

RESEARCH

Open Access



Genome of the invasive North American *Haemaphysalis longicornis* tick as a template for bovine anti-tick vaccine discovery

Mohamed Abdallah Mohamed Moustafa¹, Miranda M. Barnes¹, Nicole E. Wagner¹, Deanna Bodine², Kylie Bendele², Pete D. Teel³, Perot Saelao² and Dana C. Price^{1*}

Abstract

Background The ixodid tick *Haemaphysalis longicornis* Neumann, commonly referred to as the Asian longhorned tick, has expanded its range outside of East Asia into countries such as Australia, New Zealand, and the United States. Since the first U.S. detection in 2017, *H. longicornis* has spread to 21 states and the District of Columbia and has been implicated as a vector of various human and animal pathogens including *Theileria orientalis* Ikeda genotype, a causal agent of bovine theileriosis. Facilitated in part by the parthenogenetic nature of invasive populations, this tick has become a paramount threat to agricultural rangelands and U.S. livestock production. Reliance on traditional acaricides for vector control selects for resistant individuals, reducing the effectiveness of many chemical tools over time. Thus, focus has shifted to alternative control mechanisms including anti-tick vaccine development. To further such research, here we sequence and assemble a high-quality *H. longicornis* genome and robust gene catalog from invasive North American ticks while also providing an organ-specific transcriptomic expression catalog and in-depth informatic screening of the tick proteome for potential bovine antigenic molecules with potential utility as vaccine candidates.

Results Using a combination of PacBio HiFi single-molecule sequencing and Hi-C chromosome conformation capture data, our genome assembly contains 270 scaffolds and spans a haploid genome size of 3.09 Gbp with an N50 of 213.4 Mbp. Gene prediction identified 21,947 high-confidence gene structures containing 96.2% of the core Arthropoda odb10 orthologs. Our organ-specific transcriptome library comprising salivary glands, midgut, ovaries, foreleg and hindleg additionally highlights potential anti-tick vaccine candidates and metabolic pathways to target for future in vitro trials.

Conclusions Single-molecule sequencing of a triploid, parthenogenetic North American *Haemaphysalis longicornis* tick allowed for the generation of a highly contiguous genome assembly that, when coupled with extensive transcriptome profiling, resulted in a robust gene catalog containing multiple candidates for further study as anti-tick vaccine antigens.

Keywords Asian longhorned tick, *Haemaphysalis longicornis*, Vector, Pest, Genome assembly, Reverse vaccinology

*Correspondence:

Dana C. Price

d.price@rutgers.edu

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

The ixodid tick *Haemaphysalis longicornis* Neumann (Acari: Ixodidae), commonly known as the Asian long-horned tick, is native to temperate Asia, including China, Japan, the Republic of Korea (ROK), and southeast Russia [1, 2] with invasive populations in Australia, New Zealand, and on several Pacific Islands including New Caledonia, Fiji, Western Samoa, Vanuatu, and the Kingdom of Tonga [3]. While *H. longicornis* can be found in both bisexual and parthenogenetic populations in its native range, all known invasive lineages in Australia, New Zealand and the United States are exclusively parthenogenetic [1]. It plays an important role in the dynamics of vector-borne diseases and livestock management [4] and has garnered international attention due to the ability to transmit a variety of pathogens: for example, *Anaplasma phagocytophilum* [5], *Borrelia burgdorferi sensu stricto* [6] and the highly pathogenic Ikeda genotype of *Theileria orientalis* [7]. In North America, *H. longicornis* feeds on a diverse range of vertebrate hosts, including birds, sheep, raccoons, opossums, and white-tailed deer [8, 9], potentially enabling the spread of pathogens within and between reservoir species.

The rapid spread of *H. longicornis* across diverse habitats in the U.S. has raised concerns about its ability to adapt to new environments, potentially including pasture and rangelands where the health of economically important livestock may be at stake. In the United States, it has been transiently identified and isolated on imported livestock for nearly 50 years by the United States Department of Agriculture (USDA) quarantine protocol, but its established presence outside quarantine was only confirmed in 2017 [4, 8, 10]. Subsequent reports have identified populations in 21 states and the District of Columbia (predominantly in the eastern U.S.), with evidence suggesting its establishment as early as 2010 [4]. In August of 2024, *H. longicornis* were detected on cattle in Craig and Mayes Counties, Oklahoma [11], further accentuating the southerly and westerly expansion of this tick in the direction of the highest cow-calf growing region of the US beef cattle industry [12].

Global economic losses resulting from tick-borne diseases are estimated at US \$22–30 billion annually [13, 14], and although acaricides are currently the most effective tick control measure [15], reliance on them can lead to the development of resistant populations [16]. Novel control strategies are thus needed to create comprehensive management programs that do not rely strictly on chemical application. One such technique is anti-tick vaccination, an emerging discipline that involves host immunization with recombinant tick antigens that prime the host to mount an immune response and produce anti-tick antibodies in response to attachment and

feeding. Effective vaccination regimens can result in diminished tick feeding, tick mortality, and/or diminished reproductive capacity.

Only two tick vaccines (TickGARD(PLUS)[™] and Gavac[™]) based on recombinant *Rhipicephalus (Boophilus) microplus* Bm86 midgut antigens have been commercialized. However, both exhibited limited efficacy and TickGARD is no longer marketed, while the availability of Gavac is limited [17]. No such vaccine currently exists for *H. longicornis*; recent trials aimed at developing such a vaccine involve two approaches: one using Pmy recombinant protein (rPmy; an immunomodulatory protein that inhibits complement activation) and a peptide vaccine designed from the protective epitope (KLH-LEE) [18], and the other utilizing recombinant *H. longicornis* subolesin (HLSu; a candidate protective antigen involved in several critical pathways) expressed in *Escherichia coli* [19]. Both trials involved immunization, challenge with female ticks, and evaluation of tick fitness. Results from the Pmy and KLH-LEE trials indicate reductions in engorgement weight, oviposition, and larval hatch, with 60.37% and 70.86% efficacy, respectively. In the HLSu trial, vaccination significantly affected blood-feeding and reproduction, demonstrating 37.4% efficacy as calculated based on reduction in body weight and egg mass [18, 19]. Furthermore, studies on other potential antigens, including triosephosphate isomerase (TIM) in rabbits [20], lipocalins (LIP) in rabbits [21], serine proteinase inhibitors (serpins) in vitro [22], and hemalin in rabbits [23] showed promising results in regard to inducing immune responses and reducing tick infestations. H1ATAQ, a homologue of *R. microplus* ATAQ in *H. longicornis*, has also been explored as a candidate. This protein is expressed across tick life stages and organs; interestingly, when used as a vaccine in rabbits, ticks blood-fed for longer, attained a higher body weight, and demonstrated increased egg mass [24].

Selection of vaccine candidates has previously been somewhat haphazard and inefficient. In fact, the current rate-limiting step to anti-tick vaccine development is the a priori identification of protective antigens: many critical factors must be considered to generate effective recombinants for a metazoan organism that encodes >20,000 genes [25]. Reverse vaccinology, rooted in molecular biology and proteomics, has evolved to take advantage of genome sequences and new -omics and informatics techniques to identify and prioritize candidates that are most likely to be successful [26]. This approach to identifying potential vaccine targets often involves analyzing the genome and derived gene catalog of a target organism to uncover novel antigens and epitopes in silico, facilitating rapid antigen discovery. Pipelines have evolved increasingly discriminatory criteria to narrow the field,

incorporating information on subcellular localization, secretion, GPI anchorage, transmembrane helix orientation, and B- and T-cell epitope prediction [14].

Reverse vaccinology works best when there is a complete genome available for the target organism. There are four distinct Australasian *H. longicornis* genomes currently sequenced, with each presenting significantly different characteristics in terms of contiguity, duplicity, gene count, and completeness as assessed by BUSCO [27] single-copy arthropoda_odb10 orthologs ($n=1,013$): 1) the Japan *H. longicornis* Oita bisexual strain genome [28], consisting of 99,353 contigs summing to 2.48 Gbp with 78.5% of core single-copy orthologs identified (2.7% duplicated); 2) the bisexual HLO strain genome [29] sourced from China with 6,516 scaffolds (3.16 Gbp) and 90.2% of the core ortholog set identified (20.0% duplicated); 3) the New Zealand parthenogenetic strain genome [30] with 34,208 contigs (7.36 Gbp) and 95.1% of core orthologs identified (however, 84.2% were duplicated due to highly redundant and distinct haplotypes in the assembly) and 4) the complete genome of strain HaeL-2018 from China (GCA_013339765; [31]) with 3,885 scaffolds inclusive of 11 chromosomes (2.56 Gbp) and a BUSCO completeness score of 87.5% (2.4% duplicated); the latter represents a marked improvement over existing assemblies and is the only current *H. longicornis* genome with a predicted proteome, however this proteome contains only 53.6% of the referenced arthropod_odb10 orthologs and thus the utility for functional analyses, epitope identification and comparisons with other ticks are limited. Notably, none of these genomes were sequenced directly from the invasive populations in North America.

Furthermore, having a genome assembly in the absence of other contextual information about tick biology can only take one so far in vaccine development. Reverse vaccinology pipelines work best when the molecular underpinnings of parasite physiology are thoroughly understood and used to hone target selection. A multitude of processes have been suggested as possible candidates for reverse vaccines in ticks, e.g., blood digestion, embryogenesis/fertility, heme metabolism/detoxification, and water balance [32]. However, a robust catalog of full-length and non-redundant gene sequences from each of these pathways – for example, as generated through various transcriptomic analyses—is required to provide the greatest diversity of targets. This can be accomplished by performing RNAseq on material derived from different tick organs. Taking this localized approach is particularly important when considering where tick antigens are mostly likely to be a) recognized by the host immune system and b) accessible to host antibodies during a blood meal. For this reason, the tick salivary glands and midgut

constitute particularly appealing places to search for vaccine targets. Tick saliva is essential for feeding lesion development, blood-feeding, and transmission of tick-borne pathogens, encompassing a sophisticated blend of effectors injected into hosts over the course of the feeding cycle [33]. Assorted secreted proteins play a vital role in the tick's survival strategy by countering host hemostasis and immune functions [34–36]. The tick midgut is responsible for blood digestion and absorption of hemoglobin through specialized cells [37] and also serves as the primary portal for pathogen entry and proliferation [38]. If a host immune response primed by vaccination were to succeed in disrupting either these secreted salivary proteins or digestive processes in the midgut, tick feeding success could be greatly reduced and thus identifying the proteins produced in each tick organ represents an important first step in the reverse vaccinology process.

