Manuscript Details

Manuscript number TTBDIS_2018_79_R1

Title De novo assembly and analysis of midgut transcriptome of the argasid tick *Ornithodoros erraticus* and identification of genes differentially expressed after blood feeding

Article type Research Paper

Abstract

Ticks are hematophagous vectors of great medical and veterinary importance because they transmit numerous pathogenic microorganisms to humans and animals. The argasid *Ornithodoros erraticus* is the main vector of tick-borne human relapsing fever and African swine fever in the Mediterranean Basin. Tick enterocytes express bioactive molecules that perform key functions in blood digestion, feeding, toxic waste processing and pathogen transmission. To explore new strategies for tick control, in this work we have obtained and compared the midgut transcriptomes of *O. erraticus* female ticks before and after a blood meal and identified the genes whose expression is differentially regulated after feeding. The transcript sequences were annotated, functionally and structurally characterised and their expression levels compared between both physiological conditions (unfed females and fed females at 2 days post-engorgement). Up to 29,025 transcripts were assembled, and 9,290 of them corresponded to differentially expressed genes (DEGs) after feeding. Of these, 4,656 genes were upregulated and nearly the same number of genes was downregulated in fed females compared to unfed females. BLASTN and BLASTX analyses of the 29,025 transcripts allowed the annotation of 9,072 transcripts/proteins. Among them, the most numerous were those with catalytic and binding activities and those involved in diverse metabolic pathways and cellular processes. The analyses of functional groups of upregulated DEGs potentially related to the digestion of proteins, carbohydrates and lipids, and the genes involved in the defence response and response to oxidative stress, confirm that these processes are narrowly regulated in ticks, highlighting their complexity and importance in tick biology. The expression patterns of six genes throughout the blood digestion period revealed significant differences between these patterns, strongly suggesting that the transcriptome composition is highly dynamic and subjected to important variation along the trophogonic cycle. This may guide future studies aimed at improving the understanding of the molecular physiology of tick digestion and digestion-related processes. The current work provides a more robust and comprehensive understanding of the argasid tick digestive system.

Keywords Argasid, *Ornithodoros erraticus*, transcriptome, RNA-seq, midgut, blood digestion.

Corresponding Author Ana Oleaga

Corresponding Author's Institution IRNASA (CSIC)

Order of Authors Ana Oleaga, Prosper Obolo-Mvoulouga, Raul Manzano Roman, Ricardo Pérez-Sánchez

Submission Files Included in this PDF

File Name [File Type]

Cover letter.docx [Cover Letter]

Response to reviewers.docx [Response to Reviewers]

ManuscriptR1.docx [Manuscript File]

Figure 1.TIF [Figure]

Figure 2.TIF [Figure]

Figure 3.TIF [Figure]

Figure 4.TIF [Figure]

Figure 5.TIF [Figure]

Figure 6.TIF [Figure]

Figure 7.TIF [Figure]

Figure 8.TIF [Figure]

Table 2.docx [Table]

Table 3.docx [Table]

Table 4.docx [Table]

Table 5.docx [Table]

Table 6.docx [Table]

Table 1_R1.docx [Table]

Supplementary Figure 1.tif [e-Component]

Supplementary Figure 2.tif [e-Component]

Submission Files Not Included in this PDF

File Name [File Type]

Supplementary Table 1.xlsx [e-Component]

Supplementary Table 2_R1.xlsx [e-Component]

Supplementary Table 3.xlsx [e-Component]

Supplementary Table 4.xlsx [e-Component]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

***De novo* assembly and analysis of midgut transcriptome of the argasid tick *Ornithodoros erraticus* and identification of genes differentially expressed after blood feeding**

Ana Oleaga*¹, Prosper Obolo-Mvoulouga¹, Raúl Manzano-Román¹, Ricardo Pérez-Sánchez¹

¹ Parasitology Laboratory, Institute of Natural Resources and Agrobiology (IRNASA, CSIC). Cordel de Merinas, 40-52, 37008 Salamanca, Spain.

Corresponding author:

Dr. Ana Oleaga

Parasitology Laboratory

Instituto de Recursos Naturales y Agrobiología (IRNASA, CSIC)

Cordel de Merinas, 40-52

37008 Salamanca, Spain

Tel.: +34 923219606; fax: +34 923219609

E-mail address: ana.oleaga@irnasa.csic.es

List all author emails:

Ana Oleaga: ana.oleaga@irnasa.csic.es

Prosper Obolo-Mvoulouga: obolo2nd@gmail.com

Raúl Manzano-Román: raul_romanz@yahoo.es

Ricardo Pérez-Sánchez: ricardo.perez@irnasa.csic.es

Abstract

Ticks are hematophagous vectors of great medical and veterinary importance because they transmit numerous pathogenic microorganisms to humans and animals. The argasid *Ornithodoros erraticus* is the main vector of tick-borne human relapsing fever and African swine fever in the Mediterranean Basin. Tick enterocytes express bioactive molecules that perform key functions in blood digestion, feeding, toxic waste processing and pathogen transmission. To explore new strategies for tick control, in this work we have obtained and compared the midgut transcriptomes of *O. erraticus* female ticks before and after a blood meal and identified the genes whose expression is differentially regulated after feeding. The transcript sequences were annotated, functionally and structurally characterised and their expression levels compared between both physiological conditions (unfed females and fed females at 2 days post-engorgement). Up to 29,025 transcripts were assembled, and 9,290 of them corresponded to differentially expressed genes (DEGs) after feeding. Of these, 4,656 genes were upregulated and nearly the same number of genes was downregulated in fed females compared to unfed females. BLASTN and BLASTX analyses of the 29,025 transcripts allowed the annotation of 9,072 transcripts/proteins. Among them, the most numerous were those with catalytic and binding activities and those involved in diverse metabolic pathways and cellular processes. The analyses of functional groups of upregulated DEGs potentially related to the digestion of proteins, carbohydrates and lipids, and the genes involved in the defence response and response to oxidative stress, confirm that these processes are narrowly regulated in ticks, highlighting their complexity and importance in tick biology. The expression patterns of six genes throughout the blood digestion period revealed significant differences between these

119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177

51 patterns, strongly suggesting that the transcriptome composition is highly dynamic and
52 subjected to important variation along the trophogonic cycle. This may guide future
53 studies aimed at improving the understanding of the molecular physiology of tick
54 digestion and digestion-related processes. The current work provides a more robust and
55 comprehensive understanding of the argasid tick digestive system.

56
57 *Keywords:* Argasid, *Ornithodoros erraticus*, transcriptome, RNA-seq, midgut, blood
58 digestion.

1. Introduction

Ticks are hematophagous vectors of great medical and veterinary importance. There are two main families of ticks, Ixodidae and Argasidae, which differ in many morphological, biological and ecological properties. Typically, ixodid ticks are exophilic organisms that live in open environments on the soil or vegetation waiting for a suitable animal host. After accessing the host, they slowly feed for several days and, once they are engorged, they return to the ground to moult (larvae and nymphs) or lay eggs and die (females). In contrast, argasid ticks are endophilic organisms that live inside the burrows and nests of their hosts, including human dwellings and domestic animal premises, where they enjoy a more regular access to their host's blood. They feed quickly, usually in less than one hour, and after engorgement, they moult and reproduce inside their shelters. Adult argasids can complete up to 10 trophogonic cycles and are able to survive for several years without feeding (Oleaga-Pérez et al., 1990; Sonenshine and Roe, 2014).

