RESEARCH

Mining the secreted and membrane transcriptome of *Hyalomma dromedarii* ticks for identification of potential protective antigens

Nahla A. Hussein[1], Asmaa S. El-Shershaby[1], Shaimaa Abdel-Moez[1], Amr E. El-Hakim[1] and Yasser E. Shahein[1]*

Abstract

Background Members belonging to the tick genus *Hyalomma* function as a multi-host reservoir for several pathogens and important parasites infesting large animals, such as camels, goats, cattle and sheep. In Egypt, there is a high risk of pathogen transmission as camels and cattle are imported from Sudan and Ethiopia and shipped to slaughterhouses and animal markets located in populated areas. *Hyalomma dromedarii* ticks are semi-desert vectors and, similar to other members of the genus *Hyalomma*, characterized by long-term feeding. During this process, different physiological, biochemical and immunological interactions occur within both the feeding ticks and their hosts. These biological changes affect the different tick developmental phases. The aim of this study was to explore the transcriptome of mixed messenger RNAs (mRNAs) collected from *H. dromedarii* eggs, larvae, nymphs and fed and unfed adults, using the Gateway cDNA library prepared in pCMV sport6.1 vector

Methods The clones were sequenced and searched for potential secreted, membrane-associated or transmembrane (SMaT) sequences. The identified SMaT sequences were compared to the National Center for Biotechnology Information (NCBI) non-redundant protein sequence database using Blastx. Annotation and functional classification were achieved by comparison to sequences in the UniProtKB/Swiss-Prot and VectorBase databases and to the publicly available annotated proteomes of six hard tick species (*H. asiaticum*, *Rhipicephalus sanguineus* sensu lato, *Dermacentor silvarum*, *Rhipicephalus microplus*, *Ixodes scapularis* and *Haemaphysalis longicornis*) in addition to the published *H. dromedarii* sialotranscriptome. For the common sequences, we predicted the physicochemical properties, secondary structures and antigenicity of the fragments similar to matched sequences in the UniProtKB/Swiss-Prot database using three different methods.

Results The quality-trimmed sequences from the cDNA library revealed 319 SMaT transcripts among 1248 sequenced clones. Annotation of the SMaT sequences using the UniProtKB/Swiss-Prot database revealed only 232 non-redundant sequences with at least one match. According to the UniProtKB/Swiss-Prot and Vectorbase databases, the SMaT sequences were either secreted (extracellular) (29 sequences) or cellular (transmembrane and membrane-associated) (203 sequences). These were classified into 10 functional classes: biogenesis (49 sequences), defense (9 sequences), development (36 sequences), signal transduction (28 sequences), transport (15 sequences), protein modification (33 sequences), homeostasis (6 sequences), metabolism (45 sequences) and miscellaneous/uncharacterized (11 sequences). A total of 60 sequences were shared between *H. dromedarii* SMaT, the sialotransciptome and six

*Correspondence:
Yasser E. Shahein
yassershahein_nrc@yahoo.com; ye.shahein@nrc.sci.eg

Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 2 of 17

other hard tick species. The peptide fragments of these sequences that aligned to proteins from the UniProtKB/Swiss-Prot database were predicted to be promising epitopes and mapped to 10 functional classes at different ratios.

Conclusions Our immuno-informatics analysis identified 60 sequences common among hard tick species and encoded by *H. dromedarii* salivary glands. These annotated SMaT sequences of *H. dromedarii* will pave the way for the identification and discovery of novel potential protective antigens that are either secreted, membrane-associated or transmembrane.

Keywords *Hyalomma dromedarii*, Transcriptome, Vaccines, Membrane associated proteins, Secreted proteins, cDNA library

Background

Hyalomma dromedarii, a multi-host hard tick similar to other haematophagous arthropods, has major medical significance to both humans and animals. Similar to 44 other species of the Ixodidae family of hard ticks in Egypt, *H. dromedarii* has been reported to infest livestock, both domestic and imported animals for the Egyptian market, including buffaloes, camels, sheep, cattle and goats, as well as dogs and migratory birds crossing the country [1, 2]. Several studies have shown that *Theileria annulata*, *Anaplasma marginale*, *Coxiella burnetii* (Q-fever), *Rickettsia africae* and the Crimean-Congo haemorrhagic fever (CCHF) virus are *H. dromedarii*-borne pathogens that can infect both animals and humans [3–9]. These pathogens cause severe diseases and considerable economic losses.

Traditionally, chemical acaricides have been the most widely and commonly used anti-tick control strategy. However, the use of acaricides has several drawbacks, including harmful effects on the environment, precipitation in milk and meat and, more seriously, the development of acaricide-resistant ticks [10, 11]. As alternatives to the use of these acaricides, numerous biotechnological approaches, including immunization-based methods, have been developed to elicit the host immune response and control tick infestations. These immunological models are based on an understanding of tick biology, especially tick behaviour during blood-feeding. Anti-tick vaccines rely on host immunization with tick proteins found in the mid-gut, salivary glands or eggs or with those associated with Malpighian tubules or cement-like secretions [12–18]. Accordingly, a deeper understanding of tick biology (development, growth, reproduction) may lead to better identification of novel anti-tick vaccines with the potential to efficiently disrupt the tick life-cycle.

Several strategies can be used to identify novel protective antigens that may serve as potential vaccine candidates. Among these, -omics approaches, such as proteomics, transcriptomics and immunomics, have several advantages and have been frequently investigated in the last decade [19]. These provide a generalized overview of ticks proteins/ transcripts that can be upregulated or play essential roles in feeding and reproduction. Examples of such studies include the study of the changes of gene expression during the life-cycle of *Ixodes ricinus* using whole-body RNA [20], the observed clinical signs of *Ornithodoros brasiliensis*-bitten humans using salivary glands RNA [21] and the proteomics of potential epidemiological markers in the sialome of *Hyalomma excavatum* [22]. In addition, the presence of tick-host interacting molecules during blood-feeding and the differential expression of transcripts between male and female ticks in *H. dromedarii* sialome have been reported [23]. In parallel, screening of complementary DNA (cDNA) libraries has aided in the identification of several protective antigens in different ticks species such as *Ixodes scapularis* [24] and *Haemaphysalis longicornis* [25].

However, most studies conducted to date aimed at identifying potential protective antigens have investigated soluble proteins, including cytosolic or secreted extracellular proteins [26]. Membrane-associated and transmembrane proteins, on the other hand, remain poorly investigated, with the notable exception of the prototype Bm86 that is a membrane-anchored protein [18]. This gap could be attributed to the technical challenges encountered when assessing their heterologous expression, such as their low solubility, the need for engineered bacterial expression strains and their tendency to form aggregates [27]. However, cytosolic or secreted extracellular proteins represent promising candidates as they harbor exposed epitopes recognized by the immune system. Their use as antigens may disrupt the functional activities of adhesion, recognition, signaling and regulation [17].

Given this context, there is a potential for using an immuno-informatics approach to predict potential secreted, membrane- associated and transmembrane (SMaT) *H. dromedarii* antigens from the cDNA library. Immuno-informatics plays a crucial role in vaccine development by defining potential antigens from large datasets of proteins and predicting their antigenicity [28]. In this regard, immuno-informatics bypasses the time-consuming steps of in vitro screening and antigen identification.

Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 3 of 17

The study reported here is the first to use an immuno-informatics approach to study the secreted, membrane-associated and transmembrane (SMaT) transcriptome of mixed messenger RNAs (mRNAs) collected from eggs and unfed adults of the hard tick *H. dromedarii*, as well as from *H. dromedarii* at three feeding phases (larvae, nymphs and fed adults). Analysis of the transcripts will give insights into the gene expression profile during growth, development and feeding processes of this tick and will aid in the identification of promising antigens.

Methods

Collection and colony establishment of *H. dromedarii* ticks

Engorged females of *H. dromedarii* were collected from naturally infested camels from Toukh City (30°21′11.6″N, 31°11′ 31.5″E), Qalyubia governorate, Egypt. Tick samples were collected at monthly intervals into glass tubes containing strips of filter paper. All collected ticks were examined under a stereomicroscope (model DM 750; Leica Microsystems, Wetzlar, Germany) and identified according to the key of Walker et al. [29]. For oviposition, fully engorged females were placed separately in plastic cups (1 female/cup) that were tightly closed off with a muslin cloth and then incubated at 25 ± 1 °C and 75–80% relative humidity (RH) in a FRIOCELL incubator (MMM Medcenter Einrichtungen GmbH, Planegg, Germany). It is known that fully engorged females begin to oviposit about 5 days after their collection from camels and that females take nearly 1 month to complete oviposition and for larvae to hatch. One week after hatching, the unfed larvae were allowed to feed on healthy rabbits using the capsule technique [30]. After nearly 2 weeks of feeding, fully engorged nymphs were collected from the rabbits and incubated at 25 ± 1 °C and 75–80% RH for 21 days till molting to unfed adults. After 1 week of molting, an equal number of unfed males and females were allowed to feed on healthy rabbits for 10 days to obtain fully fed females from which to collect eggs and ticks at other life stages. This process ensures that all stages used in the library originate from the same mother to avoid polymorphism and diversity.

Construction of the *H. dromedarii* complementary DNA library

The pCMV-SPORT6.1 vector (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) was used to ligate complementary DNA (cDNA) fragments ≥ 800 bp for construction of the *H. dromedarii* cDNA library in pCMV-SPORT6.1 according to the manufacturer's instructions (Invitrogen, Thermo Fisher Scientific) and applying the basic protocols reported in Sambrook and Russell [31]. RNA was extracted from different tick life stages using TRIzol reagent (Invitrogen, Thermo Fisher

Scientific), and the quality of the RNA was examined by electrophoresis, which showed no degradation of the RNA. mRNA was isolated using the Dynabeads Oligo(dT)$_{25}$ mRNA purification kit (Thermo Fisher Scientific). The first-strand cDNA was synthesized using the Invitrogen SuperScript III Reverse Transcriptase kit; The second-strand cDNA was synthesized using the *E. coli* RNase H, *E. coli* DNA polymerase and *E. coli* DNA ligase following the reaction protocol of the Second Strand cDNA Synthesis Kit (A48570; Invitrogen, Thermo Fisher Scientific). Following size selection of the cDNAs by electrophoresis, the cDNAs were directionally cloned into the vector between NotI and EcoRV restriction sites. The EcoRV site is destroyed after cDNA cloning. Electroporation was performed using DH10B-T1R phage-resistant cells (Thermo Fisher Scientific). The primary library was then amplified and stored in suspension medium of 80% SOC and 20% glycerol. The insert size was detected using the NotI and EcoRI restriction enzymes. To carry out sequencing, we used 2×LB broth to dilute an aliquot of the library followed by spreading of each aliquot on LB agar containing 100 µg/ml ampicillin and incubation overnight at 37 °C. Subsequently, colonies were randomly picked into 96-well plates containing 100 µl of 2×LB broth, 8% glycerol and 100 µg/ml ampicillin, and the plates were incubated overnight at 37 °C. After the incubation period, the plates were stored at − 80 °C until subsequent applications. Colonies from these plates were used to inoculate 1-ml cultures of 2×LB broth containing 100 µg/ml ampicillin, and the cultures were incubated for 16 h at 37 °C. Finally, plasmids were purified and sequenced from the 5' ends using the SP6 primers.

