รายงานฉบับสมบูรณ์ของโครงการวิจัย

โครงการ

Transcriptome และ Proteome ของต่อมน้ำลาย *Simulium nodosum* ตัวเต็มวัยเพศเมีย

โดย ผู้ช่วยศาสตราจารย์ดร.นริศรา จริยะพันธุ์

กันยายน พ.ศ. 2550

รายงานฉบับสมบูรณ์ของโครงการวิจัย

โครงการ
Transcriptome และ Proteome ของต่อมน้ำลาย
Simulium nodosum ตัวเต็มวัยเพศเมีย

โดย ผู้ช่วยศาสตราจารย์ดร.นริศรา จริยะพันธุ์

กันยายน พ.ศ. 2550

สัญญาเลขที่ MRG4880037

รายงานวิจัยฉบับสมบูรณ์

โครงการ **Transcriptome และ Proteome ของต่อมน้ำลาย *Simulium nodosum* ตัวเต็มวัยเพศเมีย**

ผู้วิจัย	สังกัด
ผศ.ดร.นริศรา จริยะพันธุ์	ภาควิชาปรสิตวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยเชียงใหม่

สนับสนุนโดยสำนักงานคณะกรรมการการอุดมศึกษา และสำนักงานกองทุนสนับสนุนการวิจัย

กิตติกรรมประกาศ

ขอขอบพระคุณสำนักงานกองทุนสนับสนุนการวิจัยซึ่งเป็นผู้สนับสนุนทุนวิจัยในเรื่องนี้ ขอขอบพระคุณ อธิการบดีมหาวิทยาลัยเชียงใหม่, คณบดีคณะแพทยศาสตร์ มหาวิทยาลัยเชียงใหม่ และ หัวหน้าภาควิชาปรสิตวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยเชียงใหม่ ที่อนุญาตและสนับสนุนการวิจัย

ขอขอบพระคุณ ศาสตราจารย์เวช ชูโชติ, ศาสตราจารย์ยง ภู่วรวรรณ, Prof. Dr. Hiroyuki Takaoka, Dr. Yasushi Otsuka, Dr. Paul Bates, Dr. Mark Wilkinson ที่ช่วยให้คำปรึกษาแนะนำและร่วมทำงานวิจัยอย่างดียิ่ง

ผู้ช่วยศาสตราจารย์ดร.นริศรา จริยะพันธุ์

ABSTRACT

ภาษาไทย:

ริ้นดำ *Simulium nodosum* เป็นหนึ่งในพาหะที่คาดว่านำโรคฟิลาเรียในสัตว์ที่พบในประเทศไทย การกัดของริ้นดำยังเป็นสาเหตุของการสูญเสียเลือดในบริเวณที่ถูกกัด และยังก่อให้เกิดอาการแพ้เฉพาะที่ได้ นอกจากนี้ความรู้ความเข้าใจในแง่อณูชีววิทยาของริ้นดำยังมีอยู่น้อย ในการศึกษาครั้งนี้ จึงทำการวิเคราะห์ทรานสคริปโตมและโปรติโอมของผลผลิตจากต่อมน้ำลายของริ้นดำ *S. nodosum* เพศเมีย โดยได้สุ่มเลือกโคโลนีที่มาจากต่อมน้ำลายจำนวน 128 โคโลนี ไปทำการหาลำดับนิวคลีโอไทด์ และจัดกลุ่มตามค่าความเหมือนของแต่ละสายนิวคลีโอไทด์ ซึ่งสามารถแบ่งได้เป็น 36 กลุ่ม ทั้งนี้ยังจัดกลุ่มตามประเภทของโปรตีนที่คาดว่าได้จากการแปลรหัสของสายนิวคลีโอไทด์ดังกล่าว เป็น 4 ประเภทได้แก่ โปรตีนที่ถูกหลั่งออกมา โปรตีนที่เกี่ยวกับเซลล์ โปรตีนที่ไม่ทราบชื่อแต่พบในฐานข้อมูลทางด้านพันธุกรรม และโปรตีนที่ไม่พบในฐานข้อมูลทางด้านพันธุกรรม โดยมีสายนิวคลีโอไทด์จำนวน 7 กลุ่ม ที่เมื่อถอดรหัสแล้วได้โปรตีนที่จัดอยู่ในประเภทโปรตีนที่ถูกหลั่งออกมา ซึ่งมีบทบาทเกี่ยวกับกระบวนการแข็งตัวของเลือด ปฏิกิริยาระหว่างโฮสต์กับโปรตีน การหล่อลื่นและการกินอาหาร มีสายนิวคลีโอไทด์จำนวน 9 กลุ่ม ที่เมื่อถอดรหัสแล้วได้โปรตีนที่จัดอยู่ในประเภทโปรตีนที่เกี่ยวกับเซลล์ ซึ่งทำหน้าที่เกี่ยวกับการสังเคราะห์โปรตีน การเจริญเติบโต และหน้าที่อื่นที่เกี่ยวกับเซลล์ นอกจากนี้ยังมีสายนิวคลีโอไทด์อีก 15 กลุ่ม และ 5 กลุ่ม ซึ่งเมื่อถอดรหัสแล้วพบว่าเป็นโปรตีนที่ไม่ทราบชื่อแต่พบในฐานข้อมูลทางพันธุกรรม และโปรตีนที่ไม่พบในฐานข้อมูลทางพันธุกรรม ตามลำดับ จากการวิเคราะห์ด้วย SDS-PAGE พบว่าในต่อมน้ำลายริ้นดำเพศเมียนี้มีแถบโปรตีนหลัก 7 แถบ และยังมีแถบโปรตีนย่อยอีกหลายแถบ แต่อย่างไรก็ตามการหาลำดับกรดอะมิโนส่วนปลายด้านอะมิโนของแถบโปรตีนหลักนั้น ไม่ได้ลำดับกรดอะมิโนที่เป็นประโยชน์ ผลที่ได้จากการศึกษาครั้งนี้ ทำให้ได้ข้อมูลพื้นฐานทางด้านอณูชีววิทยา สำหรับการนำไปใช้ศึกษาบทบาทที่เป็นไปได้ของผลผลิตจากต่อมน้ำลายของ *S. nodosum*

ภาษาอังกฤษ:

Simulium nodosum black fly is one of suspected vectors of veterinary filariasis in Thailand. Their bites cause bleeding at the sites and local hypersensitivity reactions. Little is known about molecular aspects of its biology. In this study, transcriptomic and proteomic analyses were performed on the salivary gland products of female *S. nodosum*. A total of 128 randomly selected, salivary gland-derived cDNAs were sequenced and assembled based on their identity into 36 clusters. The putative translated proteins were classified into four categories: secreted proteins, cellular proteins, conserved unknown proteins, and unknown proteins. Seven clusters encoded probably secreted proteins involving homeostasis, host interaction, lubricating and feeding. Nine clusters encoded cellular proteins associated with protein synthesis, signal transduction and other cellular function. Fifteen and five clusters encoded conserved unknown proteins and unknown proteins, respectively. SDS-PAGE analysis revealed that at least 7 major and several minor protein bands were found in the female salivary glands. However, amino-terminal sequencing of some major protein bands did not yield useful amino acid sequences. These results obtained in this study provide basic information of molecular biology for further study of possible role of these salivary gland products in *S. nodosum*.

EXECUTIVE SUMMARY

Hematophagous black flies (Diptera: Simuliidae) cause a wide range of problems for human and other vertebrates. In humans, clinical manifestations of simuliosis range from localized hypersensitivity reactions to life-threatening systemic hemorrhagic syndromes. Moreover, black flies serve as vectors of *Onchocerca volvulus*, the causative agent of human onchocerciasis or "river blindness" in Africa, Latin America and Arabian peninsular, and other *Onchocerca* spp. In parts of USA, Canada and Europe these flies are a long-standing economic problem for live-stock farms. The fly bites cause severe symptoms or even death through a hypersensitization induced by zootoxins according to the biting density of the fly and the host animal's physical and immunological condition.

During blood feeding, black flies introduce many biologically active molecules in saliva into the feeding lesion that enable them to obtain blood meals while avoiding the host defense of hemostasis, inflammation, and immunity. Salivary components of black flies have been identified to date, including apyrase , anticoagulation factor Xa, anticoagulation factor V, antithrombin of *S. vittatum* Zetterstedt, *S. argus* Williston, *S. ochraceum* Walker (species A), *S. metallicum* Bellardi sensu lato, and *S. exiguum* Roubaud (Cayapa cytotype); a variety of bioamines, including histamine of *S. erythrocephalum* de Geer, *S. lineatum* Meigen and *S equinm* Linne; SV erythema protein of *S. vittatum* Zetterstedt; and hyaluronidase of *S. vittatum* Zetterstedt. So far, only complementary DNA (cDNA) encoding SV erythema protein has been cloned and characterized.