Ticks transmit a huge range of pathogenic microorganisms to humans and animals, being second after mosquitoes as vectors of arthropod-borne diseases (de la Fuente et al., 2008; Jones et al., 2008). Among argasids, several species of the genus *Ornithodoros* are vectors of the causative agents of severe diseases, such as the tick-borne human relapsing fever (HRF) and the African swine fever (ASF). *Ornithodoros erraticus*, reported from the Iberian Peninsula, north and west Africa and western Asia, is the main vector of the causative agents of these diseases in the Mediterranean Basin (Cutler, 2010; Boinas et al., 2014; Sánchez-Vizcaíno et al., 2015). Moreover, *O. erraticus* is the type species of the '*O. erraticus* complex', which includes several species that are distributed through the Middle East, the Caucasus and the Russian Federation, where they transmit different species of HRF-causing *Borrelia* (Rebaudet

237
238
239 85 and Parola, 2006) and where ASF has recently spread out of control (Gallardo et al.,
240
241 86 2015). Although not hitherto demonstrated, if these species belonging to the *O.*
242
243 87 *erraticus* complex were vectors of ASF virus, this would have an enormous impact on
244
245
246 88 the spread and long-time persistence of ASF in that area.

247
248 89 The prevention and control of human and animal tick-borne diseases would
249
250 90 require eliminating tick populations. However, complete elimination of *Ornithodoros*
251
252 91 populations may be difficult, as the ticks are long-lived organisms, can survive long
253
254 92 periods without a blood meal, and hide deeply in cracks and fissures that chemical
255
256 93 control agents fail to reach (Boinas et al., 2014). Given that chemical acaricides are
257
258 94 inefficient against these argasids, other measures for tick control were sought, and anti-
259
260 95 tick vaccines have emerged as a cost-effective and environmentally sustainable method
261
262 96 for the control of tick infestations and tick-borne diseases (Willadsen, 2008; de la
263
264 97 Fuente et al., 2016; Šmit and Postma, 2016). With this purpose, our team has done
265
266 98 significant work in the development of an anti-*O. erraticus* vaccine, testing the
267
268 99 protective effects of two types of antigens: salivary antigens and concealed antigens
269
270 100 from the tick midgut (Díaz-Martín et al., 2015). Salivary antigens provided rather
271
272 101 variable protective responses in inhibiting tick feeding and reproduction, reaching up to
273
274 102 50–70% vaccine efficacy (Astigarraga et al., 1995; Díaz-Martín et al., 2015).

275
276
277 103 On the other hand, midgut antigens induced protective responses that reduced
278
279 104 the feeding and reproduction performance of females by up to 50%, and the survival
280
281 105 rate of the nymphs by up to 80%. Protection was a consequence of the complement-
282
283 106 mediated damage to the tick midgut wall (Manzano-Román et al., 2006), in a similar
284
285 107 way to that observed for the tick vaccines directed towards the ixodid tick
286
287 108 *Rhipicephalus microplus* that are based on the intestinal antigen Bm86 (Popara et al.,
288
289 109 2013).

The antigens responsible for this protection were membrane proteins – supposedly located on the luminal side of the enterocytes – whose expression increased upon blood ingestion, peaking between 24 and 72 h post feeding (Manzano-Román et al., 2006, 2007). These proteins remain unidentified, but the above results confirmed the protective potential of the antigens from *O. erraticus* midgut, establishing this organ as a suitable source of antigens for argasid tick vaccine development, in agreement with that suggested by other authors for ixodid ticks (Guerrero et al., 2012; Maritz-Olivier et al., 2012).

Unlike other hematophagous arthropods, the digestion of blood in ticks is intracellular and takes place within the digestive lysosomal vesicles of intestinal cells. The tick midgut is an essential organ for tick survival since it is where the host blood digestion and the absorption of the released nutrients take place (Sojka et al., 2013), and together with tick saliva, the midgut forms part of the host-tick-pathogen interface. Hence, the tick midgut expresses numerous proteins that play essential functions in the digestion-related processes and in the infection and transmission of blood-borne pathogens (Kocan et al., 2004).

In ixodids, the blood digestion process, the multi-enzymatic digestive machinery involved and the iron homeostasis and metabolism have been studied in depth, and they are currently fairly well established (Sojka et al., 2013; Sonenshine and Anderson, 2014; Hajdusek et al., 2016). Extensive characterisation of tick midguts has been carried out in various ixodid tick species at both transcript and protein levels (Rachinsky et al., 2008; Kongsuwan et al., 2010; Schwarz et al., 2014; Cramaro et al., 2015; Kotsyfakis et al., 2015; Perner et al., 2016a; Xu et al., 2016). In argasids, this kind of information has been quite scant until the recent analysis and publication of the midgut proteomes of *O. erraticus* and *Ornithodoros moubata*, and the midgut transcriptomes of *O. moubata* and

135 *Ornithodoros mimon* (Oleaga et al., 2015; Landulfo et al., 2017; Oleaga et al., 2017a,
136 2017b).

137 The initial study of the midgut proteome of *O. erraticus* provided the first data
138 hitherto about the argasid machinery involved in blood digestion, and showed the
139 proteomic changes that take place in response to feeding and blood digestion (Oleaga et
140 al., 2015).

141 More recently, the extensive characterisation of the *O. moubata* mialome
142 (transcriptome and proteome) from unfed and fed females (Oleaga et al., 2017a, 2017b)
143 has provided a wealth of novel information on the argasid genes and proteins involved
144 in blood digestion, nutrient metabolism, intracellular transport mechanisms and
145 defensive, antioxidant and detoxifying responses triggered by blood ingestion.
146 Moreover, it was also observed that both the argasid and the ixodid haemoglobinolytic
147 enzyme systems are fairly similar despite the different feeding strategies used by these
148 two tick families.

149 Regarding *O. erraticus*, the aim of this work was to complete the molecular
150 information previously obtained from the midgut proteome by Oleaga et al. (2015) and
151 provide a more robust and comprehensive viewpoint of this biological system. With this
152 in mind, in the current study, Illumina RNA sequencing (RNA-seq) was performed in
153 order to obtain the midgut transcriptomes of *O. erraticus* females before and after a
154 blood meal, and to identify the genes whose expression is differentially regulated after
155 feeding. We have applied an experimental approach similar to that used previously by
156 Oleaga et al. (2017a) for *O. moubata*, and have performed a similar *in silico* data
157 analysis, including the more in-depth scrutiny of some selected groups of upregulated
158 genes that were predicted to participate in the blood-meal digestion processes, such as

the genes associated with the digestion and transport of proteins, carbohydrates and lipids, and those involved in defence responses and responses to oxidative stress.

2. Material and Methods

2.1. Ticks and midgut collection.

The *O. erraticus* specimens used came from a laboratory colony that was established from field specimens captured in Salamanca province (western Spain) in 1990. Ticks were routinely fed on rabbits and kept in a culture chamber at 28 °C, 85% relative humidity, and a 12 h light-dark cycle. The laboratory animals were maintained and manipulated according to the rules of the Ethical and Animal Welfare Committee of Institute of Natural Resources and Agrobiology (IRNASA, CSIC) where the experiments were conducted, following the corresponding Spanish rules and regulations.