Identification of potential secreted, membrane- associated or transmembrane (SMaT) sequences

A total of 1248 sequences were analysed using the BLAST2Go annotation tool of OmicsBox Bioinformatics software [32]. We manually searched the output file for all sequences predicted to encode a signal peptide and/or transmembrane helix (based on InterPro classification of protein families and domains which is a part of the BLAST2Go package). Based on the BLAST2Go results, the identified gene ontology (GO terms) were analysed using the GO categorizer (available at https://www.animalgenome.org/bioinfo/tools/countgo [33]) based on root GO annotation followed by Slim2 annotation.

Sequence annotation and functional prediction

To annotate the *H. dromedarii* SMaT sequences, we used the blastx alignment tool against the UniProtKB/Swiss-Prot database [34, 35] (with default e-value parameters of e-value=0.05 and maximum hits of 2) and the annotated proteomes of *Hyalomma asiaticum*, *Dermacentor*

Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 4 of 17

silvarum, *Rhipicephalus sanguineus* sensu lato (*R. sanguineus* s.l.), *Rhipicephalus microplus*, *Haemaphysalis longicornis* and *Ixodes scapularis* from the VectorBase database release 64 (with a cut-off e-value of 10^{-3}) [36, 37]. For sequences common with the six tick species, we conducted an alignment with the UniProtKB/Swiss-Prot database with more stringent blastx parameters (e-value $= 10^{-3}$) to determine the best sequences to be used in subsequent studies for epitope prediction.

Identification of *H. dromedarii* SMaT salivary proteins

To identify the SMaT sequences encoded by the *H. dromedarii* salivary glands, we used the publicly available sialotranscriptome of *H. dromedarii* [23]. We created a local database of *H. dromedarii* sialotranscriptome sequences and used tblastx (with a cut-off e-value of 10^{-3}) for comparison and alignment.

Physicochemical properties, secondary structure and antigenic prediction:

The physicochemical properties of sequences of interest were predicted using the Multiple Protein Profiler (MPP) 1.0 tool (https://mproteinprofiler.microbiologyand immunology.dal.ca/) [38] and Phyre2 (http://www.sbg. bio.ic.ac.uk/phyre2/html/page.cgi?id=index) [39].

The potential of the identified *H. dromedarii* SMaT sequences to serve as a potential antigen was predicted using three different tools: Vaxigen- v2 (https://www. ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html) [40, 41], BepiPred-2 (http://tools.iedb.org/bcell/) [42] and the tool generated by the Immunomedicine group of University Complutense of Madrid (http://imed.med.ucm.es/Tools/ antigenic.pl) [43]. The average score was calculated based on the average of the three prediction scores.

Graphical representations

Graphical representations were created using GraphPad Prism version 8 (GraphPad Software, San Diego, CA, USA), MS Excel (Microsoft Corp., Redmond, WA, USA), Venn diagrams using the software available at https:// bioinformatics.psb.ugent.be/webtools/Venn/ and Heat maps using http://www.heatmapper.ca/ [44].

Results and discussion

Construction of *H. dromedarii* cDNA library and sequencing of random clones

In order to gain insights into the *H. dromedarii* transcriptome and identify promising vaccine candidates, we followed the workflow described in Fig. 1. Different life stages of *H. dromedarii* ticks (eggs, larvae, nymphs and adults) were collected after the necessary colony has been established, and aliquots of 100 µg of total RNA from each stage were pooled to further extract the mRNA. The

pattern of tick total RNA and ribosomal RNA (rRNA) was similar to the control RNA. The rRNA from ticks (Fig. 2a, and lane 2 in Fig. 2b) showed a pattern of two large bands and one small band, although the sizes of the bands were smaller than those of the control tissue rRNA (Fig. 2b, lane 4). The polyA mRNA from *H. dromedarii* looked similar to that in the human tissue control (Fig. 2b, lane 5), but it contained much more rRNA (Fig. 2b, lane 3) indicating that it bonds to poly-dT oligo. We subsequently synthesized cDNA from poly-dT priming-based reverse transcription, and the size range was much smaller than the expected size, which may reach $\geq$ 2 kb. Cap-antibody selection trials revealed that the yield was much lower (5- to 10-fold; first-strand cDNA/ mRNA) and the size range also smaller (Fig. 3a, lane 2) than those from control tissue. After using fourfold more starting mRNA (2 µg) material for the first-strand cDNA synthesis, Cap-antibody and size selection were conducted and we completed the library following the protocol of the manufacturer using the pCMV-SPORT6.1 (Invitrogen, Thermo Fisher Scientific). The titer of the primary library was 0.39×10^6 cfu/ml and the average size was approximately 1.2 kb; 96% of the clones contained inserts.

Sequencing of 23 randomly selected primary clones showed that 57% of clones (13 clones) partially or perfectly matched the 16S or 28S rRNAs of *H. dromedarii*, *Dermacentor* sp. or *Dermacentor andersoni*, that 50% of clones were AT-rich (>95%) RNA sequences that included many of the 16S rRNAs and that 26% of clones from non-AT-rich sequences contained short open reading frame (ORFs). Most of these sequences were novel with no hits (7 clones) using NCBI nucleotide blast analysis, or they showed a partially match to *R. sanguineus* s.l. mitochondrial DNA (GenBank accession: AF081829.1) (1 clone) or were homologues to the *H. dromedarii COI* gene for cytochrome oxidase subunit I (Genebank accession AJ437075.1) (1 clone) and *Haemaphysalis qinghaiensis* clone HqL08 unknown mRNA (Genebank accession EF605264.1) (1 clone). In brief, low yield of cDNA from Cap-antibody selection directed us to enrich the non-ribo mRNA (Fig. 2b, lane 3) and use high-temperature oligo-dT priming for reverse transcription. The library thus produced showed a total of 3.12×10^6 cfu, with 88% of clones containing inserts with an average insert size of 0.8 kb. The library was then amplified, and we reached 5.1×10^9 cfu/ml total titer; 87% of the clones contained inserts. The size of the inserts was detected using NotI and EcoRI restriction enzymes.

Customized classic cDNA libraries and novel next-generation sequencing (NGS) are two methods that are available to identify genetic information of specific RNA. However, both technologies have advantages
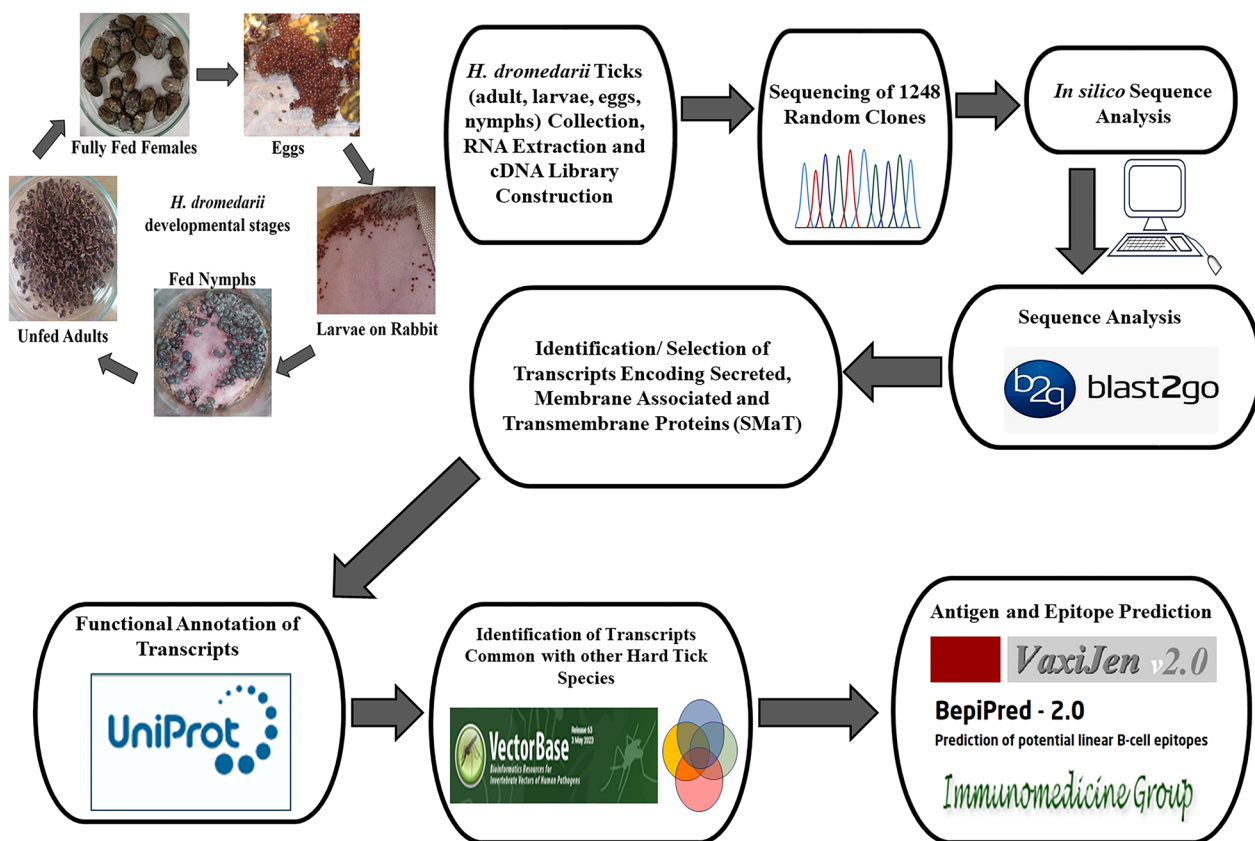
Fig. 1 Schematic illustration of the workflow for *Hyalomma dromedarii* transcriptome mining starting from tick colony preparation, collecting the developmental stages, RNA isolation followed by complementary DNA (*cDNA*) library construction and finally sequencing and analysis of transcripts
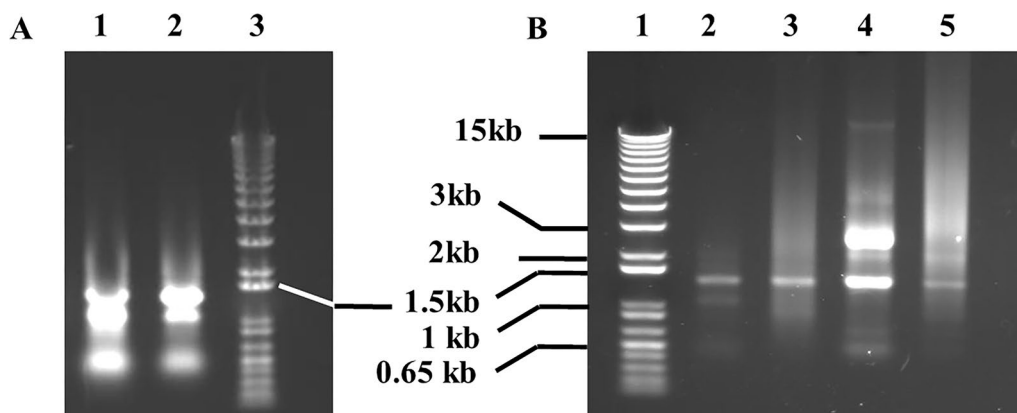