In the past five years, strategy to understand and study salivary gland proteins and genes has changed from testing one salivary molecule at a time, to studying the whole complex of genes and secreted proteins in blood-feeding arthropods using a comprehensive approach. This approach combines massive sequencing protocol of high quality full-length salivary gland cDNA libraries, a proteomic approach to isolate a large set of salivary proteins, and high-throughput computational biology and functional assays to analyze and test the biologic activities of these novel molecules. By the approach, several sets of salivary transcriptomes and proteomes of ticks, mosquitoes and sand flies have been reported. However, little is known regarding the transcriptome and proteome in salivary glands of black flies particularly simuliid species found in Thailand.

Recently, at least 56 black-fly species has been identified in Thailand. At Ban Pan Fan, Chiang Mai Province in northern Thailand, a considerable number of *S. nodosum* Puri are attracted to feed on both humans and buffalos and 2.4% of them are naturally infected with developing filarial larvae. These larvae resemble *O. suzukii*, a parasite from a wild Japanese bovid, suggesting that an unknown *Onchocerca* species from ruminants has been transmitted in Thailand. *S. nodosum* has also been found at Doi Inthanon, Chiang Mai province, one of the most tourist attract places in Thailand. Because of its medical importance, we therefore would like to obtain an insight into salivary transcriptome and proteome of this simuliid species. Generation of a set of *S. nodosum* salivary gland cDNAs and deduced proteins will provide indispensable tools for the systematic and comprehensive analysis of molecules that may play an active role in simulium blood feeding, host immunity and parasite transmission.

In this study, transcriptomic and proteomic analyses were performed on the salivary gland products of female *S. nodosum*. A total of 128 randomly selected, salivary gland-derived cDNAs were sequenced and assembled based on their identity into 36 clusters. The putative translated proteins were classified into four categories: secreted proteins, cellular proteins, conserved unknown proteins, and unknown proteins. Seven clusters encoded probably secreted proteins involving homeostasis, host interaction, lubricating and feeding. Nine clusters encoded cellular proteins associated with protein synthesis, signal transduction and other cellular function. Fifteen and five clusters encoded conserved unknown proteins and unknown proteins, respectively. SDS-PAGE analysis revealed that at least 7 major and several minor protein bands were found in the female salivary glands. But amino-terminal sequencing of some major protein bands did not yield useful amino acid sequences. These results obtained in this study provide basic information of molecular biology for further study of possible role of these salivary gland products in *S. nodosum*.

เนื้อหางานวิจัย

CHAPTER 1

INTRODUCTION

Hematophagous black flies (Diptera: Simuliidae) cause a wide range of problems for human and other vertebrates. In humans, clinical manifestations of simuliosis range from localized hypersensitivity reactions (Frese and Thiel, 1974; Farkas, 1984; Hoffman, 1987) to life-threatening systemic hemorrhagic syndromes (Pinheiro et al., 1974). Moreover, black flies serve as vectors of *Onchocerca volvulus*, the causative agent of human onchocerciasis or "river blindness" in Africa, Latin America and Arabian peninsular, and other *Onchocerca* spp (Crosskey, 1990; Davies and Crosskey, 1991; Mead et al., 1997; Stalling et al., 2002). In parts of USA, Canada and Europe these flies are a long-standing economic problem for live-stock farms. The fly bites cause severe symptoms or even death through a hypersensitization induced by zootoxins according to the biting density of the fly and the host animal's physical and immunological condition (Stokes and Arbor 1914; Millar and Rempel, 1944; Farkas, 1984).

During blood feeding, black flies introduce many biologically active molecules in saliva into the feeding lesion that enable them to obtain blood meals while avoiding the host defense of hemostasis, inflammation, and immunity. Salivary components of black flies have been identified to date, including apyrase (Cupp et al., 1995), anticoagulation factor Xa, anticoagulation factor V, antithrombin of *S. vittatum* Zetterstedt, *S. argus* Williston, *S. ochraceum* Walker (species A), *S. metallicum* Bellardi sensu lato, and *S. exiguum* Roubaud (Cayapa cytotype) (Jacobs et al., 1990; Abebe et al., 1994; 1996); a variety of bioamines, including histamine of *S. erythrocephalum* de Geer, *S. lineatum* Meigen and *S equinm* Linne (Wirtz, 1988; 1990); SV erythema protein of *S. vittatum* Zetterstedt (Cupp et al., 1998); and hyaluronidase of *S. vittatum* Zetterstedt (Ribeiro et al., 2000). So far, only complementary DNA (cDNA) encoding SV erythema protein has been cloned and characterized.

In the past five years, strategy to understand and study salivary gland proteins and genes has changed from testing one salivary molecule at a time, to studying the whole complex of genes and secreted proteins in blood-feeding arthropods using a comprehensive approach. This approach combines massive sequencing protocol of high quality full-length salivary gland cDNA libraries, a proteomic approach to isolate a large set of salivary proteins, and high-throughput computational biology and functional assays to analyze and test the biologic activities of these novel molecules. By the approach, several sets of salivary transcriptomes and proteomes of ticks, mosquitoes and sand flies have been reported. However, little is known regarding the transcriptome and proteome in salivary glands of black flies particularly simuliid species found in Thailand.

Recently, at least 56 black-fly species has been identified in Thailand. At Ban Pan Fan, Chiang Mai Province in northern Thailand, a considerable number of *S. nodosum* Puri are attracted to feed on both humans and buffalos and 2.4% of them are naturally infected with developing filarial larvae. These larvae resemble *O. suzukii*, a parasite from a wild Japanese bovid, suggesting that an unknown *Onchocerca* species from ruminants has been transmitted in Thailand. *S. nodosum* has also been found at Doi Inthanon, Chiang Mai province, one of the most tourist attract places in Thailand (Choochote et al., 2005). Because of its medical importance, we therefore would like to obtain an insight into salivary transcriptome and proteome of this simuliid species. Generation of a set of *S. nodosum* salivary gland cDNAs and deduced proteins will provide indispensable tools for the systematic and comprehensive analysis of molecules that may play an active role in simulium blood feeding, host immunity and parasite transmission.

The objectives of this proposed study were as follows;

(1) To analyze the salivary gland proteins of *S. nodosum* using SDS-PAGE.

(2) To obtain N-terminal sequences of the major salivary gland proteins.

(3) To analyse and generate a set of *S. nodosum* salivary cDNA sequences (transcriptome).

CHAPTER 2

LEITERATURE REVIEWS

The salivary glands of arthropods play an important role in food ingestion and digestion as well as transmission of pathogens. The saliva produced by the arthropod salivary glands is the vehicle that carries pathogens and it may also enhance or facilitate infectivity during the blood meal (Stark and James, 1996a). Biochemically active protein molecules contained in the saliva counteract vertebrate hemostasis, allowing the arthropods to feed successfully (Ribeiro and Francischetti, 2003). In addition, the saliva contains immunogens that cause human IgE and lymphocyte-mediated hypersensitivities (Arlian, 2002).

Almost all female arthropods feed on blood. A blood meal is required for the promotion of ovarian development in anautogenous arthropods. They often probe many times, sometimes changing site and even host, before locating a blood vessel. Once located, they begin ingestion as rapidly as possible prior to the initiation of hemostasis by the host.

Hemostasis is an important phenomenon, where platelet aggregation, vasoconstriction and blood coagulation act together to prevent blood loss from injured tissues. Therefore, a major role of the salivary glands is to secrete a number of powerfully acting proteins that counteract components of the vertebrate hemostatic response.

Platelet aggregation is the first line of defense in avoiding blood loss during tissue injury. Platelets activated by diverse agonists including thrombin, collagen, and adenosine diphosphate (ADP), aggregate, promote clotting, and release vasoconstrictor substances to form the platelet plug. Haematophagus arthropods inhibit the aggregation of platelets in diverse ways. For example, by inhibiting the action of thrombin, as in the case of the antihrombin anophelin from the saliva of *Anopheles albimanus*, and the effect of collagen in platelets, where *Rhodnius prolixus* aggregation inhibitors are the first lipocalin described that inhibits collagen-induced platelet aggregation by binding to ADP (Francischetti et al., 2000). A molecule with similar sequence and structure (pallidipin) was described in the saliva of *Triatoma pallidipennis* (Noeske-Jungblut et al., 1994). A strategy used by most haematophagus arthropods is to block platelet aggregation by targeting ADP via salivary apyrase. Apyrase hydrolyses the phosphodiester bonds of nucleoside tri- and di-phosphates, but not monophosphates. Apyrase is believed to be important for blood-finding because its activity is inversely proportional to the probing time in mosquitoes. Probing duration is defined as the time spent locating vasculature before beginning engorgement. Apyrase from *Ae. aegypti* belongs to the gene family that includes 5'-nucleotidases (Champagne et al., 1995). Apyrases are enzymes that catalyze the degradation of ADP and adenosine triphosphate (ATP) to adenosine monophosphate (AMP) and inorganic phosphates. They do not hydrolyze AMP, which distinguishes apyrase activity from general 5'-nucleotidases. ADP, a component of the granules in platelets, is also released by damaged cells and is one of the important signal molecules that lead to platelet recruitment and platelet aggregation. Examples of platelet inhibitors from arthropods are shown in Table 2.1.