The *O. erraticus* midgut tissue samples used for this study were obtained and processed as described by Oleaga et al. (2017a) for *O. moubata*. Briefly, tick midguts were dissected from unfed females and from fed females 48 h post feeding. The unfed females had remained without feeding for three months. The midgut tissues, free of any other contaminant tissue, were collected in batches of 40 midguts/batch and preserved in RNAlater solution (Ambion) until the RNA extraction was performed. For each physiological condition, two biological replicates were processed: two from unfed females (UF-1 and UF-2) and two from fed females (48h-1 and 48h-2).

2.2. RNA extraction and Illumina sequencing.

The four midgut tissue samples (UF-1 and UF-2; 48h-1 and 48h-2) were processed equally, following the procedure described by Oleaga et al. (2017a). Briefly, the tissues were lysed and homogenised using the TissueLyser II (Qiagen) and their total RNA was extracted using the Mini Kit RNeasy (Qiagen) and RNAprotect® (Qiagen) added for stabilisation of nucleic acids. Contaminant DNA was eliminated by treating the samples with Turbo DNase-free (Ambion), and the quality and integrity of the RNA preparations were tested in the Bioanalyzer 2100 (Agilent).

After that, the mRNA was purified using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (Biolabs) and the corresponding RNA libraries were produced using the NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (Biolabs) following standard protocols. Libraries were combined in equimolar amounts and sequenced using a Genome Analyzer IIx sequencer (Illumina). A total of four cDNA libraries were constructed and run under a pair-end 2×125 bp protocol for *de novo* sequencing.

2.3. Bioinformatic analysis and tools used for midgut transcriptome processing and characterisation.

The assembly of the sequences and their bioinformatic analysis were carried out as described previously Oleaga et al. (2017a). First, the sequences of adaptors and the nucleotides having a PHRED quality index below 30 were removed using the software Trimgalore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Next, sequences larger than 300 nucleotides sharing overlapping sequences of at least 25 bases were grouped using the software Oases contig assembly (Schulz et al., 2012) and the Trinity package (Grabherr et al., 2011). Then, reads from each library were aligned with non-

redundant transcriptomic data using Bowtie 2 (Langmead and Salzberg, 2012). Finally, the relative abundance of each transcript was analysed and normalised with software RSEM (Li and Dewey, 2011) into fragments per kilobase of transcript per million mapped reads (FPKM).

Once assembled, the midgut transcriptome of *O. erraticus* was characterised using the same *in silico* approach and tools used by Oleaga et al. (2017a). Transcripts were annotated by searching homologous sequences ($E < 10^{-05}$) by BLASTx and BLASTn in the protein and nucleotide databases of the National Center for Biotechnology Information (NCBI) (<https://blast.ncbi.nlm.nih.gov>), UniProt knowledgebase (UniProtKB) (<http://www.uniprot.org>) and *Ixodes scapularis* genome (http://metazoa.ensembl.org/Ixodes_scapularis/Info/Index). Coding sequences in each transcript were translated into amino acids with Transdecoder (<https://transdecoder.github.io/>) and the resulting peptide sequences were analysed as follows: (i) functional annotation of biological process, molecular function, cellular component and protein family were performed according to Gene Ontology (<http://www.geneontology.org>), Uniprot (<http://www.uniprot.org>) and Pfam (<http://pfam.xfam.org>; Finn et al., 2016) databases; (ii) the annotated genes were then classified by Protein Class using the Panther Classification System (<http://www.pantherdb.org>); (iii) the active metabolic pathways were predicted using the eggNOG database (<http://eggnogdb.embl.de/#/app/emapper>; Huerta-Cepas et al., 2017); and (iv) the presence of transmembrane domains, glycosylphosphatidylinositol (GPI) anchors and signal peptide were predicted using the tools TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>), kohgpi-1.5 (<http://gpi.unibe.ch/>) (Fankhauser and Mäser, 2005) and signalP-4.1 (<http://www.cbs.dtu.dk/services/SignalP/>) (Petersen et al., 2011).

2.4. Differential expression analysis of the midgut transcriptome after blood feeding.

The differentially expressed genes (DEGs) were detected by comparing their average transcript abundance in fed females compared to unfed females (basal condition). Comparison was performed using the DEseq algorithm, which is based on the negative binomial distribution (Love et al., 2014).

Genes were taken as differentially expressed when the abundance of their transcripts varied more than two-fold between both physiological conditions tested (fed *versus* unfed) (fold-change > 2 or < -2). Statistical significance of the change in expression level was estimated using the Bonferroni method, taking a false discovery rate lower than 0.05 (FDR < 0.05).

2.5. Validation of RNA-Seq analysis by reverse transcription-quantitative PCR (RT-qPCR)

In order to confirm and validate the RNA-seq results, the expression levels of six particular genes were assessed by quantitative real-time PCR (RT-qPCR). These genes were selected among the genes upregulated upon feeding, namely, organic cation/carnitine transporter, integrin alpha 9, hydroxyprostaglandin dehydrogenase, chitinase, Niemann–Pick type C1 domain-containing protein, and glucose transporter. Additionally, we extended the time point of expression-level analysis by inspecting the expression profiles of these genes up to 8 days after feeding: specifically, in unfed females (time 0) and in fed females at 24, 48 and 72 h and at 5 and 8 days post feeding.

RNA purification and the corresponding cDNA synthesis were performed, for each time point sampled, in duplicate from midgut tissue samples taken from different

batches of ticks. Total RNA was extracted with Mini Kit RNeasy (Qiagen), and reverse transcription (RT) was performed using the 1st Strand cDNA Synthesis kit (Roche) and the oligo-p(dT)₁₅ primer.

For the amplification of the target sequences, between 100 and 150 nucleotides long, specific primers with a T_m close to 60 °C were designed using the Primer3Plus software (Untergasser et al., 2007). The reaction mixtures, of 20 µl total volume, included the Power SYBR Green PCR Master Mix (Applied Biosystems), the cDNA and the gene-specific primers. Each reaction was performed in triplicate in the ABI PRISM 7900HT Fast Real-Time PCR System (Applied Biosystems). The primer sequences and the RT-qPCR conditions used are indicated in the Supplementary Table 1.

The mRNA levels of the analysed genes were normalised using the gene of *O. erraticus* actin, whose sequence was also obtained in the current study (ci000018243; GFWV000000000), and the 16S rRNA gene of *O. sonrai* (GenBank: DQ234726), which has already been used by Manzano-Román et al. (2012) as a suitable reference gene for *O. erraticus*. Differential gene expression between unfed females (time 0) and the different time-points after feeding was calculated using the $\Delta\Delta C_t$ method (Cikos et al., 2007).

3. Results and discussion

3.1. Illumina sequencing of the midgut transcriptome

In order to obtain a comprehensive understanding of the midgut genes of *O. erraticus* that were affected by changes after a blood meal, biological duplicates of RNA samples from midgut were prepared from unfed and engorged female ticks and

sequenced using the Illumina sequencing platform. Since this study focuses mainly on identifying the genes involved in blood digestion, the samples to be analysed were obtained from unfed female ticks (basal condition) and from engorged female ticks at 48 h post feeding (fed) because at this time-point argasid ticks start the phase of intense blood digestion (Sonenshine and Anderson, 2014).

