International Journal for Parasitology 45 (2015) 187–196

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Characterisation of the secretome of the clam parasite, QPX

Ewelina Rubin [a], Emmanuelle Pales Espinosa [a], Antonius Koller [b], Bassem Allam [a,*]

[a] School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY 11794-5000, USA
[b] Proteomics Center, Stony Brook University, Stony Brook, NY 11794-5000, USA

ARTICLE INFO

Article history:
Received 5 August 2014
Received in revised form 21 October 2014
Accepted 22 October 2014
Available online 31 December 2014

Keywords:
Mercenaria mercenaria
Thraustochytrid
Virulence
Secretome
Mass spectrometry

ABSTRACT

Secreted and cell surface-associated molecules play a major role in disease development processes and host-pathogen interactions, and usually determine the virulence of invading organisms. In this study, we investigated proteins secreted by quahog parasite unknown, a thraustochytrid protist that infects the hard clam, *Mercenaria mercenaria*. In silico analysis of quahog parasite unknown transcripts predicted over 1200 proteins to possess an amino-terminal signal peptide which directs proteins into the classical eukaryotic secretory pathway. Proteomic analysis using LC/MS technology identified 56 proteins present in the extracellular secretion of quahog parasite unknown cells grown in vitro, including six mucin-like molecules, four glycosyl hydrolases and eight peptidases. Transcription levels of 19 quahog parasite unknown extracellular proteins were investigated in clam tissue lesions (in vivo) using quantitative PCR. The overexpression of six of these extracellular proteins in clam tissues compared with in vitro cultures suggests that they are involved in interaction with the clam host.

1. Introduction

Quahog parasite unknown (QPX) is a thraustochytrid protist that infects hard clams, *Mercenaria mercenaria*, along the northwestern coasts of the Atlantic Ocean from Prince Edward Islands, Canada, to Virginia, United States of America (USA) (Whyte et al., 1994; Ragone Calvo et al., 1998). The histopathological examination of tissue lesions from diseased clams typically shows QPX cells surrounded by a thick layer of muco-polysaccharide secretions characterised by positive Alcian blue staining (Ragone Calvo et al., 1998; Smolowitz et al., 1998). In addition, parasite cells, when cultivated in vitro, produce a very dense mucus secretion containing QPX extracellular products (QPX ECPs) (Perrigault and Allam, 2009). It has been suggested that the mucus layer surrounding QPX may function as a virulence factor by facilitating the infection of clam tissue and providing protection against the clam immune response by limiting parasite phagocytosis and encapsulation by clam hemocytes (Smolowitz et al., 1998). The protective role of QPX secretions has also been supported by in vitro experiments showing that the mucoid layer surrounding the parasite provides protection against the antimicrobial activity of clam plasma (Anderson et al., 2003). A preliminary study also showed that QPX ECPs contain unidentified peptidases which are capable of breaking down proteins and possibly used by the parasite to destroy clam connective tissue (Anderson, R.S., Luskey, T.M., Strauss, M.A., 2006. In vitro protease production by QPX, Abstracts of the 98th Annual Meeting of the National Shellfisheries Association, Monterey, California, USA). Further, exposure to QPX ECPs causes the death of clam hemocytes (Perrigault and Allam, 2009), suggesting that these secretions may also contain cytotoxic molecules. The specific biochemical composition of QPX ECPs and the potential role of their specific components in QPX virulence remain unknown and represent the focus of this study.

It is well documented that pathogenic organisms secrete molecules that enable host invasion, help establish infection foci and cause damage to host tissues (Kamoun, 2006; Kale and Tyler, 2011). These extracellular and cell surface molecules include hydrolysing enzymes (peptidases, glycosidases and lipases), receptors, adhesins and other recognition and attachment proteins. In silico predictions of pathogenic secretomes are often used to evaluate the pathogenic capabilities of microorganisms (Raffaele et al., 2010; Denoeud et al., 2011; Jiang et al., 2013). Proteomic approaches such as mass spectrometry (MS) represent powerful tools for large-scale identification of proteins and are also being used successfully to investigate extracellular proteins of pathogenic protozoans e.g. *Trypanosoma brucei* (Geiger et al., 2010) and *Leishmania donovani* (Silverman et al., 2008). In this study, we used a combination of transcriptomics-based prediction and MS to identify proteins secreted by QPX and to provide new insight into the molecular basis of QPX pathogenicity. We also investigated the transcription levels of QPX putative virulence factors within the

* Corresponding author. Tel.: +1 631 632 8745; fax: +1 631 632 8915.
E-mail address: Bassem.Allam@stonybrook.edu (B. Allam).

http://dx.doi.org/10.1016/j.ijpara.2014.10.008
0020-7519/

clam infection sites (also known as clam nodules) to reveal their possible importance during the infection process.

2. Materials and methods

2.1. In silico prediction of the QPX secretome

The QPX transcriptome sequence libraries were generated as part of the Marine Microbial Eukaryotic Transcriptome Sequencing Project (MMETSP) under the sample ID numbers: MMETSP0098 and MMETSP0100 (Keeling et al., 2014). The two transcriptome libraries were generated for QPX cells cultivated in (i) minimal essential medium (MEM) supplemented with 10% (v/v) FBS (MMETSP0098) and (ii) MEM supplemented with 10% FBS and 2% clam adductor muscle homogenates, adjusted to a concentration of 7 mg/ml of protein (MMETSP0100), since clam homogenates are thought to increase the virulence of QPX cells (Perrigault and Allam, 2009). For sample MMETSP0100, 12,579 transcriptomic sequences were used to find the longest open reading frames (ORFs) using the online Virtual Ribosome Tool Version 1.1 (Wernersson, 2006). The ORFs were scanned for the N-terminal signal peptide sequences using the SignalP version 4.1 server (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen et al., 2004b; Petersen et al., 2011), and then the ORFs with signal peptide sequence were searched for cell localisation using the Protein Subcellular Localisation Prediction tool WoLF PSORT (available from http://wolfpsort.org/) (Horton et al., 2007). The sequences related to products predicted to be secreted outside the cell were annotated using a BLASTp search, gene ontology and conserved protein domains tools within the Blast2Go server tool (Gotz et al., 2008). For sample MMETSP0098, peptide sequences provided by the MMETSP were used to find peptides possessing the N-terminal signal peptide sequence using the Signal P tool and to predict extracellular localisation using the WoLF PSORT tool. Local BLASTp searches between the two datasets were performed using the Geneious bioinformatics tool http://www.geneious.com/) to find the total number of unique proteins expected to be secreted by QPX based on transcriptomic data.

2.2. MS on QPX ECP

QPX cells (isolate NY0313808B (Qian et al., 2007)) were grown in 100 ml of MEM supplemented with 0.3% yeastolate (BD Biosciences, USA; Cat. 255772). In addition, the media contained salts, HEPES and antibiotic solution according to previously described methods (Kleinschuster et al., 1998). The culture was incubated at room temperature in a 400 ml Falcon tissue culture flask (BD Biosciences) on an orbital shaker at 80 rpm for 6 days. At the time of collection, the QPX cell concentration reached 8×10^7 cells ml^{-1}. The culture was centrifuged at 2500*g* and 4 °C for 70 min and cell-free supernatant was collected. Because QPX ECP includes a dense muco-polysaccharide secretion, a few steps were applied to facilitate mucus solubility. First, EDTA (10 mM final concentration) was added to the supernatant and incubated for 2 h at room temperature to dissolve the EDTA-soluble fraction (Smith et al., 2009). Second, QPX ECP was further liquefied by several passages through a 10 ml syringe without a needle. Next, the suspected large polysaccharide fraction was separated from a smaller protein fraction by ultrafiltration through a 100 kDa column (Amicon, USA; Cat. UFC910008). A 5 ml subsample of the collected filtrate was further concentrated by a second ultrafiltration step through 3 kDa columns (Amicon, Cat. UFC500324) and a final lyophilization step. The final protein sample, a 50× concentrate of the original QPX ECP (by volume), was visualised by 12% SDS PAGE stained with silver nitrate (Fig. 1). A second gel stained with Coomassie blue G-250

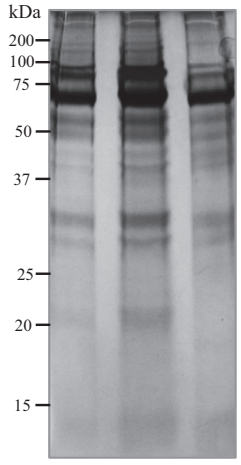

Fig. 1. SDS–PAGE silver staining of extracellular products ($n = 3$) of clam quahog parasite unknown grown in minimal essential medium supplemented with 0.3% (w/v) yeastolate. One sample was analysed by LC/MS. MW size markers are indicated.

