

Identification and isolation of cDNA clones encoding the abundant secreted proteins in the saliva proteome of *Culicoides nubeculosus*

C. L. Russell*, K. J. Heesom†, C. J. Arthur‡, C. R. Helps*, P. S. Mellor§, M. J. Day*, S. Torsteinsdottir¶, T. S. Björnsdóttir¶ and A. D. Wilson*

*University of Bristol, School of Clinical Veterinary Sciences, Langford, UK; †University of Bristol, Department of Biochemistry, School of Medical Sciences and ‡University of Bristol, School of Chemistry, Bristol, UK; §Institute for Animal Health, Pirbright Laboratory, Pirbright, Surrey, UK; and ¶Institute for Experimental Pathology, University of Iceland Keldur, Reykjavik, Iceland

Abstract

Culicoides spp. are vectors of several infectious diseases of veterinary importance and a major cause of allergy in horses and other livestock. Their saliva contains a number of proteins which enable blood feeding, enhance disease transmission and act as allergens. We report the construction of a novel cDNA library from *Culicoides nubeculosus* linked to the analysis of abundant salivary gland proteins by mass spectrometry. Fifty-four novel proteins sequences are described including those of the enzymes maltase, hyaluronidase and two serine proteases demonstrated to be present in *Culicoides* salivary glands, as well as several members of the D7 family and protease inhibitors with putative anticoagulant activity. In addition, several families of abundant proteins with unknown function were identified including some of the major candidate allergens that cause insect bite hypersensitivity in horses.

Keywords: culicoides, blue tongue, allergy, salivary proteins.

Introduction

The blood of vertebrates presents a potentially rich source of nutrients that a range of invertebrates from annelid and nematode worms (Salzet, 2001; Harrison *et al.*, 2002) to numerous species of arthropods (Champagne, 2004, 2005) have evolved the ability to exploit. To overcome the physical barriers presented by epithelial and connective tissue or keratinized skin, as well as the coagulation pathways, and inflammatory and immune responses that protect the host, the secretions of blood feeding invertebrates contain a remarkably diverse array of pharmacologically active components. These include digestive enzymes that break down tissue barriers, vasodilators, inhibitors of platelet function, thrombolytic enzymes and inflammatory or immune modulators (Salzet, 2001; Harrison *et al.*, 2002; Champagne, 2004, 2005; Bishop *et al.*, 2006; Calvo *et al.*, 2006). Blood feeding behaviour in arthropods also provides an opportunity for the transmission of infectious agents. Typically this involves a complex series of host–parasite interactions in which components of insect saliva that modify the host's inflammatory and immune responses can have a significant effect on the outcome (Billingsley *et al.*, 2006; Schneider *et al.*, 2006, 2007; Titus *et al.*, 2006).

Insects in the genus *Culicoides* Latreille, 1809 are important vectors of animal parasites such as *Onchocerca* spp. (Dohnal *et al.*, 1990), or virus infections including bluetongue virus (BTV) and African horse sickness virus (AHSV) (Mellor *et al.*, 2000). There is mounting evidence that components in *Culicoides* saliva play a role in enhancing the transmission of these infections. For example, partial proteolytic degradation of BTV and AHSV greatly increases virus infectivity for adult *Culicoides* as well as for insect and mammalian cells in culture (Cowley & Gorman, 1990; Marchi *et al.*, 1995; Mertens *et al.*, 1996). *In vivo*, partial virus proteolysis could be a direct effect of insect salivary and gut proteases, or may result from indirect effects such as up-regulation of host proteases as a result of the pro-inflammatory actions of *Culicoides* saliva. A further mechanism through which salivary gland secretions of *Culicoides* influence disease transmission is by the

First published online 16 April 2009.

Correspondence: A. D. Wilson, University of Bristol, School of Clinical Veterinary Sciences, Langford BS40 5DU, UK. Tel.: +44 117 33 19067; e-mail: doug.wilson@bris.ac.uk

reactivation of latent virus infection. *Culicoides* bites are known to trigger viraemia from latent BTV infections in cattle (Akey *et al.*, 1985) and can also stimulate reactivation of latent BTV from bovine $\gamma\delta$ T-cells. Latent BTV infection of $\gamma\delta$ T-cells in cattle has been proposed as a mechanism of virus over-wintering (Takamatsu *et al.*, 2003), but the components of *Culicoides* saliva which mediate these effects have not been identified. A 60 kDa glycoprotein present in *Culicoides sonorensis* saliva has been shown to suppress the proliferation of murine lymphocytes in response to Concanavalin-A or lipopolysaccharide (LPS) and to reduce nitric oxide (NO) production by activated monocytes (Bishop *et al.*, 2006). Although these immunosuppressive effects could influence the success of virus transmission, the identity of this glycoprotein remains unknown.

Culicoides bites are also a major cause of allergic dermatitis in horses (McCraig, 1974; Braverman *et al.*, 1983; Halldorsdottir & Larsen, 1991; Anderson *et al.*, 1993). Hypersensitivity to *Culicoides* has also been reported in cattle and sheep (Connan & Lloyd, 1988; Yeruham *et al.*, 1993). The bites of *Culicoides* are notoriously irritant and although there is considerable anecdotal evidence of hypersensitivity to *Culicoides* in humans, published scientific evidence for allergy in humans is limited (Okuda *et al.*, 2002).

The importance of *Culicoides* as a cause of hypersensitivity and as a disease vector has stimulated research to identify the saliva components that may modulate the immune response leading to allergy or enhanced disease transmission. Early investigations on the functional components of *Culicoides* saliva identified an apyrase enzyme, a peptide with vasodilator activity and an inhibitor of factor Xa (Perez de Leon & Tabachnick, 1996; Perez de Leon *et al.*, 1997, 1998). Analysis of over 1000 sequences from a cDNA library of *C. sonorensis* salivary glands produced a catalogue of over 80 deduced proteins. (Campbell *et al.*, 2005), from which several putative abundant proteins with predicted functions in sugar-feeding, anticoagulation and defence have been identified. Two separate studies of the *Culicoides nubeculosus* saliva proteome have used a combination of 2D electrophoresis and mass spectrometry to analyse and identify the proteins either in secreted saliva collected from feeding insects (Langner *et al.*, 2007) or from dissected salivary glands (Wilson *et al.*, 2008). These papers compared the mass spectrometry and N-terminal sequence data of *C. nubeculosus* proteins with predicted proteins based on cDNA sequences from *C. sonorensis* (Campbell *et al.*, 2005); they confirmed the presence of abundant proteins with significant homology to maltase, members of the D7 family and several other abundant transcripts of unknown function. However, not all proteins could be identified by this method, a result attributed in part to low levels of sequence homology between the two *Culicoides* species (Wilson *et al.*, 2008). Furthermore, there were no

predicted sequences in the *C. sonorensis* database that could match several of the most abundant larger proteins with observed isoelectric point (pI) pIs of 7.5–10 and mass 40–50 kDa which may reflect differences in salivary proteins between *Culicoides* species or indicate that the existing *C. sonorensis* cDNA library is incomplete. To address these issues we report here on the construction of a cDNA library from *C. nubeculosus* salivary glands linked to further mass spectrometry analysis of *C. nubeculosus* salivary proteins.