The replicate samples of unfed female midguts (UF-1 and UF-2) provided 55,970,317 bp and 57,606,191 bp, and the two samples taken at 48 h post feeding (48h-1 and 48h-2) provided 68,013,513 bp and 53,831,445 bp, respectively. After filtering out adaptors, poor-quality data and ambiguous reads, 50,973,679 bp (UF-1), 52,383,093 bp (UF-2), 61,619,429 bp (48h-1) and 49,194,262 (48h-2) bp clear reads were obtained, which accounted for between 90.60 and 91.39% of the initial raw reads.

The assembly of the reads resulted in 29,025 transcripts, 172 of which were considered contaminants and excluded from further characterisation and analysis. They consisted of 15 transcripts annotated as rabbit proteins, which were likely to have come from the ingested blood, and transcripts annotated as proteins from fungi (117), bacteria (28), viruses (10) and archaea (2), which may belong to the tick intestinal microbiome (not shown).

The coding sequences were deposited in DDBJ/ENA/GenBank under accession number GFWV000000000 as a Transcriptome Shotgun Assembly project (BioProject: PRJNA401392). The version described in this paper is the first version, GFWV01000000.

3.2. Functional annotation and characterisation of the transcripts.

Database searching allowed the annotation of up to 9,072 transcripts, covering 31.26% of the sequenced transcriptome (Supplementary Table 2). The remaining 19,953

transcripts (68.74%) did not produce significant homology matches within any sequence in the searched databases, suggesting that many novel sequences were obtained (Xu et al., 2016). These results were in line with our previous study on the *O. moubata* midgut transcriptome and with similar studies in other tick species (Gibson et al., 2013; Landulfo et al., 2017; Oleaga et al., 2017a). According to Tautz and Domazet-Lošo (2011) and Mans et al. (2016), some possible reasons accounting for the low ratio of homologues found might be some artefacts of the assembly process, low assembly quality, taxonomic isolation of the target species, and emergence of new genes through duplication and/or transposition mechanisms. Additional causes might also be the scarcity of this type of study (and data) on these soft ticks, or even particular features of their genetic composition that may not be present in other animals (Landulfo et al., 2017).

The annotated sequences were next functionally characterised using the information available in Gene Ontology, Protein families (Pfam) and Protein Class (Panther Classification System) databases. Initially, genes were classified according to their molecular function and biological process using the GO terms from the Gene Ontology (GO) database (Figs. 1 and 2, Supplementary Table 2).

Fig. 1 shows the classification of genes into categories according to their molecular function. Only the categories represented by more than 25 genes are depicted, each including the number and ratio of the genes assigned. In this way, the most abundantly annotated functions were binding activity (n = 3187), catalytic activity (n = 3011) and genes with 'unknown function' (n = 1796), which coincided with those observed in the midgut transcriptomes of the argasid ticks *O. moubata* and *O. mimon* (Landulfo et al., 2017; Oleaga et al., 2017a) as well as in the midgut transcriptomes of several ixodid tick species (Heekin et al., 2013; Villar et al., 2014; Cramaro et al., 2015;

Xu et al., 2016). The remaining annotated functions were as follows: transporter activity (n = 389), molecular function regulator (n = 182), structural molecule activity (n = 170), sequence-specific DNA binding (n = 129), signal transducer activity (n = 114), receptor activity (n = 87) transcription factor activity (n = 73), antioxidant activity (n = 32) and electron carrier activity (n = 25).

Similarly, gene classification according to biological process resulted in up to 14 categories represented by more than 25 genes (Fig. 2). Notably, the genes without biological process assigned (unknown) were the most abundant (n = 3527). The following, more abundant, categories corresponded to proteins involved in cellular processes (n = 2543), metabolic processes (n = 1934), single organism processes (n = 1220) and biological regulation (n = 1069). The remaining functional annotations were less numerous and were distributed into nine categories, the most abundant being involved in localisation processes (n = 574), cellular organisation or biogenesis (n = 528), response to stimulus (n = 294) and developmental process (n = 192) (Fig. 2). Supplementary Table 2 shows the GO terms for molecular function, biological process and cellular compartment assigned to all the genes annotated.

Next, in order to characterise more deeply the midgut transcriptome, the annotated sequences were analysed in the Pfam database of protein families and classified by Protein Class according to the Panther database.

The analysis of the annotated sequences in the Pfam database identified 12,248 different protein domains (Supplementary Table 3). Table 1 lists the top 30 protein domain families found in the transcriptome, indicating the number of those families that were differentially expressed (up- or downregulated) upon feeding (see section 3.3.). The more frequent protein families included up to 11 different types of zinc finger domains (zf-C2H2, zf-C2H2_4, zf-H2C2_2, zf-C2H2_jaz, zf-met, zf-C2H2_2, zf-

H2C2_5, zf-C3HC4, zf-C3HC4_2, zf-BED, zf-C3HC4_3) and up to five Ankyrin repeat domains (Ank, Ank_2, Ank_3, Ank_4, Ank_5) (Table 1). These two types of protein motifs are particularly common in nature. Zinc finger proteins shown very diverse structures and functions and are classified into different structural families. These proteins may be involved in numerous functions, such as DNA recognition, RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, several protein interactions and lipid binding (Laity et al., 2001; Gamsjaeger et al., 2007). Ankyrin repeat domains are involved in specific protein–protein interactions and in many cellular processes (Sedgwick et al., 1999).

Other interesting protein domains, already identified in the midgut transcriptomes of *O. moubata* and several ixodid ticks, are the ABC and Major Facilitator Superfamily (MFS) transporters (Table 1) (Lara et al., 2015; Perner et al., 2016a; Oleaga et al., 2017a). These are two protein super-families involved in the uptake and transport across membranes of a wide variety of substrates, including ions and macromolecules. Specifically, at least in ixodids, the ABC transporters mediate the transport of the haem group in the digestive cells and participate in pesticide detoxification (Lara et al., 2015).

Classification of the transcriptome into Protein Classes allowed the categorising of up to 2018 genes/proteins into 24 categories (Fig. 3). The Protein Classes containing more than 100 hits were the genes with nucleic acid binding activity (n = 357; 17.7%), hydrolase (n = 295; 14.6%), transferase (n = 227; 11.2%), transporter (n = 165; 8.2%), enzyme modulator (n = 163; 8.1%), transcription factor (n = 139; 6.9%) and oxidoreductase (n = 133; 6.6%). In general terms, this distribution profile was very similar to that found for the midgut proteome of *O. erraticus* (Oleaga et al., 2015).

3.3. Genes that were differentially expressed in the tick midgut after feeding.

We compared the gene expression levels, as measured in FPKM, between both physiological conditions (fed *versus* unfed), and assumed that the expression level of a particular gene increased or decreased significantly in fed ticks when its fold-change with respect unfed ticks was either higher than 2 (>2) or lower than -2 (<-2).

This analysis showed that there were up to 9290 DEGs in fed female ticks compared to the unfed females: 4,656 transcripts were upregulated and nearly the same gene numbers (4,634 transcripts) were downregulated (Supplementary Table 2). Fig. 4 represents the expression level (ordinate axis) of each transcript in respect of the change in their level of expression ($\log_2$ fold-change, abscissa axis). The fold-change distribution of both the upregulated and downregulated genes is shown in Supplementary Fig. 1.