Fig. 2 *Hyalomma dromedarii* total RNA and polyA messenger RNA (mRNA). **A** Lanes: *1*, *2* tick total RNA, *3* 1 kb+ DNA ladder. **B** Lanes: *1* 1 kb+ DNA ladder, *2* tick total RNA, *3* tick polyA mRNA, *4*, *5* human kidney total RNA and polyA mRNA as a control, respectively

and disadvantages depending on the target study project or a specified biological, medical or biotechnological application. NGS exhibits higher throughput, more flexibility and greater sensitivity than the sequencing results obtained with the classic cDNA library, and it has

the advantage of full exome sequencing and differential gene expression analyses [45]. However, even with the longest reads, in silico assembly is an indispensable step to obtain contigs and scaffolds. Nevertheless, in organisms like *H. dromedarii* where no reference genome/
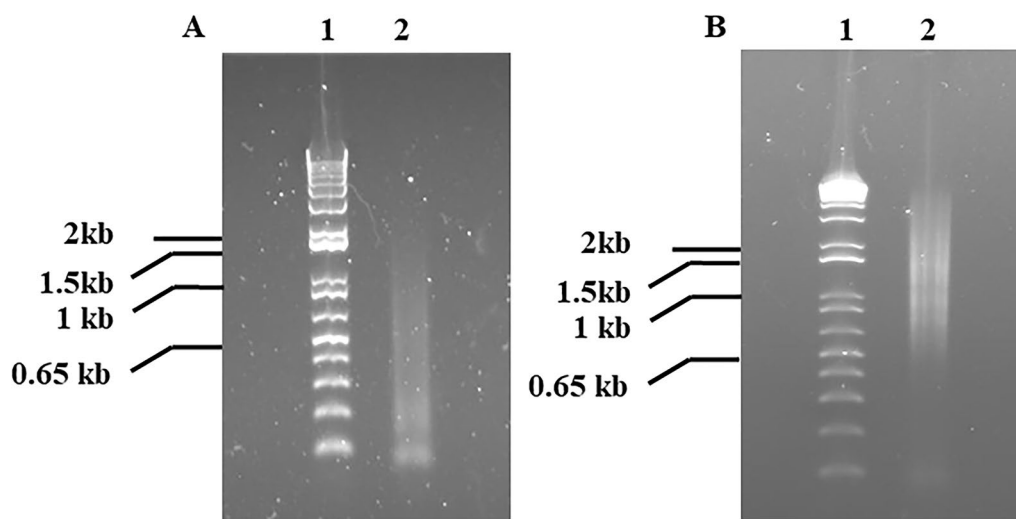
Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 6 of 17

Fig. 3 *Hyalomma dromedarii* complementary DNA (cDNA) after Cap-antibody selection. **A** Lanes: *1* 1 kb+ DNA ladder, *2* Cap-antibody-selected cDNA from tick. **B** Lanes: *1* 1 kb+ DNA ladder, *2* Cap-antibody-selected cDNA from human kidney as a control

exome is available, in vitro assembly may lead to chimeric sequence generation. Accordingly, we believe a cDNA library is a promising alternative approach if our aim is to identify potential antigenic epitope(s) encoded within single ORFs.

An additional advantage of the classic cDNA library is that it enabled us to obtain stable clones containing fragments of expressed genes from mature mRNA in a vector ready for use in several applications. The applications that can be carried out on our cDNA library may be positive selection of cDNA, direct PCR amplification of the gene, in vitro transcription analysis, screening of plates using a specified probe and/or functional analysis using the eukaryotic CMV promoter.

The novelty of the current work has three aspects. First, the identified transcripts are from different developmental stages of *H. dromedarii*. Second, the identified RNA sequences are in stable constructs ready for subsequent applications. Finally, we focused on the identification of three types of proteins, namely the secreted, membrane-associated or transmembrane proteins. It is important to mention that eggs, larvae, nymphs and adults originated from a stable colony to avoid polymorphism.

In the current study, we faced several challenges during the construction of the *H. dromedarii* library. First, the smaller size range of cDNAs was around 800 nucleotides, which may be attributed to the premature termination of reverse transcription as a result of the secondary structure in rRNAs that dominated in the reaction [46, 47]. Moreover, reactions showed low cDNA yield from Cap-antibody selection, possibly due to the lack/less/different cap structure from tick RNA since most known rRNAs

do not contain a regular cap structure. It is worth mentioning that some non-specific RNA (e.g. 16S rRNA) binding to Cap-antibody beads could be carried over, especially when high amounts of beads and mRNA are used . Finally, we could not normalize the library, which may be attributed to the high percentage of AT-rich sequences; instead we performed the non-ribo mRNA enrichment (Fig. 2b, lane 3).

Identification and preliminary analysis of SMaT sequences within the *H. dromedarii* cDNA library

The sequences obtained from the 1248 clones using the OmicsBox Bioinformatics software [32] were analysed, and the output for all sequences identified as secreted, membrane- associated or transmembrane was searched manually. A total of 448 sequences were identified. All sequences < 250 bp and sequences showing high similarity to the cloning vector pCMV, including the promoter sequences, or sequences encoding the antibiotic resistance marker beta-lactamase were excluded, resulting in 319 sequences used for further analysis, which were submitted to GenBank under accession numbers PP584923–PP585243) (Additional file 1: Table S1 and Additional file 2: Fasta 1).

Regarding the length of the *H. dromedarii* SMaT sequences, 106 sequences ranged from 700 to 800 bp, 134 ranged from 600 to 700 bp, 57 ranged from 500 to 600 bp, 14 ranged from 400 to 500 bp, five ranged from 300 to 400 bp and three ranged from 200 to 300 bp (Fig. 4a). Sequence alignment by Blastx against the non-redundant protein sequence database of NCBI using OmicsBox Bioinformatics software [32] revealed that
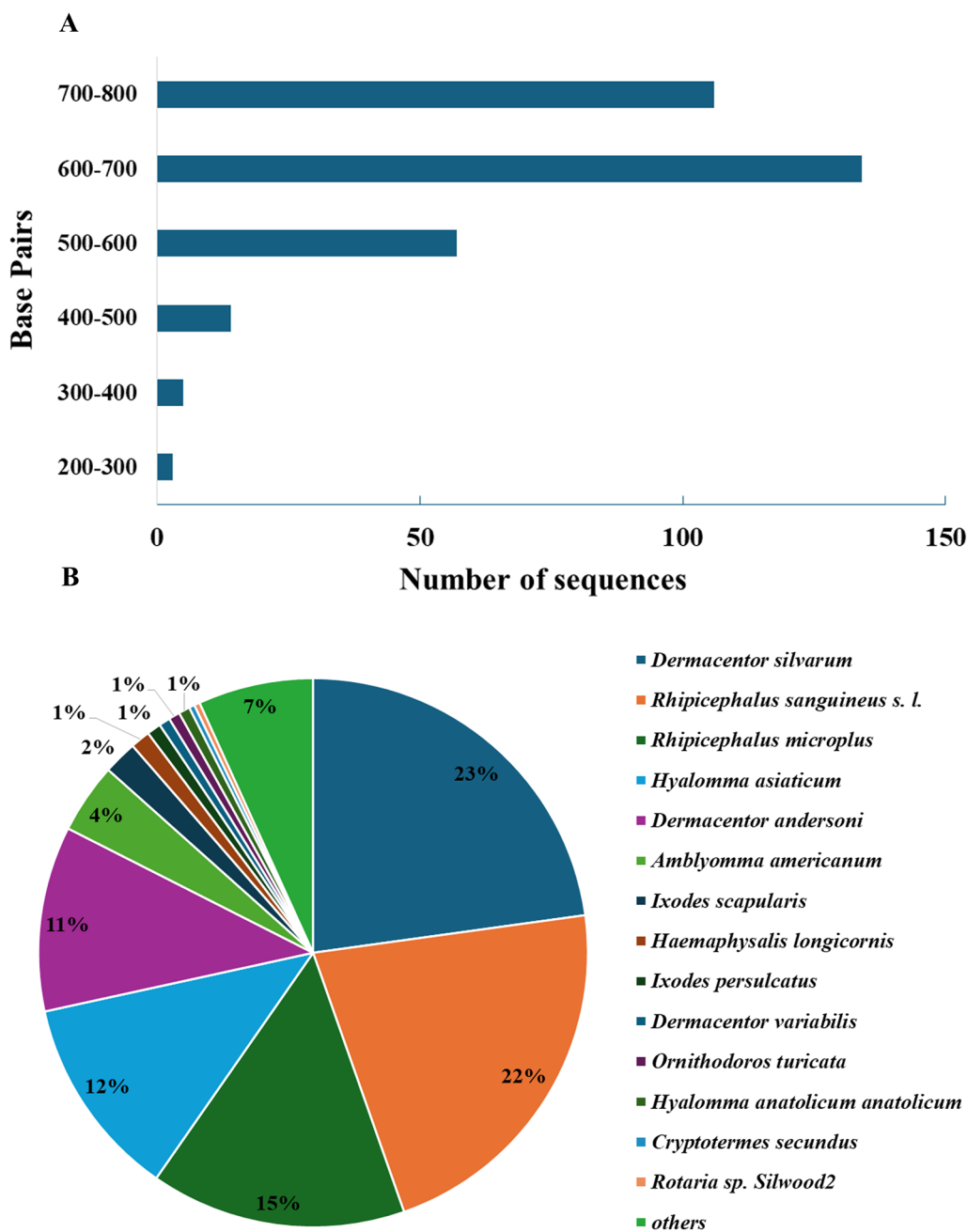
Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 7 of 17

A

B

Fig. 4 Abundance and length of SMaT-encoding sequences in the *Hyalomma dromedarii* complementary DNA (cDNA) library. **A** Sequence length distribution of the 319 SMaT-encoding sequences used in the analysis. **B** Percentage of *H. dromedarii* SMaT-encoding sequences that had matched hits with other ticks according to alignment using blastx. SMaT, Secreted, membrane-associated or transmembrane

only 126 sequences matched NCBI sequences (including 18 uncharacterized proteins and 58 hypothetical proteins). *Hyalomma dromedarii* SMaT sequences matched hits from several tick species, including *D. silvarum* (23%), *R. sanguineus* s.l. (22%), *R. microplus* (15%), *H. asiaticum* (12%), *D. andersoni* (11%), *Amblyomma*

americanum (4%) and other species (Additional file 1: Table S1; Fig. 4b).

