ACCELERATED PUBLICATION
Transcriptomic analysis in the leech Theromyzon tessulatum: involvement of cystatin B in innate immunity

Christophe LEBEBRE*, Claude COQUERELLE*, Franck VANDENBULCKE*, David HOT†, Ludovic HUOT†, Yves LEMOINE*† and Michel SALZET*†
*Laboratoire de Neuroimmunologie des Annélides, UMR CNRS 8017, SN3, Université des Sciences et Technologies de Lille, 59655 Villeneuve d’Ascq, France, and †Laboratoire des Biopuces, Institut Pasteur de Lille, 59019 Lille Cedex, France

At the present time, there is little information on mechanisms of innate immunity in invertebrate groups other than insects, especially annelids. In the present study, we have performed a transcriptomic study of the immune response in the leech Theromyzon tessulatum after bacterial challenge, by a combination of differential display RT (reverse transcriptase)–PCR and cDNA microarrays. The results show relevant modulations concerning several known and unknown genes. Indeed, theeine deaminase, malate dehydrogenase, cystatin B, polyadenylate-binding protein and α-tubulin-like genes are up-regulated after immunostimulation. We focused on cystatin B (stefin B), which is an inhibitor of cysteine proteinases involved in the vertebrate immune response. We have cloned the full-length cDNA and named the T. tessulatum gene as Tt-cysb. Main structural features of cystatins were identified in the derived amino acid sequence of Tt-cysb cDNA: namely, a glycine residue in the N-terminus and a consensus sequence of Gln-Xaa-Val-Xaa-Gly (QXVGX) corresponding to the catalytic site. Moreover, Tt-cysb is the first cystatin B gene characterized in invertebrates. We have determined by in situ hybridization and immunocytochemistry that Tt-cysb is only expressed in large coelomic cells. In addition, this analysis confirmed that Tt-cysb is up-regulated after bacterial challenge, and that increased expression occurs only in coelomic cells. These data demonstrate that the innate immune response in the leech involves a cysteine proteinase inhibitor that is not found in ecdysozoan models, such as Drosophila melanogaster or Caenorhabditis elegans, and so underlines the great need for information about innate immunity mechanisms in different invertebrate groups.

Key words: cDNA microarray, cystatin B, differential display reverse transcriptase (RT)–PCR, innate immunity, leech, Theromyzon tessulatum.

INTRODUCTION
The major characteristic of innate immunity is a fast and efficient response. Its study has generated further interest due to data obtained from invertebrate models [1–5] showing that mammalian innate immunity could be considered as a mosaic of invertebrate immune responses [5]. The main events following septic injury in both vertebrates and invertebrates are based on many mechanisms, such as the recognition of a pathogen’s patterns, activation of specific receptors depending on the nature of the pathogens (Toll-like receptors), phagocytosis, antimicrobial peptide production, systemic response and cellular recruitment. Some mechanisms have been well studied in vertebrates and in a few invertebrates, e.g. the ecdysozoan model Drosophila melanogaster [6–10]. However, experiments performed on various invertebrate models have shown that this response is based on both basic events and specific signalling molecules. These are dependent on the environment, the life cycle and the external anatomy of each animal model.

In the light of these findings, it is important to study other invertebrate models that are different from ecdysozoans, such as annelids, which belong to the lophotrochozoan branch. This will allow a better understanding of the different defence strategies that adaptation has taken during the course of evolution [4].

In this context, the purpose of our study was to identify genes that are regulated after a bacterial challenge in the leech Theromyzon tessulatum. Using a combination of molecular techniques, including differential display (DD) RT (reverse transcriptase)–PCR and DNA microarrays, several immune-system-induced genes have been identified. Among them, we focused our studies on the leech homologue of the mammalian cystatin B gene (named Tt-cysb).

MATERIALS AND METHODS

Animals
T. tessulatum leeches were maintained in our laboratory as described elsewhere [11]. In each experimental set, total RNA of 40 leeches was pooled.

Bacterial challenge
Experimental leeches were injected with a mixture containing Escherichia coli and Micrococcus luteus. After 12 h of culture, centrifuged bacteria were resuspended in PBS, before mixing with PBS (1:1, v/v) to yield a final concentration of 10⁸ bacteria/ml. The leeches were inoculated subepidermally by injection of 2 × 10⁶ bacteria (2 µl). Control animals were injected with a sterile PBS solution (2 µl). All leeches were killed 24 h after injection.

RNA preparation and DD RT–PCR
Total RNA was extracted from complete animals with TRIzol® (Invitrogen) according to the manufacturer’s instructions, before...
DNA image kit, mRNA Differential Display system (GenHunter, Nashville, TN, U.S.A.), according to the manufacturer’s instructions and as described by Liang and Pardee [12]. PCR amplifications were achieved using 24 5 ’-arbitrary oligonucleotides (HAP 1–24), each one combined with three 3 ’-oligo(dT) oligonucleotides (HT11M). Duplicate samples of each set were analysed in order to limit the number of false positives. PCR products were electrophoresed on a denaturing 4.5% polyacrylamide gel (100 W, 2500 V, 50 °C for 3.5 h) using the Genomyx LR electrophoresis system (Beckman Coulter). After autoradiography ( Biomax; Kodak), detected differential products were excised from the gel and eluted in 50 µl of sterile water at room temperature overnight. They were amplified again by PCR with the same set of initial primers, before cloning into pGEM T-easy vector (Promega) before analysis in the DNA microarray control.