(Bio-Rad, USA) was used for MS identification of the protein content.

Gel bands were excised, de-stained, reduced, alkylated and digested with trypsin (Trypsin Gold, Mass Spectrometry Grade, Promega, USA) essentially as described by Shevchenko et al. (1996) with minor modifications. The resulting concentrated peptide extract was diluted into a solution of 2% acetonitrile (ACN), 0.1% formic acid (FA) (buffer A) for analysis. Ten microlitre of the peptide mixture were analysed by automated microcapillary LC/MS-MS. Fused-silica capillaries (100 µm inner diameter (i.d.)) were pulled using a P-2000 CO2 laser puller (Sutter Instruments, Novato, CA, USA) to a 5 µm i.d. tip and packed with 10 cm of 5 µm Magic C18 material (Agilent Technologies, Santa Clara, CA, USA) using a pressure bomb. Ten microlitre of the resulting 20 µl of concentrate were pressure-loaded onto a 10 cm 100 µm i.d. fused-silica capillary packed with 3 µm Magic C18 reverse phase (RP) particles (Michrome, USA) which had been pulled to a 5 µm i.d. tip using a P-2000 CO2 laser puller (Sutter Instruments). This column was then installed in-line with a Dionex 3000 HPLC pump running at 300 nL min^{-1}. Peptides were loaded with an auto-sampler directly onto the column and were eluted from the column by applying a 30 min gradient from 5% buffer B to 40% buffer B (98% ACN, 0.1% FA). The gradient was switched from 40% to 80% buffer B over 5 min and held constant for 3 min. Finally, the gradient was changed from 80% buffer B to 100% buffer A over 0.1 min, and then held constant at 100% buffer A for 15 min longer. The application of a 1.8 kV distal voltage was used to electro-spray the eluting peptides directly into an LTQ XL ion trap mass spectrometer equipped with a nano-liquid chromatography electrospray ionisation source. Full MS were recorded on the peptides over 400–2000 m/z, followed by five MS/MS fragmentation events on the five most intense ions. MS scan functions and HPLC solvent gradients were controlled by the Xcalibur data system (Thermo Finnigan, San Jose, CA, USA). MS/MS spectra were extracted from the RAW data file with ReAdW.exe (http://sourceforge.net/projects/sashimi). The resulting mzXML file contained all of the data for all MS/MS spectra and could subsequently be read by the analysis software.

The MS/MS data were searched with Inspect (version 20101210; Tanner et al., 2005) against a database containing protein sequences for QPX sequences (MMETSP098, 11,005 peptides in total), common contaminants (84 proteins) and all proteins in reverse order for Decoy analysis with possible modifications of +16 on methionine and tryptophan (oxidation) and a fixed modifi-

cation of +57 on cysteine (carbamidomethyl), with ion fragment tolerance set at 0.5 Da and trypsin specified protease. Only peptides with $P \geqslant 0.01$ were analysed further. Summary.py was used to assemble peptides to their respective protein. Only proteins with at least two unique peptides were considered as identified.

2.3. In vivo versus in vitro QPX transcript levels

Transcript levels for genes encoding several of the identified extracellular proteins were investigated in QPX cells in vivo (in clam lesions/nodules). For this purpose, diseased adult clams were obtained from an enzootic area in Wellfleet (Massachusetts, USA) in spring 2012. Five infection nodules were excised from the mantle edges of diseased clams (Fig. 2) and kept frozen at −80 °C until RNA extraction. To test the QPX primer specificity, one sample piece of mantle from a healthy clam was treated as a negative control for QPX gene amplification. The healthy clam did not display infection nodules and the absence of QPX in tissues was confirmed using a previously described quantitative PCR (qPCR) method (Liu et al., 2009). For the in vitro comparison, seven different QPX isolates (four New York (NY), SA and three Massachusetts (MA), USA isolates) were cultivated in clam adductor muscle homogenates adjusted to a total protein concentration of 2 mg/ml (Perrigault et al., 2009). The QPX strain selection included two NY isolates and one MA isolate which have been sub-cultured since 2003, and two NY and two MA isolates obtained from diseased clams collected from the field in 2012.

Infected clam tissues (nodules) containing QPX cells and QPX cell pellets from in vitro cultures were separately homogenised in Trizol (Invitrogen, USA) reagents using a mechanical homogenizer. RNA extractions were performed following the manufacturer's protocol (Invitrogen). Total RNA was precipitated using 7 M LiCl and kept at −20 °C overnight. The RNA pellets were washed using 70% molecular grade RNase-free ethanol. cDNA was synthesized using Moloney-murine leukaemia virus reverse transcriptase (Promega) in reactions containing 0.5 µg of oligod T primers, 2 µg of total RNA and 0.5 mM final dNTPs concentration.

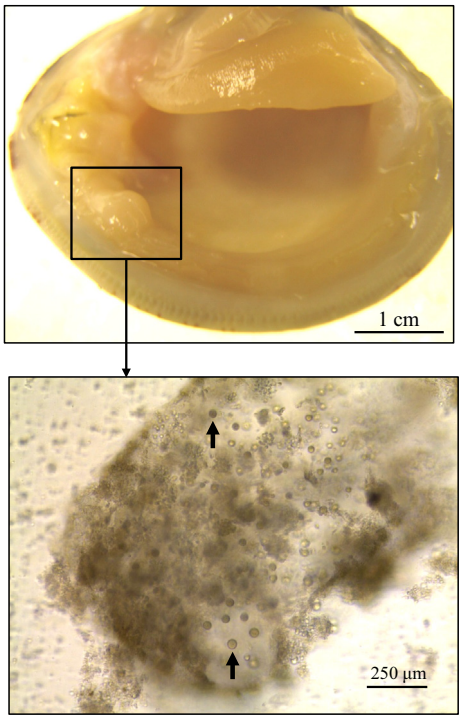

Fig. 2. A quahog parasite unknown infection nodule at the edge of the mantle of *Mercenaria mercenaria*. Arrows point to quahog parasite unknown cells viewed microscopically in a fresh biopsy preparation. Scale bars = 1 cm and 250 µm respectively. (For interpretation to colours in this figure, the reader is referred to the web version of this paper.)

Primers (all with the annealing temperature of 60 °C) for the selected QPX transcripts were designed using the PrimerQuest design tool available from the Integrated DNA Technology website (http://www.idtdna.com/Primerquest/Home/Index). All possible primer pairs were mapped onto the corresponding full cDNA sequences within Geneious software and then local BLASTn searches were performed against a clam cDNA library (B. Allam, unpublished data) to find possible clam sequences matching the QPX transcripts and to avoid selecting primers which would have more than 50% nucleotide sequence similarities to any clam sequences. In addition, to ensure QPX primer specificity, the selected primers were tested using the PrimerBLAST National Center for Biotechnology Information (NCBI) tool against the non-redundant nucleotide database and were confirmed not to amplify any sequence other than targeted transcripts. Only primers which showed amplification of QPX genes and no amplification when placed with clam-derived cDNA were chosen, and the final list of selected genes and nucleotide sequences of their corresponding primers is presented in Table 1. qPCR reactions were carried out in 96-well plates using 10 µl volume reactions containing: 1× Brilliant II Syber Green master mix (Agilent Technologies), 0.2 mM concentration of each primer and 15 ng of cDNA. The PCRs were performed on a Mastercycler ep realplex PCR machine (Eppendorf, Germany) using the following temperature cycling programme: an initial denaturation step at 95 °C for 10 min, 40 cycles of denaturation at 95 °C for 30 s and annealing plus extension at 60 °C for 1 min, and a 20 min melting curve cycle. The mRNA levels of the selected transcripts were normalised to the glyceraldehyde 3-phosphate dehydrogenase (GADPH) gene, and relative expression levels were calculated using the ΔΔcycle threshold (Ct) method (Livak and Schmittgen, 2001).