The 319 SMaT sequences matched 164 GO identities, which fell into the categories of cellular component (27%), molecular function (27%) and biological process (46%), (Fig. 5a). These were further categorized
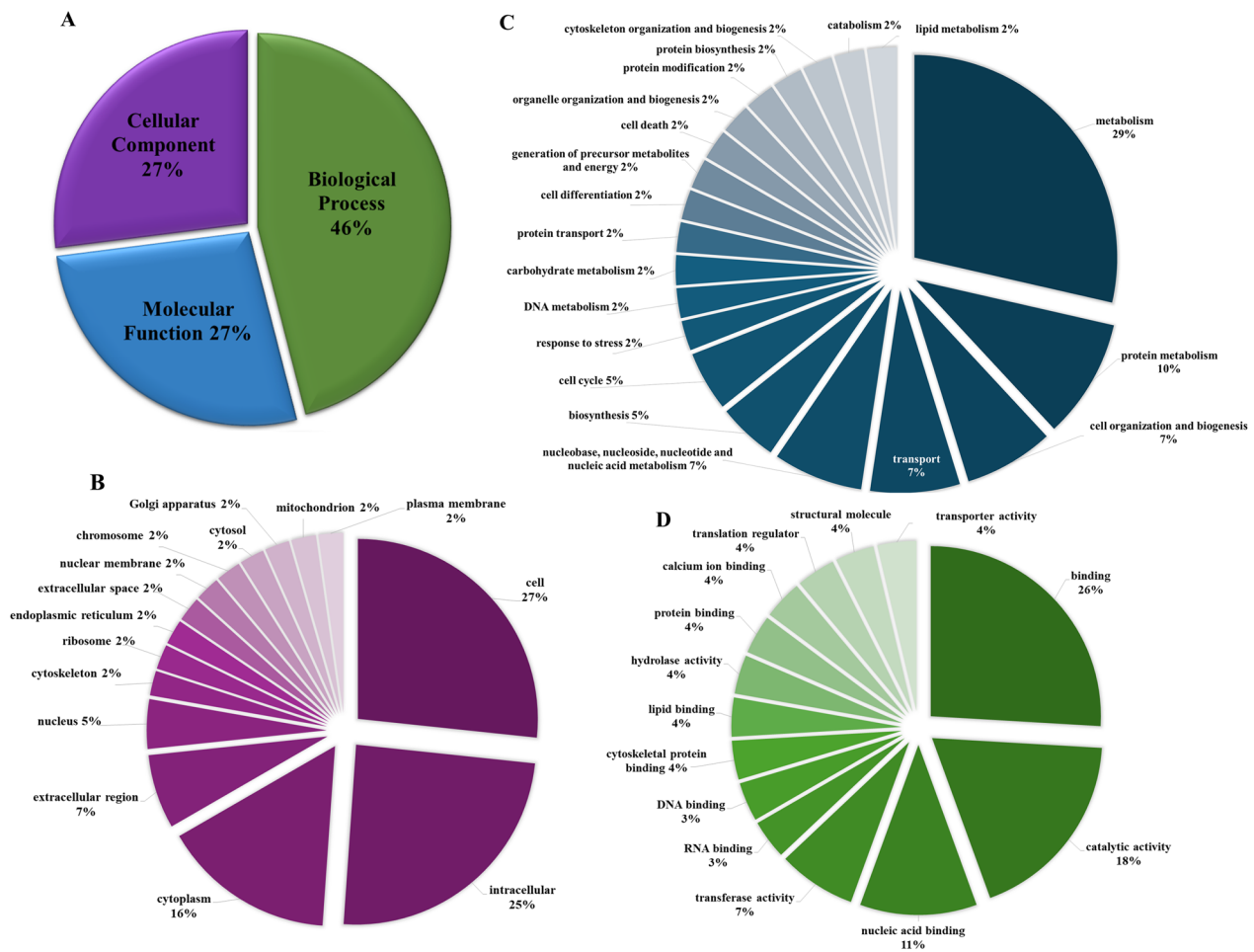
Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 8 of 17

Fig. 5 Gene ontology (GO) analysis of *Hyalomma dromedarii* SMaT sequences and their subcategories. **a** Abundance of GO root terms in *H. dromedarii* SMaT sequences identified in this study. **b** Subcategories of GO cellular component terms as detected by the Slim2 framework. **c** subcategories of GO molecular function terms as detected by the Slim2 framework. **d** subcategories of GO biological process terms as detected by the Slim2 framework. SMaT, Secreted, membrane-associated or transmembrane

into subcategories according to the Slim2 framework as depicted in Fig. 5B–D.

Subcellular and functional classification of the *H. dromedarii* SMaT transcriptome

To classify and predict the function of the identified 319 SMaT sequences, we compared them by blastx to the manually curated database UniProtKB/Swiss-Prot [34, 35], with a maximum of two hits per sequence. Our results revealed that 232 sequences matched hits from UniProtKB/Swiss-Prot (where 48 sequences matched only 1 hit, 184 sequences matched 2 hits (Additional file 3: Table S2) and 87 sequences did not match any hit). We noted the high similarity between sequences P7_H07 and P7_C07 (both matching UniProtKB/Swiss-Prot entry A1KEV7) and between sequences P11_B12 and P13_C08 (both matching UniProtKB/

Swiss-Prot entries A4D998 and Q63515), suggesting that each identical pair might originate from the same ORF. Overall, the 232 sequences were predicted to be either secreted (29 sequences) or cellular, including transmembrane and membrane-associated (203 sequences). They were classified into 10 functional classes: biogenesis (including DNA, RNA and protein biogenesis; 49 sequences), protein modification (33 sequences), metabolism (45 sequences), development (including cell division and skeleton formation; 36 sequences), defense (9 sequences), signal transduction (28 sequences), transport (15 sequences), homeostasis (6 sequences) and miscellaneous and uncharacterized (11 sequences) (Fig. 6). The importance and role of some of these functional classes to which the identified *H. dromedarii* SMaT sequences belong are discussed briefly in the following sections.
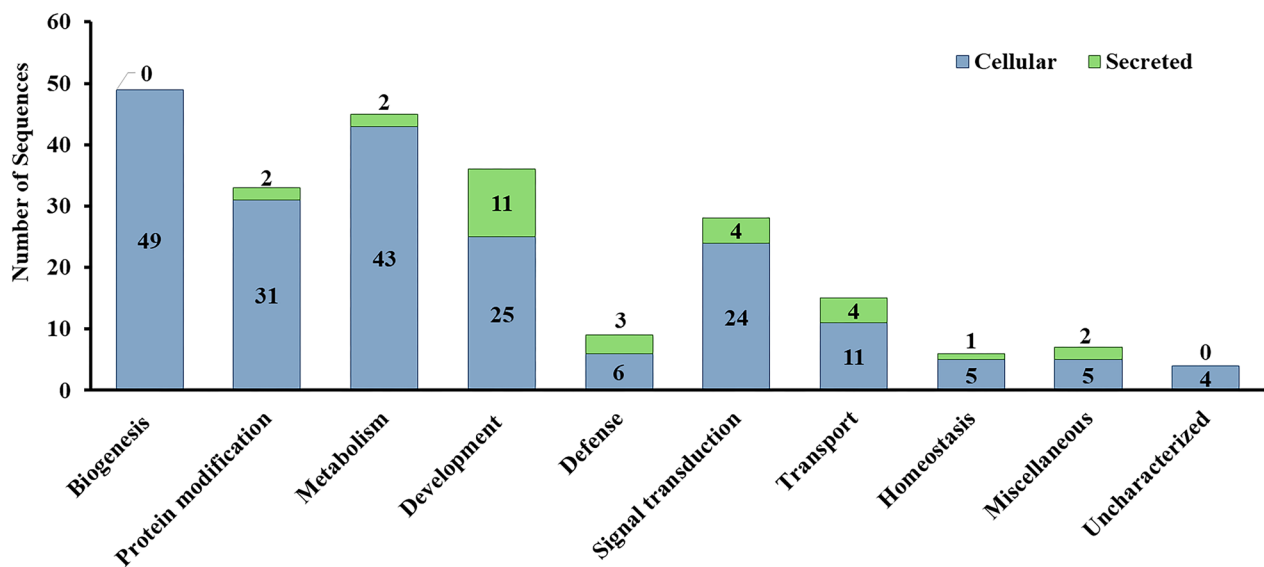
Fig. 6 Functional classification of the 232 *Hyalomma dromedarii* SMaT sequences matching UniProtKB/Swiss-Prot entries. SMaT, Secreted, membrane-associated or transmembrane

Biogenesis class

This class includes *H. dromedarii* SMaT sequences encoding the synthesis of biomolecules (DNA, RNA and proteins). Most of these sequences are predicted to encode potential ribosomal proteins or ribosomal protein subunits that are involved in ribosome biogenesis and protein translation, such as clones (P1_E07, P2_C02, P3_E08, P4_A03 and P7_E01). In ticks, ribosomal biogenesis is necessary for cell viability, survival and feeding [48]. Ribosomal proteins are involved in many extra-ribosomal functions, including genomic stability, apoptosis, cell proliferation and development [49, 50]. They have also been shown to associate with membrane proteins decorating the cell surface of other parasites, such as *Plasmodium* and *Toxoplasma gondii* [51]. Several studies have highlighted the potential of ticks' ribosomal proteins as promising antigens if their sequences are conserved among many tick species and are significantly different from the host ribosomal proteins [52, 53]. The authors of one study used *R. microplus* ribosomal protein as a single antigen in bovines and observed reduced tick engorgement, egg-laying and hatching [52]. However, previous studies on *Ha. longicornis* L24 ribosomal protein showed that it is expressed chiefly in ovaries and salivary glands but provides no protection efficiency against tick infestation [54].

Defense/immunomodulation

Proteins involved in protective and defense functions play key roles in tick physiology since ticks are exposed to various host blood or skin-borne pathogens, which in turn activates the tick immune system to protect against harmful infections. In this class, we grouped sequences expected to modulate the host immune system or protect ticks against host response.

Among the cellular proteins, we identified tick salivary membrane RING finger protein 150-like (clones P9_D07 and P3_H03). Proteins containing RING domains are characterized by the presence of the zinc binding motif containing seven nonconsecutive cysteines and one histidine. Many of these proteins function as ubiquitin ligase, thus targeting proteins to the proteasome pathway. They contribute also to cell cycle regulation, DNA repair, chromatin remodeling and signal transduction [55, 56]. The sequences of both clones P9_D07 and P3_H03 matched sequences in the annotated proteomes of *R. microplus*, *H. asiaticum* and *Ha. longicornis* (Additional file 4: Table S3). Interestingly, the homologues in *R. microplus* are expected to inhibit the host adaptive immune responses [57].

A third clone, P13_D03, is predicted to encode a polymeric immunoglobulin receptor (PIgR). This receptor functions in anchoring and facilitating the transport of immunoglobulin (Ig)A and IgM across mucosal membrane [58]. PIgRs have not been specifically studied in ticks (to the best of our knowledge) but they may play a role in the neutralization of host antibodies or provision of tick protection against potential pathogens [58].

