FULL PAPER *Immunology*

Random Sequencing of cDNA Library Derived from Partially-Fed Adult Female *Haemaphysalis longicornis* Salivary Gland

Chie NAKAJIMA[1)], Itabajara da SILVA VAZ Jr.[1,2)], Saiki IMAMURA[1)], Satoru KONNAI[1)], Kazuhiko OHASHI[1)] and Misao ONUMA[1)]*

[1)]*Laboratory of Infectious Disease, Department of Disease Control, Graduate School of Veterinary Medicine, Hokkaido University, Kita 18, Nishi 9, Sapporo, Hokkaido 060–0818, Japan and* [2)]*Centro de Biotecnologia do Estado do Rio Grande do Sul, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, C.P. 15005, 91501–970 Porto Alegre RS, Brazil*

(Received 6 May 2005/Accepted 12 July 2005)

ABSTRACT. A cDNA library was constructed from salivary glands of partially-fed adult female *Haemaphysalis longicornis* (hard tick). Randomly selected clones were sequenced and a total of 633 sequences were analyzed by bioinformatic programs. The sequences were grouped into 213 clusters, with each cluster being considered to be composed of mRNAs derived from the same gene or closely related genes. About 36% of the mRNA sequences showed significant similarity to known proteins in the non-redundant protein database by the NCBI blastx program and appeared to be coding for functional predicted proteins, whereas the remaining 64% had no similar sequences. Two thirds of the predicted proteins were annotated as basic cellular proteins (housekeeping proteins). Among the functional predicted protein sequences, other than the housekeeping proteins, several protease inhibitors including anticoagulants, two metalloproteases and a potential immunosuppressive protein could be identified. These proteins may play important roles during tick feeding and could be novel anti-tick vaccine candidates.

KEY WORDS: cDNA library, EST analysis, *Haemaphysalis longicornis*, salivary gland, tick.

J. Vet. Med. Sci. 67(11): 1127–1131, 2005

Ticks are important exo-parasites in the field of veterinary and human hygiene. They affect animals' health not only by feeding on the blood but also by transmitting a range of pathogens. Currently, acaricidal chemicals are used as the most effective agents to control tick infestation, however, they also give rise to problems associated with chemical contamination of the food chain and environmental pollution. Acquisition of resistance to acaricides by ticks has been observed and is becoming a serious problem [4]. As an alternative strategy, anti-tick vaccines using tick recombinant antigens have been tried and developed [12, 24]. For effective application of this promising technique, searching for more vaccine candidate antigens is advocated.

Blood feeding organisms, including ticks, secrete bioactive substances that modify the hosts' physiological and immunological reactions during infestation [10, 17, 23]. A large proportion of these substances are produced in the salivary glands and secreted into the host during blood feeding. Many different types of bioactive substances have been isolated, with some of them showing characteristic functions of anti-coagulation [6, 9], anti-platelet agglutination [11], anti-inflammation [16] anti-complement [21] or immune suppression [2, 3]. These bio-reactors derived from tick saliva may support the transmission of tick borne disease agents [23].

Since these substances seem to play important roles in tick blood feeding, they could be potential tick vaccine antigens [12, 22]. In this study, we constructed a cDNA library from salivary glands derived from partially fed ticks, *Haemaphysalis longicornis*, with the aim of obtaining information of the expressed proteins during blood feeding. EST (expressed sequence tag) approach followed by bioinformatical analysis was used to survey the profile of expressed genes and to determine likely useful and interesting genes [13]. By this approach, we could obtain several novel tick genes encoding bioactive-like proteins.

*CORRESPONDENCE TO: Prof. ONUMA, M., Graduate School of Veterinary Medicine, Hokkaido University, Kita 18-Jyo Nishi 9-Chome, Kita-ku, Sapporo, Hokkaido 060–0818, Japan.