Results and discussion

In total the cDNA inserts from 127 randomly selected cloned plasmids were sequenced yielding a number of complete open reading frames (ORFs) and several partial protein sequences (Table 1). Mass spectrometry data were collected from 21 of the most abundant protein spots in 2D gels (Fig. 1) and the results of the mass spectrometry are summarized in Table 2. All translated protein sequence

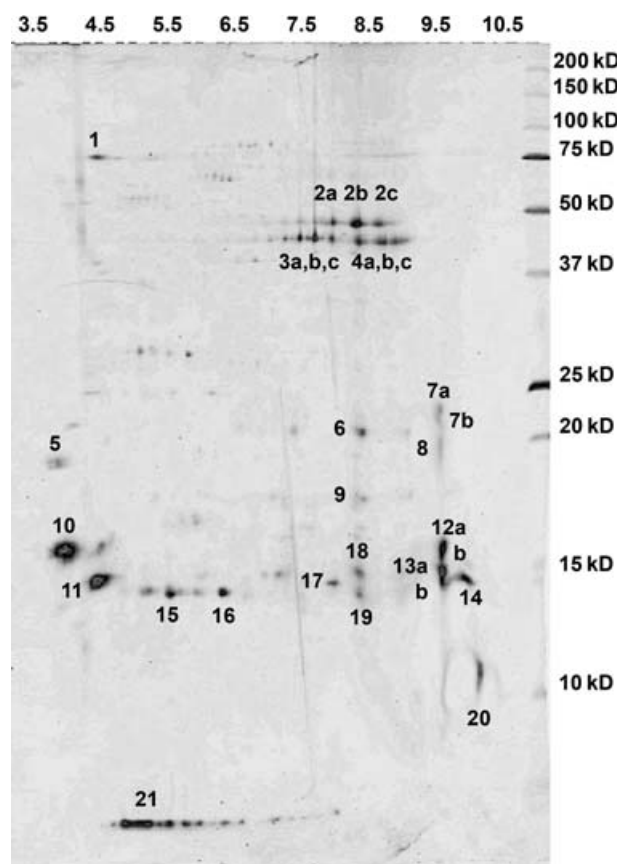


Figure 1. Gel of salivary proteins extracted from 150 *Culicoides nubeculosus* glands dissected at 4 days old. The proteins were separated by 2D electrophoresis, using a broad range linear gradient from pH 3.0–11.0 and a 15% sodium dodecyl sulphate polyacrylamide gel electrophoresis gel. The approximate pH of the strip is indicated at the top of the gel and the size of the molecular weight markers is indicated on the right hand side. The figure is a negative image of a Syproruby stained gel.

Table 1. A list of all the unique proteins sequences identified from *Culicoides nubeculosus* salivary gland cDNA

Clone number*¶	Nucleotide/ protein accession number	Domain/feature	e-value	Putative function	Signal peptide†	MW/PI‡	Best match <i>Culicoides sonorensis</i>	e Value	% identity
Ribosomal									
CNSG2	EU978881 /ACM40870	Ribosomal_L14e	2.80E-32	Ribosomal protein	N	20236/11.7			
Coagulation inhibition									
CNSG23	EU978890/ACM40879	Kunitz_BPTI	0.00083	Serine protease inhibitor	Y/18	14451/4.8	AAU06532.1	1.00E-44	58
CNSG103	EU978922/ACM40911	Kunitz_BPTI	5.60E-05	Factor Xa inhibitor	Y/22	21614/9.9	AAU06530.1	1.00E-85	74
CNSG87	EU978918/ACM40907	Kunitz_BPTI	6.00E-05	Factor Xa inhibitor	Y/19	22793/9.9	AAU06530.1	1.00E-92	73
CNSG8	EU978884/ACM40873	Kunitz_BPTI	0.00068	Serine protease inhibitor	Y/19	20684/9.6	AAU06496.1	2.00E-86	73
CNSG34	EU978894/ACM40883	Kunitz_BPTI	0.00068	Serine protease inhibitor	Y/19	20773/9.5	AAU06496.1	1.00E-87	74
CNSG99	EU978921/ACM40910	Kunitz_BPTI	0.00068	Serine protease inhibitor	Y/19	20713/9.5	AAU06496.1	8.00E-88	74
CNSG115	EU978927/ACM40916	Kunitz_BPTI	1.10E-17	Tissue factor pathway inhibitor	Y/23	17839/9.8	AAU06551.1	2.00E-73	76
D7 Family									
CNSG81	EU978915/ACM40904	PBP_GOBP	5.60E-05	D7 odorant binding protein	Y/19	14613/4.7	AAU06541.1	1.00E-56	73
CNSG20 ²	EU978889/ACM40878	PBP_GOBP	0.0031	D7 odorant binding protein	Y/17	13527/4.2	AAU06484.1	1.00E-66	81
CNSG105	EU978923/ACM40912	PBP_GOBP	0.011	D7 odorant binding protein	Y/17	13560/4.2	AAU06484.1	4.00E-65	79
CNSG58 ³	EU978907/ACM40896	PBP_GOBP	0.011	D7 odorant binding protein	Y/17	13574/4.3	AAU06484.1	1.00E-65	80
CNSG78 ³	EU978913/ACM40902	PBP_GOBP	0.0055	D7 odorant binding protein	Y/17	13600/4.3	AAU06484.1	2.00E-66	80
CNSG55	EU978905/ACM40894	PBP_GOBP	0.00068	D7 odorant binding protein	Y/26	15098/4.7	AAU06517.1	3.00E-54	60
CNSG125	EU978933/ACM40922	PBP_GOBP	0.00056	D7 odorant binding protein	Y/20	14267/9.1	AAU06535.1	4.00E-51	62
CNSG118	EU978929/ACM40918	PBP_GOBP	0.00047	D7 odorant binding protein	Y/20	14103/9.1	AAU06535.1	2.00E-51	64
CNSG27	EU978893/ACM40882	PBP_GOBP	1.50E-05	D7 odorant binding protein	Y/23	15278/9.6	AAU06486.1	2.00E-54	72
Enzymes									
CNSG15	EU978886/ACM4087	Trypsin	1.50E-64	Late trypsin	Y/21	27854/10.2	AAU06478.1	1.00E-108	67
CNSG76	EU978912/ACM40901	Trypsin	1.10E-05	granzyme_B 3e-09	Y/18	35037/7.1	AAU06507.1	4.00E-35	69
CNSG60	EU978908/ACM40897	Peptidase_C36	0.024	papain like cysteine endopeptidase	Y/23	42347/8.7	AAU06548.1	1.00E-60	54
							AAU06490.1	1.00E-48	62
CNSG79	EU978914/ACM40903	Amylo-alpha-1,6-glucosidase	0.28	Unknown function	Y/20	45161/8.8	AAU06503.1	1.00E-81	77
							AAU06499.1	1.00E-67	80
CNSG113	EU978925/ACM40914	Alpha-amylase	1.10E-102	Maltase	Y/20	66349/4.8	AAU06480.1	0	91
CNSG114	EU978926/ACM40915	Glyco_hydro_56	8.30E-113	Hyaluronidase	Y/24	44122/8.7	AAU06477.1	2.00E-90	95
Antigen 5 family									
CNSG45	EU978899/ACM40888	SCP	4.30E-27	Ca++ chelating serine protease	Y/20	23171/10.3	AAU06471.1	2.00E-99	85
CNSG97	EU978920/ACM40909	SCP	1.80E-29	Ca++ chelating serine protease	Y/20	27991/10.1	AAU06471.1	7.00E-99	85
Sugar binding									
CNSG16	EU978887/ACM40876	XRCC1_N	0.057	GDB clan (related to laminin_N)	Y/20	21012/4.4	No match		
CNSG26	EU978892/ACM40881	Laminin_II	0.82	Extra-cellular matrix protein	Y/20	52547/8.8	AAU06485.1	2.00E-116	81
		FIVAR	0.17	sugar binding domain			AAU06511.1	1.00E-81	84
CNSG96	EU978919/ACM40908	Ricin_B_lectin	0.27	Trefoil protein (related to Kunitz and IL1)	Y/20	18755/8.8	AAU06543.1	1.00E-37	41