Another challenge to haematophagy is presented by vasoconstriction. In addition to platelet aggregation, vasoconstriction is an important hemostatic response, particularly in a small wound. Every haematophagous arthropod examined secretes a vasodilator. Vasodilatory substances appear to be unique for each group of haematophagous arthropods (Table 2.2). Vasodilators are molecules that increase blood flow by antagonizing vasoconstrictors produced by the hemostatic system following tissue injury from insect mouthparts. They act directly or indirectly on smooth muscle cells activating intracellular enzymes such as adenylate cyclase and guanylate cyclase that lead to the formation of cyclic adenosine monophosphate (cAMP) or cyclic guanosine monophosphate (cGMP). For examples, the salivary gland vasodilator of *Ae aegypti*, the small peptide sialokinin, acts directly on the epithelium activating production of nitric oxide (Champagne and Ribeiro, 1994), which activates a guanylate cyclase in smooth-muscle cells resulting in vasorelaxation. Maxadilan, the vasodilator from the sand fly, *Lutzomyia longipalpis*, binds and activates the pituitary adenylate cyclase activating polypeptide type I receptor resulting in the formation of intracellular cAMP. Another example of a salivary vasodilator is the 60 kDa peroxidase from *An. albimanus*. This enzyme destroys norepinephrine, a potent vasoconstrictor (Ribeiro and Valenzuela, 1999).

The blood-coagulation cascade ends in the production of active thrombin, that cleaves fibrinogen to fibrin, which in turn polymerizes and forms the blood clot. Coagulation at the wound site is evident within 3 – 4 minutes, so the effects of coagulation on haematophagy depend on the feeding characteristics of the arthropod. Ticks feed for several days and have to battle continuously against host hemostasis at the wound site. In contrast, because most mosquitoes and other flies feed in less than four minutes, coagulation probably does not limit their initiation of feeding. However, coagulated blood meals in these insects could cause problems with digestion and processing of blood by limiting access of proteolytic and other digestive enzymes (Stark and James, 1996b). Anticoagulants are possibly necessary to maintain the flow of blood through the mouthparts during feeding. Also, anticoagulants could prevent a clot from forming within a vector-induced haematoma. Anticoagulants vary in their specific activity according to the haematophagous species (Table 2.3). Inhibitors of thrombin, for instance anophelin from the saliva of *An. albimanus* (Valenzuela et al., 1999), and inhibitors of factor Xa such as the salivary serpin from the saliva of *Ae. aegypti* (Stark and James, 1998), are some examples of anticoagulants described in the saliva of blood feeders.

Table 2.1 Examples of platelet inhibitors found in different arthropod groups.

Arthropods	**Platelet inhibitors**	**References**
Mosquitoes		
Aedes aegypti	Apyrase	Ribeiro et al. (1984); Champagne et al. (1995)
Anopheles stephensi	Apyrase	Ribeiro et al. (1985b)
An. freeborni	Apyrase	Ribeiro et al. (1985b)
An. gambiae	Apyrase	Arca et al. (1999)
An. albimanus	Apyrase, anophelin	Francischetti et al. (1999)

Table 2.1 Examples of platelet inhibitors found in different arthropod groups. (Cont.)

Arthropods	**Platelet inhibitors**	**References**
Sand flies		
Lutzomyia longipalpis	Apyrase	Ribeiro et al. (1986); Charlab et al. (1999)
Phlebotomus argentipes	Apyrase	Ribeiro et al. (1989a)
P. perniciosus	Apyrase	Ribeiro et al. (1989a)
P. papatasi	Apyrase	Ribeiro et al. (1989a); Valenzuela et al. (2001)
Blackfly		
Simulium vittatum	Apyrase	Cupp et al. (1995)
Triatomine bugs		
Rhodnius prolixus	Apyrase	Sarkis et al. (1986); Ribeiro et al. (1990)
Triatoma pallidipennis	Novel ADP-binding protein, pallidipin	Noeske-Jungblut et al. (1994)
Ticks		
Ornithodorus moubata	Apyrase	Ribeiro et al. (1991)
Ixodes dammini	Apyrase	Ribeiro et al. (1985a)

Table 2.2 Examples of vasodilators found in different arthropod groups.

Arthropods	Vasodilators	References
Mosquitoes		
Aedes aegypti	Sialokinin I, Sialokinin II	Champagne and Ribeiro (1994)
Anopheles albimanus	Catechol oxidase/peroxidase	Ribeiro and Nussenzveig (1993); Ribeiro and Valenzuela (1999)
An. gambiae	Catechol oxidase/peroxidase	Ribeiro et al. (1994)
Sand flies		
Phlebotomus papatasi	AMP and adenosine	Ribeiro et al. (1999)
Lutzomyia longipalpis	Maxadilan	Ribeiro et al. (1989b); Lerner et al. (1991); Lerner and Shoemaker (1992)
Blackfly		
Simulium vittatum	SV erythma protein	Cupp et al. (1998)

Table 2.2 Examples of vasodilators found in different arthropod groups. (Cont.)

Arthropods	**Platelet inhibitors**	**References**
Triatomine bugs		
Rhodnius prolixus	Nitrophorin	Champagne et al. (1995)
Triatoma infestans	Endothelium dependent vasodilator	Ribeiro et al. (1988)
Ticks		
Ixodes dammini	Prostaglandin, 6-keto-PGF1α (prostacyclin)	Ribeiro et al. (1985a); Ribeiro et al. (1988)

Table 2.3 Examples of anticoagulants found in different arthropod groups.

Arthropods	Anticoagulants	References
Mosquitoes		
Aedes aegypti	Anti-FXa	Stark and James (1995)
Ae. albopitus	Anti-FXa	Stark and James (1996c)
Armigeres subalbatus	Anti-FXa	Stark and James (1996c)
Culex quinquefasciatus	Anti-FXa	Stark and James (1996c)
Anopheles albimanus	Anti-thrombin	Stark and James (1996c)
An. quadrimaculatus	Anti-thrombin	Stark and James (1996c)
An. freeborni	Anti-thrombin	Stark and James (1996c)
An. gambiae	Anti-thrombin	Stark and James (1996c)
An. stephensi	Anti-thrombin	Stark and James (1996c), Waidhet-Kouadio et al. (1998)
Blackflies		
Simulium argus	Anti-thrombin, Anti-FXa	Abebe et al. (1994)
S. metallicum	Anti-FXa	Abebe et al. (1994)
S. ochraceum	Anti-FXa	Abebe et al. (1994)
S. vittatum	Anti-thrombin, Anti-FXa	Abebe et al. (1994); Abebe et al. (1996)
Triatomine bug		
Rhodnius prolixus	Anti-thrombin, Anti-FVIIIa (Prolixin S)	Hellmann and Hawkins (1965); Friedrich et al. (1993)
Ticks		
Ornithodorus moubata	Anti-FXa	Waxman et al. (1990)
Dermacentor andersoni	Anti-FV, Anti-FVII	Gordon and Allen (1991)

During blood feeding, arthropod saliva directly affects the physiology of the host at the bite site, possibly benefiting the parasite or pathogen transmission (Titus and Ribeiro, 1990). Titus and Ribeiro (1998) demonstrated that saliva of the sand fly, *L. longipalpis*, enhanced

Leishmania major infection when the parasite was co-inoculated with sand fly saliva. Similar results were obtained with tick-virus (Jone et al., 1989; Labuda et al., 1993) and mosquito-virus (Edwards et al., 1998; limesand et al., 2000). Based on these observations, the saliva or molecule(s) exacerbating parasite infection might be a target for vaccine development to control pathogen transmission.

Because saliva contains an array of proteins, it is not surprising that repeated bites of blood-sucking arthropods stimulate immune responses to salivary antigens in the vertebrate host. The response is mainly a CD4+ Th1 type induction, which mediates delayed-type reactions, but antibodies are also produced. In addition to homocytotropic antibodies, circulating antibodies reactive with protein components of salivary glands are also induced (Sandeman, 1996). For example, exposures to bites from the *P. papatasi* sand fly in either mice or humans generate strong antibodies to salivary proteins as well as a strong delayed-type hypersensitivity response. This strong cellular reaction is primarily a T cell CD4+-dependent immune response, with accumulation of proinflammatory cells at the site of the bite, 24 – 48 hours after antigen presentation to 'primed' T cells. Paradoxically, this strong cellular response helps the sand fly to probe and feed faster, as the blood flow is greater in delayed-type hypersensitivity sites than in areas of normal skin (Belkaid et al., 2000). Most components of sand fly saliva are highly immunogenic. Volf and Rohousova (2001) demonstrated that 3 species of different phlebotomine subgenera; *P. (Phlebotomus) papatasi*, *P. (Larroussius)*, *perniciosus* and *P. (Adlerius) halepensis*, have species-specific salivary antigens. In total, 4 – 9 species-specific antigens were detected in each species by immunobloting.