MATERIALS AND METHODS

Tick salivary gland collection: Adult female ticks (*Haemaphysalis longicornis*) were fed on a rabbit and partially engorged ticks were collected after five days. Thirty pairs of salivary glands were collected and total RNA was extracted by TRIZOL (Invitrogen, U.S.A.) reagent according to the manufacturer's protocol. Dissected salivary glands were washed three times in the ice cold phosphate buffered saline before soaking in TRIZOL reagent to minimize the possibility of host or other organ derived RNA contamination.

cDNA library construction and random sequencing: A cDNA library was constructed using the Creator SMART cDNA Library Construction Kit (Becton, Dickinson and Company, U.S.A.) according to the manufacturer's protocol. Briefly, mRNA was reverse transcribed to cDNA and *Sfi*I site possessing adaptors attached to both prime ends. The cDNA was PCR amplified and digested by *Sfi*I followed by column fractionation. cDNA fragments longer than 400 base pairs were collected and ligated into pDNR-LIB plasmid vectors. The gene ligated plasmid was transfected into DH5α *E. coli* and clones were isolated on a Chloramphenicol selection plate. Colonies were randomly

picked and the plasmid was extracted by alkali mini-prep method. Sequences of the genes were read from 5 prime end by single passage using a specially designed primer (5'-ATA CGA AGT TAT CAG TCG ACG-3'). Sequencing was performed by CEQ2000 (Beckman Coulter Inc., U.S.A.) according to the manufacturer's instruction.

Bioinformatic analysis: Vector and adaptor sequences were removed from the target gene sequence and the sequence quality was examined. Short sequences that were less than 150 base pairs or low quality sequences including more than 5% anonymous bases (which were given by the sequencing program when the sequencer could not call the base A,T,G or C) were omitted. Sequences were compared to those in the non-redundant protein database for homology using the NCBI blastx server program (http://www.ncbi.nlm.gov/BLAST). When the "no hit" result was obtained, the base pair sequence was compared with the nr-nucleotide database through the NCBI blastn server. Conserved domains of the translated sequences were also searched for by the NCBI rpsblast server program in the Conserved Domains Database (CDD). The secretory signal sequence was searched for using the SignalP server program (http://www.cbs.dtu.dk/services/SignalP) [15] for prediction of secretory protein when the 5 prime end of the open reading frame (ORF) of the sequence was obtained. Clustering of the sequences was performed by the standalone blastn program [1] with cut off score at 1E-60. The consensus sequence within a cluster was obtained using CAP3 program [8]. Sequence comparison was carried out by ClustalW [19]. Sequences were served and manipulated by the BioEdit program (http://www.mbio.ncsu.edu/BioEdit) [7].

RESULTS

A total of 826 clones were randomly selected and sequenced out of which 157 clones were short or low quality sequences and therefore omitted. The remaining sequences were subjected to homology search out of which 36 clones were found to be mitocondorial DNA or ribosomal RNA. Six hundred and thirty three sequences were assumed to be mRNA and used for further analysis.

Out of the 633 sequences, 225 clones (36%) had homologous sequences in the nr-protein database (E value < 1E-04) and most of their functions were predicted (Fig. 1). The remaining 408 clones (64%) had no similar sequences in the database and their functions could not be predicted. Among the annotated genes, about two thirds (142 clones) were characteristically housekeeping genes coding for essential proteins for cell survival, e.g. ATP synthase, cytochrome c oxidase, ribosomal protein and histone [22]. The remaining 83 clones appeared to have some other bioactive functions.

Sequences were clustered by blastn algorism with cut off E-value at 1E-60 to reduce redundancy and the 633 clones were grouped into 213 clusters. Each cluster consisted of gene transcripts presumed to be derived from the same gene or closely related genes [22]. Some clusters had many copies suggesting that they were derived from highly expressed genes [13] while most of the clusters had only one or a few copies. Out of the 213 clusters, 157 clusters were singleton and had only one copy.