Table 1. (Continued)

Clone number*¶	Nucleotide/ protein accession number	Domain/feature	e-value	Putative function	Signal peptide†	MW/PI‡	Best match <i>Culicoides sonorensis</i>	e Value	% identity
Other									
CNSG52	EU978903/ACM40892			Unknown function	Y/19	18702/8.9	AAU06543.1	5.00E-33	39
CNSG83	EU978917/ACM40906			Unknown function	Y/20	13746/8.3	AAU06536.1	3.00E-54	68
CNSG48	EU978901/ACM40890			Unknown function	Y/20	13629/6.6	AAU06536.1	9.00E-53	65
CNSG4	EU978883/ACM40872	PSS	0.4	Unknown function	Y/20	12955/7.0	AAU06536.1	1.00E-07	26
CNSG71	EU978911/ACM40900			Unknown function	Y/20	13484/9.7	AAU06536.1	4.00E-06	24
CNSG39 ²	EU978896/ACM40885			Unknown function	Y/20	13514/9.6	AAU06536.1	2.00E-07	25
CNSG1	EU978880/ACM40869	DUF1049	0.19	Unknown function	Y/20	14394/9.3	AAU06494.1	3.00E-71	82
CNSG46	EU978900/ACM40889	DUF1049	0.19	Unknown function	Y/20	14486/9.4	AAU06494.1	3.00E-73	84
CNSG70	EU978910/ACM40899	DUF1049	0.19	Unknown function	Y/20	14548/9.5	AAU06494.1	3.00E-72	82
CNSG82	EU978916/ACM40905	DUF1049	0.19	Unknown function	Y/20	14518/9.4	AAU06494.1	2.00E-72	82
CNSG19	EU978888/ACM40877	GASA	0.14	Unknown function	Y/22	15882/10.4	AAU06523.1	2.00E-68	74
CNSG49 ²	EU978902/ACM40891	GASA	0.31	Unknown function	Y/22	15927/10.5	AAU06523.1	4.00E-68	73
CNSG127	EU978934/ACM40923	Zn_dep_PLPC	0.22	Unknown function	Y/22	15913/10.5	AAU06523.1	3.00E-69	74
CNSG3	EU978882/ACM40871			Unknown function	Y/22	15899/10.5	AAU06523.1	6.00E-71	76
CNSG13	EU978885/ACM40874	Sex_peptide	0.8	Unknown function	Y/23	14914/9.7	AAU06539.1	2.00E-80	88
CNSG67	EU978909/ACM40898	Elf1	0.91	Unknown function	Y/23	12243/9.9	AAU06546.1	8.00E-55	80
CNSG25 ²	EU978891/ACM40880	Elf1	0.91	Unknown function	Y/22	14365/10.0	AAU06545.1	3.00E-72	84
CNSG117	EU978928/ACM40917	Elf1	0.91	Unknown function	Y/22	14395/10.0	AAU06545.1	2.00E-71	83
CNSG36	EU978895/ACM40884	Elf1	0.91	Unknown function	Y/23	14379/10.1	AAU06545.1	7.00E-72	82
CNSG41	EU978897/ACM40886	5.3 kDa proline histidine rich		Unknown function	Y/20	5440/9.8	AAU06529.1	3.00E-03	60
CNSG57	EU978906/ACM40895	5.3 kDa proline histidine rich		Unknown function	Y/20	5370/8.8	AAU06529.1	3.00E-03	60
CNSG122	EU978931/ACM40920	5.3 kDa proline histidine rich		Unknown function	Y/20	4942/8.8	AAU06529.1	2.00E-03	60
CNSG43	EU978898/ACM40887	4.8 kDa serine/threonine rich		Unknown function	Y/23	4873/11.9	AAU06488.1	8.00E-11	67
CNSG53	EU978904/ACM40893	4.8 kDa serine/threonine rich		Unknown function	Y/23	4873/11.9	AAU06488.1	4.00E-11	69
CNSG121	EU978930/ACM40919			Unknown function	Y/16	7998/9.4	No match		
CNSG107	EU978924/ACM40913			Unknown function	Y/21	20201/9.8	AAU06502.1	4.00E-84	85
CNSG123	EU978932/ACM40921			Unknown function	Y/20	7382/9.1	No match		

*superscripts indicate the number of duplicate copies of individual genes detected.

†Y, presence of a signal peptide and the number of the first residue following the signal sequence.

‡The calculated values for molecular weight (MW) and isoelectric point (pI) of the mature sequence excluding the signal peptide.

¶CNSG, *Culicoides Nubeculosus* Salivary Gland.

Table 2. Identification of *Culicoides nubeculosus* clones to abundant protein spots by mass spectroscopy.

Spot no.	Clone no.‡	Protein score	CI%†	Predicted Mw/pl*	Observed Mw/pl	Function
1	CNSG113	367	100	66.3/4.97	75/4.3	Maltase
2a	CNSG60	220	100	42.4/8.66	47/8.1	Unknown
	CNSG79	64	99.989	45.2/8.80		
2b	CNSG60	197	100	42.4/8.66	47/8.5	Unknown
	CNSG79	111	100	45.2/8.80		
2c	CNSG60	122	100	42.4/8.66	47/8.9	Unknown
	CNSG79	232	100	45.3/8.80		
3a,b,c	NI				44/7.6–8.1	
4a,b,c	NI				43 8.5–8.9	
5	CNSG23	90	100	14.5/4.76	19/3.7	Protease inhibitor
6	CNSG96	121	100	18.8/8.78	21/8.9	Ricin B lectin
7a	CNSG87	68	99.996	22.8/9.84	23/9.9	Protease inhibitor
7b	CNSG99	85	100	20.7/9.46	22/9.9	Protease inhibitor
	CNSG8	85	100	20.7/9.59		
	CNSG34	61	99.978	20.8/9.46		
8	CNSG107	40	97.411	20.2/9.75	19.5/9.9	Unknown
9	CNSG52	108	100	18.7/8.85	17.5/8.5	Unknown
10	CNSG20	46	99.287	13.5/4.15	16.0/3.8	D7 family
	CNSG58	46	99.287	13.6/4.29		
	CNSG78	46	99.287	13.6/4.29		
11	CNSG55	80	100	15.1/4.71	14.0/4.2	D7 family
12a	CNSG13	136	100	14.9/9.67	15.5/9.8	Unknown
12b	CNSG27	106	100	15.3/9.55	15.0/9.8	D7 family
13a	CNSG39	98	100	13.5/9.63	14.5/9.8	Unknown
	CNSG71	84	100	13.5/9.73		
13b	CNSG82	95	100	14.5/9.43	14.0/9.8	Unknown
	CNSG70	95	100	14.5/9.46		
	CNSG46	81	100	14.5/9.43		
	CNSG1	80	100	14.4/9.30		
14	CNSG117	91	100	14.4/10.04	14.2/10.1	Unknown
	CNSG36	78	100	14.4/10.10		
	CNSG25	66	99.994	14.4/10.10		
	CNSG67	47	99.495	12.2/9.92		
15*	CNSG4*			13.0/6.99	14.0/5.3	Unknown
16*	CNSG4*			13.0/6.99	14.0/6.2	Unknown
17	CNSG48	228	100	13.6/6.48	14.0/8.0	Unknown
18	CNSG39	105	100	13.5/9.63	15.5/8.4	Unknown
	CNSG71	105	100	13.5/9.73		
19	CNSG39	52	99.821	13.5/9.63	14.0/8.4	Unknown
	CNSG71	52	99.821	13.5/9.63		
20	NI				11.0/10.3	
21	NI				5.0/4.7–5.5	Unknown

*Predicted molecular weight (Mw) and isoelectric point (pl) of protein without its signal peptide. NI, not identified.