Thus, in order to provide resources for a future reverse vaccinology-facilitated approach to *H. longicornis* control, we have sequenced and assembled de novo a genome from an invasive North American population of parthenogenetic ticks using Pacific Biosciences long-read sequencing and Hi-C chromosome conformation capture protocols. In addition, we employed RNA sequencing (RNA-seq) to generate a robust transcriptome and predicted proteome containing 95.0% of core arthropoda_odb10 orthologs. Using these gene models, we examine gene expression within *H. longicornis* across a spectrum of organs including the ovaries, salivary glands, midgut, and legs. The predicted proteome was screened for potential antigens using informatic analyses and transcript localization. Through the analysis of organ-specific gene expression profiles, this research extends our understanding of *H. longicornis* biology beyond the genomic context while also establishing a starting point for the development of anti-tick vaccines to aid in the mitigation of this threat to North American livestock.

Materials and methods

Nuclear DNA extraction and PacBio library preparation

DNA was extracted from a single adult female *H. longicornis* individual derived from a colony initiated in 2019 from ticks collected in New Brunswick, New Jersey, USA, and maintained at the USDA-ARS Animal Disease Research Unit (Pullman, WA). PCR amplification and sequencing of the mitochondrial cytochrome c oxidase subunit 1 (COX1) locus (not shown) indicated these ticks were of haplotype H1 provenance [39]. The sample was homogenized in a 1.5 mL tube in a liquid nitrogen bath with a liquid nitrogen-cooled pestle. The Qiagen MagAttract HMW DNA Kit (Qiagen, Hilden, Germany) was used to obtain sufficiently high molecular weight DNA. Immediately following homogenization, 200 μ L

1X PBS, 20 μ L Proteinase K, 8 μ L RNase A, and 150 μ L Buffer AL were added to the sample. The tube was slowly inverted to mix. The sample was incubated for two hours at 25°C and 900 rpm. Lysate was transferred to a 2.0 mL tube using a wide-bore pipette tip following brief centrifugation. Qiagen's fresh or frozen tissue protocol was followed, starting with the addition of Buffer MB. All washes and rinses were incubated for 1 min. DNA was eluted with 105 μ L Buffer AE and incubated at room temperature for five minutes. Following isolation, genomic DNA was quantified using the Broad Range dsDNA Qubit assay (Thermo Fisher Scientific, Waltham, MA, USA). Size distribution was assessed with the Genomic DNA 165 kb kit run on the Femto Pulse System (Agilent Technologies, Santa Clara, CA, USA). A PacBio SMRTbell library was prepared using the SMRTbell prep kit 3.0 and the associated guide for preparing whole genome libraries (Pacific Biosciences, Menlo Park, CA, USA). The prepared library was bound and sequenced at the USDA-ARS Veterinary Pest Genetics Research Unit in Kerrville, Texas, on four Pacific Biosciences 8M SMRT Cells with a Sequel IIe system (Pacific Biosciences, Menlo Park, CA, USA) beginning with a 2-h preextension followed by a 30-h movie collection time. After sequencing, circular consensus sequences from the PacBio Sequel IIe subreads were obtained using the SMRTLink v11.0 software.

Hi-C Library preparation

A Hi-C (chromosome conformation capture) library was generated using the Phase Genomics Proximo Animal kit version 4.0. Approximately 500 mg of adult female ticks derived from the USDA-ARS colony above were finely chopped and then crosslinked for 15 min at room temperature with end-over-end mixing in 1 mL of Proximo crosslinking solution. The crosslinking reaction was terminated with a quenching solution for 20 min at room temperature again with end-over-end mixing. Quenched tissue was rinsed once with 1X Chromatin Rinse Buffer (CRB). Tissue was transferred to a liquid nitrogen-cooled mortar and ground to a fine powder. The powder was resuspended in 700 μ L Proximo Lysis Buffer 1 and incubated for 20 min with end-over-end mixing. A low-speed spin was used to clear the large debris, and the chromatin-containing supernatant was transferred to a new tube. Following a second higher speed spin, the supernatant was removed and the pellet containing the nuclear fraction of the lysate was washed with 1X CRB. After removing 1X CRB wash, the pellet was resuspended in 100 μ L Proximo Lysis Buffer 2 and incubated at 65°C for 15 min. Chromatin was bound to Recovery Beads for 10 min at room temperature, placed on a magnetic stand, and washed with 200 μ L of 1X CRB.

Chromatin bound on beads was resuspended in 150 μ L of Proximo fragmentation buffer, and 2.5 μ L of Proximo fragmentation enzyme was added and incubated for 1 h at 37°C. The reaction was cooled to 12°C and incubated with 2.5 μ L of finishing enzyme for 30 min. Following the addition of 6 μ L of Stop Solution, the beads were washed with 1X CRB and resuspended in 100 μ L of Proximo Ligation Buffer supplemented with 5 μ L of Proximity ligation enzyme. The proximity ligation reaction was incubated at room temperature for 4 h with end-over-end mixing. To this volume, 5 μ L of Reverse Crosslinks enzyme was added, and the reaction was incubated at 65°C for 1 h.

After reversing crosslinks, the free DNA was purified with Recovery Beads and Hi-C junctions were bound to streptavidin beads and washed to remove unbound DNA. Washed beads were used to prepare paired-end deep sequencing libraries using the Proximo Library preparation reagents.

Library quality control

Library fragment size and concentration were determined using TapeStation capillary electrophoresis and quantitative PCR. Prior to production sequencing, library performance was determined by low-pass sequencing, generating ~200,000 read pairs of sequencing data on an Illumina iSeq System. Resulting reads were mapped to a reference assembly and analyzed with Phase Genomics' open-source QC tool, `hic_qc` (GitHub—phasegenomics/hic_qc: A (very) simple script to QC Hi-C data). Libraries receiving a 'Sufficient' quality judgment from `hic_qc` were advanced to production sequencing and sequenced on four SMRT cells of a PacBio Sequel IIe system.

Genome assembly and gene prediction

PacBio HiFi sequence data were adapter- and quality-trimmed and assembled using HiCanu v2.2 [40] with an estimated genome size of 2.5 Gbp based on the estimate of Jia et al. Haplotigs [31] (or scaffolds comprising redundant haplotypes of our triploid tick) were removed with two successive runs of `purge_dups` (https://github.com/dfguan/purge_dups) with raw read mappings re-generated upon each run. Further scaffolding of the assembly using the Hi-C short-read data was performed with YaHS v1.2 [41] following the Arima genomics mapping pipeline (https://github.com/ArimaGenomics/mapping_pipeline retrieved April 2024); briefly, Hi-C reads were mapped to the Canu scaffolds using BWA-MEM v0.7.17 (<https://github.com/lh3/bwa>), screened for chimeras with `filter_five_end.pl`, and re-paired prior to PCR duplicate removal with Picard Tools v3.1.1 (<https://github.com/broadinstitute/picard>).

Gene structures were predicted on the Hi-C scaffolded assembly using BRAKER3 v3.0.8 [42]. Gapped short-read mappings of all 12 organ-specific Illumina RNAseq libraries (below) and two libraries generated from eight whole-body adult and twelve whole-body nymphal *H. longicornis*, respectively, were generated using STAR v2.7.10b [43] and used as evidence by BRAKER3. In addition, the 4.3M protein sequences that compose the Arthropoda OrthoDB v10 ortholog set [27] were used as protein evidence. A set of high-confidence genes was created by mapping all RNAseq data generated (below) to the primary isoform of each predicted transcript coding sequence using BBmap v38.94 (<https://jgi.doe.gov/data-and-tools/software-tools/bbtools/>) and retaining those with at least one mapped read. Functional annotation of predicted proteins was performed using eggNOG-mapper v2 [44]. In addition, proteins were screened for reverse-vaccine utility using Vacceed v1.1 [45], VaxiJen v2.0 [46], IEDB MHC I and MHC II v2.24 prediction scores [47], and BepiPred v3.0 [48] and selected based on the criteria described below.

Organ dissection and RNAseq library construction

To construct a representative and organ-specific RNA catalog, 114 adult female *H. longicornis* ticks were collected from the field via flagging in August of 2022 from two sites within Rutgers University, New Brunswick, NJ with dense populations (ASB2: 40.47797, -74.42521 and EcoPreserve: 40.51690, -74.43535); the ASB2 site was sampled to initiate the USDA colony from which the individual was selected for genome sequencing. Both sites have historically harbored ticks predominantly of mitochondrial haplotype H1, however haplotype H2 (currently differentiated by a single SNP within the 612bp cytochrome c oxidase subunit 1 barcoding locus [39]) has been identified at low numbers as well. As we did not assay each of the 114 individuals collected, the possibility that haplotype H2 ticks contributed RNA to the sample exists. This is unlikely to affect the gene prediction process as the hidden Markov model derived from the RNAseq mapping predicts gene structures on the reference assembly directly.

Each individual tick was embedded live in paraffin wax, and the dorsal idiosoma was dissected using a microscalpel. The ovaries, midgut, and salivary glands were removed and pooled separately in individual microcentrifuge tubes on ice containing RNAlater. In addition, the distal segment of the foreleg (containing the Haller's organ) and the hind leg were removed and placed in RNAlater. Of the 114 individuals, successful dissections of intact organs were made from $n=99$ ticks. Three pools of 33 organs each were created prior to RNA extraction of each using the Zymo Quick-RNA Microprep Kit

(Zymo Research, Irvine, CA) and eluted in 7 μ L of elution buffer. Illumina libraries were generated for each extraction template using the TruSeq RNA Library Prep Kit v2 (Illumina Inc, San Diego, CA) and sequenced on an Illumina MiSeq instrument using a 150-cycle single-end kit. Two RNAseq libraries were also generated from respective pools of whole-body adult ($n=18$) and nymphal ($n=25$) ticks collected from the ASB2 site (above), homogenized live in Buffer RLT + β -mercaptoethanol on a Qiagen TissueLyser w/ 4mm steel bead and extracted using the Qiagen RNeasy Mini kit (Qiagen, Hilden, Germany). Libraries were prepared as above and sequenced on a 300-cycle kit in a 150 bp \times 150 bp paired-end run.