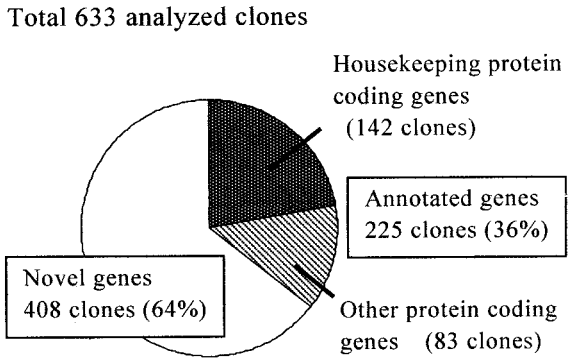

Fig. 1. Results of the homology search for the sequences obtained from the cloned cDNAs. More than 60% of the sequences had no match in the protein/nucleotide database and were assumed to be novel genes. Two thirds of the annotated genes appeared to code for housekeeping proteins. For the remaining 83 annotated genes, further analysis was done as shown in Table 1.

Sequences for which the 5 prime end of the ORF was confirmed were analyzed for the presence of a secretory signal using the SignalP server program. Two thirds of the searched sequences had a signal peptide sequence and were predicted to be encoding secretory proteins. Most of these genes had unknown functions.

The clones which had significantly similar proteins in the databases, other than housekeeping proteins, are shown in Table 1. The most remarkable proteins were protease inhibitors. Out of 83 sequences, 60 sequences were predicted to be having protease inhibitory function.

Proteins predicted to have bioactivity are as follows;

Madanin; thrombin inhibitor: Twenty-four sequences, were annotated as thrombin inhibitor madanin 1 or 2, anticoagulant small proteins derived from *Haemaphysalis longicornis* [9]. The 24 sequences could be divided into three slightly different sequence groups, A, B and C. The 3 sequence groups' comparison with madanin 1 (accession number AAP04349) and 2 (accession number AAP04350) is depicted in Fig. 2. Base pair identities and amino acid similarities among Sequence A, Sequence C, madanin 1 and madanin 2 were around 85% (83–88%) and 60% (56–66%), respectively. Sequence B had a 17 amino acid insert between the 13th and 14th residue of the predicted mature protein for Sequence A and appeared to be an intron variant of Sequence A. The origin of these 5 genes seemed to be same.

Other protease inhibitors: In this study, 17 sequences which coded for Kunitz type protease inhibitor were obtained. One of them (Clone #369) showed high similarity (E-value of 1E-19) to bovine tissue factor pathway inhibitor (TFPI), an extrinsic blood coagulation inhibitor having

Table 1. Predicted functions of expressed proteins (other than housekeeping)

Features (Predicted function)	Number of clones	Annotations (Example of high homology match in blastx)
Protease inhibitor[a)]	60	
	(24)[b)]	Thrombin inhibitor madanin1, 2 [*Haemaphysalis longicornis*]
	(17)	Kunitz type protease inhibitors Tissue factor pathway inhibitor (TFPI), Sea anemone toxin Canine hookworm protease inhibitor, Snake toxin
	(14)	von Willebrand factor [*Ixodes ricinus*], Ixodidine [Boophilus microplus]
	(5)	Trypsin Inhibitor like cysteine rich domain (TIL)
Antigenic protein	5	HL34, HL35 [*Haemaphysalis longicornis*]
Protease	2	Zn metalloprotease [*Ixodes scapularis*]
Immune-suppression	1	Immunosuppressant protein p36 [*Dermacentor andersoni*]
Others	15	Peritorophin like protein, Others
Total	83	