†CI, Confidence Interval.

‡CNSG, *Culicoides Nubeculosus* Salivary Gland.

data were aligned to the National Center for Biotechnology Information (NCBI) GenBank database using BLAST-P and included in the initial mass spectrometry analysis. In cases where a partial protein sequence had a significant match to a protein spot by mass spectrometry or aligned with *C. sonorensis* sequences, the remainder of the sequence was acquired by library screening or by 5' rapid amplification of cDNA ends (5' RACE). Several of the sequenced plasmids were detected as identical copies, a total of 54 unique sequences with complete ORFs including 5' untranslated region (UTR) sequence, 3' UTR and poly-A tails are reported here (Table 1). It is striking that apart from CNSG2, which is a ribosomal protein, every complete ORF has a signal sequence, indicating that secretory proteins

which may have a role in blood feeding are by far the most abundant products in the salivary gland proteome.

The majority of these unique sequences are similar to one or more proteins reported previously from *C. sonorensis*. Table 1 lists the unique *C. nubeculosus* sequences with their accession numbers and shows the accession numbers of the highest matching protein sequences from *C. sonorensis*. Typically there was about 70% amino acid identity between homologous sequences from the two species. Based on multiple alignments using CLUSTALW, the *C. nubeculosus* sequences displayed in Table 1 have been grouped according to similarity of amino acid sequence. The proteins within these groupings often share Pfam motifs and are highly similar to the same *C. sonorensis*

sequence. In some cases the members of these closely related groups (eg the D7-like CNSG 20, 105, 58 and 78), differ by only one or two residues and have nearly identical molecular weights (MWs)/pIs. As the library was derived from a pool of several hundred *C. nubeculosus* salivary glands, these minor amino acid differences may represent polymorphs of the same abundant proteins. In contrast, other sequences within the D7 group match to different *C. sonorensis* proteins with distinct MWs/pIs and probably represent separate related members of the D7 protein family. Similarly the seven Kunitz protease inhibitor sequences can be placed into four subgroups, suggesting that they may target different proteases involved in the coagulation cascade. Several other groups made up of related proteins with unknown function were also recognized. Overall we have identified 32 distinct groups or individual secretory proteins by this analysis. In addition to the similarities within groups there are relationships between some of the groups, such that CNSG3, 19, 49, 127 were most similar to CNSG13, 25, 36, 17, 117 and both these groups had similarity to CNSG1, 70, 42, 86. This pattern of a hierarchy amongst related groups of proteins is similar to that reported for mosquitoes (Calvo *et al.*, 2007; Ribeiro *et al.*, 2007) and is consistent with the proposed rapid evolution of salivary proteins by gene duplication and radiation driven by the need to avoid host immune responses that could inactivate a single conserved protein and thereby inhibit blood feeding (Calvo *et al.*, 2006; Arca *et al.*, 2007).

Both of the previous studies on *C. nubeculosus* salivary proteins (Langner *et al.*, 2007; Wilson *et al.*, 2008) noted a discrepancy between the mass and pI of protein bands detected in *C. nubeculosus* saliva and those predicted from known *C. sonorensis* sequences. In particular, no known sequences accounted for the mass of abundant 40–45 kDa proteins in *C. nubeculosus* secreted saliva (Langner *et al.*, 2007) or proteins with a similar mass and high pI that were identified as the putative major allergens in horses (Wilson *et al.*, 2008). Nevertheless both papers matched proteins of about 40 kDa to the same *C. sonorensis* sequence, gb|AAU06548.1| a protein with a predicted mass of only 26.5 kDa. In this analysis mass spectrometry identified spots 2a and 2b as *C. nubeculosus* clone CNSG60 (MW 44 759 kDa, pI 8.65), which also aligns at its amino-terminal to gb|AAU06548.1|. CNSG60 has a secretory signal peptide and contains peptide sequences that were identified in the mass spectrometry of a 40 kDa band reported by Langner *et al.* (2007), confirming that it is the same protein as that found in secreted saliva. Furthermore, in addition to having significant similarity to gb|AAU06548.1| at the amino terminus, CNSG60 also aligns with a second *C. sonorensis* protein gb|AAU06490.1| at its carboxy terminus, suggesting that the two corresponding *C. sonorensis* cDNAs are in fact the 5' and 3' ends of a single larger sequence. Two other *C. nubeculosus*

proteins (CNSG26 and CNSG79) also matched distinct pairs of *C. sonorensis* proteins at either terminus, further increasing the number of larger proteins that have been identified.

We had previously reported only limited success in matching mass spectrometry data from *C. nubeculosus* proteins with the available predicted protein sequences from the cDNA library of *C. sonorensis* (Wilson *et al.*, 2008). The data presented here provided a much larger number of matches when the mass spectrometry output from *C. nubeculosus* saliva extract was compared to the predicted protein sequences from a *C. nubeculosus* cDNA library, apparently confirming our earlier suggestion that homologous proteins from the two species are not sufficiently conserved to be easily identified by mass spectrometry.

Table 2 shows the clone numbers of proteins that were matched to particular spots. In several cases the spots match up to abundant transcripts. Spots 10 and 11 are D7s, Spots 5, 6, 7a and 7b are related to Kunitz protease inhibitors, whereas spots 13 and 14 match to sequences of unknown function which were also grouped together by CLUSTALW (Table 1). Not all protein spots have been identified as abundant transcripts. Spots 2a and 2b match to CNSG60 and CNSG79, two related sequences which were detected as only single copies in our library and to date, no *C. nubeculosus* transcripts have been identified that can account for spots 3 or 4. However, we do have a sequence derived from wild caught-midges of the *Culicoides pulicaris* group (CPSG113, C. L. Russell, unpubl. data) for a protein of around 40 kDa pI 8.9 that gives an almost significant mass spectrometry score for spot 4, although whether this represents a true homologue of spot 4 has yet to be confirmed. Similarly, spot 3 is possibly homologous to a protein from our *C. pulicaris* group library (CPSG47, C. L. Russell, unpubl. data). On BLAST CPSG47 is also related to SG60, which itself has a low score for spot 3 in mass spectrometry.