In the case of mosquitoes, multiple bites cause immediate and delayed local cutaneous reactions and, occasionally, systemic reaction (McCormack et al., 1995; Peng et al., 1996; Simons and Peng, 1999). The immediate cutaneous reaction consists of a pruritic wheal and surrounding flare that appears within several minutes, peaks at 20 minutes, and than subsides. The delayed reaction consists of an indurated papule that appears a few hours after the bite, reaches a peak at 24 – 36 hours, then diminishes over several days (Peng et al., 1996). There are more than 20 polypeptides in mosquito saliva separated by SDS-PAGE (Racioppi and Spielman, 1987) and 26 – 37 major polypeptides in mosquito whole body extracts (Wu and Lan, 1989). Salivary allergens in 10 mosquito species including *Ae. aegypti*, *Ae. vexans*, *Ae. albopictus*, *Ae. togoi*, *Ae.*

triseriatus, *Cx. quinquefasciatus*, *Cx. pipiens*, *Cx. tarsalis*, *An. sinensis*, and *Culiseta inornata* were analyzed using SDS-PAGE and immunoblotting techniques. Three to sixteen salivary allergens, with molecular masses ranging from 16 to 95 kDa, were found in each species. Species-shared and species-specific allergens that cause IgE responses in subjects allergic to mosquitoes were immunologically identified. Species-shared allergens are the most important for potential use in diagnosis and immunotherapy (Peng et al., 1998).

Many salivary proteins activities have been isolated and characterized following a biochemical approach. However, literature about cDNAs or genes encoding these proteins is limited compared with the variety of activities reported for the saliva of various blood-feeding arthropods (Valenzuela, 2002).

Recently, the revolutionary techniques of molecular biology as well as the miniaturization of high-performance liquid chromatography, mass spectrometry, and Edman degradation techniques allow the exploration of detailed knowledge of such molecules and their role in increasing blood feeding at a rapid pace (Ribeiro and Francischetti, 2003). The genomic and proteomic approaches have resulted in the discovery of genes and proteins not previously reported in blood feeders. These new comprehensive approaches allow not only the isolation of salivary factors involved in hemostasis and inflammation, but also the identification of novel salivary gland proteins, which for many, of which the biological function is unknown. Examples of salivary genes and proteins discovered from different arthropods using these approaches are listed in Table 2.4.

Table 2.4 Examples of novel salivary genes and proteins identified by high-throughput strategies.

Arthropods	**Salivary genes and proteins**	**References**
Mosquitoes		
Aedes aegypti	Three novel members of the D7 family	Valenzuela et al. (2002b)
	Four novel members of the anti 5 family	
	Novel member of 30kDa salivary allergen	
	Protein containing c-type lectin motif	
	Angiopoietin like	
	Sphingomyelin phosphodiesterase	
	Adenosine daeminase	
	Purine nucleosidase	
Anopheles gambiae	Novel member of the SG family	Francisshetti et al. (2002)
	Novel member of the antigen 5 family	
	Novel 30 kDa protein	
	Novel member of the D7 family	
An. stephensi	Antigen 5-related protein	Valenzuela et al. (2003)
	Novel member of the D7 family	
	Mucin like	
	GE rich	
	Novel member of the SG family	
An. darlingi	Antigen 5-related protein	Calvo et al. (2004)
	Member of the D7 family	
	Mucin like	
	Novel member of the SG family	

Table 2.4 Examples of novel salivary genes and proteins identified by high-throughput strategies. (Cont.)

Arthropods	**Salivary genes and proteins**	**References**
Culex pipiens quinquefasciatus	Novel member of the D7 family Antigen 5-related protein Gambicin 30 kDa salivary gland allergen GDE rich	Ribeiro et al. (2004b)
An. funestus	D7 gSG6 family gSG2 family gSG7 family 30-kDa allergen Mucins Anophelin Antigen 5 SG1/Trio family	Calvo et al. (2007)
Ae. albopictus	D7 30-kDa allergen Maltase Glucosidase 56-kDa mosquito family Antigen 5 Mucins Apyrase C-type lectin Serpins	Arca et al. (2007)

Table 2.4 Examples of novel salivary genes and proteins identified by high-throughput strategies. (Cont.)

Arthropods	Salivary genes and proteins	References
Sand flies		
Lutzomyia longipalpis	Adenosine deaminase 5'-nucleotidase/phosphodiesterase Hyaluronidase α-amylase C-type lectin anticoagulant Yellow protein Antigen 5-related protein, Sl1 PpSP15-related protein D7-related protein	Charlab et al. (1999); Valenzuela et al. (2004)
Phlebotomus papatasi	apyrase D7 I D7 II Antigen 5-related protein Yellow protein I Yellow protein II Silk-related protein PpSP12 PpSP14 PpSP15	Valenzula et al. (2001)
Triatomine bug		
Rhodnius prolixus	Lipocalins of the nitrophorin (NP) family Lipocalin of the Triabin family SCP / antigen 5 family Inositol phosphatase	Ribeiro et al. (2004a)

Table 2.4 Examples of novel salivary genes and proteins identified by high-throughput strategies. (Cont.)

Arthropods	Salivary genes and proteins	References
Triatomine bug		
Triatoma brasiliensis	Lipocalins antigen 5 protein Inositol phosphatase Mucin Serine protease	Santos et al. (2007)
Tick		
Ixodes scapularis	Novel member of the salivary 16 kDa protein Novel member of Thrombospondin	Valenzuela et al. (2002a)
Biting midges		
Culicoides sonorensis	Hyaluronidase 5'-nucleotidase D7Bclu1 OBP-like PBP-like Serine protease inhibitor	Campbell et al. (2005)

In blackflies (Diptera: Simuliidae), adult females of some simuliid species have been well known as pests to humans and animals and also vectors of some parasites and pathogens, like *Onchocerca* spp. in humans and cattle and *Leucocytozoon* spp. in birds (Crosskey, 1990; Mead et al., 1997; Stalling et al., 2002). The saliva of blackfiles is responsible for localized hypersensitivity reactions (Frese and Thiel, 1974; Farkas, 1984; Hoffman, 1987) and life-threatening systemic hemorrhagic syndromes in humans (Pinheiro et al., 1974). In animals, the fly bites cause severe symptoms or even death through a hypersensitization induced by zootoxins according to the biting density of the fly and the h animal's physical and immunological condition (Stokes and Arbor 1914; Millar and Rempel, 1944; Farkas, 1984).

Cupp and Cupp (1997) have described the activities and relative abundances of several types of salivary molecules within the genus, *Simulium*, compared with other hematophagous arthropods, and suggested possible roles that these molecules might play in promoting transmission of pathogens and causing disease. For example, salivary glands of *S. ochraceum*, a highly anthropophilic vector of *O. volvulus*, contain greater vasodilator activity than other New World *Simulium*, including *S. metallicum*, a secondary zoophagic vector of human onchocerciasis (Cupp et al., 1994). *S. vittatum* saliva affects immune cell responses and cytokine production. The ability of the saliva to modulate components of the host immune system provides an opportunity for enhancing transmission of pathogens during blood feeding (Cross et al., 1993a, b; 1994).

However, information on *Simulium* salivary components has been limited to general pharmacological observations. Salivary components of black flies have been identified to date, including apyrase (Cupp et al., 1995), anticoagulation factor Xa, antithrombin of *S. vittatum*, *S. argus*, *S. ochraceum*, *S. metallicum*, and *S. exiguum* (Jacobs et al., 1990; Abebe et al., 1994; 1996); a variety of bioamines, including histamine of *S. erythrocephalum*, *S. lineatum*, and *S. equinm* (Wirtz, 1988; 1990); SV erythema protein of *S. vittatum* (Cupp et al., 1998); and hyaluronidase of *S. vittatum* (Ribeiro et al., 2000). So far, only cDNA encoding SV erythema protein has been cloned and characterized (Cupp et al., 1998). Little is known regarding the transcriptome and proteome in the salivary glands of black flies.

In the Oriental region, several man-biting simuliid species have been known, for example, *S. indicum*, which is widely distributed at the southern foot of the Himalayas, from Pakistan in the west, through India and Nepal, to Myanmar and southern China in the east (Lewis, 1974; Datta, 1992), *S. asishi*, *S. himalayense*, *S. nodosum*, *S. tenuistylum* in India (Datta, 1992), and *S. japonicum* in Japan (Takaoka, 1997). In Thailand, at least 56 black-fly species have been identified (Takaoka and Choochote, 2004a, 2005a, b, c, d, e, f, g, h, i). Some of them have been studied in many areas including population cytogenetic studies (Kuvangkadilok et al., 1998; 1999a, b, c; 2003), taxonomic studies (Takaoka and Kuvangkadilok, 1999; Kuvangkadilok and

Takaoka, 2000; Takaoka, 2001; Takaoka and Choochote, 2002; Otsuka et al., 2003; Takaoka and Choochote, 2004a, b), and filarial transmission (Takaoka et al., 2003; Fukuda et al., 2003).

At Ban Pan Fan, Chiang Mai Province in northern Thailand, a considerable number of *S. nodosum* are attracted to feed on both humans and buffalos and 2.4% of them are naturally infected with developing filarial larvae (Takaoka et al., 2003). These larvae resemble *O. suzukii*, a parasite from a wild Japanese bovid, suggesting that an unknown *Onchocerca* species from ruminants has been transmitted in Thailand. *S. nodosum* has also been found at Doi Inthanon, Chiang Mai province, one of the most tourist attract places in Thailand (Choochote et al., 2005). Because of its medical importance, we therefore wish to obtain an insight into salivary transcriptome and proteome of this simulium species. Generation of a set of *S. nodosum* salivary gland cDNAs and deduced proteins would provide indispensable tools for the systematic and comprehensive analysis of molecules that may play an active role in simulium blood feeding, host immunity and parasite transmission.