3. Results

3.1. Predicted secretome of QPX

After scanning predicted QPX peptide sequences using the SignalP bioinformatics tool, a total of 806 proteins and 641 proteins with an N-terminal secretory signal peptide were identified in MMETSP0098 and MMETSP0100 libraries, respectively (Supplementary Tables S1, S2). Among these, 192 sequences were found to have 95% or higher amino acid similarity between the two transcriptome libraries, representing the same proteins or highly similar isoforms. The 806 signal peptide-containing proteins from sample MMETSP0098 were searched using the tblast algorithm against the entire MMETSP0100 transcriptome database (and vice versa) and results showed that all sequences used in the study were found in both libraries. Thus the above difference between the two samples reflects different numbers of contigs with complete 5′ untranslated regions (UTRs) and does not represent expression differences in response to the two culture media. For samples MMETSP0098 and MMETSP0100, 522 and 400 proteins, respectively, were also predicted to be extracellular using WoLF PSORT (Supplementary Tables S1 and S2). The BLASTp homology searches against the NCBI database returned BLASTp hit matches for 347 sequences (54%) for MMESTP0100 and 520 sequences (64%) for sample MMETSP0098 (Supplementary Tables S1 and S2). A summary of predicted extracellular proteins of QPX which may play a role in hard clam disease and/or clam-QPX interactions are presented in Table 2. The two groups represented by the high-

Table 1
Nucleotide sequences of primers used for the quantitative PCR estimate of transcript levels in quahog parasite unknown cells found in vivo relative to quahog parasite unknown cells grown in vitro.

Gene name	Gene ID	Transcript ID	Primer sequences
Glyceraldehyde 3-phosphate dehydrogenase	GADPH	MMETSP0100_6578	F: GGAGATAACGGTGTGGAATAC R: TTGACTCCCATAACGAACATAG
Mucin-related 1	Muc-1	MMETSP0100_3380	F: CTTGCGACTGTGTGTAGTAG R: CACATGCACTGGTGAGAATA
Mucin-related 2	Muc-2	MMETSP0098_3037	F: GGAGATGCTGGTGCTAAAC R: GTGGGTTGTAGTGGGATTTC
Mucin-related 3	Muc-3	MMETSP0100_3085	F: AACTACACTTGCACCGACTAC R: TGGAGAATGGGTTGGAGAATG
Mucin-related 4	Muc-4	MMETSP0098_18146	F: GGTGCTACCAACTGGATTAGAG R: GGAAGTCGTAAGGGTACTTTGG
Integrin	Integrin	MMETSP0098_115	F: CAGTCGTGCAAATGGTAGGT R: GTCAGCAGGGTTCTCAATAGTC
Lectin	Lectin	MMESTP0100_3133	F: TTGAAGCAGGGCATGTTCACAAGG R: ACCTTCGTGGAAACGGTAATGGGA
Heat shock protein 70	Hsp70	MMETSP0100_1067	F: AACGGTAAGGAGCCAAGTAAG R: GGGCAACATCGAGAAGAAGA
Haemolysin	Haemolysin	MMETSP0098_300	F: CCACGCTCAACTCAACTATTA R: ATTTCGTCATAGCCGTTCTC
Elicitin-like 1	El-like 1	MMETSP0098_3214	F: ATGCTCCAACACCCAATC R: CAACACTCGGGTTTCTAGTT
Elicitin like 2	El-like 2	MMETSP0098_17402	F: CTTCACCAACAGGGTCTGTATC R: CACAGTCGTAGCCCAGAAAG
Endotoxin	Endotoxin	MMETSP0098_1515	F: ACTCATCATCGGACCTATCT R: GACCCGTACCTCACATTAAC
Necrosis inducing protein	Nec Ind P	MMETSP0100_1986	F: TAACAGTGGTGGGAGAGTAG R: GCCGCCATGTTGGTAATA
Subtilase 1	S8–1	MMETSP0100_1574	F: AAGAGCGGTTGGGAATATGGGAGT R: AACAACAAGCATGCCCTCTTCTGC
Subtilase 2	S8–2	MMETSP0100_4754	F: TATGGCTACTCCATTTGTCGCTGG R: ACGAGCAAATTGGGAGATTCGTGC
Subtilase 3	S8–3	MMETSP0100_1744	F: TTGATGCAAGGCGTTTCCGATGTC R: GCGTCGATCGAAATGGCAAGTTGT
Serine carboxy-peptidase	S10	MMETSP0100_3910	F: AAACAGCTTGCCGGGTTCATTGAG R: AACTACAGCCTGGACCACCATTCA
Cysteine peptidase 1	C1–1	MMETSP0100_2357	F: AAGGAGCAGGAAGCACTCGAGAAA R: GCCAACTAAAGTTCGCAGGCAAGT
Cysteine peptidase 2	C1–2	MMETSP0098_1002	F: AGCACAAGCCATGTAATCTCGGTG R: TACCTTGGCAAAGCCGTTGTTTCC
Metallopeptidase	M35	MMETSP0100_1424	F: ATCCACCAACAAGTACGCCAGAGA R: GGTTGTTCAATGCAAGGGAGCGAA

est number of proteins in the predicted QPX secretome were hydrolases and peptidases (Table 2). These two groups included enzymes commonly associated with pathogenicity such as lipases, phospholipases, haemolysins, subtilases, cathepsins and disintegrins (Table 2). QPX was also predicted to secrete antioxidants such as thioredoxins, ascorbate (plant-like) peroxidase and glutathione peroxidase, and one superoxide dismutase. Other QPX proteins which contain the N-terminal signal peptide included heat shock proteins, lectins, cell membrane-bound receptors, glycotransferases, cyclophilins, fibronectin- and integrin-related proteins. In addition, the predicted secretome included five proteins possibly related to QPX virulence: elicitin-like proteins, endotoxins and a necrosis-inducing protein (Table 2).

3.2. Extracellular proteins identified by LC/MS

Fifty-six proteins present in QPX ECP were identified by MS ($P < 0.01$; minimum of two peptides matching the predicted sequence; Table 3, Supplementary Data S1). Based on in silico predictors, 31 of the identified proteins possess the signal peptide for secretion via the classical vesicle-mediated pathway, and 14 proteins are possibly secreted via the non-classical pathway (Table 3). The remaining 11 proteins represent either proteins of intracellular contamination, ORFs with missing N-terminal sequence ends, or proteins for which the currently available bioinformatics tools cannot predict their secretion pathway. Twelve of these extracellular proteins are unique to QPX as there were no homologous sequences found in the NCBI database. Several proteins known to be involved in host-pathogen interactions were identified in QPX ECPs using LC/MS analyses. These include six mucin-like or glycosylated proteins, eight proteolytic enzymes, five glycosyl hydrolases, an integrin-related protein, a lectin and heat shock protein 70 (Table 3). Six extracellular proteins identified in the QPX ECPs had amino acid motifs known to be sites for the attachment of carbohydrate chains. They all contained repeating sequence motifs rich in proline, threonine, serine and alanine residues which comprise between 36% and 58% of the total amino acids in these proteins (data not shown). The BLASTp results showed homology of the proline-threonine (PT) rich fragments to many different molecules in a variety of organisms including a LPXTG-motif cell wall anchor domain protein of *Lactobacillus reuteri* (WP_003664141), an elicitin-like protein of the oomycete *Phytophthora medicaginis* (ABH11745), a proline-threonine-rich repeat protein of the fungus *Trichophyton rubrum* (XP_003237106), and a mucin of the *Dictyostelium fasciculatum* (XP_004360105) (Fig. 3).

The MS data of QPX ECP also revealed that QPX secretes a legume-type lectin, an integrin-related protein, and a heat shock protein of approximately 70 kDa molecular weight (Table 3). All three proteins were predicted to be secreted via the classical endoplasmic reticulum (ER)/Golgi vesicle-mediated secretion pathway (Table 2). QPX extracellular lectin is similar to the leguminous plant lectins named L-type lectins (InterPro protein domain num-

Table 2
Number of proteins predicted to be secreted by quahog parasite unknown, grouped by categories of common virulence factors.