Development

This class includes proteins implicated in ovulation, egg development, cell-cycle progression and cytoskeleton/

Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 10 of 17

exoskeleton formation. An interesting example is clone P4_A06, predicted to encode a secreted protein involved in egg yolk development, similar to *R. sanguineus* s.l., vitellogenin-6 protein and UniProtKB/Swiss-Prot entries Q868N5 and Q9U8M0. In ticks, vitellogenin is secreted into the haemolymph and then taken up by the developing oocytes. The main trigger for vitellogenin synthesis in female fat bodies and midgut is engorgement and mating [59]. Vitellins and related proteins have been proven to have a great potential as protective antigens [28]. Interestingly, vitellogenin receptors, which are surface receptors capturing vitellogenin from the haemolymph and promoting their transport and conversion to the active form, were found to present potential targets for tick control [60].

Protein modification
Our analysis of *H. dromedarii* SMaT sequences with Swiss-Prot matches revealed 31 sequences involved in protein modification. This class included peptidases, proteases, chaperon-related proteins and some classes of ubiquitin-ligases. These are involved in a variety of physiological processes, including feeding, cellular homeostasis and regulation of cell division [61].

Serine proteases (SP) were among the proteins annotated from the *H. dromedarii* cDNA library. SP are involved in several processes, such as coagulation of blood, inflammation, immune reactions, processing of zymogen, apoptosis, and modeling of matrix [62–64]. For example, clone P11_D05 matched neprilysin-like protein (according to the VectorBase annotation, *R. microplus* proteome, see below), a membrane metalloendopeptidase found to be abundant in the *H. dromedarii* sialotranscriptome [23]. The salivary metalloproteases of ticks are known to play a central role during tick feeding and are involved in several processes, such as coagulation of blood, inflammation, immune reactions, processing of zymogen, apoptosis and remodeling of matrix [62–64]. Interestingly, Neprilysin-like peptidase has been identified as an important antigen eliciting host immune response in *I. ricinus* [65].

Additionally, clone P3_C11 is predicted to encode a thiol protease with dipeptidyl peptidase activity. In addition to its potential role in feeding, it might also play a role in activating other proteases, such as elastases and cathepsins, all of which are involved in feeding, embryogenesis and tissue remodelling [66].

P5_F06 is predicted to encode an ubiquitin conjugating enzyme, which is well known for its role in protein degradation through the proteasome complex. However, accumulating evidence suggests its involvement is immunomodulation both intra-and extra-cellularly [67]. It has also been suggested that ubiquitin-conjugating enzymes are involved in the immunomodulatory effect of *Amblyomma variegatum* saliva on bovine cells [68].

Comparison of *H. dromedarii* SMaT sequences with proteomes of other hard ticks and *H. dromedarii* salivary gland proteins
From the results described above, we noted that some *H. dromedarii* SMaT sequences had homologues in the annotated proteomes of other tick species (in particular *R. microplus* and *I. scapularis*). As the aim of this study was to predict potential antigens, we reasoned that transcripts common between different tick species might encode a potential or multi-species protective antigen. Accordingly, to identify *H. dromedarii* SMaT sequences that are common and/or similar to those of other hard tick species, we compared the identified *H. dromedarii* SMaT sequences with the annotated proteomes of the four different species showing the most common hits in the NCBI blastx results (*H. asiaticum*, *R. sanguineus* s.l., *D. silvarum* and *R. microplus*) using the corresponding annotated proteome from VectorBase [36, 37] with an e-value threshold of 10^{-3} (Additional file 4: Table S3). Simultaneously, we compared *H. dromedarii* SMaT sequences to the annotated transcripts of the *H. dromedarii* sialotranscriptome (using tblastn) due to the crucial role of the tick feeding process and the salivary proteins [23]. This comparison revealed 82 common sequences (Fig. 7a; see Additional file 5: List 1 for their IDs). We then compared the 82 common sequences with the annotated proteomes of two other hard tick species, *Ha. longicornis* and *I. scapularis*. All together, these tick species in addition to the *H. dromedarii* sialotranscriptome shared 60 common sequences (the IDs are listed in Additional file 6: List 2). These 60 common sequences were compared to the corresponding homologous sequences from *D. silvarum*, *H. asiaticum*, *Ha. longicornis*, *I. scapularis R. microplus*, *R. sanguineus* s. l. and *H. dromedarii* sialotranscriptome. For all tick species included in this analysis, the number of *H. dromedarii* SMaT sequences showing > 70% similarity to their homologues was larger than those showing < 70% similarity, with the exception of *I. scapularis* and *Ha. longicornis* (Fig. 7b; Additional file 4: Table S3). The annotation of these 60 common sequences was confirmed by alignment with sequences in the UniProtKB/Swiss-Prot database using more stringent blastx parameters (e-value = 10^{-3}) (Additional file 4: Table S3) and were classified into the functional categories shown in Fig. 7c. Among these, 21 sequences (3%) had no matches to those in UniProtKB/Swiss-Prot and were excluded from subsequent analysis, leaving 39 *H. dromedarii* SMaT sequences for further analysis.
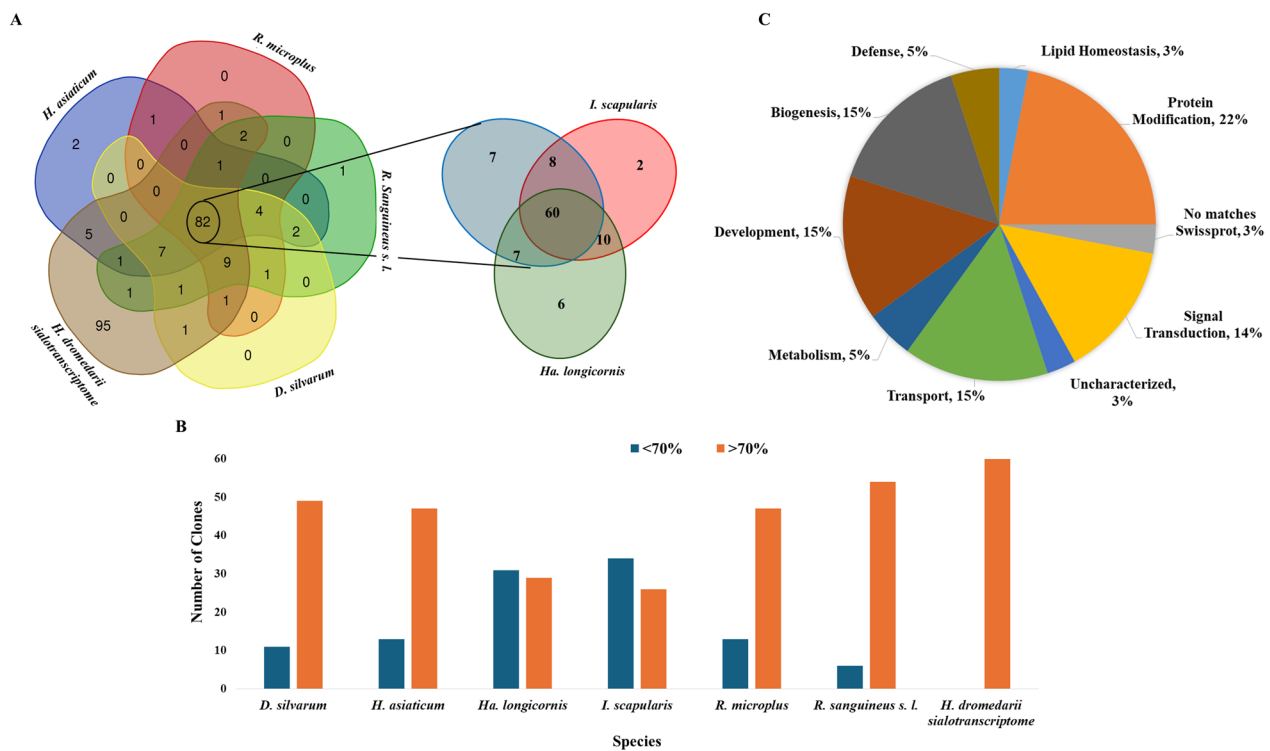
Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 11 of 17

Fig. 7 Identification of *Hyalomma dromedarii* SMaT sequences shared and /or similar to those of other tick species. **A** Number of SMaT transcriptome sequences shared between *Hyalomma asiaticum*, *Rhipicephalus sanguineus* s.l., *Dermacentor silvarum*, *Rhipicephalus microplus*, *Hyalomma dromedarii* sialotranscriptome, *Haemaphysalis longicornis* and *Ixodes scapularis*. **B** Number of clones from the 60 common sequences similar to sequences from *D. silvarum, H. asiaticum, Ha. longicornis, I. scapularis R. microplus, R. sanguineus* s.l. and *H. dromedarii* sialotranscriptome with identity lower or higher than 70%. **C** Functional classes of the 60 sequences shared between *H. dromedarii* SMaT and the tick species mentioned above. SMaT, Secreted, membrane-associated or transmembrane

Physicochemical properties, secondary structure and epitope prediction of *H. dromedarii* SMaT potential antigens

To predict the specific amino acid stretches that may have antigenic properties, we extracted the sequence stretches of the 39 SMaT sequences aligning with their UniProtKB/Swiss-Prot homologues (Additional file 7: Fasta 2) and (1) investigated their physicochemical/secondary structure properties and (2) predicted their antigenic potential.

The MPP 1.0 tool was used to predict the physicochemical properties (Additional file 8: Table S4), including the grand average of hydropathicity (GRAVY) index, aliphatic index, instability index, stability, molecular mass, aromaticity, theoretical isoelectric point (pI) and charge at pH 7 [38]. In addition, the percentage of amino acids forming alpha (α) and beta helices (β), respectively, and the percentage of amino acids constituting transmembrane domains (if present) were calculated using the fold recognition server Phyre2 [39]. Based on the analysis, 17 sequences were considered stable since their dipeptide instability weight values (DIWV) were predicted to be <

40; 22 sequences were predicted to be unstable (Fig. 8a). Nevertheless, the presence of certain dipeptides in a protein structure is one element which contributes to the stability or instability of that protein; other sequence-dependent factors, such as disulphide bonds, type of ligands and their binding and mechanisms of protease recognition, would determine overall protein stability [69].