Cystatin B cDNA isolation

Control DNA-free total RNA (2 µg) was reverse-transcribed with 200 units of Superscript™ II RNase H − reverse transcriptase (Life Technologies) using 500 ng of an oligo(dT)12−18 under the following conditions: 50 min at 42 °C in 50 mM Tris/HCl, pH 8.3, containing 75 mM KCl, 3 mM MgCl2, 10 mM dithiothreitol and 0.5 mM dNTPs, followed by enzyme inactivation for 15 min at 70 °C.

The Marathon™ cDNA amplification kit (Clontech) was used according to the manufacturer’s instructions in order to obtain the complete 5 ’-end sequence of the cystatin gene [13]. A specific primer was designed from the partial cDNA sequence to clone full-length cDNA: 5 ’-GAGTTCAAGTTGGCCCCAC-3 ’ (a specific reverse primer was used with AP1 Marathon Primer to amplify the 5 ’-end). All PCR products were cloned into the pGEM T-easy vector.

Sequencing

All DNA products were sequenced using the BigDye Terminator v3.0 polymerization kit before detection on the ABI Prism 310 Genetic Analyzer (Applied Biosystems). Sequence analysis used BLAST programs with E = 1000 and no filter [14,15].

Northern blot analysis

Total RNA (15 µg) was diluted in loading buffer [3.7% formamide/20 mM phosphate buffer (pH 7.2)/0.004% ethidium bromide/0.2 µM EDTA/0.025% Bromophenol Blue/0.25% Xylene Cyanol/5% glycerol] and loaded on to a denaturing 1% (w/v) agarose gel [20 mM phosphate buffer (pH 7.2)/7.2% formaldehyde]. Total RNA was electrophoresed according to its size in electrophoresis buffer [20 mM phosphate buffer (pH 7.2)/3.7% formaldehyde] for 3 h at 130 V. Total RNA was then transferred on to an uncharged nylon membrane (Hybond: Amersham Biosciences) by upward capillary transfer overnight. After an incubation for 2 h at 80 °C, the membrane was ready to use.

The full-length Tt-cysb cDNA clone and the partial 18 S rRNA clone were cultured in Luria–Bertani medium (with 100 µg/ml ampicillin). The extracted plasmid DNA was digested with 1 unit of EcoRI and then electrophoresed on a 1% (w/v) agarose gel. DNA fragments were finally purified after excision from the gel (Qiagroup gel extraction; Qiagen). Samples (50 µg) of the fragments were used for the 32P-DCTP-labelled probe synthesis, according to the manufacturer’s instructions (DNA labelling kit, Ready To Go system; Amersham Biosciences). Probes were purified by gel filtration on a G-25 Sephadex column, according to the manufacturer’s instructions (Quick Spin Column; Boehringer Mannheim).

The membrane was pre-incubated for 4 h at 42 °C in 50 ml of hybridization buffer [5 × SSPE (where 1 × SSPE is 0.15 M NaCl/10 mM sodium phosphate, pH 7.4)/1 mM EDTA containing 50% formamide, 5 × Denhardt’s solution (where 1 × Denhardt’s solution is 0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% BSA), 1% SDS and 40 µg/ml of salmon-sperm DNA]. Quantified, labelled DNA was then denatured for 5 min at 95 °C and used at 2 × 106 c.p.m/ml in hybridization buffer (incubated at 42 °C overnight).

Membranes were washed sequentially in 2 × SSC (where 1 × SSC is 0.15 M NaCl/0.015 M sodium citrate)/0.1% SDS, 0.5 × SSC/0.1% SDS and 0.1 × SSC/0.1% SDS for 20, 5 and 5 min respectively, and then exposed for 24–48 h to X-ray film (XAR-5, X-Ommat; Kodak) at −80 °C with an intensifying screen. Finally, autoradiograms were analysed on a densitometer (GS-710 calibrated imaging densitometer; Bio-Rad).

DNA microarrays

Slide preparation

Differential candidates from 300–600 bp were amplified with aminated oligonucleotides complementary to pGEM vector sequences flanking the DNA insert (candidates of a more extensive size are too long to be spotted on the chip). They were also amplified with internal oligonucleotides to generate expected size products up to 600 bp, and were cloned into the pGEM T-easy vector. They were then amplified with aminated oligonucleotides, as described above. Purified PCR products (50 pl of each) were spotted on to microarray slides (Genetix) by a pin-and-ring method (GMS417 Arrayer). The slides were then treated as described by Schena et al. [16].

Fluorescent labelling and hybridization reactions

Total RNA was reverse-transcribed in the presence of Cy3-dCTP or Cy5-dCTP (Amersham Biosciences) as follows: 15 µg of total RNA from naive or challenged animals was incubated at 42 °C for 110 min in 30 µl of SuperScript™ II containing 1 µg of oligo(dT)15 (Hoffmann-La Roche Ltd, Basel, Switzerland), 500 µM each of dATP, dGTP and dTTP, 200 µM dCTP, 100 µM Cy3-dCTP or Cy5-dCTP, 16 units of RNasine, 10 mM dithiothreitol and 400 units of SuperScript™ II reverse transcriptase (Invitrogen). The samples were then treated with 1.5 µl of 1 M NaOH for 10 min at 65 °C to degrade the RNA and purified on a YM 30 Centricon centrifugal filter unit (Millipore) in a final volume of 5 µl. Hybridization buffer (10 µl) was then added to a final concentration of 40% formamide, 2.5 × Denhardt’s solution, 0.5% SDS and 4 × SSC [16]. After incubation at 95 °C for 5 min, the probe was applied on to the microarray slides under a coverslip. They were hybridized with the probe and 30 µl of hybridization buffer for 14–16 h at 42 °C, washed sequentially in 0.1 × SSC/0.5% SDS and 0.1 × SSC, and then dried.