In order to understand the functions of all these DEGs, they were scrutinised in the databases eggNOG for metabolic pathways, Pfam for protein domains, and GO for molecular function and biological process.

After searching the eggNOG database, we detected up to 172 different active pathways, which were grouped into 26 major classes of biological processes (Supplementary Table 4, Fig. 5). The biological pathways and cellular processes that seem to be more active during blood digestion (i.e., those with a greater number of sequences involved) were signal transduction (305 sequences), amino acid metabolism (230 sequences), carbohydrate metabolism (211 sequences), folding, sorting and degradation (193 sequences), transport and catabolism (192 sequences), lipid metabolism (179 sequences), and cell growth and death (177 sequences).

As can be observed in Fig. 5, in virtually all metabolic pathways and cellular processes, the number of upregulated genes is significantly higher than that of

downregulated ones. Only three exceptions were observed: the genes involved in signal transduction, translation and transcription pathways, in which the downregulated genes were more numerous (Fig. 5, Supplementary Table 4).

Searching the Pfam database identified 5912 protein families in the upregulated genes and 5407 protein families in the downregulated genes (Supplementary Table 3).

Among the most abundant protein domains identified in upregulated genes were the following: the ABC transporter, MFS, the Inclusion Membrane Protein (IncA) family, WD40 and Pkinase domain. All of them were more numerous in the upregulated DEGs than in the downregulated ones (Supplementary Table 3 and Table 1). As commented previously, during the digestion process, the ABC and MFS transporters would be likely to be involved in the transport of nutrients and solutes from the midgut lumen to the interior of the enterocytes and their trafficking inside the cell (Reddy et al., 2012, Lara et al., 2015). The IncA protein family includes several proteins identified in chlamydial species that show two regions homologous to eukaryotic SNARE (soluble *N*-ethylmaleimide-sensitive factor attachment receptor) domain (Ronzone et al., 2014). The SNARE complex, in association with other proteins, is essential for the specific fusion between intracellular vesicle membranes and the vesicle-mediated intracellular transport (May et al., 2001; Han et al., 2017; Oleaga et al., 2017b). In the digestion of blood, the host proteins (including haemoglobin) and other nutrients are incorporated into the enterocyte cell by an endocytic process inside endosomes, which fuse with the lysosomes that contain the proteolytic enzymes forming digestive vesicles (Sojka et al. al., 2013, 2016; Sonenshine and Anderson, 2014). In this context, the IncA protein family might be associated with the trafficking and fusion of intracellular vesicles and the transport of proteins and other molecules inside the digestive cell. Regarding WD40 and Pkinase domains, these are two large protein families, very abundant in eukaryotic

organisms, which may participate in a multitude of cellular processes (Manning et al., 2002, Smith, 2008).

Among the most abundant protein domains identified in the downregulated genes are several zinc finger domains and Ankyrin repeats, in agreement with the identification profile obtained from the complete intestinal transcriptome discussed above. In addition, these domains were more numerous among the downregulated DEGs than in the upregulated ones (Table 1 and Supplementary Table 3).

Gene Ontology analysis of the molecular function showed that DEGs codifying catalytic and binding proteins were the most abundant. Among the other functional categories detected, all of them were significantly less represented, the most numerous was that of the genes that encode transporter proteins, the upregulated genes being notably more abundant (135) than the downregulated ones (77) (Fig. 6A). Some of these genes are involved in lipid transport (i.e. vitellogenins and Niemann–Pick type C1 domain-containing protein coding genes), carbohydrate transport (several glucose transporters) and other substrate transport (i.e. ABC transporters and ion transporters) (Supplementary Table 2).

The classification of DEGs by biological process revealed that the most numerous were those involved in cellular processes, metabolic processes and single organism processes, the upregulated genes being significantly more numerous in all three categories (Fig. 6B). The other categories, all rather less represented, were the genes involved in biological regulation, localisation, response to stimulus, cellular component organisation or biogenesis and developmental processes.

3.4. Transcripts involved in blood digestion, stress response and defence.

It has been suggested that most of the proteins involved in blood digestion in ticks are upregulated after feeding (Anderson et al., 2008). Accordingly, after annotating and functionally characterising the midgut transcriptome of *O. erraticus*, we analysed some selected groups of upregulated DEGs potentially related to feeding and blood meal digestion, namely, the genes associated with the digestion of proteins, carbohydrates and lipids and the genes involved in the defence response and response to oxidative stress.

3.4.1. Protein digestion.

Up to date, it was considered that, in ticks, blood haemoglobin was the main source of amino acids and indispensable for protein synthesis and vitellogenesis. However, it has recently been demonstrated in ixodids that although haemoglobin is absolutely essential as a source of haem group for the development of the embryo, it may be replaced as a source of amino acid for vitellogenesis, suggesting that other blood proteins are also important nutrients in the diet of ticks (Perner et al., 2016a). Moreover, these same researchers also show that ticks do not obtain bioavailable iron from haemoglobin, but it is acquired most likely from a serum transferrin, which is a major iron transporter protein from mammalian blood (Perner et al., 2016b).

The intracellular digestion of the proteins ingested with blood is performed by a set of lysosomal proteolytic enzymes that act sequentially and are known in considerable detail in ixodid ticks (Sojka et al., 2013, 2016; Sonenshine and Anderson, 2014).

In argasids, the scant information about their digestive proteolytic system has been significantly expanded recently after the analysis of the intestinal transcriptomes of *O. moubata* and *O. mimon* and the intestinal proteomes of *O. erraticus* and *O. moubata*

(Oleaga et al., 2015; Landulfo et al., 2017; Oleaga et al., 2017a, 2017b). These studies have found genes and proteins belonging to the same classes and families of proteases as those described for ixodids (cysteine peptidases, aspartic endopeptidases, metallopeptidases and serine peptidases) suggesting that the proteolytic system in hard and soft ticks can be quite similar (Sojka et al., 2013, 2016; Sonenshine and Anderson, 2014; Oleaga et al., 2017a, 2017b).

In the midgut transcriptome of *O. erraticus* we have identified up to 29 upregulated transcripts coding for different cysteine peptidases (cathepsins B, B2, C, L1 and L, cysteine protease and A protease inhibitor, midgut cystein proteinase 2 and legumain), aspartic endopeptidases (cathepsin D), metallopeptidases, and several serine peptidases (Table 2). The highest levels of expression, both before and after feeding, corresponded to four transcripts annotated as cathepsin B (ci|000022293), cathepsin D (ci|000016629), a serine carboxypeptidase (ci|000077749) and a serine protease (ci|000085239). The fold-change in expression of these genes ranged between 2.07 and 19.51, the highest value corresponding to a transcript (ci | 000028094) that encodes a serine protease (Q6U8A8) (Table 2, Supplementary Table 2). This particular serine protease was also identified in the intestine of *O. moubata* (Oleaga et al., 2017b). Besides its digestive functions, this serine could play some role in the antimicrobial defence responses, since it has a histidine- and aspartic-acid-rich motif, similar to that found in hebraein, an antimicrobial protein of the ixodid tick *Amblyomma hebraeum* (Lai et al., 2004). Similarly, cathepsin D could also intervene in defensive responses by generating antimicrobial peptides (Cruz et al., 2010). The former study on the midgut proteome of *O. erraticus* identified up to 15 proteins with peptidase activity, including cathepsin D2 and legumain (Oleaga et al., 2015). Thus, the current study significantly

expands and strengthens the previous proteomic information on the proteolytic system for protein digestion in *O. erraticus*.