Regarding the relative volume of a protein occupied by its aliphatic side chains (aliphatic index), the analysis showed that 59% (23 sequences) had high values of > 80 (Fig. 8a). This index is important as a means to control the thermal stability of a protein, where the higher the aliphatic index, the more stable the protein [70]. Contrary to Ikai [70], Panda and Chandra [71] considered cytotoxins and short neurotoxins from *Naja annulifera* and *Naja naja*, with aliphatic indices ranging between 66.5 and 84.33, to be thermostable. In our study, out of 39 SMaT sequences, 10 clones showed similarities to membrane homologues and all of these showed high aliphatic indices, with the exception of P1_C12 clone which showed an aliphatic index of 45.82 (Additional file 8:

Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 12 of 17

A

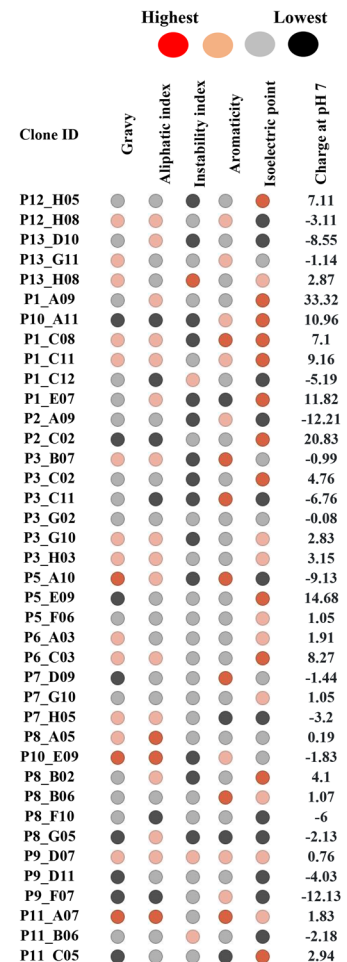| Clone ID	α helix%	β strand%	TM%	Confidence%	Coverage%
P12_H05	40	3	0	100	64
P12_H08	23	21	14	99.9	81
P13_D10	41	18	0	100	90
P13_G11	0	57	0	100	98
P13_H08	0	45	0	99.3	100
P1_A09	44	11	0	100	100
P10_A11	65	0	0	100	96
P1_C08	67	0	38	100	70
P1_C11	28	30	0	99.9	96
P1_C12	0	10	16	99.7	91
P1_E07	48	13	0	100	100
P2_A09	60	0	0	99.4	100
P2_C02	46	16	0	100	100
P3_B07	40	14	40	93.3	83
P3_C02	0	55	0	99.7	94
P3_C11	12	32	0	100	95
P3_G02	20	25	0	100	100
P3_G10	34	18	0	97.6	66
P3_H03	34	13	13	99.5	39
P5_A10	61	0	53	99.9	72
P5_E09	7	42	0	99.7	50
P5_F06	31	18	0	100	90
P6_A03	33	13	0	99.9	93
P6_C03	59	2	35	99.9	99
P7_D09	74	0	0	100	100
P7_G10	31	18	0	100	90
P7_H05	21	44	0	99.9	99
P8_A05	71	0	0	99.8	79
P10_E09	67	0	29	98	40
P8_B02	0	39	0	100	94
P8_B06	25	26	0	100	100
P8_F10	0	32	0	99.5	46
P8_G05	86	0	0	4	88
P9_D07	36	8	17	99.7	51
P9_D11	44	16	0	99.8	55
P9_F07	52	0	0	100	79
P11_A07	80	0	59	7.8	35
P11_B06	0	53	0	99.9	98
P11_C05	34	4	0	99.5	100

B

Highest ● ● ● ● Lowest (color-coded dots for Gravy, Aliphatic index, Instability index, Aromaticity, Isoelectric point)

Clone ID	Charge at pH 7
P12_H05	7.11
P12_H08	-3.11
P13_D10	-8.55
P13_G11	-1.14
P13_H08	2.87
P1_A09	33.32
P10_A11	10.96
P1_C08	7.1
P1_C11	9.16
P1_C12	-5.19
P1_E07	11.82
P2_A09	-12.21
P2_C02	20.83
P3_B07	-0.99
P3_C02	4.76
P3_C11	-6.76
P3_G02	-0.08
P3_G10	2.83
P3_H03	3.15
P5_A10	-9.13
P5_E09	14.68
P5_F06	1.05
P6_A03	1.91
P6_C03	8.27
P7_D09	-1.44
P7_G10	1.05
P7_H05	-3.2
P8_A05	0.19
P10_E09	-1.83
P8_B02	4.1
P8_B06	1.07
P8_F10	-6
P8_G05	-2.13
P9_D07	0.76
P9_D11	-4.03
P9_F07	-12.13
P11_A07	1.83
P11_B06	-2.18
P11_C05	2.94

Fig. 8 Multiple physicochemical properties of all 39 SMaT sequences. **A** Properties were calculated using the fold recognition server Phyre2. **B** GRAVY index, aliphatic index, instability index, stability, molecular mass, aromaticity, theoretical isoelectric point (pI) and charge at pH 7 were calculated using the MPP 1.0 tool. Gravy, Grand average of hydropathicity (index); MPP, Multiple Protein Profiler; SMaT, secreted, membrane-associated or transmembrane

Table S4). Based on Phyre2 prediction tool (Fig. 8b), this transcript is similar to the low-density lipoprotein receptor (Extracellular domain, Fold library ID c1n7dA) with a confidence of 99.7% and coverage of 91%. Although the sequence of the clone contains 13% serine plus threonine residues and a low content of aliphatic amino acids (24.4%), this predicted receptor binds to metal ions (calcium). It is possible that the presence of large numbers of residues forming hydrogen bonds may contribute to the thermal stability of such proteins regardless of the aliphatic index [39].

The GRAVY values of the 39 SMaT sequences are shown in Additional file 8: Table S4. Of these 39 SMaT sequences, 64.1% (25/39) had negative values, indicating that these proteins are hydrophilic in nature. The most

potent hydrophilic peptides were from clones P2_C02 and P9_F07, with GRAVY values of − 1.02 and − 1.05, respectively. Both clones were predicted based on our annotation to play important roles in protein biogenesis or modifications.

Next, using the same sequence stretches aligning to sequences in the Uniprot/Swissprot database, we predicted the antigenicity of 39 *H. dromedarii* SMaT sequences using three prediction tools: Vaxigen (threshold 0.5), Immunomedicine group tool and BepiPred tools [40–43]. The use of three different algorithms for antigenic prediction combines the advantage of relying on the physicochemical properties of the sequences (Vaxigen), on machine-learning algorithms for prediction of B cell epitopes (BepiPred) and on algorithm-based

Hussein *et al. Parasites & Vectors* (2024) 17:462

Page 13 of 17

experimentally identified epitopes (immunomedicine group) [40–43]. Based on the prediction results, we created a mean antigenicity score computed from the mean of the three scores (Fig. 9; Additional file 9: Table S5) to select the transcripts with the highest score.

The five sequences displaying highest antigenicity score were (in descending order): P1_C12 (average score 0.805), P9_D11 (average score 0.786), P3_G10 (average

score 0.771), P8_A05 (average score 0.77) and P8_B02 (average score 0.767).

P1_C12 is predicted to encode a protein with MAM domain from the low-density lipoprotein (LDL) class. Other than their role in lipid and cholesterol homeostasis, vitellogenin receptors also belong to the LDL-receptor class superfamily. Functional vitellogenin receptors are indispensable for oocyte development and
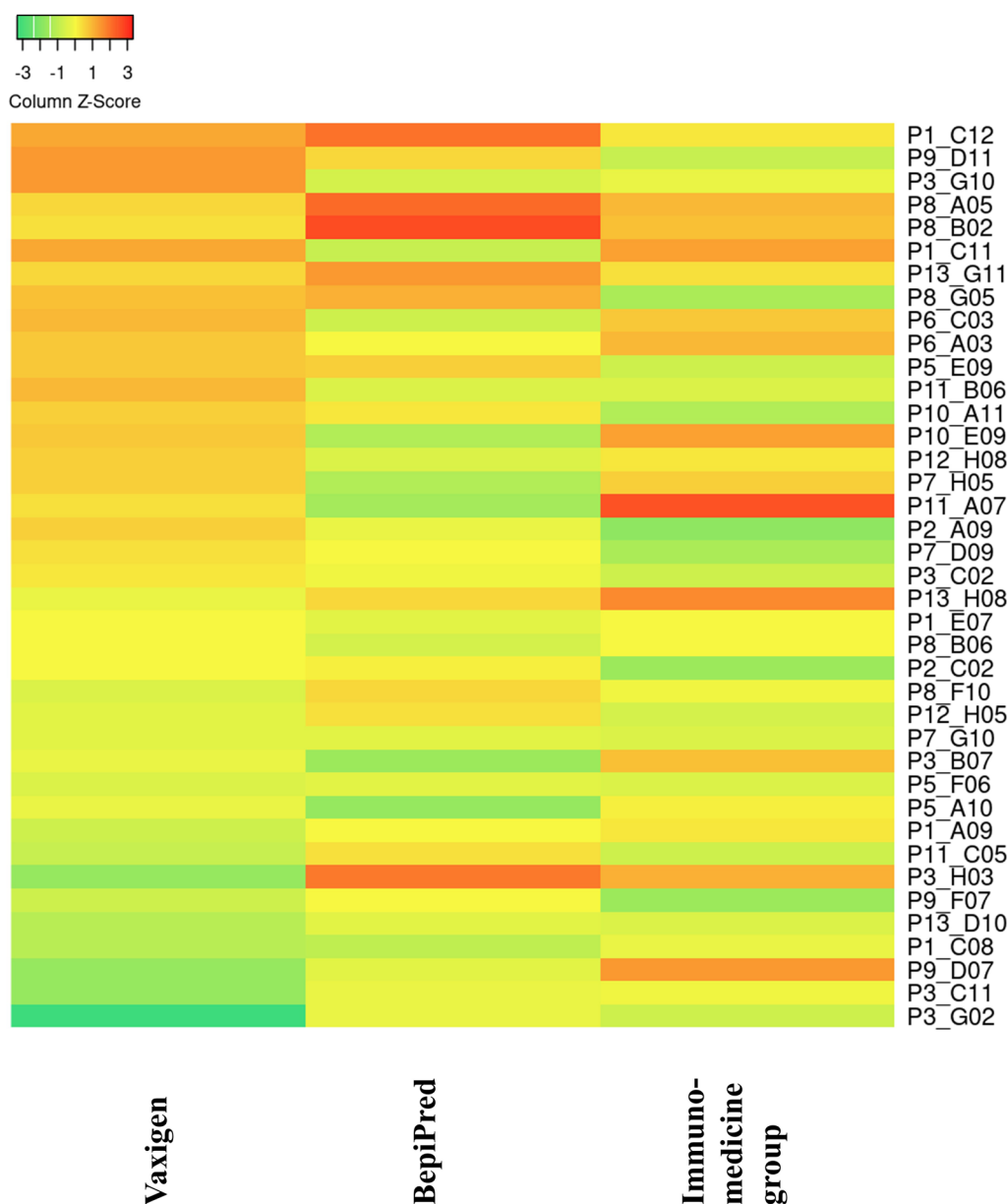

Fig. 9 Heat map of antigenic prediction scores of 39 *Hyalomma dromedarii* SMaT sequences using three different prediction tools: Vaxigen (threshold 0.5), Immunomedicine group tool and BepiPred tools [40–43]. Red colouration indicates the highest scores, yellow indicates intermediate scores and green indicates the lowest scores. Prediction was done for the chosen SMaT sequences (common with the tick species mentioned above) using the stretches aligning to those in the Uniprot/Swissprot database. SMaT, Secreted, membrane-associated or transmembrane

Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 14 of 17

oviposition. Silencing of their encoding gene(s) by RNA interference (RNAi) affects egg-laying capacity [60, 72].

