Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 349 (2005) 173–185
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Gene expression in the salivary complexes from *Haementeria depressa* leech through the generation of expressed sequence tags

Fernanda Faria[a], Inácio de L.M. Junqueira-de-Azevedo[b], Paulo L. Ho[b,c,d], Misako U. Sampaio[e], Ana M. Chudzinski-Tavassi[a,*]

[a]*Laboratório de Bioquímica e Biofísica, Instituto Butantan, Av. Vital Brazil, 1500, CEP: 05503-900, São Paulo, Brazil*
[b]*Centro de Biotecnologia, Instituto Butantan, São Paulo, Brazil*
[c]*Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil*
[d]*Instituto de Química, Universidade de São Paulo, São Paulo, Brazil*
[e]*Departamento de Bioquímica, Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil*

Received 3 September 2004; received in revised form 15 November 2004; accepted 15 December 2004
Available online 10 March 2005
Received by W. Makalowski

Abstract

A survey of the transcriptional profile of *Haementeria depressa* Ringuelet, 1972 (Annelida, Hirudinea) salivary complexes was produced through expressed sequence tag (EST). Sequences from 898 independent clones were assembled in 555 clusters, representing the transcript profile of this tissue. The repertoire of possible proteins involved in feeding and host interaction processes of the leech corresponded to 10.6% of all identified transcripts (67 clusters), being the carbonic anhydrases (30%), several coagulation inhibitors (25%) and hemerythrin-like molecules (19%), the major components. Among the 387 clusters matching cellular proteins, the majority represents molecules involved in gene and protein expression, reflecting a high specialization of this tissue for protein synthesis. Our *H. depressa* dbEST was also compared to those from other blood-feeding organisms, providing evidences that among the secreted proteins, the coagulation inhibitors present a profile very characteristic of this animal class.
Published by Elsevier B.V.

Keywords: Leeches; Transcriptome; Expressed sequence tag; Blood-feeding; Coagulation inhibitors

1. Introduction

Saliva of blood-sucking animals contains powerful substances able to prevent blood clotting during their feeding. Anticoagulant components of the leech digestive systems have been identified through biochemistry and molecular biology methods (Salzet, 2001). New compounds have been discovered in different families of Hirudinea class, mostly being anticoagulant molecules such as fibrino(geno)lytic enzymes (Chudzinski-Tavassi et al., 1998), inhibitors of coagulation factors (Salzet, 2001; Faria et al., 1999) and platelet aggregation inhibitors (Chudzinski-Tavassi et al., 2003; Salzet, 2001).

Interestingly, sanguivorous leeches have been the focus of pharmaceutical companies seeking to expand the repertoire of anticoagulants, aiming their clinical use in microsurgeries to prevent blood clot formation. Hirudin from *Hirudo medicinalis*, the most studied thrombin inhibitor (Markwardt, 1970) and antistasin from *Haementeria officinalis* (Schaffer et al., 1992), a potent factor Xa inhibitor which forms a new family of inhibitors (antistasin family), can be cited as examples.

Abbreviations: EST, expressed sequence tag; dbEST, database expressed sequence tag; bp, base pair; amp, Ampicilin; ORF, open reading frame; CDD, conserved domain database; HmCRIP, *Hirudo medicinalis* cysteine-rich intestinal protein; *Tt-cysb*, *Teromyzon tessulatum* cystatin b; CA, carbonic anhydrase; LAPP, leech antiplatelet protein precursor; FXIIIa, factor XIII active.

* Corresponding author. Tel.: +55 11 37267222x2109; fax: +55 11 37261505.
E-mail address: amchudzinski@butantan.gov.br (A.M. Chudzinski-Tavassi).

0378-1119/$ - see front matter. Published by Elsevier B.V.
doi:10.1016/j.gene.2004.12.022

Haementeria (Glossiphoniidae, Rhynchobdellida), like *Theromyzon*, *Helobdella*, *Placobdella*, and among others, is a jawless kind of leech. The *Rhyncobdellida* leeches insert a needle-like protrusion called proboscis into the body of the host, secreting some substances that assists in the feeding. The proboscis is linked to two pairs of salivary glands and this feeding apparatus is named salivary complexes (Sawyer, 1986). *Haementeria ghilianii*, *H. officinalis* and *Haementeria depressa* are mainly found in South America (Sawyer, 1986) and they have been widely studied concerning their anticoagulant compounds (Salzet, 2001). *H. ghilianii*, distinct to the others species of *Haementeria* genera, is known as a giant leech offering fundamental advantages for the studies of salivary cell excitation–secretion coupling mechanisms using its isolated cell (Jones et al., 1985). From the salivary complexes of the *H. depressa*, a powerful Factor Xa inhibitor, named lefaxin (Faria et al., 1999), and a nitridergic platelet aggregation inhibitor with fibrino(geno)lytic properties (Hementerin) were characterized (Chudzinski-Tavassi et al., 2003; Chudzinski-Tavassi et al., 1998). The N-terminal of lefaxin presents 68% of similarity with myohemerythrin of the *Nereis diversicolor* (Faria et al., 1999), being probably a member of the hemerythrin family of molecules. Myohemerythrin expressed in salivary complexes after two alimentation stages has also been identified in different tissues of *Theromyzon tessulatum* and is thought to be involved in oxygen transport and storage (Coutte et al., 2001).