The hybridization intensities for the different channels were normalized using four hop gene controls (see below). Before each labelling reaction, the samples received the four different in vitro-synthesized control RNAs (hop genes), for which four corresponding cDNAs were spotted on to the slide among the bacterial genes.

Microarray analyses

The slides were scanned using the Affymetrix 418 Array Scanner (Affymetrix). Separate images were acquired for each fluorophore.
at a resolution of 10 µm/pixel. To normalize the two channels with respect to signal intensity, the photomultiplier and laser-power settings were adjusted such that the signal ratios of the hop control genes were as close to 1.0 as possible. The average fluorescence intensity was determined by using the ScanAlyze program (http://rana.lbl.gov/EisenSoftware.htm). Background fluorescence was calculated as the median fluorescence signal of non-target pixels around each gene spot. Spots showing a signal value lower than three times the local background in both Cy3 and Cy5 channels were not considered. For every experiment, the DE (differential expression), which represents the degree of induction or repression between the control and the experimental animals, was defined for each of the genes and expressed as the natural logarithm value of the ratio. Statistically significant differences in gene expression between the different sets of animals were assessed by a Student’s t test implemented on the Cyber-T interface (http://www.genomics.uci.edu). In addition, a Bonferroni correction value of 0.25 was applied to all the data before hybridization.

In situ hybridization

Probe synthesis

A plasmid containing the original DD RT–PCR partial cystatin sequence (512 bp) was used as a template for the preparation of the probes. [35S]UTP-labelled antisense and sense probes were generated from linearized cDNA plasmids by in vitro transcription using [35S]UTP (Amersham Biosciences) and an RNA labelling kit (Hoffmann-La Roche Ltd.), according to the manufacturer’s instructions.

Tissue preparation

Control and challenged stage 2 animals were fixed in a solution containing 4 % (w/v) paraformaldehyde in 0.1 M phosphate buffer, pH 7.4, for 16 h. After dehydration, tissues were embedded in Paraplast and 8 µm sections were cut, mounted on to poly(l-lysine)-coated slides and stored at 4 °C until use.

Hybridization

[35S]UTP-labelled riboprobes (100 ng or 1 × 10⁶ c.p.m./slide) were hybridized to tissue sections, as described previously [17,18]. The [35S]UTP hybridization signal was visualized using autoradiography. Samples were coated by dipping in LM1 liquid emulsion (Amersham Biosciences), immediately dried and exposed for a 28-day period. At the end of the exposure period, the autoradiograms were developed in D19b (Kodak), fixed in 30 % (v/v) sodium thiosulphate (10 min at room temperature), stained with Toluidine Blue and mounted with XAM (Merck). Slides were observed under a Zeiss Axioskop microscope, and images were captured with a system monitored with the Leica IM1000 software. Quantification of the radiolabelling at the cellular level was performed using a Zeiss Axiosphot microscope and a BIOCOM quantification system, as described previously [18].

Controls

Controls for in situ hybridization consisted of replacing antisense riboprobe with sense riboprobe. RNase control sections were obtained by adding a pre-incubation step with 10 µg/ml RNase A before hybridization.

Immunocytochemical procedures

Antibodies

The potentially immunogenic region of the Tt-CYSB protein (Tyr-Val; derived from the Tt-cysb gene) was chemically synthesized coupled with ovalbumin, and then used for the immunization of five Balb/c mice, according to the protocol of Agrobio (La Chavannerie, La Ferté St Aubin, France).

Immunofluorescence

Tissue sections obtained as described above were pre-incubated for 1 h in PBS (Tris-buffered saline; 0.1 M Tris/HCl (pH 7.4)/0.9 % NaCl) containing 1 % normal goat serum, 1 % ovalbumin and 0.05 % Triton X-100. Sections were then incubated overnight at 20 °C in PBS containing mouse anti-cystatin B antiserum (1:400 dilution), 1 % normal goat serum, 1 % ovalbumin and 0.05 % Triton X-100. After two washes in PBS, sections were incubated for 2 h with a goat anti-mouse FITC-tagged antiserum (Jackson Immunoresearch Laboratories) diluted 1:100 in PBS. After two more washes, slides were mounted in glycerol containing 25 % TBS and 0.1 % p-phenylenediamine. Labelled cells were observed using a Leica laser-scanning microscope (TCS NT) equipped with a Leica (DMIRBE) inverted microscope and an argon/krypton laser. FITC signal was detected using a 488 nm band pass excitation filter and a 575–640 nm pass barrier filter. Images were acquired sequentially as single transcellular optical sections and averaged over 16 scans/frame.

Controls were incubations of anti-cystatin B immunoserum pre-adsorbed by the 14-amino-acid synthesized peptide [corresponding to the potentially immunogenic region of the Tt-CYSB protein (Tyr-Val; see above)].