CHAPTER 3

MATERIALS AND METHODS

3.1 Collection of adult black flies

The breeding place of *S. nodosum* at an open grassland along the stream (6-10 m wide) with moderate current, running in the village of Ban Pan Fan (18° 50'E, 99° 20' N and 250 m in altitude), Doi Saket District, Chiang Mai Province (Takaoka et al, 2003) was the site for adult black fly collection. The collection was performed during 1700-1900 hours, which is the biting peak of *S. nodosum*. Female black flies landing (or try to feed) on or flighting around human attractant was caught by an insect net. They were kept in a paper cup with a pad of cotton wool soaked with 5% sucrose solution placed on the top-screen, and stored in a humid chamber. Then, they were transported to an insectarium of the Department of Parasitology, Faculty of Medicine, Chiang Mai University for species identification. Identification of black fly species was followed the standard keys of Takaoka and Choochote (2004a).

3.2 Salivary gland dissection

Salivary gland dissection was performed as the method described for mosquitoes by Suwan et al. (2002), with some modification. The identified female black flies were cold anaesthetized on ice before salivary gland dissection. Salivary glands of the adult flies (Figure 3.1) were dissected in RNase-free PBS, transferred to a microcentrifuge tube with a small volume of PBS, and then kept at -80°C until use.

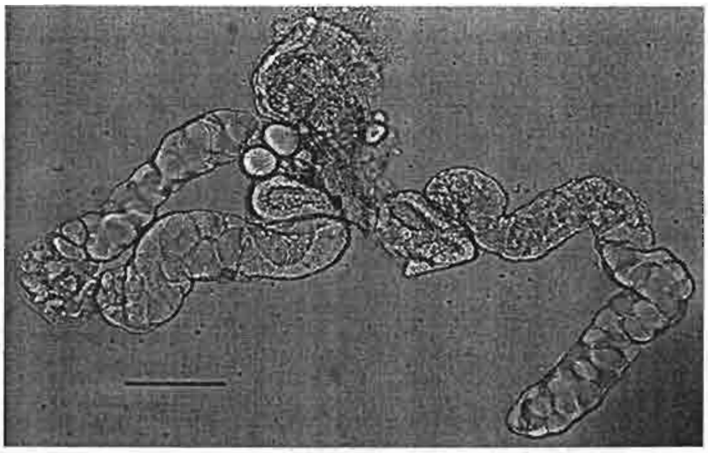

Figure 3.1 Salivary glands of female *Simulium nodosum*. Bar = 100 μm.

3.3 SDS-PAGE and N-terminal sequencing

Salivary gland samples were thawed on ice and mixed in 1:2 (v/v) 1XSDS gel loading buffer [50mM Tris-HCl (pH 6.8), 100mM DTT, 2% (w/v) SDS, 0.1% (w/v) Bromphenol blue, 10% (v/v) glycerol]. Then, the samples were heated for 5 minutes in a bath of boiling water and loaded on 15% SDS polyacrylamide gels. Molecular weight markers (Bio-rad, USA) were applied in each gel. Following the electrophoresis, The gels were stained Coomassie Brilliant Blue (CBB). First, the gels were fixed in 50% methanol and 10% acetic acid for 30 min, then stained with 1% CBB in 10% methanol and 5% acetic acid for 2 hrs, and finally de-stained in 10% methanol and 5% acetic acid until dark protein bands or spots were visible. Digital images of SDS-PAGE CBB-stained gels were captured by scanning at 600 dpi using a color scanner. The images were stored and manipulated in PDF and TIFF formats using Photoshop^{TM} 6.0 graphic software (Adobe Systems Inc., CA, USA).

Proteins that express abundantly in the female salivary glands of the flies were be chosen for N-terminal sequencing. Briefly, fifty female salivary glands (10 glands per lane) were separated on a 15% SDS-polyacrylamide gel. Following transfered to the PVDF membrane, N-terminal sequences were determined by Edman degradation on a Model 471A Protein Sequenator

(Applied Biosystems, Cheshire, UK) at School of Biological Science, the University of Liverpool, Liverpool, United Kingdom.

3.4 Isolation of messenger RNA, construction of cDNA library and DNA sequencing

Approximately 2-3 μg of poly(A) RNA were isolated from 250 pairs of *S. asakoae* female salivary glands using a Micro FastTrack^{TM}2.0 kit (Invitrogen, USA) and used as a template for double-stranded (ds) cDNA synthesis using cDNA Synthesis Kit (Phamarcia Biotech Inc., The Netherlands). A Zero Background^{TM}/Kan Cloning Kit (Invitrogen, USA) was used to construct the female salivary gland cDNA library. *Eco*RI/*Not*I adaptors were added into the blunt-ended ds cDNA before ligating into a pZEro^{TM}-2 vector (Invitrogen, USA) and then 2 μl of ligation mixture (from a total volume of 10 μl) were transformed into TOP 10 Competent cells. Kanamycin was used for colony selection. Transformants per 1 μg of mRNA were calculated. A pool of the bacterial colonies was mixed with glycerol and this cDNA library will be stored at -20°C.

For DNA sequencing, approximately 800 colonies (Ribeiro et al., 2004a, b; Calvo et al., 2004) were picked randomly from the female *S. asakoae* salivary gland cDNA library. Plasmid DNA of each clone was purified using the alkaline lysis method (Sambrook et al., 1989) and then screened with Polymerase Chain Reaction (PCR) method using M13Foward and M13Reverse primers. PCR products were electrophoresed through a 1% agarose gel, and visualized with ethidium bromide staining to determine the size of the inserted plasmid DNA. Recombinant plasmids with an insert size larger than 400 base pair (bp) were purified using the QIAGEN miniprep before sequencing by use of an automated sequencing system at Macrogen Inc., Seoul, Korea.

3.5 Complementary DNA sequence clustering and bioinformatics analysis

Sequence editing and translation were carried out using the DNAStar program. Analysis of sequence data by comparison to the Genbank sequence databases was performed by using the BLAST program at the National Centre for Biotechnology Information (NCBI) server (http://www.ncbi.nlm.nih.gov/). Signal peptides were predicted by submission of sequences to the SignalP server, allowing the identification of putative secretory peptides of each sequence (Nielsen et al., 1997). Sequence alignments were performed using the ClustalW package (Thompson et al., 1997).

CHAPTER 4

RESULTS

4.1 Analysis of the female salivary glands of *Simulium nodosum* : morphological feature, electrophoretic protein profile and N-terminal sequences

Morphological feature

Salivary glands of female *S. nodosum* are paired organs. Each of the paired salivary glands, lying in the dorsal part of the prothorax, consists of a cylindrical U-shaped secretory lobe and a more or less spherical reservoir. The lobe is about 100 μm in cross section at its U-shaped secretory lobe, whereas the reservoir is about 200 μm at its widest diameter. The whole glands of female *S. nodosum* are approximately 1,000 μm in length. The U-shaped secretory lobe is composed of a single layer of large secretory cells. The wall of the reservoir consists of a single layer of small flattened cells. No salivary duct is found (Figure 4.1). Figure 4.2 shows the ultrastructure of the U-shaped secretory lobe.

Figure 4.1 Representative adult salivary glands of female *S. nodosum*. R: reservoir; U: U-shaped secretory lobe. Bar represents 100 μm.