Group name	SignalP v. 4.0	SignalP and WoLF PSORT	Conserved protein domains (InterPro database)		
Hydrolases					
Glycosyl hydrolases	34	18	IPR001547	IPR001139	IPR001382
Lipases	14	13	IPR002921	IPR008139	
Phosphatases	14	12	IPR018946	IPR017849	IPR004843
Phospholipases	5	5	IPR017946	IPR007000	IPR002641
Haemolysins	5	4	IPR027018		
Chitinases	3	3	IPR011583		
Pectinase	1	0	IPR012334		
Peptidases					
Serine	27	22	IPR000209	IPR001563	IPR001314
Cysteine	23	22	IPR000668	IPR001096	IPR011697
Metallo	15	11	IPR000834	IPR001506	IPR024079
Aspartic	8	5	IPR001461	IPR021109	
Threonine	2	2	IPR000426	IPR001353	
Antioxidants					
Thioredoxins	9	7	IPR005746		
Ascorbate peroxidases	2	2	IPR000823	IPR012336	
Thioredoxin reductase	1	1	IPR013027		
Glutathione peroxidase	1	1	IPR000889		
Superoxide dismutase	1	1	IPR001424		
Receptors					
G-protein-coupled receptors	11	0	IPR002455	IPR001828	IPR000337
Mannose-6-phosphate receptor	3	2	IPR009011		
Transmembrane receptor	2	0	IPR009637		
Rhodopsin-like GPCR	2	0	IPR019336		
MD-2 lipid-recognition domain	2	2	IPR003172		
Heat shock proteins					
Heat shock proteins 40	7	4	IPR001623		
Heat shock proteins 70	3	1	IPR013126		
Heat shock protein 20	1	1	IPR002068		
Cyclophilins	8	6	IPR002130	IPR001179	
Glycotransferases	7	3	IPR021067	IPR007657	
Fibronectin-related	6	2	IPR003961		
Integrin-related	5	5	IPR013519	IPR002035	IPR013517
Lectins	5	3	IPR005052	IPR001304	IPR013320
Elicitin like proteins	2	1			
Endotoxins	2	1	IPR005639		
Necrosis inducing protein	1	1	IPR008701		

ber: IPR005052). The integrin related protein contains the conserved protein repeat (IPR013519, IPR013517 protein domains) found in alpha integrins. QPX ECP was shown also to contain three proteins having oxido-reductase activities which might participate in antioxidant processes: flavin-dependent oxireductase, peroxidase and glutathione reductase (Table 3).

QPX was also found to secrete eight different peptidases including four serine-type peptidases, two cysteine peptidases and two metallopeptidases (Table 3). Three of the serine peptidases belong to the subtilase family (S8 family, IPR015500 and IPR000209), one serine peptidase is a carboxypeptidase (S10 family, IPR001563), and the two cysteine peptidases belong to the papain family (C1 family, IPR000668). The two metallopeptidases (IPR024079) could not be assigned to specific families. Eight different types of hydrolases were also identified in the QPX secretome within which the three most common (based on the numbers of spectra) were QPX extracellular beta-amylase (or glycosyl hydrolase family 14), beta-1,3-glucanase (glycosyl hydrolase family 17) and an unassigned family of glycosyl hydrolase (Table 3).

3.3. In vivo versus in vitro QPX transcript levels

To reveal the possible involvement of selected extracellular molecules of QPX during interactions with its clam host, relative amounts of their mRNA levels were compared between parasite cells in clam tissue lesions (in vivo) and in vitro cultures. Transcript levels of 19 QPX ECP, which showed similarity to virulence factors of other pathogenic microorganisms or were hypothesised to be involved in host-pathogen interactions, were selected and estimated using qPCR. The expression of six extracellular proteins was shown to be significantly higher for QPX cells inside the infection nodules than for QPX cells cultivated in vitro (t-test, $P < 0.05$, Fig. 4). The higher expression was recorded for two mucin-like proteins, two peptidases, the haemolysin E and the elicitin-like protein (Fig. 4). The expression of one additional peptidase (S10) was also up-regulated for the in vivo QPX, however the difference was not statistically significant due to high variability between the nodule samples (Fig. 4).

4. Discussion

QPX was predicted to secrete 1255 proteins outside its cell based on the presence of a eukaryotic signal peptide at the N-terminal of proteins. The number of proteins predicted to be secreted by QPX is within the same range as predictions for other stramenopilan pathogens such as *Phytophthora infestans*, *Saprolegnia parasitica* and *Pythium ultimum* which, respectively, have 1415, 970 and 747 proteins predicted to be secreted based on the presence of eukaryotic signal peptides (Levesque et al., 2010; Raffaele et al., 2010; Jiang et al., 2013). To predict the secretome of QPX, we used the SignalP v. 4, which in addition to the previous version's cleavage site score, C, and the signal peptide score, S, contains an additional discriminant score, D (Petersen et al., 2011). The D score has been created so that it discriminates between

Table 3
Proteins identified in quahog parasite unknown extracellular secretions using liquid chromatography coupled to MS.

Protein name	MS results			Length	Predicted MW (kDa)[a]	SignalP D score[b]	SecretomeP NN score[c]
	Peptides	Spectra	Coverage (%)				
Mucus component							
Mucin-like 1	13	94	25.4			0.407	0.609
Mucin-like 2	22	62	30.0			0.400	0.151
Mucin-like 3	7	67	17.7	277	27.87	0.554	
Mucin-like 4	16	74	14.3	760	80.83	0.691	
Mucin-like 5	9	79	12.2	784	80.82	0.439	0.585
Mucin-like 6	9	24	26.8	331	34.70	0.542	
Extracellular matrix protein	10	17	20.1	726	78.22	0.288	0.523
Adhesion and recognition							
Integrin-related	20	96	51.8	440	47.25	0.819	
Lectin	3	3	11.5	434	49.40	0.558	
Coagulation factor	2	3	3.4	591	64.99	0.683	
Antioxidant activity							
Glutathione reductase	4	6	9.9	482	51.53	0.171	0.685
Oxireductase	15	36	35.5	537	58.35	0.679	
Peroxidase	2	3	9.8	348	38.18	0.151	0.851
Peptidase activity							
Subtilase 1	18	140	48.3	400	41.46	0.537	
Subtilase 2	5	9	17.5	407	43.73	0.777	
Subtilase 3	3	3	17.2	402	42.38	0.722	
Serine carboxypeptidase 1	15	36	36.1	457	50.78	0.702	
Papain like peptidase 1	10	15	11.2	546	60.82	0.688	
Papain peptidase 2	4	8	12.6	556	60.95	0.655	
Metallopeptidase M35	3	6	4.9	508	55.99	0.800	
Metallopeptidase M12B	6	14	10.5	902	99.58	0.567	
Hydrolase activity							
Glycosidase beta-amylase	19	53	41.4	464	52.61	0.612	
Beta-glucosidase/glucanes	15	81	56.6	318	34.80	0.481	
Glycosyl hydrolase 1	14	43	46.8	449	50.26	0.682	
Glycosyl hydrolase 2	2	7	4.0	447	49.25	0.703	
Acetylesterase	8	12	21.8	499	55.86	0.458	
Aminohydrolase	2	3	6.3				
Adenosylhomocysteinase	4	7	10.2	481	52.86	0.098	0.483
Chitolectin/chitinase	2	7	1.7	1259	139.22		
Other							
Heat shock protein 70	3	5	6.8	664	72.61	0.659	
Aminotransferase 1	9	18	31.6	395	42.67	0.106	0.514
Aminotransferase 2	2	2	2.2	420	46.82	0.180	0.628
Aminotransferase 3	11	33	34.9	397	43.17	0.223	0.550
Dihydrodipicolinate synthase	10	14	50.2	305	32.22	0.235	0.585
Triosephosphate isomerase 1	13	39	57.2	258	27.74	0.105	0.546
Triosephosphate isomerase 2	6	7	17.3	256	27.69	0.156	0.549
2-Hydroxyacid dehydrogenase	15	28	49.7	426	46.96	0.179	0.724
Inorganic pyrophosphatase	3	5	13.6	274	31.23	0.135	0.621
Chitin binding domain	15	48	28.9	398	43.24	0.324	0.231
WD repeats (G-protein)	8	16	33.1	315	35.16	0.110	0.433
Dihydrolipoamide dehydrogenase	20	44	58.2	502	53.58	0.135	0.442
Ribose-5-phosphate isomerase	4	5	22.9	319	34.13	0.316	0.321
Phosphoenolpyruvate carboxykinase	4	6	10.3	575	63.53	0.140	0.415
Transaldolase	4	4	19.8	323	36.02	0.104	0.286
Unknown 1	12	40	47.7			0.603	
Unknown 2	5	27	22.8	347	36.87	0.483	
Unknown 3	7	13	26.3	413	46.08	0.766	
Unknown 4	4	5	7.9	850	94.79	0.824	
Unknown 5	2	3	10.2	381	40.40	0.606	
Unknown 6	11	20	69.2	246	27.74	0.690	
Unknown 7	5	13	9.1	616	69.11	0.566	
Unknown 8	15	58	52.0	352	38.75	0.478	
Unknown 9	13	45	14.0	1178	130.95	0.530	
Unknown 10	3	6	6.7	476	53.12	0.709	
Unknown 11	3	4	1.6				
Unknown 12	4	24	6.4	422	46.17	0.390	

[a] Molecular weight predictions using Geneious v. 5, missing values for incomplete coding sequences.
[b] SignalP v. 4.1; classically secreted proteins, with D score above 0.45; missing values for sequences without a 5′ end.
[c] SecretomeP v. 2.0 - non-classically secreted proteins should obtain an NN-score exceeding the normal threshold of 0.5, but should not simultaneously be predicted to contain a signal peptide; missing values for sequences with D score above 0.45.