The P9-D11 clone matched a sequence containing the ubiquitin-like (UBX) domain similar to FAS-associated factor 1 (hFAF1)/ UBX domain-containing protein 12)/ (UBX domain-containing protein 3A). The UBX domain-containing proteins were found to be phosphorylated salivary proteins participating in protein processing in the endoplasmic reticulum in *Ha. longicornis* [73]. Another model of UBX domain-containing proteins was identified in *I. scapularis* [74]. The results showed that the UBX domain of the p47 protein molecule was capable of binding the tick E3 ubiquitin ligase X-linked inhibitor of apoptosis (XIAP) protein that polyubiquitylates the lysine (K63) of the p47. However, upon infection of *I. scapularis* with *Anaplasma phagocytophilum* and *Borrelia burgdorferi*, the p47 is ubiquitinated via XIAP, which in turn will activate the tick immune deficiency (IMD) pathway to help in tick survival.

The third clone is the P3_G10, which was predicted to code a Pogo transposable element with the Krüppel associated box (KRAB domain). The definite function of the protein is not known, but the KRAB domain is found in Zn-finger (ZnF)-based transcription factors. The domain is considered to be a repressor domain for the transcription process. Accordingly, the KRAB-ZnF proteins are involved in the regulation of several biological processes, such as cell differentiation and proliferation, embryonic development and apoptosis [75, 76].

Clone P8_A05 matched HEAT repeat-containing protein-1 (protein BAP28/U3 small nucleolar RNA-associated protein 10 homolog). Orthologues from different ticks (*D. silvarum*, *I. scapularis*, *R. sanguineus* s.l. and *R. microplus*) were annotated and grouped at the Acari level (group number: 15322at6933/https://www.orthodb.org/?query=15322at6933). HEAT-1 protein is poorly characterized to date. Nevertheless, this protein is an important positive player in the biogenesis of ribosomes, and its deficiency will lead to the disruption of ribosome synthesis and consequently p53-independent damage of the DNA [77].

VectorBase resources (release 68, accessed July 2024) showed that *I. scapularis* 5′-AMP-activated protein kinase subunit beta-1 (AMPK subunit beta-1) corresponds to clone P8_B02, is localized in the cytoplasm and plays roles in lipid metabolism, kinase activity and regulation of catalytic processes. However, in ticks, the protein is poorly characterized. In general, the most suggested function for AMPK subunit beta-1 is in male gonad development, production and motility of spermatozoa and the regulation of necessary redox and energy processes needed for fertilization, among other processes [78].

Conclusions

Construction of the *H. dromedarii* cDNA library revealed the complexity of the *H. dromedarii* transcriptome where the rRNA was dominant during the process of library generation and the mRNA did not contain a regular cap structure as judged by cap-antibody selection. Since eukaryotic cells contain tenfold more rRNA molecules than of other RNA types, these limitations could be overcome in future studies and before direct extraction of mRNA, by depletion of the most abundant rRNA using biotinylated probe oligonucleotides designed to bind ribosomal RNA to enrich the less abundant transcripts. However, mixing RNAs from different life stages of *H. dromedarii* within a library has provided an enhanced deep-sequence ready genetic library to identify and explore novel and unique genes. The sequence data developed herein will pave the way to study gene expression during tick development, create comparative maps with other species and study genetic index of the species. This work provides a start database for *H. dromedarii* SMaT transcripts, including 60 common transcripts between hard tick species and those encoded by *H. dromedarii* salivary glands. The study shows that the secreted transcripts are more immunogenic than membrane-associated or transmembrane proteins, which may steer researchers to novel secreted-in-origin targets. Additional insights on the physicochemical properties of *H. dromedarii* transcripts with UniProtKB/ Swiss-Prot matches revealed that their structural composition affects mainly the DIWV, GRAVY and the aliphatic indices. These factors play vital roles in the stability and antigenicity of the transcripts. Nevertheless, the modelling calculations of the transcripts showed 93.3–100% confidence for 94.9% of the studied transcripts (37 transcripts out of the 39 transcripts). More in-depth analyses and experimental studies of these transcripts may help to identify and develop novel targets for vaccine development to minimize economic losses resulting from tick infestation and overcome the different drawbacks of using acaricides and other control methods.

Abbreviations
Ma   Membrane associated (sequences predicted to have a signal peptide and a transmembrane helix)
S   Secreted (sequences predicted to have an N-terminal signal peptide and predicted to have no transmembrane helix.
T   Transmembrane (sequences predicted to have one or more transmembrane domains and no signal peptides)

Supplementary Information

Additional file 1. Table S1: H. dromedarii SMaT sequences with corresponding matches from NCBI non redundant protein database.

Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 15 of 17

Additional file 2. Fasta 1: Nucleotide sequences of the 319 H. dromedarii SMaT sequences.

Additional file 3: Table S2. H. dromedarii SMaT sequences showing matches to UniProtKB/Swiss-Prot database.

Additional file 4: Table S3. Comparison of the 319 H. dromedarii SMaT sequences to other hard tick species annotated proteomes from the Vector base database.

Additional file 5: List 1. IDs of 82 sequences common between H. dromedarii SMaT, four hard tick species and H. dromedarii sialotranscriptome.

Additional file 6: List 2. IDs of 60 sequences common between H. dromedarii SMaT, six hard tick species and H. dromedarii sialotranscriptome.

Additional file 7: Fasta 2. Amino acid sequences stretch from the 39 H. dromedarii SMaT sequences aligning to the corresponding UniProtKB/Swiss-Prot matches.

Additional file 8: Table S4. Physicochemical properties and secondary structures predictions of amino acid sequences stretch described in Additional file 7.

Additional file 9: Table S5. Antigenicity score of the 39 H. dromedarii sequences stretch using three different prediction tools and their average.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the repository of the National Center for Biotechnology Information, under the project PRJEB67593, https://www.ebi.ac.uk/ena/browser/view/PRJEB 67593 and GenBank accession numbers PP584923 to PP585243.

Declarations

Ethics approval and consent to participate

All steps were approved by the Ethical Review Board of the National Research Centre, Egypt (NRC#11127082023), in accordance with the Helsinki Declaration (1975).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Molecular Biology Department, Biotechnology Research Institute, National Research Centre, Cairo, Egypt.

References

1. Abdelbaset AE, Nonaka N, Nakao R. Tick-borne diseases in Egypt: a one health perspective. One Health. 2022;15:100443.

2. Senbill H, Tanaka T, Karawia D, Rahman S, Zeb J, Sparagano O, et al. Morphological identification and molecular characterization of economically important ticks (Acari: Ixodidae) from North and North-Western Egypt. Acta Trop. 2022;231:106438.

3. Abushahba MFN, Abdelbaset AE, Rawy MS, Ahmed SO. Cross-sectional study for determining the prevalence of Q fever in small ruminants and humans at El Minya Governorate. Egypt BMC Res Notes. 2017;10:538.

4. Abdullah H, El-Molla A, Salib F, Ghazy A, Allam N, Sanad YM, et al. Molecular characterization of rickettsiae infecting camels and their ticks vectors in Egypt. Anim Husb Dairy Vet Sci. 2018;2:3–10.

5. Selim A, Abdelrahman A, Thiéry R, Sidi-Boumedine K. Molecular typing of *Coxiella burnetii* from sheep in Egypt. Comp Immunol Microbiol Infect Dis. 2019;67:101353.

6. Abbass H, Selim SAK, Sobhy MM, El-Mokhtar MA, Elhariri M, Abd-Elhafeez HH. High prevalence of *Coxiella burnetii* infection in humans and livestock in Assiut, Egypt: a serological and molecular survey. Vet World. 2020;13:2578–86.

7. Damaty HME, Yousef SG, Mahmmod YS, El-Balkemy FA, Mweu MM. Sensitivity and specificity of piroplasm indirect fluorescent antibody test and PCR for Theileria annulata infection in clinically asymptomatic large ruminants using Bayesian latent class analysis. Vet Parasitol Reg Stud Rep. 2021;24:100563.

8. Mohamed W, Mohamed A, Ali AO, Mahmoud HYAH, Omar MA, Chatanga E, et al. Exploring prokaryotic and eukaryotic microbiomes helps in detecting tick-borne infectious agents in the blood of camels. Pathogens. 2021;10:351.

9. Bendary HA, Rasslan F, Wainwright M, Alfarraj S, Zaki AM, Abdulall AK. Crimean-Congo hemorrhagic fever virus in ticks collected from imported camels in Egypt. Saudi J Biol Sci. 2022;29:2597–603.

10. Obaid MK, Islam N, Alouffi A, Khan AZ, Vaz IDS, Tanaka T, et al. Acaricides resistance in ticks: selection, diagnosis, mechanisms, and mitigation. Front Cell Infect Microbiol. 2022;12:941831.

11. de Castro JJ. Sustainable tick and tickborne disease control in livestock improvement in developing countries. Vet Parasitol. 1997;71:77–97.

12. Shahein YE. Molecular cloning and expression of a larval immunogenic protein from the cattle tick *Boophilus annulatus*. Vet Immunol Immunopathol. 2008;121:281–9.

13. Shahein YE, El-Rahim MTA, Hussein NA, Hamed RR, El-Hakim AE, Barakat MM. Molecular cloning of a small heat shock protein (sHSPII) from the cattle tick *Rhipicephalus* (*Boophilus*) *annulatus* salivary gland. Int J Biol Macromol. 2010;47:614–22.

14. Shahein YE, Abouelella AM, Hussein NA, Hamed RR, El-Hakim AE, Abdel-Shafy S, et al. Identification of four novel *Rhipicephalus annulatus* upregulated salivary gland proteins as candidate vaccines. Protein J. 2013;32:392–8.

15. El-Hakim AE, Shahein YE, Abdel-Shafy S, Abouelella AMK, Hamed RR. Evaluation of glycoproteins purified from adult and larval camel ticks (*Hyalomma dromedarii*) as a candidate vaccine. J Vet Sci. 2011;12:243–9.

16. Hussein NA, Shahein YE, El-Hakim AE, Abouelella AM, Guneidy RA, Hamed RR. Molecular cloning of Ra-sHSPI, a novel member of the HSP20 family from *Rhipicephalus annulatus* salivary glands. Int J Biol Macromol. 2014;67:7–15.

17. Abbas MN, Jmel MA, Mekki I, Dijkgraaf I, Kotsyfakis M. Recent advances in tick antigen discovery and anti-tick vaccine development. Int J Mol Sci. 2023;24:4969.

18. Rodríguez-Mallon A. The Bm86 discovery: a revolution in the development of anti-tick vaccines. Pathogens. 2023;12:231.

19. Valle MR, Guerrero FD. Anti-tick vaccines in the omics era. Front Biosci. 2018;10:122–36.

20. Charrier NP, Couton M, Voordouw MJ, Rais O, Durand-Hermouet A, Hervet C, et al. Whole body transcriptomes and new insights into the biology of the tick *Ixodes ricinus*. Parasit Vectors. 2018;11:364.