Differential expression analysis

RNAseq reads from each replicated organ-specific library were quality- and adapter-trimmed with BBduk (min. Q-score=20, length=50 bp), corrected with BayesHammer (bundled with SPAdes v3.15.5 [<https://github.com/ablab/spades>]) and mapped to the nucleotide transcripts corresponding to the predicted genes (above) using BBmap v38.94 (minimum nucleotide sequence identity=95%; reads with multiple top-scoring alignments were excluded) and the read count data was used to generate an input table for the differential expression analysis by DESeq2 v4.0.5 [49] in R.

The raw count data from RNA-seq reads were obtained for each organ-specific library, including the three pools from ovaries, midgut, salivary glands, forelegs and hindlegs collected from 114 adult female *H. longicornis* ticks. Sample counts were processed to create a sample table associating each sample with its corresponding condition. A DESeqDataSet file was generated from the count data and sample table, and the condition column was converted to a factor with specified levels. DESeq analysis was then performed to estimate size factors and dispersion for each gene, followed by differential expression analysis comparing the conditions. Results were filtered based on adjusted p -values ($padj < 0.05$) to identify differentially expressed genes in each organ with regard to the others; the ovaries, midgut and salivary glands were used in a three-way comparison, while the foreleg and hindleg were compared exclusively to each other.

Gene Ontology (GO) analysis

We tested for Gene Ontology (GO) enrichment among the set of significantly upregulated genes in each organ using GO_MWU (https://github.com/z0on/GO_MWU; accessed on May 16, 2024). Briefly, the GO terms assigned to each gene by eggNOG-mapper were parsed and divided into a test set (comprising those assigned to the significantly upregulated genes) and reference (comprising the entire gene set). A binary significance

measure (Fisher's exact test) was implemented in GO_MWU to identify over-represented ontologies in the expression data for each organ and grouped by top-level terms Biological Process (BP), Molecular Function (MF) and Cellular Component (CC). The parameters were set to exclude very large and very small GO categories (excluding terms with >10% of total number of genes [largest=0.1] and those with less than 5 genes [smallest=5]) and to merge similar terms based on gene sharing (clusterCutHeight=0.9). The results were visualized through a hierarchical clustering dendrogram of significant GO categories.

Vaccine candidate screen

To search for gene products that may provide utility for future vaccine candidates, we employed a set of multiple open-source applications. First, the predicted *H. longicornis* proteome was screened using Vacceed [45] trained with the parasitic immunogens of Goodswen et al. 2022 [50] and run using the nearest-neighbor (nn), k-nearest-neighbor (knn), random forest (rf) and adaptive boosting (ada) classifiers. Any protein with an average ML score of ≥ 0.90 was retained. This set was further analyzed with VaxiJen v2.0 [46] (target organism="Parasite", default threshold=0.5), BepiPred v3.0 [48] (default threshold=0.15) and the IEDB MHC-I and MHC-II binding predictions to bovine MHC alleles using all possible windows of 8 – 14 amino acids. A BLASTP homology search (e-value = 1×10^{-5}) of each tick protein was performed

against the proteomes of *Bos taurus* (NCBI accession GCA_002263795.3), *R. microplus* (GCF_013339725.1) and *Ixodes scapularis* (GCF_016920785.2) to ensure further vaccine candidates did not retain high homology to the existing host (bovine) proteome, and to assess the conservation between our predicted proteins and those of other ticks. The final genes and pathways of interest highlighted here were manually selected for discussion based on the annotation, gene ontology, organ-specific RNAseq count data, and prediction pipeline scores.

Results

Our PacBio HiFi sequencing generated 9.26M reads with a read N50 of 12.68 kbp and summing to 97.65 Gbp (Table 1). The final Hi-C scaffolded genome assembly comprised 270 scaffolds, summing to 3.09 Gbp, with an N50 value of 213.4 Mbp; a total of 13 scaffolds exceeded 100 Mbp in length. The genome assembly has been deposited at DDBJ/ENA/GenBank under the accession JBLBWX000000000. The version described in this paper is version JBLBWX010000000. Gene prediction initially identified 33,071 gene structures within the assembly; removal of those without any evidence of RNAseq mapping (below) reduced this set to 21,947 high-confidence gene structures (Supplemental Table 1). The evaluation of this gene set using BUSCO single-copy ortholog analysis revealed that 96.1% of the core Arthropoda odb10 orthologs were found (94.9% complete, 1.2% fragmented), with 5.7% of those duplicated and 3.9% missing; this

Table 1 Sequencing metrics for sequencing libraries generated from RNA and DNA of *H. longicornis*

Library	Read Count	Run type	Template	Accession
Foreleg 1	7,015,198	150 bp SE	RNA	SRR31075554
Foreleg 2	6,953,712	150 bp SE	RNA	SRR31075566
Foreleg 3	6,252,410	150 bp SE	RNA	SRR31075565
Hind leg 1	6,995,694	150 bp SE	RNA	SRR31075564
Hing leg 2	9,854,039	150 bp SE	RNA	SRR31075563
Hind leg 3	8,756,249	150 bp SE	RNA	SRR31075562
Midgut 1	6,676,084	150 bp SE	RNA	SRR31075568
Midgut 2	6,896,959	150 bp SE	RNA	SRR31075567
Midgut 3	6,474,363	150 bp SE	RNA	SRR31075561
Ovary 1	6,286,098	150 bp SE	RNA	SRR31075560
Ovary 2	6,207,232	150 bp SE	RNA	SRR31075559
Ovary 3	6,288,798	150 bp SE	RNA	SRR31075558
Salivary gland 1	6,573,236	150 bp SE	RNA	SRR31075557
Salivary gland 2	6,272,090	150 bp SE	RNA	SRR31075556
Salivary gland 3	10,169,027	150 bp SE	RNA	SRR31075555
Whole body (18A)	19,335,022	150 bp × 150 bp PE	RNA	SRR15089914
Whole body (25N)	19,232,682	150 bp × 150 bp PE	RNA	SRR15089913
Hi-C	520,869,992	150 bp × 150 bp PE	DNA	SRR31757894
Genome	9,259,435	PacBio HiFi	DNA	SRR31757895

score did not change when the initial (unpurged) 33,071 genes were screened in the same manner, indicating that our purging process was not likely to have removed legitimate gene structures. The average predicted CDS (coding transcript), exon, and intron lengths exclusive of UTR were 1,314 bp, 252 bp and 4,143 bp respectively for the high-confidence gene set; by comparison, the most recent assembly of Jia et al., reported average CDS length of 891 bp, exon and intron lengths of 295 and 2,754 bp respectively. We attribute the increased CDS sequence length of our gene set to be reflective of more correct and/or full-length predictions as evidenced by the much higher BUSCO proteome scores.

For gene identification, we generated >19.3M reads from 18 whole adult *H. longicornis* ticks and >19.2M reads from 15 whole nymphs using 2x150 bp paired-end Illumina sequencing. For organ-specific expression characterization, a total of 107.7M raw reads were generated from separate 150 bp single-end Illumina

RNA-seq runs conducted on triplicate pools of salivary glands, midguts, ovaries, forelegs and hindlegs of the 99 adult ticks dissected above (Table 1; 33 organs per replicate, per organ type). Differential expression analysis revealed significant gene expression differences commensurate with localization between organs (Fig. 1; Supplemental Fig. 1; Supplemental Tables 2 and 3). We identified 2,492 genes significantly upregulated in ovaries, 2,329 in salivary glands, and 1,109 in the midgut ($p < 0.05$) with respect to the other two organ types; 152 genes were up-regulated in foreleg dissections with respect to the hindleg, and an additional 16 genes upregulated in hindlegs with respect to the foreleg. The gene ontology (GO) enrichment analysis identified many significantly over-represented GO terms associated with these localized genes across organs, discussed below (Fig. 2; Supplemental Figs. 2 and 3; Supplemental Table 4).