3.4.2. Carbohydrate metabolism and transport.

Besides proteins, carbohydrates ingested with host blood also seem to be essential nutrients in tick feeding. Despite the studies on the digestion of carbohydrates by ticks being fairly scarce, the identification of genes associated with their metabolism and transport in the tick midgut supports the idea that in these parasites the digestion of carbohydrates may be a significant part of the blood digestion process. Accordingly, the detection in the midgut of several ixodid tick species of a number of enzymes which, due to their activity, could participate in carbohydrate metabolism, lends support to this notion (Anderson et al., 2008; Moreti 2013; Stutzer et al., 2013; Kotsyfakis et al., 2015; Xu et al., 2016). Additionally, the major pathways involved in carbohydrate metabolism were recently characterised in the tick *I. scapularis* (Cabezas-Cruz et al., 2017).

In argasids, the data available are limited and come from a few studies carried out in three species, *O. mimon*, *O. moubata* and *O. erraticus*, in which they have been identified some carbohydrases, sugar transporters and enzymes involved in the metabolism of chitin, possibly related to the formation and maintenance of the peritrophic membrane (Oleaga et al., 2015; Oleaga et al., 2017a, b; Landulfo et al., 2017).

Specifically, in the mialome of *O. moubata*, the authors identified up to 111 genes involved in carbohydrate metabolism (37 of them upregulated), as well as 27 proteins with various enzymatic activities – including glucosidase, hydrolase, isomerase, transferase, phosphatase, oxidoreductase, kinase- and three proteins involved in carbohydrate transport. Furthermore, some enzymes that participate in energy

metabolism and glucose transport were particularly abundant in the mialome of this species, which suggested that blood glucose is an important nutrient in ticks (Cabezas-Cruz et al., 2017; Oleaga et al., 2017a, 2017b).

In the current study on the midgut transcriptome of *O. erraticus* we have identified up to 54 genes whose expression level significantly increased after feeding, showing fold-changes between 2.01 and 67.46. According to their functional annotation, these genes would be associated with the metabolism and transport of several carbohydrates including fructose, mannose, galactose, malate and chitin (Table 3).

The highest expression levels either before or after feeding were those of transcripts annotated as beta-galactosidase (ci|000086301), galactokinase (ci|000078330), fructose 1,6-bisphosphatase (ci|000080606) and malate dehydrogenase (ci|000078112) (Table 3). The first two enzymes are involved in the galactose metabolism: galactokinase catalyses the second step of the Leloir pathway, which converts galactose to glucose (Holden et al., 2004), and β -galactosidase hydrolyses the β -glycosidic bond formed between a galactose and its organic moiety. Fructose 1,6-bisphosphatase and malate dehydrogenase participate in the energy metabolism of glucose, in particular in the gluconeogenesis and tricarboxylic acid cycle, respectively (Cabezas-Cruz et al., 2017).

We have also identified up to 11 upregulated transcripts that could be responsible for the transport of glucose and other sugars from the intestinal lumen into the enterocytes (Table 3). Similar increases in the level of expression during the process of digestion of genes involved in the transport of carbohydrates, especially glucose, has also been evidenced in *O. moubata* (Oleaga et al., 2017a) and in several species of ixodids (Stutzer et al., 2013; Kotsyfakis, et al., 2015; Xu et al., 2016).

Regarding chitin metabolism, up to 16 upregulated transcripts have been identified in the current *O. erraticus* midgut transcriptome. These are several chitinases (I4DN67, B7QD78, P36362, G3MSZ2) that hydrolyse the chitin oligosaccharides, and a set of proteins, including antibacterial peptide Cp6, chitin deacetylase 4, mucin peritrophin salivary protein, GM25314 protein, salivary mucin and TIL domain-containing protein, that all have chitin-binding activities and contain the protein domain CBM_14 (Pfam database), which is related to peritrophic matrix proteins (Table 3). These genes could potentially be involved in the formation and maintenance of the peritrophic membrane during the digestion process. The peritrophic membrane is a structure that lines the invertebrate midgut and acts as a semipermeable barrier that permits transport of nutrients, small molecules and pathogens, and is also thought to play important yet undefined roles in gut physiology (Hegedus et al., 2009; Kariu et al., 2013). These results are the first evidence of the existence of a peritrophic membrane in *O. erraticus*.

3.4.3. Lipid metabolism and transport

Numerous genes involved in lipid transport and metabolism were observed to be upregulated in tick midgut after feeding, suggesting that lipids may also play an important role as a source of nutrients for ticks (Stutzer et al., 2013; Kotsyfakis et al., 2015; van Zyl et al., 2015; Xu et al., 2016; Perner et al., 2016a; Oleaga et al., 2017a).

According to the functional annotation of the *O. erraticus* midgut transcriptome in the GO database, we selected 52 upregulated transcripts that are potentially involved in the metabolism and transport of blood dietary lipids (Table 4, Supplementary Table 2). Most of them have also been identified in the midgut transcriptome and proteome of *O. moubata* (Oleaga et al., 2017a, 2017b).

The fold-change of these genes ranged from 2.07 to 26.72, and the highest expression levels, both before and after feeding, corresponded to one transcript annotated as fatty-acid binding protein (ci|000020395, fold-change 13.78), one transcript annotated as oxysterol-binding protein (ci|000006034, fold-change 2.20) and two transcripts coding for Niemann–Pick type C1 domain-containing protein (NPC1) (ci|000080152, fold-change 2.79; ci|000086679, fold-change 3.51) (Table 4).

Fatty-acid binding proteins (FABPs) are members of the superfamily of lipid-binding proteins (LBP). This family of proteins is involved in general lipid metabolism, acting as intracellular transporters of hydrophobic metabolic intermediates and as carriers of lipids between membranes (Chmurzyńska, 2006). In the *O. erraticus* midgut, we found a highly abundant transcript annotated as fatty-acid binding protein (E2IH93) and two transcripts annotated as cellular retinol binding protein (Q8WR15) and CBN-LBP-7 protein (G0MRL4), these latter also belonging to the family of cytosolic FABPs (Table 4).

Other genes involved in lipid transport that were upregulated in the *O. erraticus* midgut are the LDL receptor-related protein, NPC1 protein, oxysterol-binding protein, sphingolipid transporter, hemelipoglycoprotein 2 and vitellogenins (Table 4).

It is known that arthropods lack the enzymatic machinery for cholesterol synthesis, so they are obliged to obtain cholesterol from their hosts. The low-density lipoprotein receptor (LDLR) is the major cholesterol-carrying lipoprotein of plasma, acting to regulate cholesterol homeostasis in cells. The LDLR binds LDL and transports it into digestive cells (Brown and Goldstein, 1986). NPC1 proteins, functioning in intracellular cholesterol cycling (Garver and Heidenreich, 2002), have been identified from the midgut of *Dermacentor variabilis*, *Ixodes ricinus*, *Rhipicephalus microplus* and *O. moubata* (Anderson et al., 2008; Stutzer et al., 2013; Perner et al., 2016a; Oleaga

et al., 2017a, 2017b). Oxysterol-binding protein is a conserved family of lipid binding/transfer proteins in eukaryotes. Some members of this protein family bind oxysterols or cholesterol, and other family members interact with phospholipid ligands to regulate lipid fluxes, organelle lipid compositions and cell signalling (Olkkonen, 2013).