21. Reck J, Webster A, Dall'Agnol B, Pienaar R, de Castro MH, Featherston J, et al. Transcriptomic analysis of salivary glands of *Ornithodoros brasiliensis* aragão, 1923, the agent of a neotropical tick-toxicosis syndrome in humans. Front Physiol. 2021;12:725635.

22. Ribeiro JMC, Slovák M, Francischetti IMB. An insight into the sialome of *Hyalomma excavatum*. Ticks Tick Borne Dis. 2017;8:201–7.

23. Bensaoud C, Nishiyama MY, Hamda CB, Lichtenstein F, Oliveira UCD, Faria F, et al. De novo assembly and annotation of *Hyalomma dromedarii* tick

Hussein *et al. Parasites & Vectors*   (2024) 17:462

Page 16 of 17

(Acari: Ixodidae) sialotranscriptome with regard to gender differences in gene expression. Parasit Vectors. 2018;11:314.

24. Almazán C, Blas-Machado U, Kocan KM, Yoshioka JH, Blouin EF, Mangold AJ, et al. Characterization of three Ixodes scapularis cDNAs protective against tick infestations. Vaccine. 2005;23:4403–16.

25. Tsuda A, Mulenga A, Sugimoto C, Nakajima M, Ohashi K, Onuma M. cDNA cloning, characterization and vaccine effect analysis of *Haemaphysalis longicornis* tick saliva proteins. Vaccine. 2001;19:4287–96.

26. Merino O, Alberdi P, José M, de la Lastra P, de la Fuente J. Tick vaccines and the control of tick-borne pathogens. Front Cell Infect Microbiol. 2013;3:30.

27. Zorman S, Botte M, Jiang Q, Collinson I, Schaffitzel C. Advances and challenges of membrane–protein complex production. Curr Opin Struct Biol. 2015;32:123–30.

28. Oli AN, Obialor WO, Ifeanyichukwu MO, Odimegwu DC, Okoyeh JN, et al. Immunoinformatics and vaccine development: an overview. Immuno-Targets Ther. 2020;9:13–30.

29. Walker A, Bouattour A, Camicas JL, Estrada-Peña A, Horak I, Latif A, et al. Ticks of domestic animals in Africa: a guide to identification of species. Edinburgh: Bioscience Reports; 2003.

30. Abdel-Shafy S. Scanning electron microscopy and comparative morphology of *Hyalomma anatolicum excavatum*, *H. dromedarii* and *H. marginatum marginatum* (Acari: Ixodidae) based on nymphs. Acarologia. 2008;40:3–18.

31. Sambrook J, Russell DW. Molecular cloning: a laboratory manual. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001.

32. Götz S, Mez JMG-G, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 2008;36:3420–35.

33. Hu Z, Bao J, Reecy JM, Hu Z-L. CateGOrizer: a web-based program to batch analyze gene on-tology classification categories. Online J Bioinforma. 2008;9:108–12.

34. UniProt Consortium. UniProt: the universal protein knowledgebase in 2023 the UniProt Consortium. Nucleic Acids Res. 2022;51:523–31.

35. Coudert E, Gehant S, De Castro E, Pozzato M, Baratin D, Neto T, et al. Annotation of biologically relevant ligands in UniProtKB using ChEBI. Bioinformatics. 2023;39:793.

36. Amos B, Aurrecoechea C, Barba M, Barreto A, Basenko EY, Bażant W, et al. VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center. Nucleic Acids Res. 2021;50:D898-911.

37. Giraldo-Calderón GI, Harb OS, Kelly SA, Rund SS, Roos DS, McDowell MA. VectorBase.org updates: bioinformatic resources for invertebrate vectors of human pathogens and related organisms. Curr Opin Insect Sci. 2022;50:100860.

38. Sganzerla Martinez G, Dutt M, Kumar A, Kelvin D. Multiple protein profiler 1.0 (MPP): a webserver for predicting and visualizing physiochemical properties of proteins at the proteome level. Protein J. 2024;43:1–7.

39. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10:845–58.

40. Doytchinova I, Flower D. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8:4.

41. Doytchinova I, Flower D. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. Vaccine. 2007;25:856–66.

42. Jespersen M, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 2017;45:W1.

43. Kolaskar A, Tongaonkar P. A semiempirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990;276:172–4.

44. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, et al. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 2016;44:W147–53.

45. Hert DG, Fredlake CP, Barron AE. Advantages and limitations of next-generation sequencing technologies: a comparison of electrophoresis and non-electrophoresis methods. Electrophoresis. 2008;29:4618–26.

46. Poulsen LD, Kielpinski LJ, Salama SR, Krogh A, Vinther J. SHAPE Selection (SHAPES) enrich for RNA structure signal in SHAPE sequencing-based probing data. RNA. 2015;21:1042–52.

47. Gutell RR, Lee JC, Cannone JJ. The accuracy of ribosomal RNA comparative structure models. Curr Opin Struct Biol. 2002;12:301–10.

48. Rahman K, Kim B, You M. Molecular cloning, expression and impact of ribosomal protein S-27 silencing in *Haemaphysalis longicornis* (Acari: Ixodidae). Exp Parasitol. 2019;209:107829.

49. Zhou X, Liao W, Liao J-M, Liao P, Lu H. Ribosomal proteins: functions beyond the ribosome. J Mol Cell Biol. 2015;72:92–104.

50. Xu X, Xiong X, Sun Y. SCIENCE CHINA The role of ribosomal proteins in the regulation of cell proliferation, tumorigenesis, and genomic integrity. Sci China Life Sci. 2016;59:656–72.

51. Sehgal A, Kumar N, Carruthers VB, Sharma S. Translocation of ribosomal protein P0 onto the *Toxoplasma gondii* tachyzoite surface. Int J Parasitol. 2003;3314:1589–94.

52. Rodríguez-Mallon A, Encinosa PE, Méndez-Pérez L, Bello Y, Fernández RR, Garay H, et al. High efficacy of a 20 amino acid peptide of the acidic ribosomal protein P0 against the cattle tick *Rhipicephalus microplus*. Ticks Tick-Borne Dis. 2015;6:530–7.

53. Pérez-Sánchez R, Manzano-Roman R, Obolo Mvoulouga P, Oleaga A. In silico selection of functionally important proteins from the mialome of *Ornithodoros erraticus* ticks and assessment of their protective efficacy as vaccine targets. Parasit Vectors. 2019;12:508.

54. Tian Z, Zhang P, Luo J, Yin H, Luo J, Xie J, et al. Cloning and characterization of a ribosomal protein L24 from *Hemaphysalis longicornis* eggs. Parasitol Res. 2010;107:1213–20.

55. Chai X, Tao Q, Li L. The role of RING finger proteins in chromatin remodeling and biological functions. Epigenomics. 2023;15:1053–68.

56. Wang M-X, Liuyu T, Zhang Z-D. Multifaceted roles of the E3 ubiquitin ligase RING finger protein 115 in immunity and diseases. Front Immunol. 2022;13:936579.

57. Paulino P, Vitari G, Rezende A, Couto J, Antunes S, Domingos A, et al. Characterization of the *Rhipicephalus* (*Boophilus*) *microplus* sialotranscriptome profile in response to *Theileria equi* infection. Pathogens. 2021;10:167.

58. Wei H, Wang J-Y. Role of polymeric immunoglobulin receptor in IgA and IgM transcytosis. Int J Mol Sci. 2021;22:2284.

59. Chinzei Y, Taylor D. Hormonal regulation of vitellogenin biosynthesis in ticks. In: Harris KF, editor. Advances in disease vector research. New York: Springer New York; 1994. p. 1–22.

60. Ali A, Khan AW, Mardan U, Termignoni C, Mitchell RD, Sonenshine DE, et al. Vitellogenin receptor as a target for tick control: a mini-review. Front Physiol. 2019;10:618.

61. Upadhyay A. Natural compounds in the regulation of proteostatic pathways: an invincible artillery against stress, ageing, and diseases. Acta Pharm Sin B. 2021;11:2995–3014.

62. Liu Y, Patricelli MP, Cravatt BF. Activity-based protein profiling: the serine hydrolases. Proc Natl Acad Sci USA. 1999;96:14694–9.

63. Cesarman-Maus G, Hajjar KA. Br J Haematol. 2005;129:307–21.

64. Zlem O, Ekici AN, Paetzel M, Dalbey RE. Unconventional serine proteases: variations on the catalytic ser/his/asp triad configuration. Protein Sci. 2008;17:2023–37.

65. Perner J, Helm D, Hatalova T, Kropackova S, Ribeiro JM, Kopacek P, et al. The central role of salivary metalloproteases in host acquired resistance to tick feeding. Front Cell Infect Microbiol. 2020;10:1.

66. McGuire MJ, Lipsky PE, Thiele DL. Generation of active myeloid and lymphoid granule serine proteases requires processing by the granule thiol protease dipeptidyl peptidase I. J Biol Chem. 1993;268:2458–67.

67. Hu H, Sun S-C. Ubiquitin signaling in immune responses. Cell Res. 2016;26:457.

68. Rodrigues V, Fernandez B, Vercoutere A, Chamayou L, Andersen A, Vigy O, et al. Immunomodulatory effects of *amblyomma variegatum* saliva on bovine cells: characterization of cellular responses and identification of molecular determinants. Front Cell Infect Microbiol. 2018;7:521. https://doi.org/10.3389/fcimb.2017.00521.

69. Guruprasad K, Reddy BVB, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng. 1990;4:155–61.

70. Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem. 1980;88:1895–8.

71. Panda DRS, Chandra G. Physicochemical characterization and functional analysis of some snake venom toxin proteins and related non-toxin proteins of other chordates. Bioinformation. 2012;8:891–6.

72. Mitchell RD III, Ross E, Osgood C, Sonenshine DE, Donohue KV, Khalil SM, et al. Molecular characterization, tissue-specific expression and RNAi knockdown of the first vitellogenin receptor from a tick. Insect Biochem Mol Biol. 2007;37:375–88.

73. Agwunobi D, Li M, Wang N, Chang G, Zhang X, Xue X, et al. Proteomic analysis suggests that monoterpenes in lemongrass disrupt Ca2+ homeostasis in *Haemaphysalis longicornis* leading to mitochondrial depolarization and cytotoxicity. Proteomics. 2022;22:2100156.

74. Carroll E, Wang X, Shaw D, O'Neal A, Oliva Chavez A, Brown L, et al. P47 licenses activation of the immune deficiency pathway in the tick *Ixodes scapularis*. Proc Natl Acad Sci USA. 2018;116:205.

75. Urrutia R. KRAB-containing zinc-finger repressor proteins. Genome Biol. 2003;4:231.

76. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. KRAB-zinc finger proteins: a repressor family displaying multiple bio-logical functions. Curr Genomics. 2013;14:268–78.

77. Turi Z, Lacey M, Mistrik M, Moudry P. Impaired ribosome biogenesis: mechanisms and relevance to cancer and aging. Aging. 2019;11:2512–40.

78. Nguyen TMD. Impact of 5′-AMP-activated protein kinase on male gonad and spermatozoa functions. Front Cell Dev Biol. 2017;5:25.

Publisher's Note