Recently, the salivary gland compounds of some blood-feeding arthropods such as insects (*Anopheles* and *Aedes* genera) (Valenzuela et al., 2003; Franscischetti et al., 2002) and ticks (*Ixodes* and *Amblyomma* genera) (Valenzuela et al., 2002; Nene et al., 2002) have been characterized using transcriptome and proteome methods, providing a general vision of the prominent salivary components. More recently, a transcriptomic study of the leech immune response was performed from *T. tessulatum*'s complete body (Lefebvre et al., 2004). However, a transcriptomic analysis by ESTs to focus the mapping of the leech salivary repertoire has not yet been determined.

Here we report the results of the sequencing of 898 clones from a cDNA library from the salivary complexes of the *H. depressa* leeches, presenting a profile of their transcripts. This is the first EST databank determining the repertoire of possible proteins involved in the feeding of a leech. We also compared the leech transcripts with those from the other blood-feeding animals, providing evidences that different anti-haemostasis molecules evolved from different classes of organisms.

2. Materials and methods

2.1. cDNA library construction and EST generation

The salivary complexes (anterior and posterior salivary glands and proboscis) of 60 specimens of *H. depressa* Ringuelet, 1972 kept in captivity in our laboratory were dissected 2 months after the second meal for total RNA extraction by Trizol reagent (Invitrogen), as recommended by the manufacturer's protocol. The integrity of the total RNA was checked by discerning the 28 S and 18 S bands of ribosomal RNA in a formaldehyde denaturing 1% agarose gel. Next, the mRNA purification was performed in a column of oligo-dT cellulose (Amersham-Pharmacia Biotech) following the manufacturer's protocol. The cDNAs were synthesized from 5 μg of mRNA using the Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies), linked to *Eco*RI adapters (Amersham-Pharmacia Biotech) and selected by size (400–800 bp and higher then 800 bp) in agarose gel electrophoresis. The adapter-linked cDNAs were directionally cloned in pGEM11Zf+ (Promega) at 5′*Eco*RI and 3′*Not*I sites. *Escherichia coli* DH5α competent cells were transformed with the two cDNA library plasmids and plated on a 2YT agarose plate containing 100 μg/ml of ampicilin.

To check the library quality, a sample of 20 aleatory clones was chosen and grown overnight in 2YT broth containing 100 μg/ml ampicilin in individual tubes. The plasmid DNA from each culture was extracted with In Concert Plasmid System (Life Technologies) and digested with cloning enzymes to verify the inserts presence and sizes.

For DNA sequencing, random clones were grown in 150 μl of 2YT-amp in 96 well plates for 16 h and stored in glycerol 15% at −80 °C. The 96 well stock plates were replicated in 1 ml CircleGrow medium (Bio 101) containing 100 μg/ml of ampicilin, in 96 well plates for 22 h. The plasmid DNA was isolated using alkaline lyses followed by filtration on MultiScreen MAGVN2250 (Millipore) filtering plates and precipitation of the filtrate by 75% isopropanol. The DNA was dissolved in water and sequenced on an ABI 3100 sequencer (Applied Biosystems) using BidDye2 dideoxyterminators (Applied Biosystems) and the standard M13 forward primer.

2.2. Cluster assembly and identification

The chromatogram file sequences were exported to a database in a Unix based workstation running a set of sequence analysis softwares in a semi-automatic way, as follows: The Phred program (http://www.phrap.com) was used to remove poor quality sequences, with the general parameter of trimming sequences using a window length of 75 bases with 75% of standard quality. The CrossMatch Program was used to remove the adapter and vector sequences from the ESTs. Sequences below 150 bp (after trimming) were also discarded. A final manual examination and base editing of each sequence were also carried out, trying to fix eventual software mistakes.

ESTs were then assembled in clusters of contiguous sequences using CAP3 program (http://genome.cs.mtu.edu/sas.html). The parameters used were basically keeping sequences with 98% or more of base identity in a high-quality