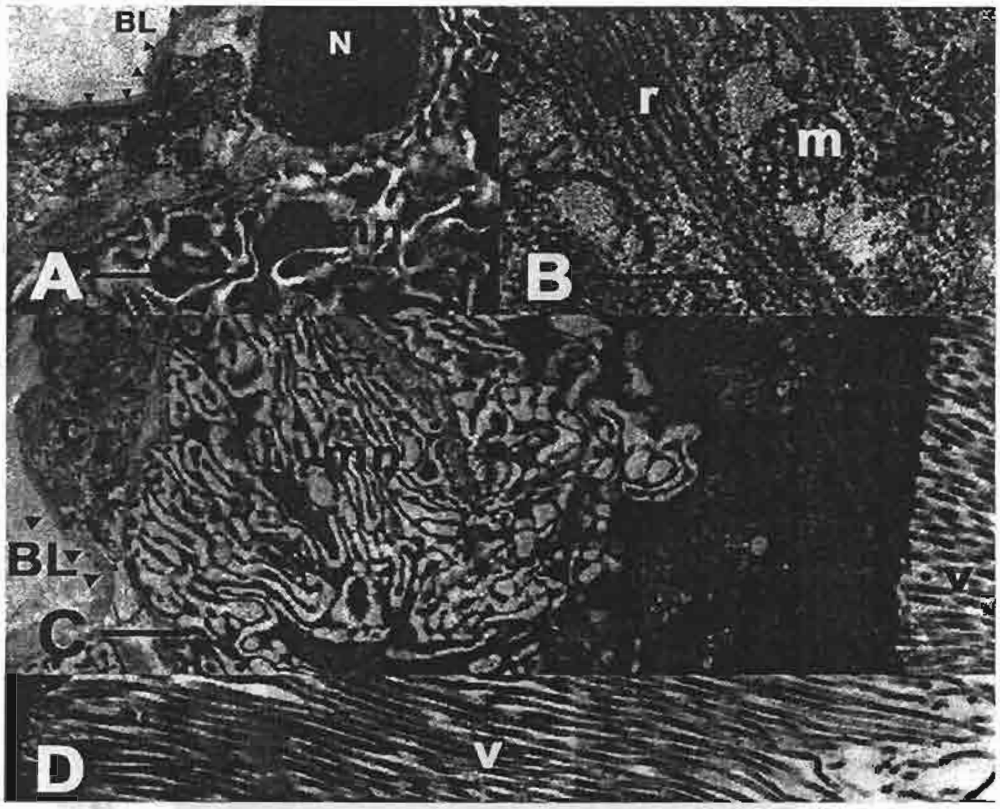

Figure 4.2 Electron micrograph of adult U-shaped secretory lobe of female *S. nodosum*. (A) transverse section showing the nuclei (N) of the cell, basal lamina (BL) (arrowheads) and membrane network (mn). Bar represents 10 μm. (B) high magnification electron micrograph of salivary gland cell cytoplasm. Reticulated rough endoplasmic reticulum (r), and numerous mitochondria (m). Bar represents 1 μm. (C) high magnification electron micrograph of salivary gland cell. Basal lamina (BL) (arrowheads), membrane network (mn) and villi (v). Bar represents 10 μm. (D) high magnification electron micrograph of villi (v). Bar represents 10 μm.

Electrophoretic protein profile and N-terminal sequences

Total proteins in the female salivary glands of *Simulium nodosum* were examined in 15% SDS polyacrylamide gels and stained with silver (Figure 4.3). At least 7 major and several minor protein bands were detected. The molecular masses of these major proteins were estimated at

63.5, 43, 40, 34, 32.5, 31 and 16 kDa. For N-terminal sequencing, proteins from 10 female salivary gland pairs were separated on a 15% SDS-polyacrylamide gel, transfered to PVDF membranes, and stained with Coomassie Brilliant Blue (CBB). Major protein bands, with a molecular mass of 40, 34 and 16 kDa, were submitted for Edman degradation. The major protein band, with a molecular mass of 16 kDa, contained at least 7 amino acids; A or SYFPVDS. However, the N-terminal sequences were too short for use in a comparison with known protein sequences in the GenBank or NR database. Other molecules did not yield successful amino acid sequences.

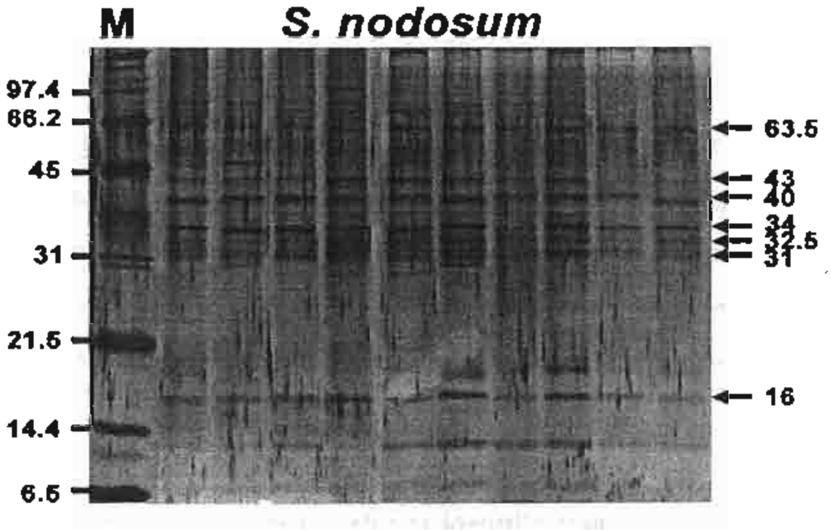

Figure 4.3 Salivary gland protein profiles of female *S. nodosum*. Salivary gland proteins of individual flies were loaded into each lane and then separated on 15% SDS-PAGE and stained with silver. Arrows indicate major protein bands. Molecular weight markers (MW) are on the left.

4.2 A female *S. asakoae* salivary gland cDNA library

A cDNA library was constructed from poly(A) RNA, isolated from 250 salivary gland pairs of female *S. nodosum*. The yield of the poly(A) RNA was determined by spectrophotometry (A260/A280). Two hundred and fifty pairs of salivary glands yielded approximately 2 μg of poly(A) RNA. Double-stranded (ds) cDNA was synthesized using approximately 2 μg of poly(A) RNA as a template. Approximately 500 μl of ds cDNA with *Eco*RI/*Not*I adaptors were collected. Two microliters of the ds cDNA with adaptors were ligated into a pZEro^{TM}-2 vector and all of the insert ligation mixture was transformed into One shot TOP 10 competent cells. The transformation reaction mix was plated onto LB-kanamycin plates. The number of transformats per plate varied from 550 - 600 colonies per 50 μl of the transformation mix, thus, the total number of colonies formed in 300 μl of the transformation mix was estimated at 3,450. Then, transformats per 1 μg of mRNA were calculated. Approximately one microgram of the *S. nodosum* salivary gland mRNA was used in the synthesis of 250 μl of the cDNA with adaptors, and the average number of colonies per 2 μl of cDNA used in the ligation mixture was 3,360. Therefore, the transformation efficiency was (250x3,450)/2. Thus, approximately 4.3×10^5 recombinants should be obtained from a salivary gland cDNA library constructed from 1 μg of mRNA.

4.3 Salivary gland cDNA fragments and cluster identification

In this study, approximately 800 colonies were picked randomly from the female *S. nodosum* salivary gland cDNA library. Plasmid DNA of each clone was purified and screened for size of the cDNA inserts by the Polymerase Chain Reaction (PCR) method using M13F and M13R primers. Three hundred and ninety recombinant plasmids, with an insert size larger than 500 base pairs (bp), were purified and sent for DNA sequencing. However, a total of 128 recombinant plasmids produced readable sequences. A polyadenylation signal sequence, AATAAA, was found at the 3'-terminus of 97.65% of sequences. The length of the cDNA fragments ranged from 266 bp to 813 bp, after removed of the adaptor and vector sequences (trimming). Subsequently, the cDNA sequences were clustered in unique sequences. Only cDNA sequences with 98% identity were clustered together, resulting in 36 clusters of contiguous sequences, including 13 composed by single clones and 23 with two or more sequences.

Therefore, these clusters were considered as putative unigenes. Sequence clusters were classified as SAS_clu01s to SAS_clu13s when referring to clusters with only one sequence reading, or SAS_clu01m to SAS_clu23m when referring to clusters containing more than one assembly reading. Assembled sequences from clusters were used to identify the expressed messages, since they were supposed to be larger than a single cDNA fragment sequence and provided a better source or comparison. Therefore, clusters were compared against GenBank using BLASTx and/or BLASTp. The results showed that from a total of 36 clusters (128 clones), 86.11% of them produced hits to known cDNA and protein sequences.

The identified clusters were separated into the four general functions (categories) expected for *S. nodosum* salivary gland transcripts : proteins probably involved in feeding and host interaction functions ("secreted proteins"); proteins probably involved in cellular functions such as general metabolism, transcription and translation, processing and sorting, degradation, structural function, cell regulation and others ("cellular proteins"); proteins with undetermined functions, but containing a conserved sequence, similar to unknown or hypothetical proteins in a nonredundant (NR) protein data base ("conserved unknown proteins"); and no-match sequences ("unknown proteins"). Table 4.1 presents a general overview of the relative abundance of these groups of clusters. Secreted proteins represented 26.56% of all cDNAs. The cellular transcripts represented 34.38% over the total clones, and the group of conserved unknown transcripts was representative of 32.03% over the total clones. However, all these matching sequences represented only 92.97%, and 7.03% of the remaining sequences, which represented transcripts with no-match sequences in the database. It should be noted that the cDNA library was constructed without any PCR or bacterial amplification, and therefore, the proportion of clones should reflect the ratio of messages in the tissue, as generally accepted in these kind of analyses (Adam et al., 1995).

After the individual analysis of all the clusters, the groups of related sequences were assembled. This assembly was conducted by BLAST, which compared of clusters between each other. The results for the related-cluster groups were compiled as shown in Table 4.2, 4.3, 4.4, and 4.5.

Table 4.1 Repesentation of the 36 clusters assembled from 128 cDNAs of the *S. nodosum* salivary gland

Categories	**Number of clusters**	**Number of clones**	**Redundancy (clones/cluster)**	**Representation over total clones (%)**
1. Similar to secreted proteins	7	34	4.86	26.56
2. Similar to cellular proteins	9	44	4.89	34.38
3. Similar to conserved unknown proteins	15	41	2.73	32.03
4. Unknown proteins	5	9	1.8	7.03
Total	36	128	-	100

4.4 Analysis of probable secreted protein clusters

Searching for known sequence similarity using BLASTx and/or BLASTp revealed 7 clusters of sequences belonging to the secretory category (Table 4.2). Two full-length clusters were obtained, SND_clu01m and SND_clu02m. After determination by the SignalP program, two signal peptides were revealed (Appendix).