RESULTS

Transcriptomic approach of leech innate immune response

Seventy-two primer combinations [of the 24 5′-arbitrary oligonucleotides (HAP 1–24) with HT11A, G and C] were used in the present study. Ninety-one differential bands with a size range from 300 to 1500 bp were excised (Figure 1). Among these 91 products, 64 show an increase in amplification signal (positive candidates) following bacterial challenge, whereas 27 show a decrease in this signal (negative candidates) under the same conditions (Table 1). All the candidate genes were submitted to BLAST analysis (Table 1) [15]. Of these, 25 positive and 10 negative products matched with known sequences; 27 positive and 14 negative products are still unknown after BLAST alignment (Table 1). Twelve positive and 3 negative products did not re-ampify after gel extraction because of their high size and/or low amount of eluted DNA (Table 1). Some identified genes, such as those for ribosomal proteins and aldehyde dehydrogenase, and paramyosin genes, are redundant and correspond to highly expressed genes (Table 1).

Because DD RT–PCR is known to generate many false positives [19], a second screening technique was undertaken with DNA microarrays. All DD RT–PCR candidates were spotted on to glass slides. Total RNA from control or challenged animals was used in order to hybridize fluorescent cDNA probes with spotted candidates. Two positive controls were used, namely the cDNAs of T. tessulatum antimicrobial peptides named theromacin and theromyzin, which are known to be induced after bacterial challenge [20]. Both antimicrobial peptide genes were selected during our statistical analysis and thus validated the DNA microarrays. The statistical analysis confirmed in part the data obtained.
by DD RT–PCR (Table 2). All the products selected by the microarray approach correspond to positive candidates in the DD RT–PCR experiments. No negative candidate was selected. Interestingly, three positive differential products from distinct amplification signals correspond to the threonine deaminase gene. They are involved in processing functions of mammalian antigen-presenting cells [22,23]. They allow class II antigenic peptide binding by degradation of Ii chaperone in the initial class II–Ii complex. Moreover, D. melanogaster cathepsin L (CP1) is present in small granules of haemocytes, and might play a role in phagocytic events [24]. Thus, owing to the major implication of cathepsins in immunity, we focused on the cystatin gene analysis.

The cystatin gene was partially characterized with the DD RT–PCR analysis and its complete cDNA sequence was obtained by PCR amplification from the Marathon® cDNA library (ClonTech). Its sequence showed an open reading frame of 306 nucleotides (Figure 2A). BLAST analysis [15] showed that T. tessulatum cystatin possesses 54% identity with human cystatin B (also named stefin B). The full amino acid sequence of this inhibitor was deduced using the SignalP V1.1 program. The analysis did not reveal the presence of a signal peptide, which is typical of cystatin B. Cystatin B (stefin B) is an intracellular, tight-binding and irreversible inhibitor of papain-like cysteine proteinases, such as cathepsins. Leech cystatin, like other cystatin Bs, exhibits an active site characterized by a glycine residue (Gly5) in the N-terminus and a putative active site located at residues Gln17-Xaa-Val-Xaa-Gly19 (QXVXG). Finally, mammalian cystatins B are unique among cystatins in having a free cysteine residue before the Gly5 residue at the N-terminus. Recent studies have shown that this cysteine is involved in binding to proteinases and seems to be specific for cystatin B [25,26]. The leech cystatin also possesses this cysteine residue. Thus the presence of all these features in the leech inhibitor confirms that we have characterized the leech cystatin B gene (named Tt-cysb), and thus Tt-cysb is the first cystatin B gene known in invertebrates. The Multalin program presents a comparison of Tt-CYSB protein with some known cystatins B (Figure 2B) [27].

According to DD RT–PCR and cDNA microarray analyses, the Tt-cysb gene is overexpressed after a bacterial challenge (Table 2). Tt-cysb mRNA and Tt-CYSB protein localizations were therefore performed using in situ hybridization and immunocytochemistry in control and immunostimulated animals (Figure 3). Interestingly, Tt-cysb is specifically expressed in large circulating coelomocytes, and are distributed throughout the whole animal. The sections hybridized with the sense Tt-cysb probe were totally devoid of labelling (Figure 3B'). In addition, pre-treatment of sections with RNase A before hybridization abolished the positive staining, providing further evidence of the signal specificity (results not shown). To determine whether Tt-cysb could be transcriptionally regulated at the level of the coelomocytes or whether changes of transcript rates could be only the results of changes in cell populations, Tt-cysb mRNA contents were quantified at the cellular level (Figure 3C). Among the tissue sections probed with a 35S-radiolabelled Tt-cysb antisense riboprobe, four different in vitro-synthesized control RNAs (hop genes) were added to the samples. Their relative levels were monitored using the four corresponding cDNAs spotted on to the slide. This normalization allows us to consider whether the experimental changes in expression levels are significant.

Cystatin B gene: a cysteine proteinase inhibitor up-regulated in leech innate immunity

Among the genes induced following septic shock that were detected in the transcriptomic studies, a leech cystatin gene was identified (GenBank® accession number AF542131). In vertebrates, cystatins are known to inhibit the cysteine proteinases belonging to the cathepsin family [21]. In immunity, cathepsins are involved in processing functions of mammalian antigen-presenting cells [22,23]. They allow class II antigenic peptide binding by degradation of Ii chaperone in the initial class II–Ii complex. Moreover, D. melanogaster cathepsin L (CP1) is present in small granules of haemocytes, and might play a role in phagocytic events [24]. Thus, owing to the major implication of cathepsins in immunity, we focused on the cystatin gene analysis.