When we compared the tandem mass spectrometry (MS/MS) peptide sequences generated by Langner *et al.* (2007) with the predicted protein sequences from *C. nubeculosus* reported here, we found several that matched 100% to our sequences for maltase and members of the D7 family. Other mass spectrometry peptide sequences gave consistent but lower scores with *C. nubeculosus* proteins, suggesting that although many of the peptide sequences derived by Langner *et al.* did not match with *C. sonorensis* proteins, they are nevertheless secretory proteins from *C. nubeculosus*. In particular, Langner *et al.* analysed three spots on 2D gels with masses of approximately 15 kDa and pIs between 6.8–7.4 from which a peptide sequence (LLADSTGVQVWLDDGK) was identified by mass spectrometry. This peptide sequence is an exact match for our clone CNSG4 and occurred along with several other

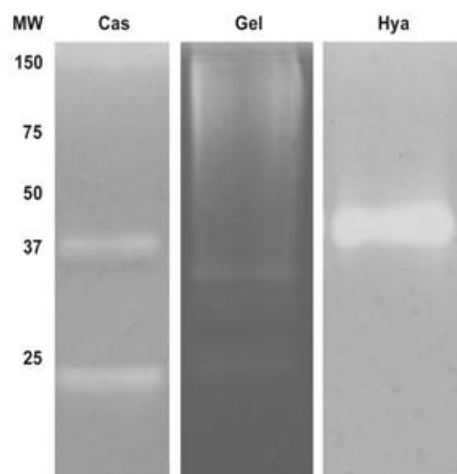


Figure 2. Gel zymography of *Culicoides nubeculosus* saliva on three different substrates casein (Cas), gelatine (Gel) and hyaluronic acid (Hya). Pale areas within each cell indicate the presence of enzyme activity specific for each substrate. Saliva extract equivalent to a single midge has been loaded onto each lane.

MS/MS peptide sequences that were highly similar to CNSG4. The predicted mass/pI of CNSG4 without its signal peptide is 12.955 kDa/pI 7.0; we have therefore tentatively designated CNSG4 as the identity for spots 15 and 16 in our analysis which correspond to the location described by Langner *et al.*

Functions can be assigned to only a minority of the abundant proteins that we have identified by mass spectrometry. These include the maltase protein (spot 1); several protease inhibitors which most likely act as anticoagulants (spots 5, 7a, b). Also, members of the D7 family (spots 10, 11, 12b) that we predict are likely to share the function of their homologues from mosquitoes which bind vaso-active amines (Calvo *et al.*, 2006). Several digestive enzymes were identified in the saliva of *C. nubeculosus* by functional gel zymography (Fig. 2). Casein gel zymography showed two distinct bands at 38 and 24 kDa. Both activities were eliminated by the serine protease inhibitor 4-(2-Aminoethyl) benzenesulfonyl fluoride (AEBSF) (data not shown), suggesting that the upper band may be CNSG76 (35 kDa) and the lower one CNSG15 (27.8 kDa) that both contain trypsin-like domains (Table 1). The predicted mass of CNSG15 is rather large, but the extra mass could be accounted for if it is synthesized as an inactive precursor that is cleaved to form the active protease. This hypothesis is supported by a degree of similarity between CNSG15's amino-terminal and the sequence of other *C. nubeculosus* Kunitz proteases inhibitors indicated by BLAST. It is interesting to speculate that the anticoagulant protease inhibitors found in *Culicoides* saliva may have originally evolved from such self-inactivating inhibitor domains of digestive enzymes with feeding functions in the ancestral midge. A different pattern of proteolysis was seen with gels in which gelatine

was used as the substrate (Fig. 2). In addition to the two bands, a smear of activity at the top of the gel was observed; this activity was also eliminated by AEBSF confirming that it was the result of a functional serine protease. We interpret this smearing as indicating that one or both of the proteases were being retarded during electrophoresis through binding to the gelatine even in the presence of sodium dodecyl sulphate (SDS), which indicates a strong affinity for collagen-like proteins. Hyaluronidase activity formed a band just below 50 kDa. This accords well with the predicted MW of 44.2 kDa and pI 8.71 for CNSG114, our hyaluronidase sequence. Although we could not associate the hyaluronidase sequence with an abundant protein spot by mass spectrometry, it was highly active, being detectable when as little as one-tenth of a single salivary gland was loaded. This observation, combined with the high activity of salivary proteases against denatured collagen (gelatine) fits well with a putative salivary function in assisting skin penetration by mouth parts and breakdown of connective tissue associated with pool feeding.

The results of staining 2D gels of *C. nubeculosus* saliva for glycoproteins and phosphorylated proteins are shown in Fig. 3: several of the larger proteins are glycosylated and in some cases phosphorylated. Secretory proteins are often glycosylated or modified by phosphorylation as they pass through the Golgi apparatus. Fig. 3 shows examples of 2D gels that have been stained to detect these modifications. The majority of the larger proteins have been glycosylated as, to a lesser extent, have some of the lower molecular weight peptides. Similarly, several proteins have been modified by phosphorylation, differences in the extent of phosphorylation can account for the pattern of multiple spots with slightly different pIs that characterize some of the proteins. A 60 kDa glycoprotein with broad immunosuppressive function has been partially purified by lectin affinity from *C. sonorensis* (Bishop *et al.*, 2006), but we have not positively identified a glycoprotein of 60 kDa that could account for these activities.

The second most abundant sequence detected in the *C. sonorensis* library reported by Campbell *et al.* (2005) was for a nonsecretory laminin-like protein (AAU06485.1). This matches the C-terminus of a *C. nubeculosus* sequence (CNSG26) that also matches a different *C. sonorensis* sequence with a secretory signal at its N-terminus (AAU06511.1). In addition to having a laminin-like domain, CNSG26 also has a potential carbohydrate binding domain (Table 1) that gives it even greater similarity to laminin. Once again, although it does contain a secretory signal peptide, we were unable to identify an abundant protein spot that could account for this protein. Furthermore, CNSG26 does not match with any of the peptide sequences identified from proteins secreted during feeding by Langner *et al.* (2007) and it remains unclear what if any function this sequence could have in blood feeding.

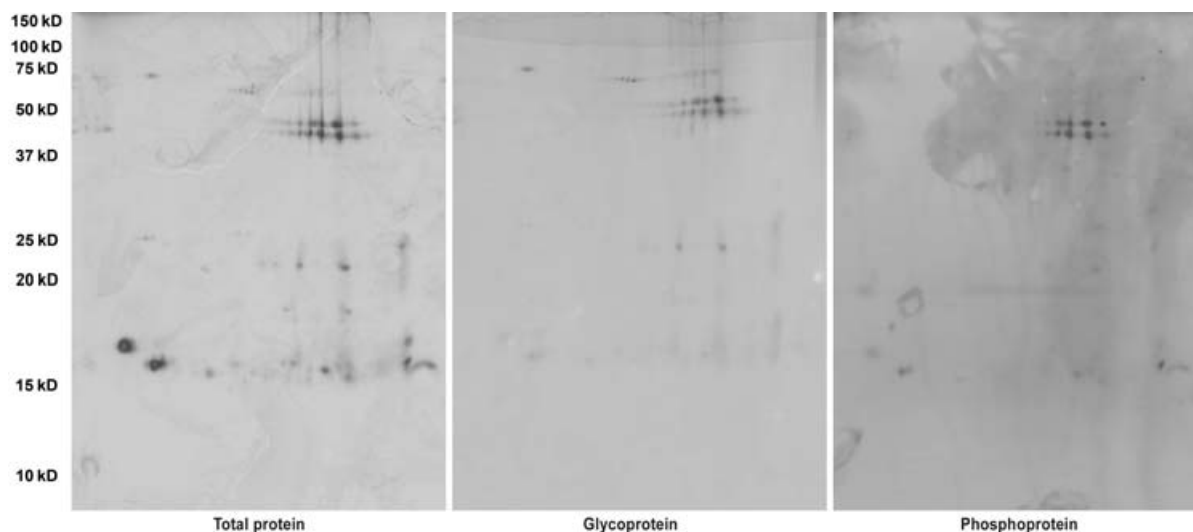


Figure 3. 2D gel electrophoresis images of 150 *Culicoides nubeculosus* salivary glands stained for total protein (Sypro Ruby™), glycoprotein (ProQ Emerald 488™) and phosphoprotein (ProQ Diamond™).