Table 4.2 Sequence clusters of salivary gland cDNAs from female *S. nodosum* coding for secreted proteins (secreted protein category)

Group and cluster name	**Number of clones**	**Best match to NR protein database (organism)**	**E-value**	**Comments**
1				
SND_clu01m	2	gb|AAC26163.1| erythema protein SVEP [*Simulium vittatum*]	2e-21	Homeostasis
2				
SND_clu02m	5	gb|AAD31875.1|AF139912_1 odorant-binding protein RpalOBP2 [*Rhynchophorus palmarum*]	0.002	Host interaction
3				
SND_clu03m	5	gb|EAT48295.1| alpha-amylase [*Aedes aegypti*]	5e-69	Feeding
4				
SND_clu04m	7	gb|AAR18428.1| salivary protein [*Culex pipiens quinquefasciatus*]	0.002	Feeding ?
5				
SND_clu05m	11	gb|ABF18029.1| putative salivary mucin [*Aedes aegypti*]	9e-04	Lubricating

6				
SND_clu06m	4	ref|XP_001089922.1| PREDICTED: similar to chymotrypsin-like [*Macaca mulatta*]	0.57	Feeding
7				
SND_clu1s	1	gb|EAT46094.1| venom allergen [*Aedes aegypti*]	6e-10	Feeding ?

SND_clu01m. The SND_clu01m cDNAs were found to be similar to *S. vittatum* erythema protein (SVEP) with 40% identity. The cDNAs consisted of 778 (SND155) and 794 (SND164) nucleotides encoding 148 amino acids. After cleavage of their predicted signal peptide (MRAFIILVVLSSTTFSATTA), a mature protein consisting 128 amino acids, with a calculated molecular mass of 16.46 kDa and the isoelectric point of 6.56, was obtained. Consensus sites for asparagines-linked glycosylation, and N*T and N*S tripeptides (N, asparagines; T, threonine; S, serine; and * to indicate any amino acid) were observed. The deduced SND_clu01m amino acid sequence lacked the glycosylation site. An alignment of the SND_clu01m amino acid sequence to the SVEP is shown in Figure 4.4. Two conserved cysteines (C) were observed in both sequences.

```
SVEP            MSITQSFFVLTLAIFGAAS-DNPIADRKCIVISDGDLVMHERKPGQEFPYYVYMIPKGTE 59
SND_clu01m      MRAFIILVVLSSTTFSATTADSPIADQQCITIQKDNLVMGELKDG-EYATYVYMATKSGQ 59
                *     :.**: : *.*:: *.****::**.*...:*** * * * *:. **** .*. :

SVEP            YDDQRWILESVGGDHYKLKNKFSGRYLVYGTFDYFLTAGAAVREMDHFKFTADGTGKYDI 119
SND_clu01m      YNDQ-WILQSVGNGFYTLKNKNSGRTLGLDSYDYFKTTSDAVGQMSHFQFVPQG-GGYEI 117
                *:** ***:***...*.**** *** *  .::*** *:. ** :*.**:*..:* * *:*

SVEP            SSKANGHPRSRGKNWGVMKDGEKHYFTVENCQE 152
SND_clu01m      INKVAKRITSMGAETMAKLTSDVQYFKVERC-- 148
                 .*.  :  * * :  .   .: :**.**.*
```

Figure 4.4 CLUSTAL alignment of SND_clu01m and *S. vittatum* erythema protein (SVEP) (GenBank accession number AAC26163). Similar amino acid residues are marked with asterisks (*).

SND_clu02m. The SND_clu02m deduced amino acid sequences showed a similarity to an odorant-binding protein RpalOBP2 of *R. palmarum*, a palm beetle, with 23% identity. SND_clu02m cDNAs encoded a 148 amino acid protein [774 (SND153) and 813 (SNd219) nucleotides]. The sequences included a predicted signal peptide (MSHLTLFILLLGLASTLA) and a mature protein with a calculated molecular mass of 16.50 kDa (128 amino acids) and the isoelectric point of 9.61. The deduced SND_clu02m amino acid sequences had 2 concensus glycosylation site, Ans^{43}, Ans^{70}. Alignment of the deduced amino acid sequence of SND_clu02m cDNA with the odorant-binding protein RpalOBP2 is showed in Figure 4.5. Six conserved cysteines were found in both aligned sequences.

```
RpalOBP2        --KFDDSIISDDIKKLLK--GLHDVCVGKIGVEEALIENLKNAEFTEDDKLKCYVHCLLI 56
SND_clu02m      MSHLTLFILLLGLASTLASDGTQPNCVKKAGKQAANIKNWK---FNDSEEVKCYFKCLLI 57
                  ::   *:  .: . *   * :  ** * * : * *:* *   *.:.:::***.:****

RpalOBP2        QVGAMDLAG-HIDAEAAIELIPEQIRVSVIQEANKCAKDKEKIENH-CSRAFATIKCLHD 114
SND_clu02m      SENIMDKNGQNFNGTTFFKFFNTQE---LKGNADRCLTYAQILNAHTCAGSYTIFKCNLD 114
                . . **  * :::. : ::::  *    :  :*::* .  : :: * *: ::: :**  *

RpalOBP2        VNPD------IYYMF------------------- 123
SND_clu02m      ANPEGVKKSMIEYFSKKSKKPKNGGKKKGKKSKV 148
                .** :     * *:
```

Figure 4.5 CLUSTAL alignment of SND_clu02m and odorant-binding protein RpalOBP2 of *R. palmarum* (GenBank accession number AF139912). Similar amino acid residues are marked with asterisks (*).

4.5 Description of probable cellular protein clusters

Three cellular protein clusters in the salivary transcriptome of *S. nodosum* corresponded to genes involved in proteins that encode ribosomal proteins, lectin precursor, protein transporter, elongation factor 1 beta, cell surface protein precursor and beta-agarase precursor (Table 4.3).

Table 4.3 Sequence clusters of salivary gland cDNAs from female *S. nodosum* coding for cellular proteins (cellular protein category)

Group and cluster name	**Number of clones**	**Best match to NR protein database (organism)**	**E-value**	**Comments**
1				
SND_clu07m	4	gb|ABF18250.1| ribosomal protein L3 [*Aedes aegypti*]	3e-122	Protein synthesis
SND_clu08m	6	gb|EAT36458.1| 60S ribosomal protein L14 [*Aedes aegypti*]	4e-06	Protein synthesis
SND_clu09m	4	gb|ABI83774.1| ribosomal protein L12 [*Anopheles funestus*]	2e-76	Protein synthesis
SND_clu10m	6	gb|ABF18043.1| 60S ribosomal protein L38 [*Aedes aegypti*]	4e-31	Protein synthesis
2				
SND_clu11m	5	gb|AAY34343.1| lectin precursor [*Viscum album* subsp. *coloratum*]	0.41	Development ?

3				
SND_clu12m	3	ref|NP_188360.3| protein transporter [*Arabidopsis thaliana*]	1.5	Protein synthesis
4				
SND_clu13m	8	ref|NP_524808.2| Elongation factor 1 beta CG6341-PA [*Drosophila. melanogaster*]	1e-23	?
5				
SND_clu14m	7	ref|NP_786369.1| cell surface protein precursor [*Lactobacillus plantarum* WCFS1]	2e-13	?
6				
SND_clu02s	1	gb|ABL06969.1| beta-agarase precursor [*Vibrio* sp. V134]	0.26	?

4.6 Description of conserved unknown protein clusters

The *S. nodosum* salivary gland cDNA library yielded 4, 2, 1 and 8 clusters of sequences that were similar to putative hypothetical proteins from *Drosophila* spp, *Ae. Aegypti*, *An. gambiae* and other organisms, respectively. These hypothetical proteins had not been characterized in molecular biology and biochemistry (Table 4.4).