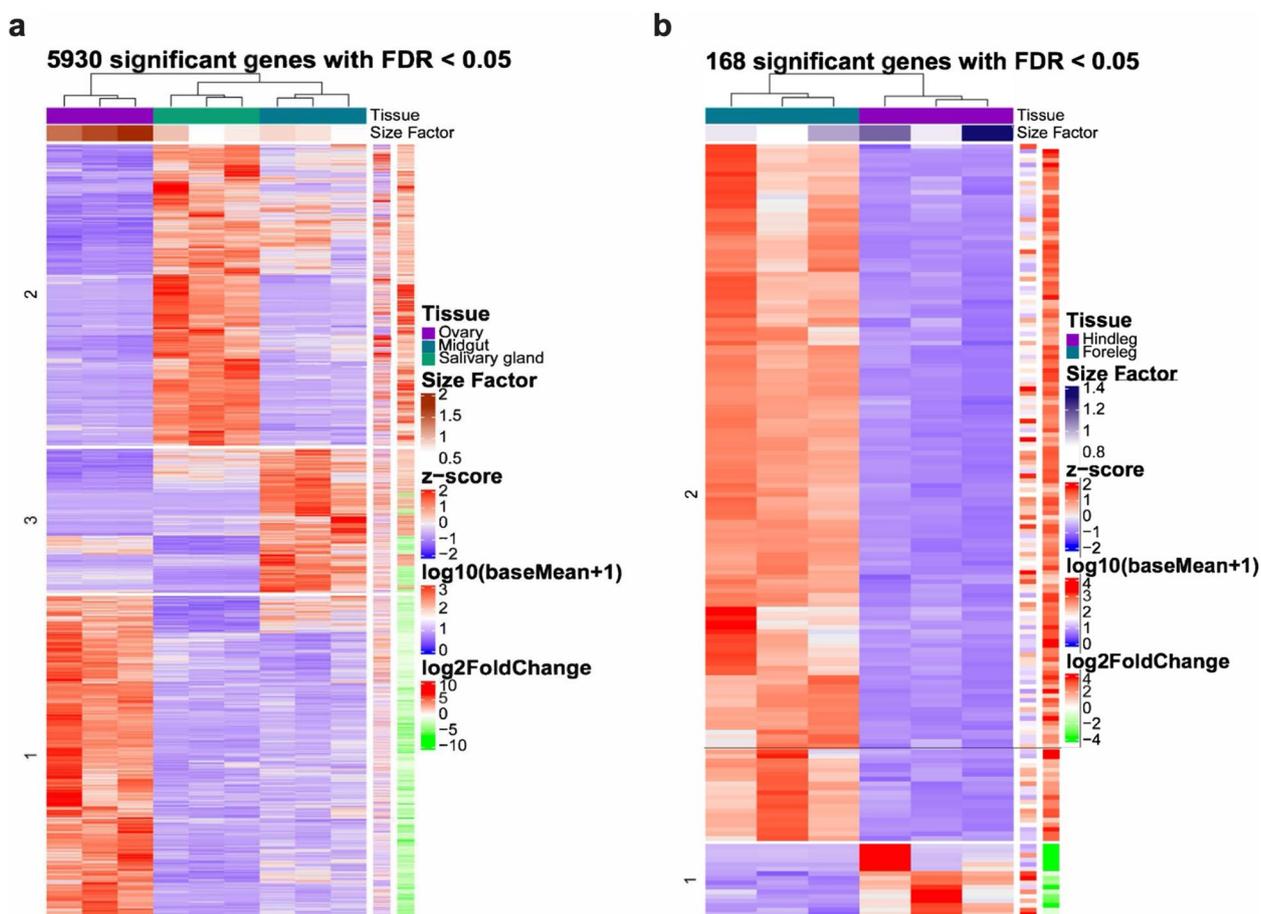


Fig. 1 Differential gene expression across *H. longicornis* organs. Heatmat illustrating gene expression observed across organ types, highlighting genes upregulated in (a) ovaries (as compared to salivary glands + midguts; 2,492 genes), salivary glands (as compared to ovaries + midguts; 2,329 genes), midgut (as compared to salivary glands + ovaries; 1,109 genes), and (b) fore leg – hind leg comparison (168 genes) as revealed through RNA-seq analysis and statistical significance ($p < 0.05$)

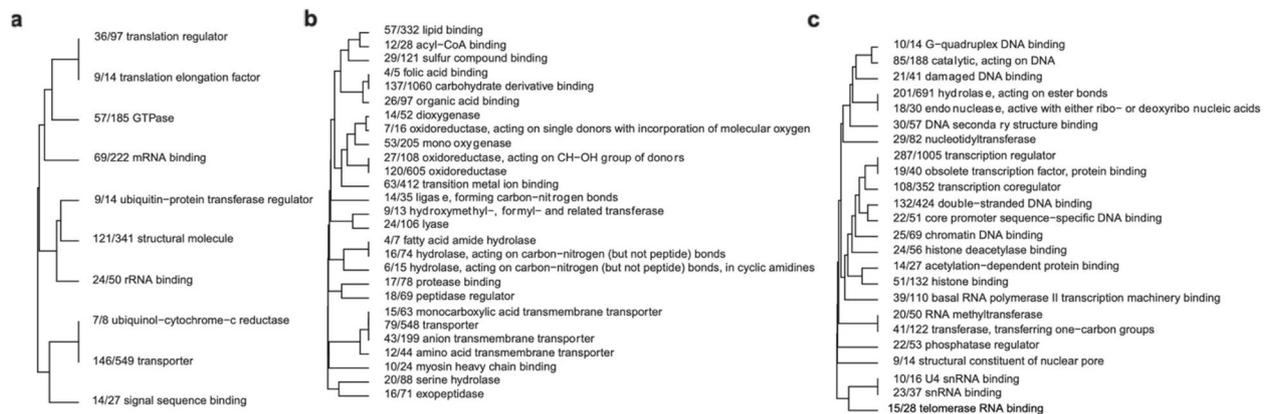


Fig. 2 Gene Ontology Enrichment (GO-MWU) analysis for molecular function (MF) ontology in *H. longicornis* organ-specific genes. This figure presents hierarchical trees of significantly over-represented Gene Ontology (GO) terms associated with molecular function for genes upregulated in: **A** salivary glands, **B** midgut, and **C** ovaries. Remaining ontologies (Cellular Component and Biological Process) are provided in Supplemental Figs. 2 and 3

Salivary glands

The salivary gland gene set was significantly enriched for various Molecular Function (MF) categories, largely related to ribosomal biogenesis, peptide synthesis, and electron transport (MF, Mann–Whitney U-test, $p < 0.05$), while Cellular Component (CC) ontologies included many related to the rough endoplasmic reticulum and Golgi-associated vesicular transport/protein targeting, cytosolic ribosomes, and the inner mitochondrial membrane. We noted multiple constituents encoded RVT_1 reverse-transcriptase and DDE_Tnp_4 endonuclease Pfam domains of eukaryote provenance, a phenomenon also observed in the sialotranscriptome of *Ixodes ricinus* [51]. The RVT_1 domain was additionally noted in multiple members of the gene set localized to ovaries and midguts (below), indicating that transposition may have a core function in tick genomes; whether this function pertains to host (eukaryote) derived gene function or the self-proliferation of mobile elements and/or endogenous viral elements [52] remains unknown. Several THAP domain-encoding proteins were noted and likely associated with ribosomal structures, while proteins encoding P450 domains are linked to oxidation–reduction processes (GO:0030684). The 7tm_1 Pfam domain, found in three assigned genes, suggests a role in cellular signaling pathways through its association with G-protein coupled receptor (GPCR) activity (GO:0008135).

Midgut

The gene set over-expressed in the midgut was enriched for MF categories including lipid/acyl-CoA binding, oxidoreductase and monooxygenase enzymes, protease and peptidase regulators, and serine hydrolases. CC categories included many gene products localized to the ER membrane, secretory vesicles, endosomes and lytic

vacuoles along with various other ion and organic compound transport functions. Pfam families were identified that contribute to nutrient absorption and detoxification processes; the P450 family, represented by nine genes, is linked to vesicle structures and cellular junctions (GO:0005777, GO:0042175). Additional Pfams, including AMP-binding (e.g., g30548.t1, g29853.t1) and RRM_1 (e.g., g18086.t1, g30419.t1), contribute to metabolic and RNA-binding activities, respectively. The Pkinase Pfam, representing protein kinases, and RRM_1, with RNA recognition motifs, are also involved in nucleic acid binding and transporter activity across various cellular pathways.

Ovaries

The ovary gene set was significantly enriched for MF categories including multiple single/double-stranded DNA binding enzymes, helicases, (endo)nucleases, DNA/RNA polymerases, and transcription factor activity reflecting an emphasis on gene regulation, protein modification, and structural binding activities critical to the reproductive processes in tick ovaries. CC categories included a large number of complexes associated with cellular and DNA replication (nuclear chromosome, pronucleus, meiotic spindle, mismatch repair complex, pole plasm) and transcription (spliceosomal tri-snRNP complex, transcription regulator complex). AMP-binding and AMP-binding_C Pfams, present in six genes, highlight the importance of resource allocation in the ovaries through their involvement in metabolic pathways and ligase activity (GO:0044427, GO:0000228). Similarly, the Mito_carr Pfam supports mitochondrial transport and energy distribution, and Trypsin, part of the peptidase S1 family, contributes to protein degradation and nutrient processing. Finally, the 7tm_1 Pfam, identified in four genes, is associated with GPCR activity (GO:0008135).

Foreleg

The expression data for the foreleg/hindleg comparison was analyzed to identify genes with potential localized expression in the Haller's organ that contribute to tick chemosensory processes [53]. Due to the small number of upregulated genes ($n=152$), we did not perform statistical tests for GO enrichment; rather, we examined the list and annotations manually. Using a protein database comprising the qualified ionotropic receptor (IR) gene set of *Ixodes scapularis* [54] we performed a BLASTp search with our predicted foreleg proteome and identified top hits to four such IRs: Ir93a (g5615.t1, $e\text{-val}=0.0$), Ir25a (g5616.t1, $e\text{-val}=0.0$), Ir107 (g19560.t1 $e\text{-val}=0.0$) and Ir142 (g19558.t1 $e\text{-val}=3.9 \times 10^{-166}$); of note, both Ir93a and Ir25a exhibited biased expression in the forelegs of *Ix. scapularis* in the analysis of Josek et al. [54]. Beyond described ionotropic receptors, we identified an otopetrin domain-containing protein (g18830.t1) encoding a multi-transmembrane proton-selective channel that functions as an acid taste receptor in *Drosophila* [55]. The top NCBI BLASTp hit for this protein was to a *Dermacentor silvarum* homolog ("proton channel OtopLc"; XP_037569237.1). Although the alignment spanned the full length of both proteins, the amino acid identity remained low at 81.7%. A second gene encoding an anoctamin calcium-activated chloride channel (g27605.1) was identified with a top BLASTp hit to "anoctamin-4 like" in *Dermacentor albipictus* (XP_065300563). Anoctamin channels have previously been shown to participate in

thermal nociception in insects [56]; however, any function with respect to *H. longicornis* remains unknown.

Informatic screening of the predicted proteome with Vacceed identified 1,594 proteins with an average ML score ≥ 0.90 . Subsequent analysis of these proteins with VaxiJen, BepiPred, and the IEDB MHCI/II binding prediction pipelines provided an annotated list for manual inspection with regard to any potential role in critical metabolic processes and further efficacy as an anti-tick vaccine (Supplemental Table 5; discussed below). It should be noted that the ticks from which RNA was extracted were questing and thus the expression data generated here may not reflect a metabolic snapshot of an actively blood-feeding tick, as would be ideal for deriving vaccines that act during this physiological state.