The cystatin gene was partially characterized with the DD RT–PCR analysis and its complete cDNA sequence was obtained by PCR amplification from the Marathon® cDNA library (ClonTech). Its sequence showed an open reading frame of 306 nucleotides (Figure 2A). BLAST analysis [15] showed that T. tessulatum cystatin possesses 54% identity with human cystatin B (also named stefin B). The full amino acid sequence of this inhibitor was deduced using the SignalP V1.1 program. The analysis did not reveal the presence of a signal peptide, which is typical of cystatin B. Cystatin B (stefin B) is an intracellular, tight-binding and irreversible inhibitor of papain-like cysteine proteinases, such as cathepsins. Leech cystatin, like other cystatin Bs, exhibits an active site characterized by a glycine residue (Gly5) in the N-terminus and a putative active site located at residues Gln17-Xaa-Val-Xaa-Gly19 (QXVXG). Finally, mammalian cystatins B are unique among cystatins in having a free cysteine residue before the Gly5 residue at the N-terminus. Recent studies have shown that this cysteine is involved in binding to proteinases and seems to be specific for cystatin B [25,26]. The leech cystatin also possesses this cysteine residue. Thus the presence of all these features in the leech inhibitor confirms that we have characterized the leech cystatin B gene (named Tt-cysb), and thus Tt-cysb is the first cystatin B gene known in invertebrates. The Multalin program presents a comparison of Tt-CYSB protein with some known cystatins B (Figure 2B) [27].

According to DD RT–PCR and cDNA microarray analyses, the Tt-cysb gene is overexpressed after a bacterial challenge (Table 2). Tt-cysb mRNA and Tt-CYSB protein localizations were therefore performed using in situ hybridization and immunocytochemistry in control and immunostimulated animals (Figure 3). Interestingly, Tt-cysb is specifically expressed in large circulating coelomocytes in control and immunostimulated animals (Figures 3A and 3B). These ‘large coelomocytes’ are only detected in coelomic cavities, and are distributed throughout the whole animal. The sections hybridized with the sense Tt-cysb probe were totally devoid of labelling (Figure 3B’). In addition, pre-treatment of sections with RNase A before hybridization abolished the positive staining, providing further evidence of the signal specificity (results not shown). To determine whether Tt-cysb could be transcriptionally regulated at the level of the coelomocytes or whether changes of transcript rates could be only the results of changes in cell populations, Tt-cysb mRNA contents were quantified at the cellular level (Figure 3C). Among the tissue sections probed with a 35S-radiolabelled Tt-cysb antisense riboprobe,
Table 1  Database search for candidates with significant similarity