Table 4.4 Sequence clusters of salivary gland cDNAs from female *S. nodosum* coding for hypothetical proteins (conserved unknown protein category)

Group and cluster name	**Number of clones**	**Best match to NR protein database (organism)**	**E-value**	**Comments**
1				
SND_clu15m	4	gi|54643131|gb|EAL31875.1| GA21762-PA [*Drosophila pseudoobscura*]	8e-30	Hypothetical protein
2				
SND_clu16m	4	emb|CAG02640.1| unnamed protein product [*Tetraodon nigroviridis*]	1.5	Hypothetical protein
3				
SND_clu17m	4	gi|66772013|gb|AAY55318.1| IP12601p [*Drosophila melanogaster*]	0.003	Hypothetical protein
4				
SND_clu18m	2	gi|54639011|gb|EAL28413.1| GA18044-PA [*Drosophila pseudoobscura*]	2.5 -1	Hypothetical protein
5				
SND_clu3s	1	ref|XP_969643.1| PREDICTED: similar to CG2206-PB, isoform B [*Tribolium castaneum*]	9e-09	Hypothetical protein

6 SND_clu4s	1	ref|XP_341506.3| PREDICTED: similar to FYVE, RhoGEF and PH domain-containing protein 3 [*Rattus norvegicus*]	9.4	Hypothetical protein
7 SND_clu5s	1	gb|AAX33317.1| secondary cell wall-related glycosyltransferase family 8 [*Populus tremula x Populus tremuloides*]	3.6	Hypothetical protein
8 SND_clu6s	1	ref|XP_534613.2| PREDICTED: similar to alpha 3 type VI collagen isoform 1 precursor [*Canis familiaris*]	2e-11	Hypothetical protein
9 SND_clu7s	1	gb|EAT36697.1| conserved hypothetical protein [*Aedes aegypti*]...	5e-45	Hypothetical protein
10 SND_clu19m	5	ref|XP_001369831.1| PREDICTED: hypothetical protein [*Monodelphis*	0.26	Hypothetical protein

		domestica]		
11				
SND_clu20m	8	ref|XP_780578.2| PREDICTED: similar to stromal cell derived factor 2-like protein [*Strongylocentrotus purpuratus*]	3.3	Hypothetical protein
12				
SND_clu8s	1	emb|CAF95522.1| unnamed protein product [*Tetraodon nigroviridis*]	0.65	Hypothetical protein
13				
SND_clu21m	3	gb|EAT34172.1| uncoordinated protein [*Aedes aegypti*]	0.010	Hypothetical protein
14				
SND_clu9s	1	ref|XP_314002.3| ENSANGP00000010049 [*Anopheles gambiae* str.]	2e-25	Hypothetical protein
15				
SND_clu22m	2	gb|AAM75007.1| GH04057p [Drosophila melanogaster]	3e-18	Hypothetical protein

4.7 Description of unknown protein clusters

This study sequenced 9 clones, which were included in 5 clusters (Table 4.5), and annotated them as unknown sequences. These sequences did not show significant similarity with known proteins and could represent novel proteins, unique to *S. nodosum* salivary glands.

Table 4.5 Sequence clusters of salivary gland cDNAs from female *S. nodosum* coding for unknown proteins (unknown protein category)

Group and cluster name	**Number of clones**	**Best match to NR protein database**	**E-value**	**Comments**
1				
SND_clu23m	5	No significance similarity found	-	-
2				
SND_clu10s	1	No significance similarity found	-	-
3				
SND_clu11s	1	No significance similarity found	-	-
4				
SND_clu12s	1	No significance similarity found	-	-
5				
SND_clu13s	1	No significance similarity found	-	-

CHAPTER 5

DISCUSSION

Insects vary in the nature of their mouthparts and method by which they blood-feed. Vessel-feeding insects have relatively long, tube-like mouths that can penetrate skin to the depth of arterioles and venules, and feed directly from vessels. Another general type has mouthparts that are shorter and unable to cannulate vessels, and thus relies on cutting smaller capillaries to form a pool of blood for feeding. For example, black flies, similar to sand flies, feed on capillaries rather than arterioles and venules as mosquitoes and triatomine bugs do (Ribeiro, 1995a). The mouthparts of black flies comprise of the labrum, mandibles, maxillae and hypopharynx. For blood feeding, they use the mandibles to penetrate the host's skin and open a wound that allows the hypopharynx and labrum to enter along with the maxillay laciniae, which anchor the mouthparts while the mandibles continue to snip the host's flesh (Sutcliffe and McIver, 1984). A subdermal hematoma is thus formed, and the pooled blood is ingested via the food canal that is formed when the mandibles fold over the labral channel. A minute tubercle on the anterior surface of each mandible abuts with the edge of the labral channel. Uptake of blood is facilitated by the tight seal formed by the membranous portion of the mouthparts; these membranes prevent both loss of blood from the wound and entry of air. The actual uptake is effected by two muscular pumps, one in the cibarium and the other in the pharynx, in conjunction with the oral and postpharyngeal constrictors (McIver and Sutcliffe, 1988). During feeding, saliva is released from the salivary glands and flows down the salivary gutter of the hypopharynx. Black flies with short mouthparts may require pharmacological, vasoactive compounds in their saliva to increase blood perfusion in superficial regions of the skin and assist in the location of blood supply. Several salivary components of some black fly species have been identified including apyrase (Cupp et al., 1995), anticoagulation factor Xa, antithrombin (Jacobs et al., 1990; Abebe et al., 1994; 1996), a variety of bioamines (Wirtz, 1988; 1990), SV erythema protein (Cupp et al., 1998), and hyaluronidase (Ribeiro et al., 2000). The salivary secretions promote vasodilation, prevent clotting, and inhibit aggregation of platelets (Cupp and Cupp., 1997).

An analysis of salivary gland proteins of four simuliid species using SDS-PAGE was reported by Cross et al. (1993b). After the analysis on SDS silver-stained gels, salivary gland profiles of *S. metallicum*, *S. ochraceum*, *S. vittatum* and *S. argus* comprised individually separated proteins of 11, 12, 19, and 20 bands, respectively. Comparing the protein profiles of salivary gland extracts revealed that those of the two zoophilic species (*S. vittatum* and *S. argus*) were similar, and emphasized the close relationship between these two species. However, these profiles differed markedly from those observed from salivary gland extracts of the two anthropophilic species (*S. metallicum* and *S. ochraceum*), both qualitatively and quantitatively (i.e. in terms of the number of individual components per µg of total soluble protein). Furthermore, salivary glands extracted from the anthropophilic species contained higher levels of total soluble protein than the zoophilic species. The study suggested that differences in saliva composition may represent modifications of this part of the more anthropophilic species, because of evolutionary adaptation to feeding on hosts. In this study, *S. nodosum*, which was reported as anthropophilic (Choochote et al., 2005), had a small number of proteins detected in the salivary glands (approximately 7 major protein bands). However, the amount and profile of salivary gland proteins of the female *S. nodosum* should be compared with other simuliid species. Moreover, further studies concerning the biochemical and pharmacological activities of individual components of salivary materials in the female *S. nodosum* are required before any further conclusions can be drawn, especially in the enzyme involved in vertebrate hemostatic response (Ribeiro and Francischetti, 2003).

In parallel to the organization of the salivary gland transcriptome of *S. nodosum*, we sought to obtain information on the abundance of proteins in the salivary glands of *S. nodosum* by sequencing the amino-terminal ends. Unfortunately, no protein bands yielded any useful information, either because they were blocked at their amino-terminal ends or the Edman degradation resulted in uninformative low signals. Separation of the salivary gland proteins using two-dimensional gel electrophoresis would successfully allow isolation of interesting proteins. After that, digestion of the proteins with trypsin, following amino acid sequencing to obtain internal peptide sequences, would help to identify and describe the corresponding cDNAs or transcriptomes.

Although many salivary proteins of black flies have been isolated and characterized following a biochemical approach, literature about cDNAs or genes encoding these proteins is limited. In this study, a female salivary gland cDNA library of *S. nodosum* was constructed. Typically, 10^5 clones per 1 µg of mRNA would be sufficient to represent a low abundance of mRNA (Jendrisak et al., 1987). The cDNA library described here was estimated to have $4.3\text{x}10^5$ clones from 1 µg of the salivary gland mRNA. From this calculation and the observation of complete polyadenylation signal sequences 3'- end of all the sequences, it was concluded that the cDNA library constructed was a good representation of the 3'-end enriched terminal genes of *S. asakoae* salivary glands. Although a total of 800 cDNA inserts were sequenced from the *S. asakoae* salivary gland cDNA library, 128 produced readable sequences and the data are summarised in Table 4.1.

To date, only one salivary gland cDNA sequence of *S. vittatum* has been submitted to Genbank (GenBank accession number U94515). The cDNA encodes the protein that has the ability to increase blood perfusion in skin capillaries. Therefore, it was named *Simulium vittatum* erythema protein (SVEP) (Cupp et al., 1998). The full-length SVEP cDNA contains 548 bp, which encode 152 amino acid residues. Post-translational processing produces a mature, secreted protein of 133 residues, with a molecular mass of 15.4 kDa. SVEP is potent in its action on the peripheral circulation, with an end point for induction of visible erythema of approximately 65 pmol in the slaved skin of rabbit. This potency equals that of maxidilan, the erythema protein of sand flies, and thus SVEP is among the most powerful vasodilator compounds yet described.

Cupp et al (1998) produced a recombinant SVEP (rSVEP) in a baculovirus expression system and isolated it to high purity using differential centrifugation and reverse phase high performance liquid chromatography (HPLC). The rSVEP relaxed rabbit aorta preparations when preconstricted with 2 µmoll^{-1} phenylephrine or 25 n moll^{-1}K^{+}, but not with 60 m moll^{-1}K^{+}. Furthermore, the SVEP-induced relaxation response of phenylephrine-constricted aorta was inhibited by glibendamide (10 µmoll^{-1}), suggesting that at least part of its action to relax smooth muscle may result from the opening of ATP-dependent K^{+} channels. Recently, Cupp et al. (2004) tested the hypothesis that the rSVEP would improve healing of surgically created wounds in the skin of beagle dogs. Results showed that healing of open wounds was enhanced by 14% at day 21, indicating a novel new wound treatment, but only if the time 0 dose was given intradermally.