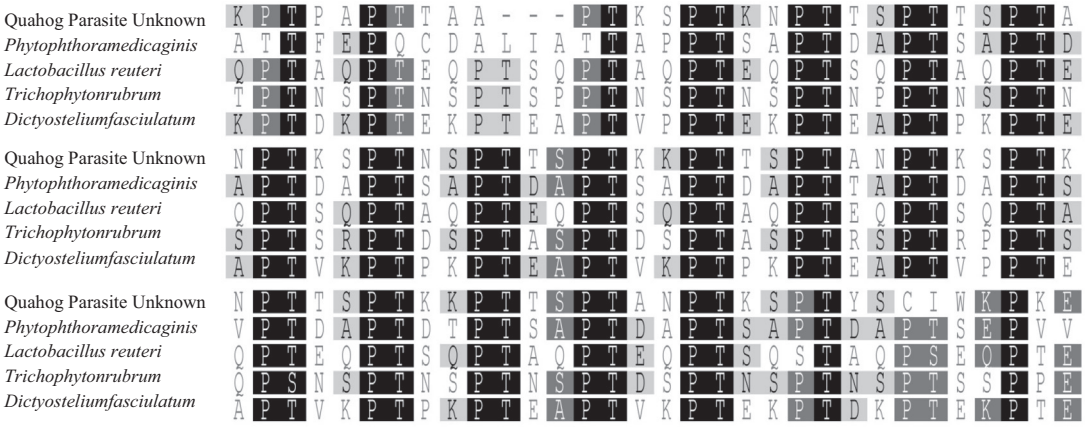

Fig. 3. Conserved proline-threonine residues in a partial amino acid sequence alignment of quahog parasite unknown mucin with mucin-related proteins from four different organisms: *Phytophthora medicaginis* (ABH11745), *Lactobacillus reuteri* (WP_003664141), *Trichophyton rubrum* (XP_003237106) and *Dictyostelium fasciculatum* (XP_004360105).Shading indicates similar amino acid residues (black, 100% similarity; dark grey, 80–100% similarity; light grey, 60–80% similarity; none, less than 60% similarity).

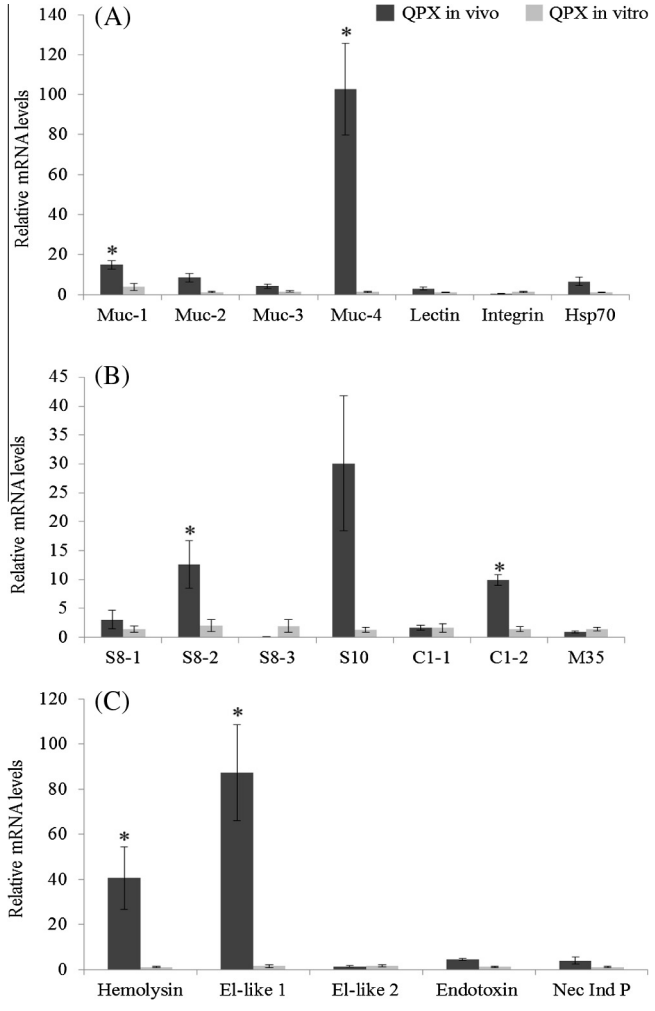

Fig. 4. Relative mRNA levels of selected proteins of clam quahog parasite unknown cells in vivo versus quahog parasite unknown cells in vitro. (A, B) Selected proteins found to be extracellular by MS; (C) predicted in silico to be extracellular. Plotted are means ± S.E. ($n = 5$ for in vivo samples, $n = 7$ for in vitro samples). *Statistically significant ($P < 0.05$; *t*-test). Muc, mucin; Hsp, heat shock protein; S8–1 to M35, peptidases; El-like, elicitin like; Nec Ind P. necrosis inducing protein.

secreted (D score $\geqslant 0.450$) and non-secreted sequences and excludes sequences which are predicted to contain N-terminal transmembrane helices (Petersen et al., 2011). Despite these improvements, the predicted secretome of QPX still contains many cell membrane-bound proteins including G-protein coupled receptors and transmembrane transporters. In the present study we also used an additional tool, WoLF PSORT, to predict protein localisation of QPX sequences possessing signal peptides. Based on this tool, only 548 proteins were predicted to be extracellular and include mostly hydrolytic enzymes which comprised approximately 21% of all proteins. Using LC/MS analysis only 31 out of the 1255 predicted proteins containing signal peptides were found in culture supernatant. This most likely reflects the response of QPX to the particular media conditions under which it was cultivated for the analysis. Medium composed of only hydrolysed yeast proteins was added to the QPX culture to avoid foreign protein contamination during sample analysis. It is likely that the numbers and types of proteins actually secreted by QPX cells change in response to different environmental or growth conditions. These 31 extracellular proteins represented over half of the proteins found by LC/MS analysis, suggesting that a large fraction of QPX extracellular proteins is secreted via the classical eukaryotic mechanism. Some extracellular proteins are known to be transported via a non-conventional secretion pathway which does not require the presence of a signal peptide (Nickel and Rabouille, 2009). In fact, for some kinetoplastid pathogens, the majority of extracellular proteins do not possess a signal peptide and are released by a micro-vesicle based secretion system (Silverman et al., 2008; Cuervo et al., 2009; Geiger et al., 2010). SecretomeP software (Bendtsen et al., 2004a) was used to determine whether some of the QPX proteins, which were found to be extracellular by LC/MS analysis but did not contain an N-terminal signal peptide, were secreted via a non-classical pathway (Geiger et al., 2010). The analysis revealed that 14 QPX extracellular proteins might be transported via non-classical mechanisms. However, SecretomeP has been created based on conserved amino acid motifs found in mammalian proteins (Bendtsen et al., 2004a) and might not reflect the secretion mechanisms in protists.