a) Including anticoagulative proteins.
b) Numbers given in parenthesis are indicative of proteins included in 60 protease inhibitors.

```
A          MKHFAILILAVVASAVVMAYPESDSAKDDGNQ----------------EKEKALL
B          MKHFAILILAVVASAVVMAYPESDSAKDDGNQAFKFSVYVPSSQEIYSHEKEKALL
madanin2   MKHFVILILAVVASAVVMAYPERDSAKD-GNQ----------------EKERALL
C          MKHFAILILAVVASAVVMAYPERDSAKD-GNE----------------EQERALP
madanin1   MKHFAILILAVVASAVVMAYPERDSAKE-GNQ----------------EQERALH
           ****.********************* ****: **:                *:*:**

A          VKVQERS--DDGDYDEYDNDETTHTPDPSAPTARPRIREHQA
B          VKVQERS--DDGDYDEYDNDETTHTPDPNAPTARPRIREHQA
madanin2   VKVQERYQGNQGDYDEYDQDETTPPPDPTAQTARPRLRQNQD
C          VNVQERGEVADADYDDYDEEGTTPTVDPTAQTARPRLQGNQS
madanin1   VKVQKR-TDGDADYDEYEEDGTTPTPDPTAPTAKPRLRGNKP
           *:**:*     :.***:*::: ** . **.* **:**:: ::
```

Fig. 2. Sequence alignment of the cloned and translated madanin-like sequences A, B and C compared with madanin 1 (accession number AAP04349) and madanin 2 (accession number AAP04350). Alignment was performed by ClastalW. Amino acid positions of identical, similar and highly conserved substitutes are marked with *, : and ., respectively. Predicted secretory signal peptide position by SignalP is underlined.

Kunitz domains at the active center. This sequence also had significant similarity to protein inhibitors derived from other Ixodidae ticks, *Ixodes ricinus, Boophilus microplus, Ixodes pacificus* and *Amblyomma hebraeum*. Comparison of these sequences with bovine TFPI around the Kunitz domain region is shown in Fig. 3. The number and position of cysteine residues which are important for the activity of the Kunitz domain were found to be conserved [6, 22]. Other putative Kunitz type protease inhibitors obtained in the current study gave significant matches to other TFPIs, a sea-anemone toxin, a canine hookworm protease inhibitor and to snake toxins.

Fourteen clones had significant similarity to other tick proteins named "von Willbrand factor" derived from *Ixodes ricinus* (accession number AAQ01562) and ixodidin derived from *Boophilus microplus* (accession number P83516). Since ixodidin had Chymotrypsin-elastase inhibitor activity, these clones could be predicted to have the same kind of bioactive function.

Five sequences appeared to encode protease inhibitor possessing a trypsin inhibitor like domain (TIL).

HL34 and HL35: HL34 and HL35 genes were previously isolated from a partially fed *Haemaphysalis longicornis* cDNA library by immuno-screening with anti tick saliva serum derived from a rabbit repeatedly fed on by ticks [20]. HL34 was reported to have anti tick efficacy when used as a tick vaccine antigen.

Metalloproteases: Two zinc metalloprotease coding sequences were found. One of them showed significant similarity (E-value of 3E-17) to the metalloprotease coding gene isolated from *Ixodes scapularis* [5]. The function may be related to matrix digestion.

Immunosuppressant protein: A sequence similar to immunosuppressant protein p36 derived from *Dermacentor*

Fig 3. Comparison of Kunitz domain site for the protease inhibitor sequences derived from 5 tick species and bovine TFPI (accession number AAO84035). Cysteine residues are boxed. (Note: Some amino acid residues in B.m. sequence were substituted in X.) Amino acid positions of identical, similar and highly conserved substitutes are marked with *, : and ., respectively. H.l.: *Haemaphysalis longicornis* (clone#369 translated), I.r.: *Ixodes ricinus* (accession number CAB55816), B.m.: *Boophilus microplus* (accession number P83609_1), A.h.: *Amblyomma hebraeum* (accession number AAR97367), I.p.: *Ixodes pacificus* (accession number AAT92116).

andersoni was found (E-value of 8E-5). This protein has been reported to have suppressive activity against T cell proliferative reaction [3]. Homologous genes have also been isolated from *Amblyomma variegatum* (accession number BAD11807) and *Rhipicephalus appendiculatus* [14].

Others: Peritropin like proteins possessing chitin binding domain were also isolated.