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Not reamplified</td>
<td>+</td>
<td>47</td>
<td>No reamplified</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>Actin 2 [CK640381]</td>
<td>+</td>
<td>48</td>
<td>No reamplified</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>Cystatin B [AF542131]</td>
<td>+</td>
<td>49</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>Phosphatidyethanolamine-binding protein [CK640382]</td>
<td>+</td>
<td>50</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>Unknown</td>
<td>+</td>
<td>51</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>Unknown</td>
<td>+</td>
<td>52</td>
<td>Alpha protein coatamer sub-unit [CK640393]</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>Unknown</td>
<td>−</td>
<td>53</td>
<td>Aldolase dehydrogenase [CK640394]</td>
<td>−</td>
</tr>
<tr>
<td>8</td>
<td>Unknown</td>
<td>−</td>
<td>54</td>
<td>Unknown</td>
<td>−</td>
</tr>
<tr>
<td>9</td>
<td>Ribosomal protein 40S S18 [CK640383]</td>
<td>+</td>
<td>55</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>Unknown</td>
<td>−</td>
<td>56</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>11</td>
<td>Ribosomal protein 40S S4 [CK640384]</td>
<td>+</td>
<td>57</td>
<td>Peptidyl-prolyl cis–trans isomerase [CK640396]</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>Unsequenced</td>
<td>−</td>
<td>58</td>
<td>Lipidic acid synthetase [CK640396]</td>
<td>−</td>
</tr>
<tr>
<td>13</td>
<td>Ribosomal protein 40S S4 [CK640385]</td>
<td>−</td>
<td>59</td>
<td>Unknown</td>
<td>−</td>
</tr>
<tr>
<td>14</td>
<td>Fructose-bisphosphate aldolase 1 [CK640386]</td>
<td>−</td>
<td>60</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>15</td>
<td>Unknown</td>
<td>−</td>
<td>61</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>Unknown</td>
<td>−</td>
<td>62</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>No reamplified</td>
<td>+</td>
<td>63</td>
<td>Unknown</td>
<td>−</td>
</tr>
<tr>
<td>18</td>
<td>Ribosomal protein 40S S4 [CK640387]</td>
<td>+</td>
<td>64</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>19</td>
<td>Ribosomal protein 40S S4 [CK640387]</td>
<td>+</td>
<td>65</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>20</td>
<td>PABP1 [CK640388]</td>
<td>+</td>
<td>66</td>
<td>Cullin-3 [CK640397]</td>
<td>−</td>
</tr>
<tr>
<td>21</td>
<td>Not reamplified</td>
<td>−</td>
<td>67</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>22</td>
<td>Unsequenced</td>
<td>−</td>
<td>68</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>23</td>
<td>Unknown</td>
<td>−</td>
<td>69</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td>Unknown</td>
<td>−</td>
<td>70</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>25</td>
<td>Unknown</td>
<td>−</td>
<td>71</td>
<td>Unknown</td>
<td>−</td>
</tr>
<tr>
<td>26</td>
<td>Unknown</td>
<td>+</td>
<td>72</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>27</td>
<td>Unknown</td>
<td>+</td>
<td>73</td>
<td>Aldehyde dehydrogenase [CK640398]</td>
<td>−</td>
</tr>
<tr>
<td>28</td>
<td>Unknown</td>
<td>+</td>
<td>74</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>29</td>
<td>Unknown</td>
<td>+</td>
<td>75</td>
<td>Growth hormone inducible transmb protein [CK640399]</td>
<td>+</td>
</tr>
<tr>
<td>30</td>
<td>Unknown</td>
<td>+</td>
<td>76</td>
<td>Growth hormone inducible transmb protein [CK640400]</td>
<td>+</td>
</tr>
<tr>
<td>31</td>
<td>Unknown</td>
<td>−</td>
<td>77</td>
<td>Threonine deaminase [CK640401]</td>
<td>+</td>
</tr>
<tr>
<td>32</td>
<td>Unknown</td>
<td>+</td>
<td>78</td>
<td>Ribosomal protein 40S SA [CK640402]</td>
<td>+</td>
</tr>
<tr>
<td>33</td>
<td>Unknown</td>
<td>+</td>
<td>79</td>
<td>α-Galactosidase A [CK640403]</td>
<td>−</td>
</tr>
<tr>
<td>34</td>
<td>Not reamplified</td>
<td>+</td>
<td>80</td>
<td>Paranyisin [CK640404]</td>
<td>−</td>
</tr>
<tr>
<td>35</td>
<td>Ribosomal protein 60S L3 [CK640389]</td>
<td>+</td>
<td>81</td>
<td>Paranyisin [CK640405]</td>
<td>−</td>
</tr>
<tr>
<td>36</td>
<td>Ribosomal protein 60S L3 [CK640389]</td>
<td>+</td>
<td>82</td>
<td>Paranyisin [CK640405]</td>
<td>−</td>
</tr>
<tr>
<td>37</td>
<td>Elongation factor 1-α 2 [CK640390]</td>
<td>+</td>
<td>83</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>38</td>
<td>Malate dehydrogenase [CK640391]</td>
<td>+</td>
<td>84</td>
<td>ATP-dependent RNA helicase [CK640406]</td>
<td>+</td>
</tr>
<tr>
<td>39</td>
<td>Ribosomal protein 60S L3 [CK640392]</td>
<td>−</td>
<td>85</td>
<td>Actin 2 [CK640407]</td>
<td>+</td>
</tr>
<tr>
<td>40</td>
<td>Not reamplified</td>
<td>+</td>
<td>86</td>
<td>α-Tubulin [CK640408]</td>
<td>+</td>
</tr>
<tr>
<td>41</td>
<td>Not reamplified</td>
<td>+</td>
<td>87</td>
<td>Tropomyosin 2 [CK640409]</td>
<td>−</td>
</tr>
<tr>
<td>42</td>
<td>Not reamplified</td>
<td>+</td>
<td>88</td>
<td>Threonine deaminase [CK640410]</td>
<td>+</td>
</tr>
<tr>
<td>43</td>
<td>Not reamplified</td>
<td>+</td>
<td>89</td>
<td>Threonine deaminase [CK640411]</td>
<td>+</td>
</tr>
<tr>
<td>44</td>
<td>Not reamplified</td>
<td>+</td>
<td>90</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>45</td>
<td>Not reamplified</td>
<td>+</td>
<td>91</td>
<td>Unknown</td>
<td>+</td>
</tr>
<tr>
<td>46</td>
<td>Ribosomal protein 60S L3 [CK640392]</td>
<td>+</td>
<td>92</td>
<td>Unknown</td>
<td>+</td>
</tr>
</tbody>
</table>

30 large coelomocytes from four individual animals (120 cells) were analysed at each time post-injection. Quantification was expressed as AU (arbitrary units) corresponding to the number of silver grains counted for every large coelocyte using Biocom autoradiography software. Silver grains are produced by contact of 35S emission with the autoradiographic emulsion. The number of grains is proportional to the hybridization signal. Background level was measured and subtracted for each slide. According to the quantification, an increase of signal between sham-treated animals and bacteria-injected animals (28.8 AU) and sham-treated animals (33.8 AU) was not significant. In contrast, the increase of silver density between sham-treated animals and bacteria-injected animals (65.8 AU) was significant (P < 0.05). These data demonstrate a 2-fold increase in Tt-cysb mRNA levels in the large coelomocytes of challenged leeches (Figure 3C). No other tissue was shown to exhibit a hybridization signal. All these results corroborated and extended the DD RT–PCR and microarray data by identifying the cells implicated in the Tt-cysb mRNA synthesis.