CNSG41 pI 9.97 MW 5311 (without signal peptide)

MKWFAFAIFLIAIIVCVMAQGPEGRKPGPPGPHHGPHGHGHGHHPRPPHPRPPKPRPSTTEEPSSTPSE

CNSG43 pI 10.19 MW 4786 (without signal peptide)

MKLLSIFFMFFILNSFLILSEGSPIFFGWGHPWRRCTRSTTTTTTTTTTTTTTPSTTTTRDPNADQE

Figure 4. Two sequences encoding unusual *Culicoides nubeculosus* salivary proteins of low molecular weight, the underlined residues form their secretory leader sequences. Homologues of these sequences have been identified as abundant transcripts in all midge species studied thus far.

Similarly, we have no protein sequences that would yield identical mass and pI for spots 20 or 21 and no peptides from these spots were identified by mass spectrometry. Two candidate proteins, one serine/threonine rich and the other proline/histidine rich with a MW less than 10 kDa were initially considered (Fig. 4). They are present in several near identical copies in our library and homologues of these two proteins have been reported as abundant in *C. sonorensis*. We have also frequently detected homologous sequences in libraries from wild caught *C. pulicaris* and *Culicoides obsoletus* group (C. L. Russell, unpubl. data). When the signal peptide is removed the predicted mass of both candidates is near 5 kDa, but they would retain a pI of around 10 although it is possible that the pI of either candidate sequence could be reduced by phosphorylation of serine, threonine or histidine residues. However, we have been unable to confirm this by staining for phosphoproteins (Fig. 3). Indeed the 5 kDa spot 21 is not always present in our protein preps. For example, it is clearly seen in Fig. 1 but is absent from Fig. 3; overall it occurs in about half of our salivary gland extracts even though they were all prepared in an identical way. As yet we have no confirmed

explanation for this observation but it may be the result of leakage of the protein through spin filters during preparation. The filters have a nominal 3 kDa cut-off, but this will still allow some loss of smaller proteins. The six repeated spins used for sample dialysis may result in substantial overall losses, whereas batch variation in filter performance could account for why only some samples are affected.

We are currently working to express representative examples of our cDNA library in baculovirus, which will enable some of these issues to be resolved and further functional studies on *Culicoides* salivary proteins to be carried out.

Experimental procedures

Extraction of *C. nubeculosus* salivary gland protein

Laboratory bred *C. nubeculosus* were from the colony maintained by the Institute for Animal Health, Pirbright Laboratory, Woking, Surrey, UK.

Live *C. nubeculosus* were kept on ice; the female midges removed and placed in a drop of Tris-saline, pH 7.4 on a glass coverslip. The salivary glands were dissected from 125 female

midges under a Leica MZ6 stereomicroscope (Leica Microsystems GmbH, Wetzlar, Germany; <http://www.leica-microsystems.com>) and collected in 500 µl of 10 mM Tris (pH 7.4) containing 5 µl each of protease inhibitor cocktail (Sigma-Aldrich, St Louis, MO, USA; <http://www.sigmaaldrich.com>). The glands were maintained at 4 °C during dissection; batches of 200–300 glands were stored at –20 °C. The extract was subject to 3× freeze-thaw episodes from 4 to –70 °C, followed by filtration using a 0.2 µm Nanosep® centrifugal device (Pall, Ann Arbor, MI, USA; <http://www.pall.com>) at 14 000 *g* for 3 min, 4 °C. The filtrate was then concentrated using a 3K Nanosep® centrifugal device at 14 000 *g* for 10 min, 4 °C. Protein concentration was measured using an Experion™ automated electrophoresis apparatus (Bio-Rad Laboratories Inc., Hercules, CA, USA; <http://www.bio-rad.com>). The protein extracts were stored at –20 °C until used.

Extraction of *C. nubeculosus* salivary gland total RNA

Salivary glands from 150 female midges were extracted as above but collected in 5 ml RNA Lysis Buffer (RLT) buffer containing β-mercaptoethanol (Qiagen, Crawley, West Sussex, UK; <http://www.qiagen.com>) and stored at –70 °C. Total RNA was extracted using RNeasy® Mini Kit (Qiagen) according to the manufacturer's instructions. RNA concentration and quality was measured using an Experion™ automated electrophoresis apparatus (Bio-Rad). RNA was stored at –80 °C.

C. nubeculosus salivary gland cDNA library

The *C. nubeculosus* cDNA library was constructed using a Creator™SMART™ cDNA library construction kit (Clontech Laboratories Inc, Mountain View, CA, USA; www.clontech.com), according to the manufacturer's instructions. cDNA inserts were amplified from individual colonies by colony PCR using primers (Libseq_fwd 5' GAGTCAGTGAGCGAGGAAGC; Libseq_rev 5' CCAGGATCTC CTAGGGAAACA) (Invitrogen, Carlsbad, CA, USA; www.invitrogen.com). PCR consisted of 95 °C for 15 min, 30 cycles of 94 °C for 10 s, 55 °C for 20 s and 72 °C for 2 min and a final extension 72 °C for 3 min. Inserts ≥ 500 bp were cleaned using NucleoSpin® Extract II (Macherey-Nagel, Düren, Germany; www.macherey-nagel.com) and sent for sequencing using Libseq_fwd primer (reaction using Libseq reverse primer invariably failed on reaching the poly-A tail). DNA sequencing was performed by The Sequencing Service (School of Life Sciences, University of Dundee, Dundee, UK; www.dnaseq.co.uk) using Applied Biosystems Big-Dye Ver 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer (Applied Biosystems, Foster City, CA, USA). Sequence data were analysed using MACVECTOR® with ASSEMBLER, Ver 9.5.2 (MacVector Inc, Cary, NC, USA; www.macvector.com). Each ABLview sequence file was examined using the Phred function in MACVECTOR and any files that showed poor quality data were re-sequenced. In the case of larger inserts where the ABLview trace became degenerate before a clear poly-A tail was reached, internal primers were synthesized and the reaction re-run to complete the 3' sequence. Where only 3' tail ends of sequences were detected, reverse primers were designed around the stop codon and the full length sequence generated from the original *C. nubeculosus* cDNA by 5' RACE using the 5' PCR Primer (AAGCAGTGGTATCAACGCAGAGT) supplied with the cDNA library kit.

Deduced protein sequences were searched against the NCBI's GenBank database using BLASTp (www.ncbi.nlm.nih.gov/

blast/). Putative secretory peptides were detected using SIGNALP 3.0 (www.cbs.dtu.dk/services/SignalP/) (Bendtsen *et al.*, 2004). Sequences were analysed for protein domains and families using PFAM (www.Pfam.sanger.ac.uk/) (Finn *et al.*, 2006). The predicted mass and pI of the deduced protein sequences was calculated with the MACVECTOR® protein analysis toolkit.

cDNA library screen

Colony lifts were performed using Hybond-N+ (Amersham, Little Chalfont, Buckinghamshire, UK; www.amersham.com). DNA was labelled using Gene Images Direct Labelling and Detection system (Amersham) according to the manufacturer's instructions.