DISCUSSION

By the EST analysis of partially fed tick salivary glands, we found some interesting protein-coding genes which seemed to have biological functions other than housekeeping. Among these sequences, there were several types of protease inhibitors, including Kunitz type protease inhibitors and protease inhibitors possessing trypsin inhibitor like domains (TIL). These groups of protease inhibitors are considered to regulate host hemostatic or inflammatory reactions [6, 10]. Most of these sequences had a signal peptide for extra-cellular secretion. Excreted substances from salivary gland presumably posses some special bioactive functions to control host reactions [13, 22]. Other than the annotated sequences, a large portion of the novel genes also possessed the signal peptide indicating that they were coding for secretory proteins possibly having special bioactivities. Sequences of essential proteins for cell survival tend to be conserved beyond the species boundary whereas the structures of species specific proteins may vary. This may be the reason why most of the annotated genes were found to be housekeeping genes. After accumulation of the tick genes' information by analyzing each protein's function, the tick gene annotation using sequence database will become more useful [13].

In the current study, we found a large diversity and redundancy of expressed genes in tick salivary glands during infestation. There were different types of proteins and varied sequences of related protein-coding genes that seem to have similar bioactivities. Our findings were consistent with the transcriptome analysis work previously done on partially fed *Ixodes scarplais* tick salivary glands [22]. This could be because of the significant long duration of tick feeding period compared to other haematophagous arthropods. They need to adapt to the host during the feeding period for up to 2 weeks, while evading the host's immune reaction [18]. A large repertory of exposed antigens may confuse the host's immune system and give the ticks increased chances of survival. Although the idea of making a tick vaccine from salivary proteins would seem reasonable because of the proteins' relevance for tick feeding and pathogen transmission [12, 17], the diversity of the antigens involved could be an obstacle to vaccine development. To overcome this problem, some strategies, for instance searching for crucial and less variant proteins that can be used in combination would be desirable [22]. Several secretory proteins were similar to salivary proteins derived from other tick species and some of them showed significantly high homology. Such conserved type of antigens, when used in a cocktail vaccine, are likely to be effective against a range of different tick species [10].

There could still be a large number of novel bioactive substances which could be obtained from tick salivary glands. These molecules could have possibilities of being new anti-tick drug target candidates as well as tick vaccine candidates. Furthermore, analysis of these molecules could provide some information for the development of new therapeutic drugs, for example, anticoagulants for thrombosis [9]. They may also help our understanding of the host-parasite relationship and basic bioactive substance biochemistry.

ACKNOWLEDGEMENTS. We are grateful to Dr. William H. Witola for critical reading of the manuscript. This work was supported by a grant-in-aid for general scientific research (15380206) from the Ministry of Education, Science, Sports and Culture, Japan.