The predicted secretome of QPX has a very similar composition to that of other pathogenic stramenopiles which have been described to secrete elicitin-like proteins, necrosis-inducing proteins, phospholipases, glycosyl hydrolases, peptidases, acid phosphatases, peptidyl-prolyl-cis–trans isomerases, lectins and glycotransferases (Levesque et al., 2010; Raffaele et al., 2010; Denoeud et al., 2011; Jiang et al., 2013). *Saprolegnia parasitica* is

the only oomycete pathogen predicted to secrete haemolysins (Jiang et al., 2013) and is one of the few oomycete species infective to aquatic animal hosts while the majority of oomycetes are pathogenic to terrestrial plants (Phillips et al., 2008). Interestingly, this study reports haemolysins to be present in the predicted secretome of QPX, another stramenopile infective to an aquatic animal. This could have implications in research on evolution of parasitism within the stramenopiles which live in an aquatic environment. Haemolysins of QPX show 18.8% amino acid sequence similarity to haemolysin E described from bacterial species (Wallace et al., 2000). A number of different types of haemolysins are known from bacteria and protozoan pathogens which are unrelated but exhibit cytolytic and/or cytotoxic activity against a wide range of host cells (e.g. erythrocytes, granulocytes, monocytes and endothelial cells). The expression of one QPX haemolysin was shown to be up-regulated in QPX cells derived from infection sites inside clam tissue, showing that this enzyme plays a role during the infection process. These findings are in agreement with those made by Perrigault and Allam (2009) who demonstrated an increase in cytotoxic activity of QPX ECP towards clam hemocytes for parasite cells supplemented with clam tissue homogenates, suggesting that the cytotoxic activity of QPX secretions may be caused by haemolysin(s). However, this scenario remains speculative without targeted studies assessing the cytotoxic activity of QPX haemolysin against clam hemocytes.

The elicitins and the necrosis-inducing proteins of pathogenic oomycetes are known to cause necrotic cell death or evoke an immune response when they are administered to their plant hosts. The necrosis-inducing protein sequences are found in bacteria, fungi and oomycetes, but the elicitins and elicitin-like protein are apparently exclusive to the oomycetes (Jiang et al., 2006). In this study, we identified two QPX sequences sharing between 33% and 36% amino acid similarity with the elicitin RAM6 from *Phytophthora ramorum* (ABB55989). Some of the elicitin-like proteins have C-terminal extensions rich in threonine, serine and proline residues as an indication of extensive glycosylation (Jiang et al., 2006). The higher expression of one QPX elicitin-like protein inside the infection nodules indicates its possible involvement in QPX interactions with the clam host.

QPX was also predicted to secrete a protein containing conserved domain coding for delta endotoxin N (IPR005639). The only endotoxin N which has been characterised in the literature is derived from *Bacillus thuringiensis*; the toxin causes lysis of the epithelial cells of the gut of infected insects (Boncheva et al., 2006; Ito et al., 2006; Srinivasan, 2008). The presence of delta endotoxin in the QPX transcriptome might be a consequence of horizontal gene transfer (HGT) from a prokaryotic donor. There have been other known examples of HGT facilitating evolution of parasitism in oomycetes (Richards et al., 2011). QPX also possesses a predicted ECP which contains a conserved protein domain named necrosis-inducing protein (IPR008701, NPP1). The NPP1 purified from a few *Phytophthora* pathogens has been shown to cause necrotic cell death on tested plant host species (Fellbrich et al., 2002; Qutob et al., 2002). The transcripts of endotoxin from QPX and the NPP were not up-regulated inside the infection nodules. These results might reflect either a lack of importance of these factors in QPX virulence, their requirement at different times during the infection, or a lack of transcriptional level regulation of these molecules. Post-translational protein modification or other means of protein activation might, for instance, be more important than the transcriptional level regulation of expression of these factors.

The present study used MS to identify 56 different proteins in QPX mucus secretion. Important components of QPX secretion are mucin-like glycoproteins which can be identified by the presence of proline- and threonine-rich motifs. These motifs (PT or serine-PT) are found in mucins from different kingdoms and are sites of extensive glycosylation (Jain et al., 2001; Lang et al., 2007). Three of the extracellular glycoproteins of QPX contain 50% of the amino acid content composed of proline, threonine, serine and alanine, which is a characteristic of mammalian mucins (Jain et al., 2001). Mucin-like molecules found in protozoan cells are either secreted into extracellular space or attached to the cell membrane by glycosylphosphatidylinositol (GPI) anchors (Hicks et al., 2000; Guha-Niyogi et al., 2001; Jain et al., 2001). The main role of mucins and mucin-like molecules in parasitic protozoa such as the kinetoplastid (*Trypanosoma*, *Leishmania*), apicomplexan (*Cryptosporidium*) and amoebic (*Entamoeba*) parasites is cell surface protection (might coat entire cells) facilitating establishment of infection sites (Jain et al., 2001). In the present study, the expression of two (out of four tested) QPX mucin-like molecules was increased under in vivo conditions compared within vitro cultures, suggesting that QPX mucins may also play a role during the infection process and are likely involved in the protection of QPX cells against clam hemocytes phagocytosis or encapsulation. These conclusions are in agreement with previous findings reporting that the mucus layer surrounding QPX cells provides protection to the parasite against clam defense factors (Anderson et al., 2003).

Mucin-like molecules are also involved in other host-parasite interactions such as attachment to host cells and tissues via protein-carbohydrate binding (Hicks et al., 2000). Some parasites which are known to secrete mucins also secrete mucin-degrading enzymes which might be used by the parasite to degrade the protective mucus lining of host epithelial tissue (Hicks et al., 2000). Extracellular glycosyl hydrolases and extracellular transferases have been hypothesised to participate in cross-linking between glycoproteins and polysaccharides of the pathogen and the host (Denoeud et al., 2011). In this study, QPX has been shown (LC/MS results) to secrete five glycosyl hydrolases and has been predicted (in silico data) to secrete over 30 glycosyl hydrolases and seven glycotransferases, which should be subjected to further investigation for their potential roles in QPX infection of clam tissue.

The in silico probing of the transcriptome of QPX allowed the identification of five proteins containing conserved amino acid motifs commonly found in integrins, one of which was detected by LC/MS in QPX ECPs. Integrin-mediated adhesion is the main adhesion mechanism of metazoan cells (Hynes, 2002). However, integrin-like molecules have also been shown to mediate adhesion in the protozoa *Dictyostelium discoideum*, which secretes a protein containing integrin and a von Willebrand factor type A domain (Cornillon et al., 2006). Investigation of the *Entamoeba histolytica* beta-integrin receptor showed that it functions in adhesion (a function similar to the integrins of metazoans) to the human fibronectin of the extracellular matrix (Talamas-Rohana et al., 1998). Similarly, a number of apicomplexan pathogens possess adhesive thrombospondins, surface molecules containing integrin-related protein domains (Yuda et al., 1999; Harper et al., 2004; Pereira et al., 2011). Additional studies are required to test the role of QPX integrins and their possible roles in QPX adhesion or recognition of clam cells and extracellular matrix.

The analysis of the QPX transcriptome also allowed identification of five lectins, one of which was detected in QPX ECPs. Lectin is a very general term describing structurally diverse proteins whose main function is binding to carbohydrates, fulfilling the role of recognition (as receptors) or of attachment (as adhesins). They bind either mono- or oligosaccharides reversibly and with high specificity, and each lectin molecule contains one or more carbohydrate-binding sites, i.e., they are mono-, di- or polyvalent (Lis and Sharon, 1998). Lectins of pathogenic organisms are often considered virulence factors due to their role in the initiation of host colonisation. In pathogenic bacteria, some fimbriae (or pili) which are important for the attachment and initiation of infection are par-

tially composed of lectins (Sharon, 2006). Pathogenic protozoans secrete lectin-like molecules that interact with host cells and host carbohydrates (Hicks et al., 2000). The predicted extracellular lectins identified in this study are excellent targets for future investigations of factors involved in the adhesion to, and interactions with, clam tissues.

The LC/MS analysis showed that eight different peptidases are secreted by QPX when it is cultivated in one type of medium. The mRNA levels of two of these peptidases (one subtilisin and one papain-like peptidase) were recorded to be higher for QPX cells present in infection nodules, suggesting possible involvement of these peptidases in QPX interactions with the host. In fact, subtilisins are universal virulence factors secreted by bacterial, fungal and protozoan pathogens (Monod et al., 2002; Withers-Martinez et al., 2004; Kennan et al., 2010; Muszewska et al., 2011). The papain-like cysteine peptidases are the most studied peptidases of parasitic protozoa and the family contains some of the best described virulence factors such as falcipains of *Plasmodium falciparum* and cruzain of *Trypanosoma cruzi* (Lecaille et al., 2002; Atkinson et al., 2009). The specific functions of the extracellular subtilisin- and papain-like peptidases of QPX in clam tissues are yet to be investigated but the data from this study suggests that they are important in disease development. Extracellular peptidases are major constituents of the QPX predicted secretome (71 different peptidases have been predicted to possess a signal peptide), so it is very likely that additional peptidases also play a role in QPX virulence, warranting further studies.