2D electrophoresis

Salivary gland extracts containing 50–100 glands were made up with 7 M urea, 2 M thiourea, 4% 3-[(3-Cholamidopropyl)dimethylammonio]-1-Propane Sulfonate (CHAPS), 0.002% bromophenol blue, 0.5% (v/v) IPG Buffer pH 3–11NL (GE Healthcare, Chalfont St. Giles, Buckinghamshire, UK; www.gelifesciences.com) and 1.2% (v/v) Destreak reagent (GE Healthcare) then concentrated by centrifugation through a 3 kDa cut-off Nanosep® (Pall, Ann Arbor, MI, USA). This was repeated five times to ensure complete removal of salts, and the sample made up to 200 µl after the final spin. Samples were left at room temperature for 1 h prior to loading onto 110 mm immobiline DryStrip first dimension immobilised pH gradient (IPG) strips pH 3–11 linear (GE Healthcare) by passive re-hydration.

Following overnight rehydration, the strips were focused using the Ettan IPGphor (GE Healthcare) according to the manufacturer's instructions (in brief, 500 V for 1 h, 1000 V for 1 h and 8000 V for 2 h). The focused IPG strips were then incubated for 15 min in 10 ml SDS equilibration buffer [50 mM Tris-HCl, pH 8.8, 6 M urea, 30% (v/v) glycerol, 2% SDS, 0.002% bromophenol blue] containing 100 mg dithiothreitol (DTT) and for a further 15 min in 10 ml equilibration buffer containing 250 mg iodoacetamide.

The equilibrated strips were applied to the surface of vertical 12.5 or 15% SDS-polyacrylamide gels and proteins separated in the second dimension using the Ettan DALT 6 separation unit (GE Healthcare) at 0.2 W/gel for 1 h, 0.4 W/gel for 1 h and then 20 W/gel until completion. Gels were fixed and stained for glycoproteins (using ProQ Emerald 488 glycoprotein stain), phosphoproteins (using ProQ Diamond phosphoprotein stain) or for total protein (using SYPRO ruby protein gel stain) according to the manufacturer's instructions (Invitrogen). The gels were then imaged using a Typhoon 9400 scanner (GE Healthcare).

Gel zymography

Enzyme activities in salivary gland extract were assayed by substrate gel zymography. In brief, 7.5 µl extract was mixed with 2.5 µl 4× nonreducing sample buffer and loaded onto either a 10% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) gel containing 0.75 mg per ml bovine gelatin or a 12% SDS-PAGE gel with 0.5 mg/ml casein. After electrophoresis, gels were washed three times in 50 ml 2.5% Triton X-100 for 10 min each then incubated for 16 h at 37 °C in 50 mM Tris/HCl, pH 7.8, 100 mM NaCl. Gels were stained with 0.1% Coomassie blue R250 in 10% acetic acid, 40% methanol for 1 h then destained in 7.5% acetic acid, 10% methanol until the zones of proteolysis or the stacking gel had cleared. For hyaluronidase, 10% PAGE gels were copolymerized

with 1 mg/ml hyaluronic acid, no SDS was added to the gel and $3 \times$ the standard concentration of ammonium persulphate (APS) and Tetramethylethylenediamine (TEMED) were used to ensure polymerization. The gel was run with 196 mM glycine 50 mM Tris-HCl pH 8.3 buffer, in which the concentration of SDS was reduced to 0.01%. After running the gels were washed for 3×10 min in 50 ml 2.5% Triton then incubated overnight in 100 mM NaCl 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid (HEPES) pH 6.5, washed for 2×10 min in distilled H₂O, stained with 0.5% Alcian blue, 3% acetic acid for 1 h then destained in 7% acetic acid (3×1 h).

Matrix assisted laser desorption ionization (MALDI) MS/MS and protein identification

Protein spots of interest were excised from the gel using the Investigator ProPic automated spot picker and digested with trypsin using the ProGest automated digestion unit (both from Perkin Elmer Life Sciences, Boston, MA, USA). Peptides generated by in-gel digestion of spots were bound onto a ZipTip® (Millipore, Billerica, MA, USA; www.millipore.com), eluted into a saturated solution of α -cyano hydroxycinnamic acid (CHCA) in 50:50 acetonitrile : water and spotted directly onto a MALDI target plate.

Mass analysis was performed using an Applied Biosystems 4700 matrix assisted laser desorption ionization time of flight/time of flight tandem (MALDI TOF/TOF) mass spectrometer. Peptide mass fingerprint data were collected in positive ion reflector mode. Peptides with a signal-to-noise ratio above 50 were selected for fragmentation by collision induced dissociation (CID); a maximum of five MS/MS were allowed per spot. The precursor mass window was 100 relative resolution (full width at half maximum). CID of the peptides was performed with air as a collision gas. The resultant spectra were not computer enhanced (smoothed, or baseline subtracted) prior to subsequent data analysis.

Combined MS and MS/MS data were searched against all known *Culicoides* sequences using GPS Explorer (ABI, Applied Biosystems, Lincoln, CA, USA) and Mascot (MatrixScience, Boston, MA, USA) with trypsin as the proteolytic enzyme. Two analyses were undertaken. In the first, methionine oxidation was considered as a variable modification, in the second, variable modifications of methionine oxidation and serine, threonine and tyrosine phosphorylation were considered. For the database search a mass tolerance of 100 ppm for precursor ions and 0.5 Da for fragment ions was used; a maximum of one missed cleavage was allowed.

Protein identification was made on the basis of a statistically significant Mowse score. The Mascot implementation of the Mowse algorithm takes account of not only accurate matching of peptide and fragment masses but also data quality, the library size and behaviour of the proteolytic enzyme (in our case trypsin). The total Mowse score is the probability that the match is a random event and is reported as $-10 \cdot \log_{10} P$, where P is the absolute probability. In this work, the commonly accepted threshold that a protein identification is significant, $P < 0.05$, has been used; where P is the probability of a match occurring randomly 5% of the time. Thus, in this work Mowse scores greater than 41 indicate a statistically significant match.

Acknowledgements

This work was supported by grant no G1106 from The Horse Trust, Princes Risborough, Bucks, HP27 0PP, UK.