Discussion

Our single-molecule sequencing of North American *H. longicornis* ticks yielded a high-quality and contiguous genome and predicted proteome. The arthropoda_odb10 BUSCO scores for the data generated here were notably positive given the triploid karyotype of North American populations, which has resulted in low completeness and/or high duplication of currently available gene sets for this tick [28–31] (summarized in Table 2). The haploid assembly size of our genome reported here (3.10 Gbp) falls within the range of those currently reported (2.48 – 3.16 Gbp, excluding the 7.4Gbp assembly of [30]),

Table 2 Assembly statistics and BUSCO scores for existing *H. longicornis* genome assemblies and predicted proteomes; the arthropoda_odb10 ortholog dataset was used for BUSCO analyses ($n=1,013$ orthologs). The highest contiguity level for each assembly ([S]caffold / [C]ontig) is reported

Reference	NCBI Assembly ID	(S)caffolds / (C)ontigs	Sequencing Technology	Assembly size	Gene Count	Complete	Duplicate	Fragmented	Missing
Jia et al. 2020 (G)	BIME_Hael_1.3	3,885 (S)+11 chromosomes	PacBio Sequel+Illumina+Illumina Hi-C	2.6 Gbp	N/A	87.50%	2.40%	6.10%	6.40%
Jia et al. 2020 (P)	BIME_Hael_1.3	N/A	N/A	22.71 Mbp	25,651	53.60%	2.00%	12.90%	33.50%
Guerrero et al. 2019 (G)	HLAgrLifeRun1	34,208 (C)	PacBio Sequel + Illumina	7.36 Gbp	N/A	95.10%	84.20%	3.50%	1.50%
Yu et al. 2022 (G)	ASM2241470v1	6,516 (S)	PacBio Sequel + Illumina	3.16 Gbp	N/A	90.20%	20.00%	4.80%	4.90%
Umemiya-Shirafuji et al. 2023 (G)	ASM2984928v1	99,353 (C)	Nanopore + Illumina	2.48 Gbp	N/A	78.50%	2.50%	13.20%	8.30%
This study (G)	GCA_048455015.1	270 (S)	PacBio Sequel II + Illumina Hi-C	3.1 Gbp	N/A	94.40%	5.70%	3.40%	2.30%
This study (P)	GCA_048455015.1	N/A	N/A	25.85 Mbp	21,947	95.00%	5.70%	1.20%	3.80%

which tend to exhibit larger proportions of duplicated BUSCOs as assembly size increases and genomic haplotypes remain uncollapsed. The earliest assembly [30] for example had high genome BUSCO completeness score of 95.10% but also likely comprised multiple chromosomal haplotypes, as 84.2% were duplicated and the scaffolds summed to 7.36 Gbp; the most recent *H. longicornis* genome assembly generated using Oxford Nanopore and Illumina data [28] by contrast summed to 2.5 Gbp however encoded 78.50% of core BUSCOs with an additional 13.2% fragmented. As sequencing technologies continue to evolve, so too does the fidelity of the data created; both assembly size and complete gene count are likely to improve as a result by reducing polymorphisms in the overlap graph that arise from sequencing errors (for example, from “noisy” Oxford Nanopore and/or early generation PacBio long read data), and commensurately creating intact, complete gene structures free of artifactual nonsense mutations. To our knowledge, this is the first application of Sequel IIe sequence technology in support of the generation of an *H. longicornis* genome.

A large number of potential vaccine candidates were predicted based strictly on informatic analyses. We highlighted a diverse set of candidates for future experimentation based on gene expression, annotation, and implication in several major metabolic pathways (cellular morphogenesis, water transport, chitin, and cholesterol metabolism, oogenesis, and neural signaling). As mentioned above, it should be noted that to act as a canonical anti-tick vaccine target, the antigenic determinant should be accessed (usually by cell surface presentation) by the antibody derived during feeding; this aspect remains to be tested for the molecules discussed below.

Among the top vaccine candidates predicted were members of the Innexin [57] and Integrin beta (ITGB1) [58] families (e.g., g17885, g17696, g16011), which have roles in cellular morphogenesis. In arthropods, Innexins are integral membrane proteins that create gap junction channels and hemichannels [59] to facilitate the swift transmission of endocrine signals among adjacent cells by enabling the direct exchange of ions and small molecules [60]. Innexin-like transcripts were down-regulated in the salivary glands of *R. microplus* ticks that fed on resistant cattle [61]. It was also identified in the salivary glands of the female tick *Amblyomma americanum* during blood feeding [62]. Integrins are transmembrane proteins composed of linked α and β subunits that act as crucial adhesion and signaling centers on the cell surface [63]. The silencing of integrin beta subunits has been shown to negatively impact the development of various arthropods, including the Oriental tobacco budworm *Helicoverpa assulta* [64] and the beet armyworm *Spodoptera exigua* [58]. In *Ixodes ricinus* ticks, reduced expression of

integrin beta led to decreased engorgement [65]. In our study, innexin was highly expressed in the midgut, while ITGB1 was highly expressed in the midgut, ovaries, and legs but moderately expressed in the salivary glands suggesting a potentially significant role in tick biology, particularly in the processes of blood-feeding and salivary gland function.

We next investigated Aquaporin (AQP12B), which exhibited high levels of expression in the legs but low expression in other organs (e.g., g16408, g4905). We identified three aquaporins, each at different loci. In ticks, aquaporins (aquaporin 1 and aquaporin 2) have a critical role in osmoregulation across cellular membranes, including the salivary gland during feeding [66]. Aquaporin 1 can be detected in the gut, rectal sac, and salivary glands, while Aquaporin 2 is solely expressed in the salivary glands [67, 68]. Although it is challenging to develop aquaporin 1-derived anti-tick vaccines due to the high structural similarities between tick and bovine host aquaporin 1 ortholog, recombinant tick aquaporin 1 antigens were successful against *I. ricinus* and *R. microplus* [69, 70]. Furthermore, vaccination with *I. ricinus* aquaporin 1 and a highly-conserved aquaporin 1 motif showed an efficacy of 32% and 80%, respectively, in controlling *I. ricinus* larvae in rabbits by reducing tick survival and molting [70]. Moreover, cattle vaccination with synthetic peptides of aquaporin 2 reduced the number of replete-feeding *R. microplus* ticks by 25% [71].

Chitin metabolic processes are critical components of various extracellular matrices in arthropods. We detected two members of the carbohydrate-binding module 14 (CBM_14) family as plausible candidates for future examination: Chitin deacetylase 1 (CDA1; g18805) and Obstructor-E (Obst-E; g17730), both containing the Chitin binding Peritrophin-A domain. We found that CDA1 exhibited high expression levels in the ovary and moderate expression in the legs. Obstructor-E showed moderate expression in both the ovary and legs. Chitin deacetylases (CDAs) are enzymes that modify chitin and play essential roles in insect metamorphosis and development [72]. For instance, in the cigarette beetle (*Lasioderma serricorne*) CDA1 is primarily expressed in late larval and pupal stages, with the highest expression in the integument; its silencing results in high larval mortality and prevents larval-pupal molting [72]. Another study detected a predominant immunogenic 60 kDa protein homologous to arthropod chitin deacetylase (IsCDA) in the *Ixodes scapularis* acellular epithelial barrier [73]. This is also known as the peritrophic membrane and serves to encapsulate the food bolus during feeding. Although silencing of this protein did not affect peritrophic membrane formation, treatment with antibodies against IsCDA significantly increased *B. burgdorferi* levels in ticks, indicating a role

in spirochete establishment and persistence within *I. scapularis* [73]. Elsewhere it has been shown that IsCDA disruption impairs the structure of the peritrophic membrane and affects tick engorgement duration and pathogen transmission [74]. By contrast, Obst-E is a secreted protein that plays a significant role in the physical regulation of body shape during metamorphosis in *Drosophila melanogaster* [75] by facilitating the oriented contractility and expandability in the larval cuticle essential for metamorphic shape change. The absence of Obst-E leads to abnormal puparium formation [75].

Like chitin, cholesterol is a critical component of many arthropod structures, including ecdysteroid hormones and egg waxes. However, this sterol cannot be synthesized directly by the tick [76], and it stands to reason that interrupting such pathways could directly affect fitness. We detected very high expression of Niemann–Pick type C2 (NPC2) protein in the midgut (g5197). This protein is important for mediating intracellular transport and regulation of cholesterol metabolism [77]. In arthropods, NPC2 proteins are also involved in chemical communication [78, 79] and may have a role in odor recognition in ticks [80]. NPC2 has been detected in *H. longicornis* [81], *I. ricinus* [82] and *A. americanum* [78]. In another study, NPC2 proteins were found to be expressed at each developmental stage of *Rhipicephalus linnaei* with the highest levels in adult males; molecular docking indicated that several amino acid residues in this protein could bind multiple ligands, suggesting that these residues might be important in chemosensation [80]. Our study also showed high expression of Sterol O-acyltransferase 1 (SOAT1; g2856) in the ovaries, midgut, and legs. SOAT1 proteins are involved in processes such as lipoprotein assembly and the absorption and metabolism of cholesterol [83]. This protein has similarly been detected in the midgut of adult female *Rhipicephalus* species [84] and in the midgut and ovaries of *R. microplus* [57, 85].

Among the proteins that we predicted as potential vaccine candidates was SPARC (g2853), which exhibited high expression in the ovaries and midgut, with nominal expression in the salivary glands and legs of the *H. longicornis* ticks in our study. SPARC is a calcium-binding glycoprotein that was first detected in calf bones [86]. In the cockroach *Blattella germanica*, it functions in oogenesis by regulating follicular cell mitosis and maintaining cytoskeleton integrity [87]. Among insects, SPARC has been thoroughly studied in *D. melanogaster*, where it also has roles in oogenesis, embryo development, maintaining tissue integrity [88], and heart development [89]. Another study detected SPARC in the proteome of *Rhipicephalus sanguineus* tick saliva, but its main function in ticks has not yet been identified [90]. Our results also showed very high expression of Follistatin-like 1 (g22970) in the

ovary, midgut, and legs, consistent with the homolog previously identified in *H. longicornis* ovaries; it may play a role in the negative regulation of cellular growth [91]. It has been shown that this gene is expressed in all tick life stages in the midgut, salivary glands, and hemocytes [91]. Silencing of follistatin causes a decrease in tick oviposition but does not affect blood feeding [91].