We then examined the time course of Tt-cysb induction by Northern blot analysis. Purification of coelomocytes was not possible, because of the high degree of adhesion to the coelomic cavity wall, and thus we performed this study on the whole animal. In spite of the weak representation of coelomocyte mRNA among the total RNA of the whole animal, the analyses nevertheless confirmed the induction of expression of Tt-cysb (Figure 3D). The data were normalized with the 18 S rRNA signal. Signals from PBS-injected animals, compared with control signals, showed a slight increase in Tt-cysb mRNA level after PBS injection. Signals from E. coli-injected animals, compared with those of PBS-injected animals, revealed a slight increase in Tt-cysb mRNA level 3 h after E. coli injection. This increase was more significant 24 h after injection. Signals from M. luteus-injected animals, compared
with those of PBS-injected animals, showed an increase in Tt-cysb mRNA level 3 h after _M. luteus_ injection. As for the _E. coli_ injection, this modulation was more important 24 h after _M. luteus_ injection. Globally, _Tt-cysb_ induction seemed to be stronger after the injection of _M. luteus_ than after the injection of _E. coli_. Otherwise, regardless of which bacteria were injected, other analyses were performed 12 h after the injections and did not detect any further significant increase in Tt-cysb mRNA levels other than those obtained 3 h after bacterial injections (results not shown).

In addition, for the animals analysed 24 h after injection, we noted that the Tt-cysb mRNA level was more markedly increased (approx. 1.2-fold after injection with a bacterial mixture (_E. coli_ and _M. luteus_)) than were those obtained after single bacteria injections (approx. 1.6-fold) after injection with a bacterial mixture (_E. coli_ and _M. luteus_). As for the animals analysed 24 h after injection, we noted that the Tt-cysb mRNA level was more markedly increased (approx. 1.2-fold after injection with a bacterial mixture (_E. coli_ and _M. luteus_)) than were those obtained after single bacteria injection (approx. 1.4-fold after _E. coli_ injection and approx. 1.4-fold after _M. luteus_ injection; Figure 3D).

_Tt-CYSB_ storing sites were investigated by immunocytochemistry on whole-animal paraffin sections. Strong immunoreactivity was observed in large coelomocytes only. _Tt-CYSB_ labelling was detected as intensely fluorescent small spots distributed throughout the cytoplasm and surrounding the nucleus (Figure 3E), suggesting that _Tt-CYSB_ is packed in granules or vesicles.

**DISCUSSION**

The present study demonstrates that DD RT–PCR and microarrays are, as for other invertebrate models [28], powerful and efficient combined techniques for transcriptomic analysis of gene expression in the leech. Furthermore, this allowed an exploration of biological responses to bacterial challenge [29]. DD RT–PCR generates many candidates with an important number of false positives [19]. The combination of this technique with DNA microarray analysis resolves this problem due to its sensitivity and its large powerful discrimination. As for many other experimental models, the _T. tessulatum_ genome remains to be sequenced. However, this approach should allow a large-scale analysis of the transcriptome in the leech, as well as for other invertebrate species [30]. In addition, these combined techniques permit us to undertake kinetic analyses in order to elucidate clustering of immune-system-induced genes.

Among the selected candidates, several appear to be unrelated to the immune response. With regard to plant immunity studies, the threonine deaminase gene is induced by methyl jasmonate or wounding [31,32]. Methyl jasmonate is involved in signalling processes, leading to immune responses [33,34]. Concerning the malate dehydrogenase gene, previous studies have reported its induction in human monocytes and macrophages after LPS (lipopolysaccharide) stimulation [35,36]. These data are comparable with events observed during bacterial challenge in leeches. This enzyme acts in gluconeogenesis, and might be crucial for the energetic requirements of immune cells. As for the ribosomal protein 60S L4 gene, several studies have shown modulations of these gene families in human monocytes after LPS injection [35,36]. Some authors reported that ribosomal-protein-derived peptides have

**Figure 2** The amino acid and nucleotide sequences of _T. tessulatum_ cystatin B

(A) Complete molecular cloning of _Ts-cysb_ cDNA reveals an open reading frame of 306 nt (show underlined) encoding a protein of 102 amino acids in length (shown by the lettering in bold typeface). The SignalP V1.1 program revealed no signal peptide. (B) The Multalin program shows a high identity with known cystatin B (54 % with human cystatin B and 50 % with murine cystatin B). It also shows similarities with a putative cystatin B in the oyster _Crassostrea virginica_ (GenBank® accession no. BG624574). However, lack of a cysteine residue (Cys6 in _Ty-CYSB_) close to a glycine residue seem to indicate that _oyster cystatin_ is not a form B cystatin [26].

Identical (high consensus) residues are shown on a dark-grey background; residues that are similar (low consensus) are shown on a light-grey background.
Leech cystatin B may be involved in innate immunity

Figure 3 Detection of Tt-cysb mRNA by in situ hybridization and Northern blotting, and immunodetection of Tt-CYSB protein in a tissue section of the leech.

Paraffin-embedded sections of leeches were hybridized with Ty-cysb antisense (A and B) or sense (B’) riboprobes labelled with [35S]UTP radiolabelling appears as dark-silver deposits. Silver grains are seen in the coelomocytes of both non-stimulated (A) and stimulated (B) animals. Comparatively, 24 h after bacterial challenge, signals in coelomocytes appear stronger. Individual coelomocyte titration for Tt-cysb mRNA levels was performed using a Biocom system. Results are expressed as AU. Tt-cysb mRNA content in coelomocytes of stimulated animals is approximately twice the level of that found in control animals (C). After normalization with 18S RNA, Northern blot analyses show that Tt-cysb mRNA from bacteria injected into the whole body is approx. 1.6-fold the level of control body mRNA (D). Sections were investigated for Tt-CYSB protein content by immunodetection with FITC-labelled secondary antibody. Strong immunoreactivity is observed in coelomocytes from non-stimulated leeches (E; shown by the white arrow). c, coelomocyte; n, nucleus. Bars represent either 100 µm (A, B and B’) or 30 µm (D).