More genes associated with cholesterol metabolism that were upregulated in *O. erraticus* midgut after feeding were insulin-induced gene protein (INSIG1) (G3N9F3), phosphomevalonate kinase (Q5RK24), sterol O-acyltransferase (B7PS11) and triacylglycerol lipase (B7PP52, B7PP53). The INSIG1 gene encodes an endoplasmic reticulum membrane protein that modulates cholesterol metabolism, lipogenesis, and glucose homeostasis (Dong and Tang, 2010). It has been demonstrated that several components of the insulin signalling pathway are necessary for adult ticks to complete the blood feeding process accompanied by the growth of internal organs (Umemiya-Shirafuji et al., 2012).

The finding of phosphomevalonate kinase (Q5RK24) enzyme catalysing a step of cholesterol biosynthesis in tick midgut could be paradoxical having in mind that arthropods do not synthesise cholesterol. However, it has recently been shown that this enzyme participates in the biosynthesis of the juvenile hormone precursor through the mevalonate pathway (Zhu et al., 2016).

Regarding vitellogenins, these are the most common transcripts identified in this group of upregulated transcripts; they affect lipid transport and are precursors of vitellin, thus being essential for egg development and oviposition. Additionally, we found hemelipoglycoprotein 2, which also belongs to the same protein superfamily as vitellogenins, and all of them contain the characteristic beta-sheet shell domains found in lipid transport proteins (Boldbaatar et al., 2010).

3.4.4. Response to oxidative stress associated with blood feeding

The digestion of the blood releases pro-oxidant compounds that in turn stimulate production of hydroxyl radicals and other toxic molecules, inducing oxidative responses. In addition, ticks ingest large amounts of blood at 37 °C, causing an increase in their body temperature, whose negative effects must be counteracted. Thus, ticks need to deploy in their midgut a range of detoxifying mechanisms, and mechanisms blocking the oxidative responses and other stressful conditions produced by the ingestion and digestion of blood (Sonenshine and Anderson, 2014).

Up to 64 upregulated transcripts coding for antioxidant proteins with oxidoreductase activities, and proteins involved in heat shock responses and protein folding have been identified after feeding in the current *O. erraticus* midgut transcriptome (Table 5). Most of these proteins have already been found in the *O. erraticus* midgut proteome (Oleaga et al., 2015), in the midgut of *O. mimon* and *O. moubata* (Landulfo et al., 2017; Oleaga et al., 2017a, 2017b), as well as in the midgut of several species of ixodid ticks (Anderson et al., 2008; Kongsuwan et al., 2010; Stutzer et al., 2013; Perner et al., 2016a).

Regarding the antioxidant proteins, by far the highest expression level found (>10,000 FPKM) was that of a transcript annotated as catalase (ci|000086362, fold-change 2.97) (Table 5), which would be in agreement with the already well-known fact that catalase is an essential enzyme in regulating hydrogen peroxide levels released in digestive cells after blood ingestion and digestion (Citelli et al., 2007).

Other upregulated antioxidant enzymes were glutathione peroxidase, glutathione S-transferases (GSTs), the oxidoreductase yrbE, superoxide dismutase (SOD) and several thioredoxins (TRX) (Table 5). Superoxide dismutase converts O_2^- into H_2O_2 ,

and thioredoxins catalyse the reversible reduction of protein disulphide bonds (Rahal et al., 2014). Moreover, these two enzymes might also participate in other biological processes, such as regulation of the size of the populations of pathogens transmitted by the tick, or the populations of the native microbiota associated with tick intestine (Budachetri and Karim, 2015; Crispell et al., 2016).

GSTs are known to play an important role in cellular stress responses and, in ixodid ticks, they are upregulated upon tick feeding, probably due to the stresses incurred during blood feeding (Dreher-Lesnick et al., 2006). A similar fact was observed in the *O. erraticus* midgut, where up to nine upregulated transcripts were annotated as GSTs, three of them showing high expression levels (>1000 FPKM) (ci|000089852, fold-change 2.87; ci|000076478, fold-change 2.39; ci|000080474, fold-change 6.66) (Table 5).

Heat shock and chaperone proteins (HSP) are molecules required for correct folding and subsequent assembly of some proteins, and are essential for normal cell growth, as they act to stabilise or protect disassembled polypeptides under heat-shock and stress conditions. HSPs are overexpressed in ticks during the blood-feeding and digestion processes, counteracting the negative effects of temperature rising and other stress situations, such as pathogen infections (Villar et al., 2014; Oleaga et al., 2017a, 2017b).

Up to 43 upregulated transcripts were found in the midgut transcriptome of *O. erraticus* encoding proteins involved in heat-shock responses, some of which were also identified in its midgut proteome (Table 5). Three of these transcripts showed high expression levels (>10,000 FPKM): a calreticulin (ci|000084379, fold-change 6.51), a protein disulfide isomerase (ci|000089827, fold-change 5.29) and a heat shock protein of the Hsp90 family (ci|000077634, fold-change 8.18) (Table 5).

Calreticulin is a calcium binding protein, widely distributed in eukaryotic cells, which functions as a chaperone to help other proteins fold correctly (Mendlovic and Conconi, 2010). It is found in a number of tick species and is involved in evading the host's immune system (Kumar et al., 2017). It is secreted in tick saliva and may facilitate tick feeding and pathogen transmission due to its anti-thrombotic and complement-inhibition activities (Kaewhom et al., 2008). Having these functions of calreticulin in mind, it could be speculated that its expression in tick midgut could contribute to maintain blood fluidity in the intestinal caeca and prevent complement activation and, in turn, any complement-mediated damage to the midgut wall.

The protein disulfide-isomerase (PDI) is a multifunctional protein that participates in protein folding, assembly and post-translational modification in the endoplasmic reticulum. In ixodid ticks, the PDI genes are upregulated after feeding and are involved in blood feeding, viability and oocyte development and pathogen infection in ticks (Liao et al., 2007, 2008).

Table 5 also includes some additional genes/proteins involved in protein-folding processes, such as the heat-shock proteins of the families HSP70 and HSP90, peptidyl-prolyl cis-trans isomerase, FK506-binding protein and several chaperonins (10 kDa heat shock protein, Grpe protein, molecular chaperone and prefolding protein).

3.4.5. Defence response

Tick midgut is the primary site of tick–pathogen interactions. Ticks need effective defence mechanisms to deal with pathogenic microorganisms ingested with blood and to maintain the intestinal microbiota at a tolerable level.

Not surprisingly, up to 18 upregulated transcripts were identified in the *O. erraticus* midgut that encoded proteins putatively involved in immunity functions,

specifically defensins and lysozyme (Table 6). These antibacterial peptides represent the major components of innate immunity in ticks and provide protection against gram-negative and gram-positive bacteria (Chrudimská et al., 2010).