Finally, we noted potential vaccine candidates that are likely to be involved in the tick nervous system; two such predicted proteins were annotated as Neurexin IV (NRX-4/CNTNAP1; g22371 and g21630), which showed high expression in the legs and moderate expression in the ovary, midgut, and salivary gland. Neurexin IV was first identified in *Drosophila* (FLYBASE:FBgn0013997) and characterized as a new integral protein of septate junctions (SJs) [92] where they act predominantly at the presynaptic terminal in neurons and play essential roles in neurotransmission and differentiation of synapses [93]. These proteins may have additional roles in the initiation of myelination and glial wrapping [94, 95]. Although integral components of synaptic terminals may not be commonly associated with reverse-vaccinology approaches due to the nature of their localization, the function of NRX-4 at the plasma membrane of paracellular diffusion barriers may still mean antigenic sites could be exposed to host immune effectors during a blood meal.

Conclusions

PacBio HiFi and Hi-C sequencing coupled with organ-specific transcriptome profiling facilitated a highly contiguous genome assembly of invasive *H. longicornis* ticks in North America. Assembly completeness with respect to gene content as assessed by BUSCO identified 962 (96.1%) core eukaryote orthologs and underlines the robust nature of this assembly and its value as a future resource for tick genomic research. Moreover, our organ-specific analyses yielded many significant localizations across key organ systems in ovaries, salivary glands, midguts, and legs; we identified several thousand up- and down-regulated genes, highlighting the complex expression networks underlying tick physiology. Using these data, our bioinformatic screening identified many potential anti-tick vaccine candidates, including key members of diverse pathways that include cellular morphogenesis, osmoregulation, chitin-binding, cholesterol metabolism, oogenesis, and neural development. We envision that these will be among many targets for further experimentation aimed at validation. Our work provides a firm foundation for genetic and functional studies of parthenogenetic *H. longicornis* in North America and will inform vaccine development and other interventions for reducing these arthropods' burden on human and animal health.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11477-1>.

Additional file 1. Table S1. Annotations corresponding to the high-confidence gene set predicted from the *H. longicornis* Hi-C scaffolded genome assembly, including functional predictions and ortholog evidence. The expression column lists the organ type(s) (salivary gland, ovary, midgut, foreleg) with upregulated expression; columns E through X contain standard eggNOG-mapper output. Table S2. DEseq2 output for organ-specific RNAseq analysis of Ovary, Midgut and Salivary glands from *H. longicornis*. The 5,930 genes with significant upregulation in any condition (2,492 localized to ovaries, 2,329 to salivary glands, and 1,109 to the midgut; adjusted p-value < 0.05) are listed. Normalized counts are given in columns I – Q. Table S3. DEseq2 output for organ-specific RNAseq analysis of Foreleg and Hind leg from *H. longicornis*. The 168 genes with significant upregulation in either condition (152 localized to fore leg, 16 to hind leg; adjusted p-value < 0.05) are listed. Normalized counts are given in columns I – N. Table S4. Gene Ontology (GO) enrichment analysis of significantly upregulated genes from *H. longicornis*. The analysis identifies GO categories enriched in upregulated genes based on the $-\log(p\text{-value})$ between salivary glands, midguts, and ovaries of *H. longicornis*. The GO terms represent functional categories, and the “Gene Ratio” indicates the ratio of significant genes ($P < 0.05$) present in a category to the total number of predicted genes associated with that category. P-values were adjusted for multiple testing using false discovery rate (FDR) correction. Table S5. Comprehensive analysis of 1,593 candidate anti-tick vaccine proteins identified through informatic screening. Column headings contain the respective software packages from which the data were derived (see Methods) and the associated descriptors. Column L contains the VaxiJen score; column M contains the average of Vaxceed and Vaxijen scores; column N contains the number of strong (IC50 < 50nM) MHC1 alleles for all possible windows of 8 – 14 amino acids with affinity to any bovine (BoLA) MHC1 allele; column O contains the number of amino acid residues with BepiPred-3.0 epitope scores above the default threshold of 0.151

Additional file 2. Figure S1. MA Plots for Pairwise RNA-Seq differential expression analysis between *H. longicornis* organ types. This figure consists of four MA plots (a-d) visualizing pairwise RNA-seq differential expression between organ types: (a) Ovary vs. Midgut, (b) Ovary vs. Salivary Gland, (c) Salivary Gland vs. Midgut, and (d) Hindlegs vs. Forelegs. In each panel, the x-axis represents the mean of normalized counts (average expression), while the y-axis shows the log2 fold change (differential expression). These MA plots highlight genes with significant up- or down-regulation between the compared organs. Figure S2. Gene Ontology Enrichment (GO-MWU) analysis for Cellular Component in *H. longicornis* organ-specific genes. This figure displays the hierarchical clustering of significantly enriched GO terms related to cellular components for genes upregulated in: (a) Salivary glands, (b) Midgut, and (c) Ovaries. Figure S3. Gene Ontology Enrichment (GO-MWU) analysis for Biological Process in *H. longicornis* organ-specific genes. This figure showcases hierarchical GO term trees for biological processes enriched among genes upregulated in: (a) Salivary glands, (b) Midgut, and (c) Ovaries

Acknowledgements

The authors wish to acknowledge continuing support from the New Jersey Agricultural Experiment Station (NJAES) and the Rutgers University Office for Advanced Research Computing (OARC). We thank Dr. Glen Scoles at the USDA-ARS Invasive Insect Biocontrol & Behavior Laboratory for discussions and helpful insight on this and other related projects.

Authors' contributions

D.C.P, P.T and P.S designed the project. D.C.P and M.A.M conducted genome assembly, gene prediction and annotations, antigen screening and drafted the initial manuscript. M.M.B collected ticks and performed tissue microdissections. M.A.M performed differential expression analyses. D.B, P.S, and K.B performed DNA extractions, HiFi library preparation, and oversaw PacBio sequencing protocols. N.E.W generated RNAseq libraries and oversaw MiSeq sequencing. All authors edited the final manuscript draft.

Funding

This work was funded by USDA-NIFA AFRI 2021–67015-34461 and USDA-NIFA Multistate NE1943 (NJ08540) awards to D. Price.

Data availability

The *H. longicornis* genome assembly, PacBio HiFi sequencing reads, and Illumina RNAseq data have been deposited in the NCBI GenBank database under BioProject PRJNA1172737. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JBLBWX000000000. The version described in this paper is version JBLBWX010000000 and can be accessed via <https://www.ncbi.nlm.nih.gov/nuccore/JBLBWX000000000.1/>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Entomology, Center for Vector Biology, The State University, 180 Jones Ave, New Brunswick, NJ 08901, USA. ²Knippling-Bushland U.S. Livestock Insects Research Laboratory, USDA-ARS, Kerrville, TX 78028, USA. ³Department of Entomology, Texas A&M AgriLife Research, 370 Olsen Blvd, College Station, TX 77843, USA.

Received: 18 December 2024 Accepted: 12 March 2025

Published online: 28 March 2025

References

1. Heath A. Biology, ecology and distribution of the tick, *Haemaphysalis longicornis* Neumann (Acari: Ixodidae) in New Zealand. *N Z Vet J.* 2016;64(1):10–20.
2. Hoogstraal H, Roberts FH, Kohls GM, Tipton VJ. Review of *Haemaphysalis (kaiseriana) longicornis* Neumann (resurrected) of Australia, New Zealand, New Caledonia, Fiji, Japan, Korea, and Northeastern China and USSR, and its parthenogenetic and bisexual populations (Ixodoidea, Ixodidae). *J Parasitol.* 1968;54(6):1197–213.
3. Heath A, Hardwick S. The role of humans in the importation of ticks to New Zealand: a threat to public health and biosecurity. *N Z Med J.* 2011;124(1339):67–82.
4. Beard CB, Occi J, Bonilla DL, Egizi AM, Fonseca DM, Mertins JW, Backenson BP, Bajwa WI, Barbarin AM, Bertone MA, et al. Multistate infestation with the exotic disease-vector tick *Haemaphysalis longicornis* - United States, August 2017-September 2018. *MMWR Morb Mortal Wkly Rep.* 2018;67(47):1310–3.
5. Price KJ, Ayres BN, Maes SE, Witmier BJ, Chapman HA, Coder BL, Boyer CN, Eisen RJ, Nicholson WL. First detection of human pathogenic variant of *Anaplasma phagocytophilum* in field-collected *Haemaphysalis longicornis*, Pennsylvania, USA. *Zoonoses Public Health.* 2022;69(2):143–8.
6. Price KJ, Graham CB, Witmier BJ, Chapman HA, Coder BL, Boyer CN, Foster E, Maes SE, Bai Y, Eisen RJ. *Borrelia burgdorferi sensu stricto* DNA in field-collected *Haemaphysalis longicornis* ticks, Pennsylvania, United States. *Emerg Infect Dis.* 2021;27(2):608.
7. Oakes VJ, Yabsley MJ, Schwartz D, LeRoith T, Bissett C, Broadus C, Schlater JL, Todd SM, Boes KM, Brookhart M, et al. *Theileria orientalis* Ikeda genotype in cattle, Virginia, USA. *Emerg Infect Dis.* 2019;25(9):1653–9.
8. Rainey T, Occi JL, Robbins RG, Egizi A. Discovery of *Haemaphysalis longicornis* (Ixodida: Ixodidae) Parasitizing a Sheep in New Jersey, United States. *J Med Entomol.* 2018;55(3):757–9.
9. Tufts DM, Goodman LB, Benedict MC, Davis AD, VanAcker MC, Diuk-Wasser M. Association of the invasive *Haemaphysalis longicornis* tick with vertebrate hosts, other native tick vectors, and tick-borne pathogens in New York City, USA. *Int J Parasitol.* 2021;51(2–3):149–57.