In conclusion, this study used a combination of complementary transcriptomic and LC/MS analyses to explore and identify proteins secreted by QPX to the surrounding environment. The proteomics analysis revealed 56 proteins secreted into the culture media by QPX cells grown in MEM supplemented with yeastolate. The analysis revealed that QPX ECPs include several mucin-like molecules predicted to represent high molecular weight glycoconjugates. In addition, results of the gene expression experiment suggest that some of the mucin-like molecules are directly involved in the infection process. QPX cells also secrete other proteins including a serine and a cysteine peptidase which are excellent targets for future investigations of QPX biology and virulence.

Acknowledgments

This research was supported by projects R/XG-19 and R/FBM-34 (to BA and EPE), funded under awards NA07OAR4170010 from the National Sea Grant College Program of NOAA to the Research Foundation of State University of New York, USA on behalf of New York Sea Grant. The study was also partially supported by project IOS 1050596 (to BA and EPE) supported by the National Science Foundation, USA. The sequencing of the QPX transcriptome was made with support from the Gordon and Betty Moore Foundation's Marine Microbial Eukaryote Transcriptome Sequencing Project (to BA).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijpara.2014.10.008.

References

Anderson, R.S., Kraus, B.S., McGladdery, S.E., Smolowitz, R., 2003. QPX, a pathogen of quahogs (hard clams), employs mucoid secretions to resist host antimicrobial agents. J. Shellfish Res. 22, 205–208.

Atkinson, H.J., Babbitt, P.C., Sajid, M., 2009. The global cysteine peptidase landscape in parasites. Trends Parasitol. 25, 573–581.

Bendtsen, J.D., Jensen, L.J., Blom, N., Von Heijne, G., Brunak, S., 2004a. Feature-based prediction of non-classical and leaderless protein secretion. PEDS 17, 349–356.

Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004b. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795.

Boncheva, R., Dukiandjiev, S., Minkov, I., de Maagd, R.A., Naimov, S., 2006. Activity of *Bacillus thuringiensis* delta-endotoxins against codling moth (*Cydia pomonella* L.) larvae. J. Invertebr. Pathol. 92, 96–99.

Cornillon, S., Gebbie, L., Benghezal, M., Nair, P., Keller, S., Wehrle-Haller, B., Charette, S.J., Bruckert, F., Letourneur, F., Cosson, P., 2006. An adhesion molecule in free-living *Dictyostelium* amoebae with integrin beta features. EMBO Rep. 7, 617–621.

Cuervo, P., De Jesus, J.B., Saboia-Vahia, L., Mendonca-Lima, L., Domont, G.B., Cupolillo, E., 2009. Proteomic characterization of the released/secreted proteins of *Leishmania* (*Viannia*) *braziliensis* promastigotes. J. Proteomics 73, 79–92.

Denoeud, F., Roussel, M., Noel, B., Wawrzyniak, I., Da Silva, C., Diogon, M., Viscogliosi, E., Brochier-Armanet, C., Couloux, A., Poulain, J., Segurens, B., Anthouard, V., Texier, C., Blot, N., Poirier, P., Ng, G.C., Tan, K.S., Artiguenave, F., Jaillon, O., Aury, J.M., Delbac, F., Wincker, P., Vivares, C.P., El Alaoui, H., 2011. Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. Genome Biol. 12, R29.

Fellbrich, G., Romanski, A., Varet, A., Blume, B., Brunner, F., Engelhardt, S., Felix, G., Kemmerling, B., Krzymowska, M., Nurnberger, T., 2002. NPP1, a Phytophthora-associated trigger of plant defense in parsley and *Arabidopsis*. Plant J. Cell Mol. Biol. 32, 375–390.

Geiger, A., Hirtz, C., Becue, T., Bellard, E., Centeno, D., Gargani, D., Rossignol, M., Cuny, G., Peltier, J.B., 2010. Exocytosis and protein secretion in *Trypanosoma*. BMC Microbiol. 10, 20.

Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36, 3420–3435.

Guha-Niyogi, A., Sullivan, D.R., Turco, S.J., 2001. Glycoconjugate structures of parasitic protozoa. Glycobiology 11, 45–59.

Harper, J.M., Hoff, E.F., Carruthers, V.B., 2004. Multimerization of the *Toxoplasma gondii* MIC2 integrin-like A-domain is required for binding to heparin and human cells. Mol. Biochem. Parasitol. 134, 201–212.

Hicks, S.J., Theodoropoulos, G., Carrington, J.C., Corfield, A.P., 2000. The role of mucins in host–parasite interactions. Part 1 – protozoan parasites. Parasitol. Today 16, 476–481.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K., 2007. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 35, W585–W587.

Hynes, R.O., 2002. Integrins: bidirectional, allosteric signaling machines. Cell 110, 673–687.

Ito, T., Bando, H., Asano, S., 2006. Activation process of the mosquitocidal delta-endotoxin Cry39A produced by *Bacillus thuringiensis* subsp. *aizawai* BUN1-14 and binding property to *Anopheles stephensi* BBMV. J. Invertebr. Pathol. 93, 29–35.

Jain, M., Karan, D., Batra, S.K., Vershney, G.C., 2001. Mucins in protozoan parasites. Front Biosci. 6, 1276–1283.

Jiang, R.H., de Bruijn, I., Haas, B.J., Belmonte, R., Lobach, L., Christie, J., van den Ackerveken, G., Bottin, A., Bulone, V., Diaz-Moreno, S.M., Dumas, B., Fan, L., Gaulin, E., Govers, F., Grenville-Briggs, L.J., Horner, N.R., Levin, J.Z., Mammella, M., Meijer, H.J., Morris, P., Nusbaum, C., Oome, S., Phillips, A.J., van Rooyen, D., Rzeszutek, E., Saraiva, M., Secombes, C.J., Seidl, M.F., Snel, B., Stassen, J.H., Sykes, S., Tripathy, S., van den Berg, H., Vega-Arreguin, J.C., Wawra, S., Young, S.K., Zeng, Q., Dieguez-Uribeondo, J., Russ, C., Tyler, B.M., van West, P., 2013. Distinctive expansion of potential virulence genes in the genome of the oomycete fish pathogen *Saprolegnia parasitica*. PLoS Gene. 9, e1003272.

Jiang, R.H., Tyler, B.M., Whisson, S.C., Hardham, A.R., Govers, F., 2006. Ancient origin of elicitin gene clusters in *Phytophthora* genomes. Mol. Biol. Evol. 23, 338–351.

Kale, S.D., Tyler, B.M., 2011. Entry of oomycete and fungal effectors into plant and animal host cells. Cell. Microbiol. 13, 1839–1848.

Kamoun, S., 2006. A catalogue of the effector secretome of plant pathogenic oomycetes. Annu. Rev. Phytopathol. 44, 41–60.

Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A., Armbrust, E.V., Archibald, J.M., Bharti, A.K., Bell, C.J., Beszteri, B., Bidle, K.D., Cameron, C.T., Campbell, L., Caron, D.A., Cattolico, R.A., Collier, J.L., Coyne, K., Davy, S.K., Deschamps, P., Dyhrman, S.T., Edvardsen, B., Gates, R.D., Gobler, C.J., Greenwood, S.J., Guida, S.M., Jacobi, J.L., Jakobsen, K.S., James, E.R., Jenkins, B., John, U., Johnson, M.D., Juhl, A.R., Kamp, A., Katz, L.A., Kiene, R., Kudryavtsev, A., Leander, B.S., Lin, S., Lovejoy, C., Lynn, D., Marchetti, A., McManus, G., Nedelcu, A.M., Menden-Deuer, S., Miceli, C., Mock, T., Montresor, M., Moran, M.A., Murray, S., Nadathur, G., Nagai, S., Ngam, P.B., Palenik, B., Pawlowski, J., Petroni, G., Piganeau, G., Posewitz, M.C., Rengefors, K., Romano, G., Rumpho, M.E., Rynearson, T., Schilling, K.B., Schroeder, D.C., Simpson, A.G., Slamovits, C.H., Smith, D.R., Smith, G.J., Smith, S.R., Sosik, H.M., Stief, P., Theriot, E., Twary, S.N., Umale, P.E., Vaulot, D., Wawrik, B., Wheeler, G.L., Wilson, W.H., Xu, Y., Zingone, A., Worden, A.Z., 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLoS Biol. 12, e1001889.

