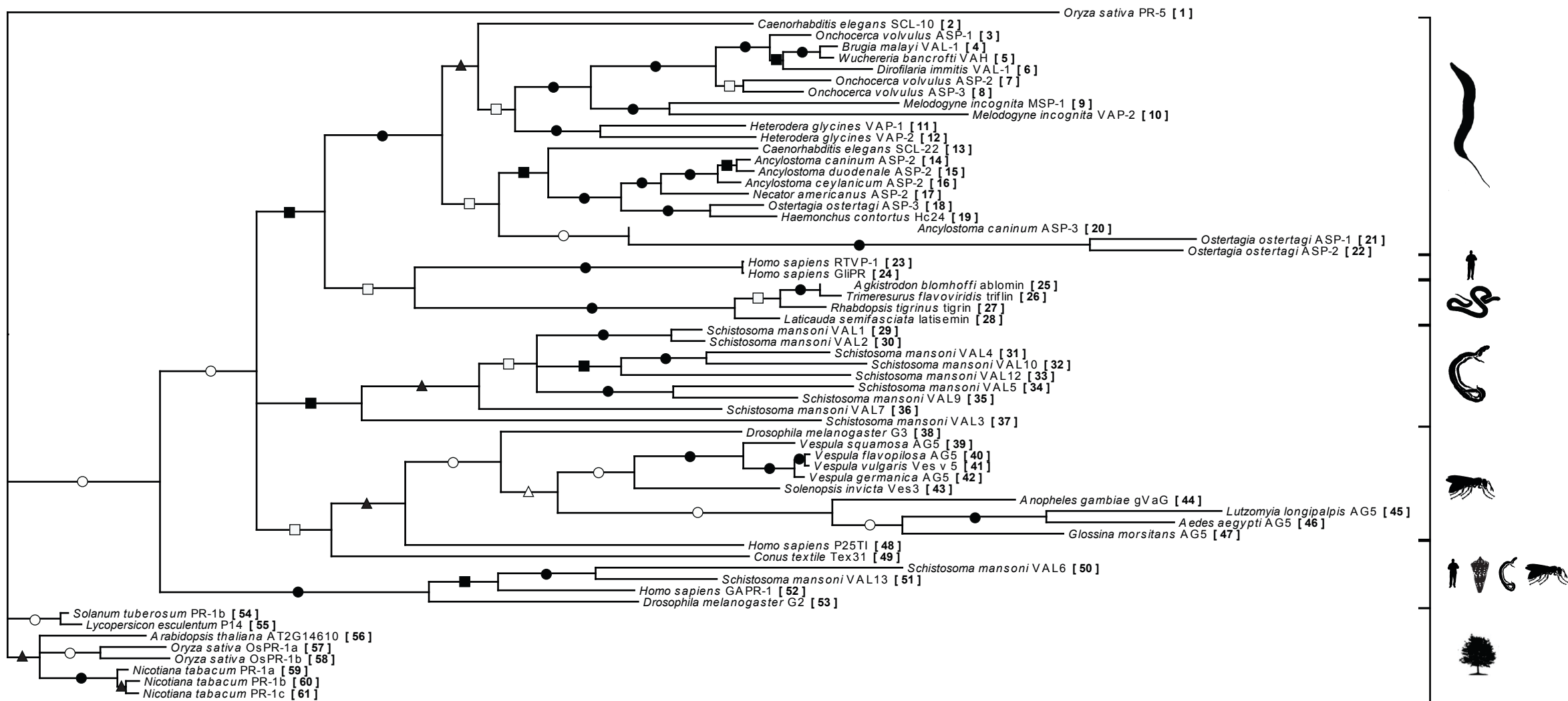
Venom allergen protein- 1

VAP-1

Q19348

Unknown

www.wormbase.org



0.1

● 1.00 ○ 0.90-0.99 ■ 0.80-0.89 □ 0.70-0.79 ▲ 0.60-0.69 △ 0.50-0.59

Nematodes
 Humans
 Snakes
 Trematodes
 Arthropods
 Cone Shell
 Plants