10. BurrIDGE MJ. Non-native and invasive ticks: Threats to human and animal health in the United States. University Press of Florida; 2011.
11. Myers SA, Scimeca RC. First Report of *Haemaphysalis longicornis* (Neumann) in Oklahoma, USA. *Pathogens*. 2024;13(10):861.
12. McBride WD, Mathews K, Jr. The Diverse Structure and organization of the U.S. Beef Cow-Calf Farms. In: Edited by EIB-73. U.S. Dept. of Agriculture ERS. Economic Research Service; 2011. <https://www.ers.usda.gov/publications/pub-details?pubid=44532>.
13. Singh K, Kumar S, Sharma AK, Jacob SS, RamVerma M, Singh NK, Shakya M, Sankar M, Ghosh S. Economic impact of predominant ticks and tick-borne diseases on Indian dairy production systems. *Exp Parasitol*. 2022;243:108408.
14. Lew-Tabor AE, Rodriguez Valle M. A review of reverse vaccinology approaches for the development of vaccines against ticks and tick borne diseases. *Ticks Tick Borne Dis*. 2016;7(4):573–85.
15. Domingos A, Antunes S, Borges L, Rosário VE. Approaches towards tick and tick-borne diseases control. *Rev Soc Bras Med Trop*. 2013;46(3):265–9.
16. Abbas RZ, Zaman MA, Colwell DD, Gilleard J, Iqbal Z. Acaricide resistance in cattle ticks and approaches to its management: the state of play. *Vet Parasitol*. 2014;203(1–2):6–20.
17. Ndawula C Jr, Tabor AE. Cocktail anti-tick vaccines: The unforeseen constraints and approaches toward enhanced efficacies. *Vaccines (Basel)*. 2020;8(3):457.
18. Wu PX, Cui XJ, Cao MX, Lv LH, Dong HM, Xiao SW, Liu JZ, Hu YH. Evaluation on two types of paramyosin vaccines for the control of *Haemaphysalis longicornis* infestations in rabbits. *Parasit Vectors*. 2021;14(1):309.
19. Lee SH, Li J, Moumouni PFA, Okado K, Zheng W, Liu M, Ji S, Kim S, Umemiya-Shirafuji R, Xuan X. Subolesin vaccination inhibits blood feeding and reproduction of *Haemaphysalis longicornis* in rabbits. *Parasit Vectors*. 2020;13(1):478.
20. Liang N, Dong HM, Fan XY, Wu YX, Yang F, Liu XY, Hu YH. Characterization and evaluation of a new triosephosphate isomerase homologue from *Haemaphysalis longicornis* as a candidate vaccine against tick infection. *Ticks Tick Borne Dis*. 2022;13(4):101968.
21. Fan XY, Xu XC, Wu YX, Liu XY, Yang F, Hu YH. Evaluation of anti-tick efficiency in rabbits induced by DNA vaccines encoding *Haemaphysalis longicornis* lipocalin homologue. *Med Vet Entomol*. 2022;36(4):511–5.
22. Feng T, Tong H, Zhang Q, Ming Z, Song Z, Zhou X, Dai J. Targeting *Haemaphysalis longicornis* serpin to prevent tick feeding and pathogen transmission. *Insect Sci*. 2024;31(3):694–706.
23. Ma H, Ai J, La Y, Zhao X, Zeng A, Qin Q, Feng S, Kang M, Sun Y, Li J. Hemalin vaccination modulates the host immune response and reproductive cycle of *Haemaphysalis longicornis*. *Vet Parasitol*. 2023;323:110051.
24. Adjou Moumouni PF, Naomasa S, Tuvshintulga B, Sato N, Okado K, Zheng W, Lee SH, Mosqueda J, Suzuki H, Xuan X, et al. Identification and characterization of *Rhipicephalus microplus* ATAQ homolog from *Haemaphysalis longicornis* ticks and its immunogenic potential as an anti-Tick vaccine candidate molecule. *Microorganisms*. 2023;11(4):822.
25. Bowman AS, Nuttall PA. *Ticks: biology, disease and control*. Cambridge University Press; 2008. ISBN-10:0521867614, ISBN-13:978-0521867610.
26. Rappuoli R. Reverse vaccinology. *Curr Opin Microbiol*. 2000;3(5):445–50.
27. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;38(10):4647–54.
28. Umemiya-Shirafuji R, Xuan X, Fujisaki K, Yamagishi J. Draft genome sequence data of *Haemaphysalis longicornis* Oita strain. *Data Brief*. 2023;49:109352.
29. Yu Z, He B, Gong Z, Liu Y, Wang Q, Yan X, Zhang T, Masoudi A, Zhang X, Wang T, et al. The new *Haemaphysalis longicornis* genome provides insights into its requisite biological traits. *Genomics*. 2022;114(2):110317.
30. Guerrero FD, Bendele KG, Ghaffari N, Guhlin J, Gedye KR, Lawrence KE, Dearden PK, Harrop TWR, Heath ACG, Lun Y, et al. The Pacific Biosciences de novo assembled genome dataset from a parthenogenetic New Zealand wild population of the longhorned tick, *Haemaphysalis longicornis* Neumann, 1901. *Data Brief*. 2019;27:104602.
31. Jia N, Wang J, Shi W, Du L, Sun Y, Zhan W, Jiang JF, Wang Q, Zhang B, Ji P, et al. Large-scale comparative analyses of tick genomes elucidate their genetic diversity and vector capacities. *Cell*. 2020;182(5):1328–1340.e1313.
32. de la Fuente J, Kopáček P, Lew-Tabor A, Maritz-Olivier C. Strategies for new and improved vaccines against ticks and tick-borne diseases. *Parasite Immunol*. 2016;38(12):754–69.
33. Nuttall PA. Wonders of tick saliva. *Ticks Tick Borne Dis*. 2019;10(2):470–81.
34. Jmel MA, Voet H, Araújo RN, Tironi L, Sá-Nunes A, Kotsyfakis M. Tick salivary kunitz-type inhibitors: Targeting host hemostasis and immunity to mediate successful blood feeding. *Int J Mol Sci*. 2023;24(2):1556.
35. Francischetti IM, Sa-Nunes A, Mans BJ, Santos IM, Ribeiro JM. The role of saliva in tick feeding. *Front Biosci (Landmark Ed)*. 2009;14(6):2051–88.
36. Martins LA, Kotál J, Bensaoud C, Chmelař J, Kotsyfakis M. Small protease inhibitors in tick saliva and salivary glands and their role in tick-host-pathogen interactions. *Biochim Biophys Acta Proteins Proteom*. 2020;1868(2):140336.
37. Lara F, Lins U, Bechara G, Oliveira P. Tracing heme in a living cell: hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. *J Exp Biol*. 2005;208(16):3093–101.
38. Narasimhan S, Rajeevan N, Liu L, Zhao YO, Heisig J, Pan J, Eppler-Epstein R, DePonte K, Fish D, Fikrig E. Gut microbiota of the tick vector *Ixodes scapularis* modulate colonization of the Lyme disease spirochete. *Cell Host Microbe*. 2014;15(1):58–71.
39. Egizi A, Bulaga-Seraphin L, Alt E, Bajwa WI, Bernick J, Bickerton M, Campbell SR, Connally N, Doi K, Falco RC et al: First glimpse into the origin and spread of the Asian longhorned tick, *Haemaphysalis longicornis*, in the United States. *Zoonoses and Public Health*. 2020, 67(6):637–650.
40. Nurk S, Walenz BP, Rhie A, Vollger MR, Logsdon GA, Grothe R, Miga KH, Eichler EE, Phillippy AM, Koren S. HiCanu: accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. *Genome Res*. 2020;30(9):1291–305.
41. Zhou C, McCarthy SA, Durbin R. YaHS: yet another Hi-C scaffolding tool. *Bioinformatics*. 2022;39(1):btac808.
42. Gabriel L, Brúna T, Hoff KJ, Ebel M, Lomsadze A, Borodovsky M, Stanke M. BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Res*. 2024;34(5):769–77.
43. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
44. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38(12):5825–9.
45. Goodswen SJ, Kennedy PJ, Ellis JT. Vacceed: a high-throughput in silico vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. *Bioinformatics*. 2014;30(16):2381–3.
46. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007;8:4.
47. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*. 2015;32(4):511–7.
48. Clifford JN, Høie MH, Deleuran S, Peters B, Nielsen M, Marcatili P: BepiPred-3.0: Improved B-cell epitope prediction using protein language models. *Protein Sci*. 2022;31(12):e4497.
49. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
50. Goodswen SJ, Kennedy PJ, Ellis JT. A state-of-the-art methodology for high-throughput in silico vaccine discovery against protozoan parasites and exemplified with discovered candidates for *Toxoplasma gondii*. *Sci Rep*. 2023;13(1):8243.
51. Chmelař J, Anderson JM, Mu J, Jochim RC, Valenzuela JG, Kopecký J. Insight into the sialome of the castor bean tick, *Ixodes ricinus* BMC Genomics. 2008;9:233.
52. Barnes M, Price DC. Endogenous viral elements in ixodid tick genomes. *Viruses*. 2023;15(11):2201.
53. Carr AL, Mitchell RD, III, Dhammi A, Bissinger BW, Sonenshine DE, Roe RM. Tick Haller's organ, a new paradigm for arthropod olfaction: How ticks differ from insects. *Int J Mol Sci*. 2017;18(7):1563.
54. Josek T, Walden KKO, Allan BF, Alleyne M, Robertson HM. A foreleg transcriptome for *Ixodes scapularis* ticks: Candidates for chemoreceptors and