References

- Akey, D.H., Luedke, A.J. and Jones, R.H. (1985) Salivary gland homogenates from the vector *Culicoides variipennis* may aid in detection of bluetongue virus in chronically infected cattle. *Prog Clin Biol Res* **178**: 135–145.
- Anderson, G.S., Belton, P. and Kleider, N. (1993) Hypersensitivity of horses in British Columbia to extracts of native and exotic species of *Culicoides* (Diptera: Ceratopogonidae). *J Med Entomol* **30**: 657–663.
- Arca, B., Lombardo, F., Francischetti, I.M., Pham, V.M., Mestres-Simon, M., Andersen, J.F. et al. (2007) An insight into the sialome of the adult female mosquito *Aedes albopictus*. *Insect Biochem Mol Biol* **37**: 107–127.
- Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**: 783–795.
- Billingsley, P.F., Baird, J., Mitchell, J.A. and Drakeley, C. (2006) Immune interactions between mosquitoes and their hosts. *Parasite Immunol* **28**: 143–153.
- Bishop, J.V., Mejia, J.S., Perez de Leon, A.A., Tabachnick, W.J. and Titus, R.G. (2006) Salivary gland extracts of *Culicoides sonorensis* inhibit murine lymphocyte proliferation and no production by macrophages. *Am J Trop Med Hyg* **75**: 532–536.
- Braverman, Y., Ungar-Waron, H., Frith, K., Adler, H., Danieli, Y., Baker, K.P. et al. (1983) Epidemiological and immunological studies of sweet itch in horses in Israel. *Vet Rec* **112**: 521–524.
- Calvo, E., Mans, B.J., Andersen, J.F. and Ribeiro, J.M. (2006) Function and evolution of a mosquito salivary protein family. *J Biol Chem* **281**: 1935–1942.
- Calvo, E., Tokumasu, F., Marinotti, O., Villeval, J.L., Ribeiro, J.M. and Francischetti, I.M. (2007) Aegyptin, a novel mosquito salivary gland protein, specifically binds to collagen and prevents its interaction with platelet glycoprotein VI, integrin $\alpha 2 \beta 1$, and von Willebrand factor. *J Biol Chem* **282**: 26928–26938.
- Campbell, C.L., Vandyke, K.A., Letchworth, G.J., Drolet, B.S., Hanekamp, T. and Wilson, W.C. (2005) Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). *Insect Mol Biol* **14**: 121–136.
- Champagne, D.E. (2004) Antihemostatic strategies of blood-feeding arthropods. *Curr Drug Targets Cardiovasc Haematol Disord* **4**: 375–396.
- Champagne, D.E. (2005) Antihemostatic molecules from saliva of blood-feeding arthropods. *Pathophysiol Haemost Thromb* **34**: 221–227.
- Connan, R.M. and Lloyd, S. (1988) Seasonal allergic dermatitis in sheep. *Vet Rec* **123**: 335–337.
- Cowley, J.A. and Gorman, B.M. (1990) Effects of proteolytic enzymes on the infectivity, haemagglutinating activity and protein composition of bluetongue virus type 20. *Vet Microbiol* **22**: 137–152.
- Dohnal, J., Blinn, J., Wahl, G. and Schulz-Key, H. (1990) Distribution of microfilariae of *Onchocerca lienalis* and *Onchocerca gutturosa* in the skin of cattle in Germany and their development in *Simulium ornatum* and *Culicoides nubeculosus* following artificial infestation. *Vet Parasitol* **36**: 325–332.
- Finn, R.D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T. et al. (2006) Pfam: clans, web tools and services. *Nucleic Acids Res* **34**: D247–251.

- Halldorsdottir, S. and Larsen, H.J. (1991) An epidemiological study of summer eczema in Icelandic horses in Norway. *Equine Vet J* **23**: 296–299.
- Harrison, L.M., Nerlinger, A., Bungiro, R.D., Cordova, J.L., Kuzmic, P. and Cappello, M. (2002) Molecular characterization of *Ancylostoma* inhibitors of coagulation factor Xa. Hookworm anticoagulant activity *in vitro* predicts parasite bloodfeeding *in vivo*. *J Biol Chem* **277**: 6223–6229.
- Langner, K.F., Darpel, K.E., Denison, E., Drolet, B.S., Leibold, W., Mellor, P.S. *et al.* (2007) Collection and analysis of salivary proteins from the biting midge *Culicoides nubeculosus* (Diptera: Ceratopogonidae). *J Med Entomol* **44**: 238–248.
- Marchi, P.R., Rawlings, P., Burroughs, J.N., Wellby, M., Mertens, P.P., Mellor, P.S. *et al.* (1995) Proteolytic cleavage of VP2, an outer capsid protein of African horse sickness virus, by species-specific serum proteases enhances infectivity in *Culicoides*. *J Gen Virol* **76** (Pt 10): 2607–2611.
- McCraig, J. (1974) A study to establish the incidence of sweet itch in ponies in the United Kingdom. *Veterinary Record* **93**: 444–446.
- Mellor, P.S., Boorman, J. and Baylis, M. (2000) *Culicoides* biting midges: their role as arbovirus vectors. *Annu Rev Entomol* **45**: 307–340.
- Mertens, P.P., Burroughs, J.N., Walton, A., Wellby, M.P., Fu, H., O'Hara, R.S. *et al.* (1996) Enhanced infectivity of modified bluetongue virus particles for two insect cell lines and for two *Culicoides* vector species. *Virology* **217**: 582–593.
- Okuda, M., Usami, A., Itoh, H. and Ogino, S. (2002) [Nationwide investigation of insect allergy in patients with allergic rhinitis]. *Nippon Jibiinkoka Gakkai Kaiho* **105**: 1181–1188.
- Perez de Leon, A.A. and Tabachnick, W.J. (1996) Apyrase activity and adenosine diphosphate induced platelet aggregation inhibition by the salivary gland proteins of *Culicoides variipennis*, the North American vector of bluetongue viruses. *Vet Parasitol* **61**: 327–338.
- Perez de Leon, A.A., Ribeiro, J.M., Tabachnick, W.J. and Valenzuela, J.G. (1997) Identification of a salivary vasodilator in the primary North American vector of bluetongue viruses, *Culicoides variipennis*. *Am J Trop Med Hyg* **57**: 375–381.
- Perez de Leon, A.A., Valenzuela, J.G. and Tabachnick, W.J. (1998) Anticoagulant activity in salivary glands of the insect vector *Culicoides variipennis* sonorensis by an inhibitor of factor Xa. *Exp Parasitol* **88**: 121–130.
- Ribeiro, J.M., Arca, B., Lombardo, F., Calvo, E., Phan, V.M., Chandra, P.K. *et al.* (2007) An annotated catalogue of salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. *BMC Genomics* **8**: 6.
- Salzet, M. (2001) Anticoagulants and inhibitors of platelet aggregation derived from leeches. *FEBS Lett* **492**: 187–192.
- Schneider, B.S., Soong, L., Girard, Y.A., Campbell, G., Mason, P. and Higgs, S. (2006) Potentiation of West Nile encephalitis by mosquito feeding. *Viral Immunol* **19**: 74–82.
- Schneider, B.S., McGee, C.E., Jordan, J.M., Stevenson, H.L., Soong, L. and Higgs, S. (2007) Prior exposure to uninfected mosquitoes enhances mortality in naturally-transmitted west nile virus infection. *PLoS ONE* **2**: e1171.
- Takamatsu, H., Mellor, P.S., Mertens, P.P., Kirkham, P.A., Burroughs, J.N. and Parkhouse, R.M. (2003) A possible overwintering mechanism for bluetongue virus in the absence of the insect vector. *J Gen Virol* **84**: 227–235.
- Titus, R.G., Bishop, J.V. and Mejia, J.S. (2006) The immunomodulatory factors of arthropod saliva and the potential for these factors to serve as vaccine targets to prevent pathogen transmission. *Parasite Immunol* **28**: 131–141.
- Wilson, A.D., Heesom, K.J., Mawby, W.J., Mellor, P.S. and Russell, C.L. (2008) Identification of abundant proteins and potential allergens in *Culicoides nubeculosus* salivary glands. *Vet Immunol Immunopathol* **122**: 94–103.
- Yeruham, I., Braverman, Y. and Orgad, U. (1993) Field observations in Israel on hypersensitivity in cattle, sheep and donkeys caused by *Culicoides*. *Aust Vet J* **70**: 348–352.