REFERENCES

1. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389–3402.
2. Anguita, J., Ramamoorthi, N., Hovius, J.W., Das, S., Thomas, V., Persinski, R., Conze, D., Askenase, P.W., Rincon, M., Kantor, F.S. and Fikrig, E. 2002. Salp15, an *ixodes scapularis* salivary protein, inhibits CD4(+) T cell activation. *Immunity* **16**: 849–859.
3. Bergman, D.K., Palmer, M.J., Caimano, M.J., Radolf, J.D. and Wikel, S.K. 2000. Isolation and molecular cloning of a secreted immunosuppressant protein from *Dermacentor andersoni* salivary gland. *J. Parasitol.* **86**: 516–525.
4. Foil, L.D., Coleman, P., Eisler, M., Fragoso-Sanchez, H., Garcia-Vazquez, Z., Guerrero, F.D., Jonsson, N.N., Langstaff, I.G., Li, A.Y., Machila, N., Miller, R.J., Morton, J., Pruett, J.H. and Torr, S. 2004. Factors that influence the prevalence of acaricide resistance and tick-borne diseases. *Vet. Parasitol*. **125**: 163–181.
5. Francischetti, I.M., Mather, T.N. and Ribeiro, J.M. 2003. Cloning of a salivary gland metalloprotease and characterization of gelatinase and fibrin(ogen)lytic activities in the saliva of the Lyme disease tick vector Ixodes scapularis. *Biochem. Biophys. Res. Commun.* **305**: 869–875.
6. Francischetti, I.M., Valenzuela, J.G., Andersen, J.F., Mather, T.N. and Ribeiro, J.M. 2002. Ixolaris, a novel recombinant tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick, *Ixodes scapularis*: identification of factor X and factor Xa as scaffolds for the inhibition of factor VIIa/tissue factor complex. *Blood* **99**: 3602–3612.
7. Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **41**: 95–98.
8. Huang, X. 1992. A contig assembly program based on sensitive detection of fragment overlaps. *Genomics* **14**: 18–25.
9. Iwanaga, S., Okada, M., Isawa, H., Morita, A., Yuda, M. and Chinzei, Y. 2003. Identification and characterization of novel salivary thrombin inhibitors from the ixodidae tick, *Haemaphysalis longicornis*. *Eur. J. Biochem.* **270**: 1926–1934.
10. Mans, B.J. and Neitz, W.H. 2004. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect. Biochem. Mol. Biol.* **34**: 1–17.
11. Mans, B.J., Louw, A.I. and Neitz, A.W. 2002. Savignygrin, a platelet aggregation inhibitor from the soft tick *Ornithodoros savignyi*, presents the RGD integrin recognition motif on the Kunitz-BPTI fold. *J. Biol. Chem.* **277**: 21371–21378.
12. Mulenga, A., Sugimoto, C. and Onuma, M. 2000. Issues in tick vaccine development: identification and characterization of potential candidate vaccine antigens. *Microbes. Infect.* **2**: 1353–1361.
13. Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M.J. and Bishop, R. 2002. AvGI, an index of genes transcribed in the salivary glands of the ixodid tick *Amblyomma variegatum*. *Int. J. Parasitol.* **32**: 1447–1456.
14. Nene, V., Lee, D., Kang'a, S., Skilton, R., Shah, T., de Villiers, E., Mwaura, S., Taylo,r D., Quackenbush, J. and Bisho, R. 2004. Genes transcribed in the salivary glands of female *Rhipicephalus appendiculatus* ticks infected with *Theileria parva*. *Insect. Biochem. Mol. Biol.* **34**: 1117–1128.
15. Nielsen, H., Engelbrecht, J, Brunak, S. and von Heijne G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**: 1–6.
16. Paesen, G.C., Adams, P.L., Harlos, K., Nuttall, P.A. and Stuart, D.I. 1999. Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure. *Mol. Cell* **3**: 661–671.
17. Ribeiro, J.M. 1987. Role of saliva in blood-feeding by arthropods. *Annu. Rev. Entomol.* **32:** 463–478.
18. Sonenshine, D.E. 1993. Biology of Ticks. Oxford University Press, Oxford.
19. Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
20. Tsuda, A., Mulenga, A., Sugimoto, C., Nakajima, M., Ohashi, K. and Onuma, M. 2001. cDNA cloning, characterization and vaccine effect analysis of Haemaphysalis longicornis tick saliva proteins. *Vaccine* **19**: 4287–4296.
21. Valenzuela, J.G., Charlab, R., Mather, T.N. and Ribeiro, J.M. 2000. Purification, cloning, and expression of a novel salivary anticomplement protein from the tick, *Ixodes scapularis*. *J. Biol. Chem.* **275**: 18717–18723.
22. Valenzuela, J.G., Francischetti, I.M., Pham, V.M., Garfield, M.K., Mather, T.N. and Ribeiro, J.M. 2002. Exploring the sialome of the tick *Ixodes scapularis*. *J. Exp. Biol.* 205: 2843–2864.
23. Wikel, S.K. 1999. Tick modulation of host immunity: an important factor in pathogen transmission. *Int. J. Parasitol.* **29**: 851–859.
24. Willadsen, P., Bird, P., Cobon, G.S. and Hungerford, J. 1995. Commercialisation of a recombinant vaccine against *Boophilus microplus*. *Parasitology* **110** (*Suppl.*): S43–50.