- binding proteins that might be expressed in the sensory Haller's organ. *Ticks and Tick-borne Diseases*. 2018;9(5):1317–27.
55. Ganguly P, Madonsela L, Chao JT, Loewen CJR, O'Connor TP, Verheyen EM, Allan DW. A scalable *Drosophila* assay for clinical interpretation of human PTEN variants in suppression of PI3K/AKT induced cellular proliferation. *PLoS Genet*. 2021;17(9): e1009774.
 56. Jang Y, Oh U. Anoctamin 1 in secretory epithelia. *Cell Calcium*. 2014;55(6):355–61.
 57. Richards SA, Stutzer C, Bosman A-M, Maritz-Olivier C. Transmembrane proteins – Mining the cattle tick transcriptome. *Ticks and Tick-borne Diseases*. 2015;6(6):695–710.
 58. Surakasi VP, Mohamed AA, Kim Y. RNA interference of $\beta 1$ integrin subunit impairs development and immune responses of the beet armyworm, *Spodoptera exigua*. *J Insect Physiol*. 2011;57(11):1537–44.
 59. Güiza J, Barría I, Sáez JC, Vega JL. Innexins: expression, regulation, and functions. *Front Physiol*. 2018;9:1414.
 60. Phelan P, Starich TA. Innexins get into the gap. *BioEssays*. 2001;23(5):388–96.
 61. Giachetto PF, Cunha RC, Nhani A, Garcia MV, Ferro JA, Andreotti R. Gene expression in the salivary gland of *Rhipicephalus (Boophilus) microplus* fed on tick-susceptible and tick-resistant hosts. *Front Cell Infect Microbiol*. 2020;9:477.
 62. Aljamali M, Hern L, Kupfer D, Downard S, So S, Roe B, Sauer J, Essenberg R. Transcriptome analysis of the salivary glands of the female tick *Amblyomma americanum* (Acari: Ixodidae). *Insect Mol Biol*. 2009;18(2):129–54.
 63. LaFoya B, Munroe JA, Miyamoto A, Detweiler MA, Crow JJ, Gazdik T, Albig AR. Beyond the matrix: The many non-ECM ligands for integrins. *Int J Mol Sci*. 2018;19(2):449.
 64. Park Y, Ahn S-J, Vogel H, Kim Y. Integrin β subunit and its RNA interference in immune and developmental processes of the Oriental tobacco budworm. *Helicoverpa assulta Developmental & Comparative Immunology*. 2014;47(1):59–67.
 65. Knorr S, Reissert-Oppermann S, Tomás-Cortázar J, Barriales D, Azkargorta M, Iloro I, Elortza F, Pinecki-Socias S, Anguita J, Hovius JW, et al. Identification and characterization of immunodominant proteins from tick tissue extracts inducing a protective immune response against *Ixodes ricinus* in cattle. *Vaccines (Basel)*. 2021;9(6):636.
 66. Ball A, Campbell EM, Jacob J, Hoppler S, Bowman AS. Identification, functional characterization and expression patterns of a water-specific aquaporin in the brown dog tick, *Rhipicephalus sanguineus*. *Insect Biochem Mol Biol*. 2009;39(2):105–12.
 67. Campbell EM, Burdin M, Hoppler S, Bowman AS. Role of an aquaporin in the sheep tick *Ixodes ricinus*: assessment as a potential control target. *Int J Parasitol*. 2010;40(1):15–23.
 68. Hussein HE, Scoles GA, Ueti MW, Suarez CE, Adham FK, Guerrero FD, Bastos RG. Targeted silencing of the Aquaporin 2 gene of *Rhipicephalus (Boophilus) microplus* reduces tick fitness. *Parasit Vectors*. 2015;8:1–12.
 69. Guerrero FD, Andreotti R, Bendele KG, Cunha RC, Miller RJ, Yeater K, Pérez de León AA: *Rhipicephalus (Boophilus) microplus* aquaporin as an effective vaccine antigen to protect against cattle tick infestations. *Parasit Vectors*. 2014;7:1–12.
 70. Contreras M, de la Fuente J. Control of infestations by *Ixodes ricinus* tick larvae in rabbits vaccinated with aquaporin recombinant antigens. *Vaccine*. 2017;35(9):1323–8.
 71. Scoles GA, Hussein HE, Olds CL, Mason KL, Davis SK. Vaccination of cattle with synthetic peptides corresponding to predicted extracellular domains of *Rhipicephalus (Boophilus) microplus* aquaporin 2 reduced the number of ticks feeding to repletion. *Parasit Vectors*. 2022;15(1):49.
 72. Yang WJ, Xu KK, Yan Y, Li C, Jin DC. Role of chitin deacetylase 1 in the molting and metamorphosis of the cigarette beetle *Lasioderma serricorne*. *Int J Mol Sci*. 2020;21(7):2449.
 73. Kariu T, Smith A, Yang X, Pal U. A chitin deacetylase-like protein is a predominant constituent of tick peritrophic membrane that influences the persistence of Lyme disease pathogens within the vector. *PLoS ONE*. 2013;8(10): e78376.
 74. Yang X, Koči J, Smith AA, Zhuang X, Sharma K, Dutta S, Rana VS, Kitsou C, Yas OB, Mongodin EF, et al. A novel tick protein supports integrity of gut peritrophic matrix impacting existence of gut microbiome and Lyme disease pathogens. *Cell Microbiol*. 2021;23(2): e13275.
 75. Tajiri R, Ogawa N, Fujiwara H, Kojima T. Mechanical control of whole body shape by a single cuticular protein obstructor-E in *Drosophila melanogaster*. *PLoS Genet*. 2017;13(1): e1006548.
 76. Bansal D, Bhatti HS, Sehgal R. Role of cholesterol in parasitic infections. *Lipids Health Dis*. 2005;4(1):10.
 77. McCauliff LA, Xu Z, Li R, Kodukula S, Ko DC, Scott MP, Kahn PC, Storch J. Multiple surface regions on the Niemann-Pick C2 protein facilitate intracellular cholesterol transport. *J Biol Chem*. 2015;290(45):27321–31.
 78. Renthal R, Manghnani L, Bernal S, Qu Y, Griffith WP, Lohmeyer K, Guerrero FD, Borges LM, Pérez de León A: The chemosensory appendage proteome of *Amblyomma americanum* (Acari: Ixodidae) reveals putative odorant-binding and other chemoreception-related proteins. *Insect Science*. 2017;24(5):730–42.
 79. Zhu J, Guo M, Ban L, Song L-M, Liu Y, Pelosi P, Wang G. Niemann-Pick C2 proteins: a new function for an old family. *Front Physiol*. 2018;9:52.
 80. Liang D, Chen H, An L, Li Y, Zhao P, Upadhyay A, Hansson BS, Zhao J, Han Q. Molecular identification and functional analysis of Niemann-Pick type C2 proteins, carriers for semiochemicals and other hydrophobic compounds in the brown dog tick, *Rhipicephalus linnaei* Pesticide Biochemistry and Physiology. 2023;193: 105451.
 81. Cui Y, Wang J, Liu Q, Li D, Zhang W, Liu X, Wang J, Song X, Yao F, Wu H. Identification and expression of potential olfactory-related genes related to Niemann-Pick C2 protein and ionotropic receptors in *Haemaphysalis longicornis*. *Exp Appl Acarol*. 2022;87(4):337–50.
 82. Iovinella I, Ban L, Song L, Pelosi P, Dani FR. Proteomic analysis of castor bean tick *Ixodes ricinus*: a focus on chemosensory organs. *Insect Biochem Mol Biol*. 2016;78:58–68.
 83. Chang T-Y, Li B-L, Chang CC, Urano Y. Acyl-coenzyme A: cholesterol acyltransferases. *American Journal of Physiology-Endocrinology and Metabolism*. 2009;297(1):E1–9.
 84. van Zyl WA, Stutzer C, Olivier NA, Maritz-Olivier C. Comparative microarray analyses of adult female midgut tissues from feeding *Rhipicephalus* species. *Ticks and Tick-borne Diseases*. 2015;6(1):84–90.
 85. Maritz-Olivier C, van Zyl W, Stutzer C. A systematic, functional genomics, and reverse vaccinology approach to the identification of vaccine candidates in the cattle tick, *Rhipicephalus microplus* *Ticks and Tick-borne Diseases*. 2012;3(3):179–87.
 86. Termine JD, Belcourt AB, Conn KM, Kleinman H. Mineral and collagen-binding proteins of fetal calf bone. *J Biol Chem*. 1981;256(20):10403–8.
 87. Irls P, Ramos S, Piulachs M-D. SPARC preserves follicular epithelium integrity in insect ovaries. *Dev Biol*. 2017;422(2):105–14.
 88. Martinek N, Zou R, Berg M, Sodek J, Ringuette M. Evolutionary conservation and association of SPARC with the basal lamina in *Drosophila*. *Dev Genes Evol*. 2002;212:124–33.
 89. Volk T, Wang S, Rotstein B, Paululat A. Matricellular proteins in development: perspectives from the *Drosophila* heart. *Matrix Biol*. 2014;37:162–6.
 90. Oliveira CJ, Anatriello E, de Miranda-Santos IK, Francischetti IM, Sá-Nunes A, Ferreira BR, Ribeiro JMC. Proteome of *Rhipicephalus sanguineus* tick saliva induced by the secretagogues pilocarpine and dopamine. *Ticks and Tick-borne Diseases*. 2013;4(6):469–77.
 91. Zhou J, Liao M, Hatta T, Tanaka M, Xuan X, Fujisaki K. Identification of a follistatin-related protein from the tick *Haemaphysalis longicornis* and its effect on tick oviposition. *Gene*. 2006;372:191–8.
 92. Baumgartner S, Littleton JT, Broadie K, Bhat MA, Harbecke R, Lengyel JA, Chiquet-Ehrismann R, Prokop A, Bellen HJ. A *Drosophila* neurexin is required for septate junction and blood-nerve barrier formation and function. *Cell*. 1996;87(6):1059–68.
 93. Reissner C, Runkel F, Missler M. Neurexins. *Genome Biol*. 2013;14(9):213.
 94. Scott LO: Regulation of Oligodendrocyte differentiation through Neurotrophin-Neurexin signalling (Master's thesis, University of Calgary, Calgary, Canada). 2013. Retrieved from <https://prism.ucalgary.ca>. <https://doi.org/10.11575/PRISM/27587>.
 95. Kneeland R. E. LSB, Folsom T.D. The role of glial pathology in autism. In: *Astrocytes*. 1st edn: CRC Press; 2012: p. 18.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.