Kennan, R.M., Wong, W., Dhungyel, O.P., Han, X., Wong, D., Parker, D., Rosado, C.J., Law, R.H., McGowan, S., Reeve, S.B., Levina, V., Powers, G.A., Pike, R.N., Bottomley, S.P., Smith, A.I., Marsh, I., Whittington, R.J., Whisstock, J.C., Porter, C.J., Rood, J.I., 2010. The subtilisin-like protease AprV2 is required for virulence and uses a novel disulphide-tethered exosite to bind substrates. PLoS Pathal. 6, e1001210.

Kleinschuster, S.J., Smolowitz, R., Parent, J., 1998. In vitro life cycle and propagation of quahog parasite unknown. J. Shellfish. Res. 17, 75–78.

Lang, T., Hansson, G.C., Samuelsson, T., 2007. Gel-forming mucins appeared early in metazoan evolution. Proc. Natl. Acad. Sci. U.S.A. 104, 16209–16214.

Lecaille, F., Kaleta, J., Bromme, D., 2002. Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chem. Rev. 102, 4459–4488.

Levesque, C.A., Brouwer, H., Cano, L., Hamilton, J.P., Holt, C., Huitema, E., Raffaele, S., Robideau, G.P., Thines, M., Win, J., Zerillo, M.M., Beakes, G.W., Boore, J.L., Busam, D., Dumas, B., Ferriera, S., Fuerstenberg, S.I., Gachon, C.M., Gaulin, E., Govers, F., Grenville-Briggs, L., Horner, N., Hostetler, J., Jiang, R.H., Johnson, J., Krajaejun, T., Lin, H., Meijer, H.J., Moore, B., Morris, P., Phuntmart, V., Puiu, D., Shetty, J., Stajich, J.E., Tripathy, S., Wawra, S., van West, P., Whitty, B.R., Coutinho, P.M., Henrissat, B., Martin, F., Thomas, P.D., Tyler, B.M., De Vries, R.P., Kamoun, S., Yandell, M., Tisserat, N., Buell, C.R., 2010. Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. Genome Biol. 11, R73.

Lis, H., Sharon, N., 1998. Lectins: carbohydrate-specific proteins that mediate cellular recognition. Chem. Rev. 98, 637–674.

Liu, Q., Allam, B., Collier, J.L., 2009. Quantitative real-time PCR assay for QPX (Thraustochytriidae), a parasite of the hard clam (*Mercenaria mercenaria*). Appl. Environ. Microbiol. 75, 4913–4918.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–408.

Monod, M., Capoccia, S., Lechenne, B., Zaugg, C., Holdom, M., Jousson, O., 2002. Secreted proteases from pathogenic fungi. Int. J. Med. Microbiol. 292, 405–419.

Muszewska, A., Taylor, J.W., Szczesny, P., Grynberg, M., 2011. Independent subtilases expansions in fungi associated with animals. Mol. Biol. Evol. 28, 3395–3404.

Nickel, W., Rabouille, C., 2009. Mechanisms of regulated unconventional protein secretion. Nat. Rev. Mol. Cell Biol. 10, 148–155.

Pereira, L.M., Candido-Silva, J.A., De Vries, E., Yatsuda, A.P., 2011. A new thrombospondin-related anonymous protein homologue in *Neospora caninum* (NcMIC2-like1). Parasitology 138, 287–297.

Perrigault, M., Allam, B., 2009. Cytotoxicity of quahog parasite unknown (QPX) toward hard clam (*Mercenaria mercenaria*) haemocytes and interactions between different pathogen isolates and host strains. Parasitology 136, 1281–1289.

Perrigault, M., Bugge, D.M., Hao, C.C., Allam, B., 2009. Modulatory effects of hard clam (*Mercenaria mercenaria*) tissue extracts on the in vitro growth of its pathogen QPX. J. Invertebr. Pathol. 100, 1–8.

Petersen, T.N., Brunak, S., Von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.

Phillips, A.J., Anderson, V.L., Robertson, E.J., Secombes, C.J., van West, P., 2008. New insights into animal pathogenic oomycetes. Trends Microbiol. 16, 13–19.

Qian, H., Liu, Q., Allam, B., Collier, J.L., 2007. Molecular genetic variation within and among isolates of QPX (Thraustochytridae), a parasite of the hard clam *Mercenaria mercenaria*. Dis. Aquat. Organ. 77, 159–168.

Qutob, D., Kamoun, S., Gijzen, M., 2002. Expression of a *Phytophthora sojae* necrosis-inducing protein occurs during transition from biotrophy to necrotrophy. Plant J. Cell Mol. Biol. 32, 361–373.

Raffaele, S., Win, J., Cano, L.M., Kamoun, S., 2010. Analyses of genome architecture and gene expression reveal novel candidate virulence factors in the secretome of *Phytophthora infestans*. BMC Genomics 11, 637.

Ragone Calvo, L.M., Walker, J.G., Burreson, E.M., 1998. Prevalence and distribution of QPX, quahog parasite unknown, in hard clams *Mercenaria mercenaria* in Virginia, USA. Dis. Aquat. Organ. 33, 209–219.

Richards, T.A., Soanes, D.M., Jones, M.M., Vasieva, O., Leonard, G., Paszkiewicz, K., Foster, P.G., Hallc, N., Talbota, N.T., 2011. Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. Proc. Natl. Acad. Sci. 108, 15258–15263.

Sharon, N., 2006. Carbohydrates as future anti-adhesion drugs for infectious diseases. Biochim. Biophys. Acta 1760, 527–537.

Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Boucherie, H., Mann, M., 1996. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. U.S.A. 93, 14440–14445.

Silverman, J.M., Chan, S.K., Robinson, D.P., Dwyer, D.M., Nandan, D., Foster, L.J., Reiner, N.E., 2008. Proteomic analysis of the secretome of *Leishmania donovani*. Genome Biol. 9, R35.

Smith, A.M., Robinson, T.M., Salt, M.D., Hamilton, K.S., Silvia, B.E., Blasiak, R., 2009. Robust cross-links in molluscan adhesive gels: testing for contributions from hydrophobic and electrostatic interactions. Comp. Biochem. Phys. B 152, 110–117.

Smolowitz, R., Leavitt, D., Perkins, F., 1998. Observations of a protistan disease similar to QPX in *Mercenaria mercenaria* (Hard Clams) from the Coast of Massachusetts. J. Invertebr. Pathol. 71, 9–25.

Srinivasan, R., 2008. Susceptibility of legume pod borer (LPB), *Maruca vitrata* to delta-endotoxins of *Bacillus thuringiensis* (Bt) in Taiwan. J. Invertebr. Pathol. 97, 79–81.

Talamas-Rohana, P., Hernandez-Ramirez, V.I., Perez-Garcia, J.N., Ventura-Juarez, J., 1998. *Entamoeba histolytica* contains a beta 1 integrin-like molecule similar to fibronectin receptors from eukaryotic cells. J. Euk. Microbiol. 45, 356–360.

Tanner, S., Shu, H., Frank, A., Wang, L., Zandi, E., Mumby, M., Pevzner, P.A., Bafna, V., 2005. InsPecT: identification of posttranslationally modified peptides from tandem mass spectra. Anal. Chem. 77, 4626–4639.

Wallace, A.J., Stillman, T.J., Atkins, A., Jamieson, S.J., Artymiuk, P.J., 2000. *E. coli* hemolysin E (HlyE, ClyA, SheA): X-ray crystal structure of the toxin and observation of membrane pores by electron microscopy. Cell 21, 265–276.

Wernersson, R., 2006. Virtual ribosome–a comprehensive DNA translation tool with support for integration of sequence feature annotation. Nucleic Acids Res. 34, W385–W388.

Withers-Martinez, C., Jean, L., Blackman, M.J., 2004. Subtilisin-like proteases of the malaria parasite. Mol. Microbiol. 53, 55–63.

Whyte, S.K., Cawthorn, R.J., McGladdery, S.E., 1994. QPX (Quahog Parasite X), a pathogen of northern quahaug *Mercenaria mercenaria* from the Gulf of St. Lawrence, Canada. Dis. Aquat. Organ. 19, 129–136.

Yuda, M., Sawai, T., Chinzei, Y., 1999. Structure and expression of an adhesive protein-like molecule of mosquito invasive-stage malarial parasite. J. Exp. Med. 189, 1947–1952.