As a result, this transcriptomic study in a lophotrochozoan animal suggests an approach for further experiments that should allow us to elucidate the functions of these induced genes in immunity. Previously, these up-regulated genes have been detected from a comparison between PBS-injected (control) animals and bacteria-injected (experimental) animals. As a single PBS injection or sterile injury can trigger immune mechanisms, we may anticipate further important modulations for these immune-induced genes in basal states. Moreover, we performed our transcriptomic analysis with total RNA from the whole organism. Further differential studies that are focused on specific immune tissues, such as coelomocytes, should identify powerful induced genes with higher modulation values.

Although our results are supported by other studies using ecdysozoans animals, some immune-system-induced genes detected in the present leech analysis are newly identified as having a role in immunity. For example, the existence of a protein related to cystatin B and its potential involvement in immunity are novel features in invertebrates. Based on its known inhibitor function, we have initially chosen to focus the analysis on this immun system-induced gene.

Tt-CYSB corresponds to the first cystatin B cloned in invertebrates. Cystatin B belongs to the first family of cystatins, i.e. the intracellular stefins [44,45]. Cystatin B is known to inhibit some cathepsins, namely cysteine proteinases [46,47]. In vertebrates, cystatin B has a broad distribution range in cells and tissues, and targets in vitro lysosomal cathepsins B, H, K and L [48,49]. The unique pathophysiological role known for cystatin B in humans, despite a lack of putative target proteinases, is a childhood form of epilepsy, Unverricht-Lundborg disease (or EPM1) [50]. The loss-of-function mutations in human cystatin B and deletion of the gene in mice gave rise to progressive myoclonus epilepsy, caused by cerebellar apoptosis [51]. The down-regulation of cystatin B triggers an increase in cathepsin activity in the cytoplasm of proliferating cells in the central nervous system [52]. Recently, some authors reported evidence of cathepsin B activity in EPM1 [53]. The EPM1 disease demonstrates that cystatin–cathepsin interactions are vital. Cystatin B may inhibit cathepsin-dependent apoptotic mechanisms [51]. Indeed, the consequence of a lack of cystatin B is the increased proteolytic activity of cathepsins, leading to apoptosis and glial activation [54].

In leeches, the Tt-cysb gene is up-regulated 24 h after a bacterial challenge in large circulating coelomic cells named large coelomocytes (by approx. 2-fold when compared with control animals by in situ hybridization). This time of analysis (24 h post-infection) seems to be appropriate for the study of immune-response-regulation processes. Interestingly, these data show that antimicrobial activity [37]. This gene family might then comprise the origin-of-defence molecules [38,39]. Finally, the PABP1 gene is also of interest. A PABP gene is induced in activated T-cells of the mouse by PMA and ionomycin [40]. All these molecules are involved in T-cell activation [41,42]. Moreover, PABPs initiate the translation of specific mRNAs corresponding to molecules involved in the same biological process [43]. We suggest that, in the leech, PABP gene up-regulation might trigger the expression of gene clusters all involved in immune response.
cystatin B inhibits cathepsins [46]. These enzymes are involved in immune mechanisms. The genetic modulation in *D. melanogaster* after infection with a pathogen has been reported [6–8]. Taken together, these data suggest that Tt-CYSB may regulate leech cathepsins and have a role in leech immune response regulation in these circulating coelomic cells. Indeed, Tt-cysb up-regulation and its localization in circulating coelomic cells allow us to hypothesize new immune functions for this proteinase inhibitor. Because cathepsins are implicated in phagosome processing in mammals [55], we suggest that the proteinase inhibitor may regulate some phagocytosis processes via the inhibition of the cysteine proteinases. As for the secreted cystatin family 2 members, various studies have shown involvement of cystatins in the modulation of the immune response [56]. In the present study, we report for the first time the potential involvement of an intracellular cystatin family 1 member, i.e. the leech cystatin B, Tt-cysb. In future studies we will consider the characterization of cathepsins in *T. tessulatum* in order to elucidate cystatin–cathepsin-interaction processes in leech immunity.

The present study provides new knowledge of the mechanisms of the innate immune response in a lophotrochozoan, which bears some resemblance with vertebrate, plant and invertebrate immune responses. Leech innate-immune-response studies are still in the early stages, but may bring new insights to the overall study of innate immunity. Our combination of differential analysis and cDNA microarrays in the leech highlights new immune-system-induced genes and the necessity to explore innate immunity in diverse models.

We thank J. C. Beauxvillain and V. Mitchell for access to the Cellular Imaging Center of the IFR 22 (Institut Fédéral de Recherche 22, Faculté de Médecine, Lille, France), and A. Desmons and M. Vandeckerkove for technical assistance (UMR 8017 CNRS). This work was supported by the CNRS, the MNERT, the Génopole of Lille, the FEDER and the Conseil Régional Nord-Pas de Calais.

REFERENCES


Leech cystatin B may be involved in innate immunity


Received 24 March 2004/April 2004; accepted 19 April 2004
Published as BJ Immediate Publication 19 April 2004, DOI 10.1042/BJ20040478