It is worth mentioning the high fold-change shown by several transcripts, reaching figures higher than 100 (ci|000083012, ci|000013443, ci|000022397) and even higher than 1000 (ci|000083008) (Table 6). The highest expression levels after feeding (>10,000 FPKM) corresponded to transcripts annotated as defensin (Q09JE6) (ci|000083012, fold-change 163.42), defensin D (B6UYJ7) (ci|000083009, fold-change 20.12; ci|000008745, fold-change 8.86) and ovatusin (B7XFT2, defensin from *Ixodes ovatus*) (ci|000013443, fold-change 137.47). Other upregulated antimicrobial peptides were defensin B, persulcatusin (defensin from *Ixodes persulcatus*) and lysozyme.

These results suggest that in *O. erraticus*, as was also observed for *O. moubata* (Oleaga et al., 2017a), the genes involved in defence responses are tightly regulated. Furthermore, considering the high expression levels reached by these genes after blood ingestion, it could be supposed that the midgut is being equipped for defence against a potential invasion of pathogens ingested with blood, or even to modulate a possible increase in the size of native bacterial populations that can occur after feeding (Narasimhan and Fikrig, 2015).

3.5. Validation of RNA-seq data by RT-qPCR.

In order to validate RNA-seq results, the expression level of six randomly selected upregulated genes were assessed by RT-qPCR, using both the actin and 16S rRNA genes as reference. The selected genes encoded the following proteins: organic cation/carnitine transporter, integrin alpha-9, hydroxyprostaglandin dehydrogenase,

chitinase, Niemann–Pick type C1 domain-containing protein and glucose transporter (Fig. 7, Supplementary Table 1).

The suitability of the primer pairs and the conditions of the RT-qPCR reaction were checked by running the amplification products in agarose gel electrophoresis. Bands of the expected size were obtained for all the genes, including the two reference genes, which confirmed the fitness of all the reactions (Supplementary Figure 2).

After that, the values of fold-change in gene expression level at 48 h post feeding obtained by RT-qPCR were correlated to the corresponding values provided by RNA-seq (Fig. 7). The results of both approaches were consistent and showed a good linear correlation ($R^2 = 0.7$) despite the differences between the two methods in the magnitude of the fold-change values. These differences were not unexpected, as they were already observed and discussed by Oleaga et al. (2017a). Perner et al. (2016a) reported that the RT-qPCR analysis seemed to overestimate the differences in transcript levels with respect to the RNA-seq data. Our current results paralleled this observation, since the values of fold-change obtained by RT-qPCR were between two and three times higher than their RNA-seq for every gene except for that of glucose transporter, which showed the inverse situation (Fig. 7).

Additionally, we investigated the expression profiles of these six genes throughout an important part of the digestion process by extending the time-points analysed up to 8 days after feeding (Fig. 8). Former studies made on *O. moubata* indicated that in argasid ticks, the initial phase of intense digestion of blood takes place between 2 and 5 days post feeding, so that at 8 days post feeding this phase has already finished and the phase of very slow digestion has started, enabling argasid ticks to survive for long time periods (years) in fasting condition (Sonenshine and Anderson, 2014).

Fig. 8 shows the differential expression profiles of the six genes at 24, 48 and 72 h and at 5 and 8 days after feeding with respect to their respective baseline levels before feeding. The six transcripts were overexpressed (fold-change > 2) all through the digestion period analysed, although different gene expression patterns emerged during this period.

Transcripts encoding organic cation/carnitine transporter (Car), chitinase (Chi), hydroxyprostaglandin dehydrogenase (HPD) and Niemann–Pick type C1 domain-containing protein (NPC1) all showed rather similar expression patterns. These genes reached the highest levels of expression in the first phases of digestion (between 24 and 48 h post feeding) and then gradually decreased, returning to values close to baseline by 8 days post feeding; except Car, which experienced a slight increase at 8 days post feeding. For Car and Chi, the maximum fold-change values were detected at 24 h post feeding (27.6 and 9.7 respectively). HPD and NPC1 were also overexpressed at 24 h, but their expression level peaked at 48 h (fold-changes 22.7 and 5.9, respectively). The expression pattern of integrin alpha-9 was quite different. It was upregulated throughout the 8 days, showing fold-change values higher than 7.2, but its expression peaked at 34.0 fold-change at 5 days post feeding and decreased significantly by 8 days post feeding. Finally, the glucose transporter showed a different expression pattern. It recorded expression levels close to the baseline during the first 72 h post feeding, which dramatically increased from 5 days post feeding onwards, apparently peaking at 8 days post feeding (fold-change 11.2). (Fig. 8)

Accordingly, Car, HPD, Chi and NPC1 could be suggested as gene markers for the intense blood digestion phase, while Int and Glu1 could be envisaged as markers indicating the end of the intense digestion phase (Int) and the starting of the slow digestion phase (Glu1).

The biological significance that these particular gene expression profiles may have along the digestion process is interesting, but it has not been examined in this work and any comment currently made on this could be too speculative, thus requiring further in-depth investigation.

Regardless of their physiological significance, the differences among these expression profiles indicate that the midgut transcriptome of *O. erraticus* female ticks is very dynamic, and its composition may vary significantly along the blood digestion process and trophogonic cycle. By contrast, the composition of the midgut proteome of the *O. erraticus* females was more stable and did not experience main variations, at least in the first 48 h post feeding (Oleaga et al., 2015). These observations paralleled what was also observed in the transcriptome and proteome of *O. moubata* females (Oleaga et al., 2017a, 2017b) as well as in the mialome of other tick species (Anderson et al., 2008; Villar et al., 2014).

4. Conclusions

The current work expands significantly the molecular information obtained about the midgut proteome of *O. erraticus* by Oleaga et al. (2015), and provides a more robust and comprehensive understanding of the argasid tick digestive system.

The midgut transcriptome of *O. erraticus* females before and after the blood meal was obtained and characterised, allowing the identification of the genes differentially expressed upon feeding. These transcriptomic data provide valuable information for the detection of novel genes, and open new ways for investigations into the physiology and biochemistry of blood digestion. The closer inspection of particular groups of upregulated genes, such as those potentially related to the digestion and transport of proteins, carbohydrates and lipids, as well as of genes involved in the

defence response and response to oxidative stress, confirms the previous observation that these processes are narrowly regulated in ticks, highlighting their complexity and importance in tick biology.

Some genes that could be involved in the formation and maintenance of the peritrophic membrane during the digestion process have been identified in the current study, which constitutes the first evidence about the existence of a peritrophic membrane in *O. erraticus*.

The inspection of the expression patterns of six genes involved in different processes throughout virtually the entire blood digestion period revealed significant differences between these patterns, strongly suggesting that the transcriptome composition is highly dynamic and subject to important variation along the trophogonic cycle. This may guide future studies aimed at improving the understanding of the molecular physiology of tick digestion and digestion-related processes.

Acknowledgement

The authors are grateful to Rocío Vizcaíno Marín, María González Sánchez and Diana Ramos Pulido from the Instituto de Recursos Naturales y Agrobiología de Salamanca (IRNASA, CSIC) (Spain), for their skilful technical assistance, and to Dr. Sergio González (Molecular Biology lab of IRNASA) for his assistance in the RT-qPCR assays. This research was funded by project AGL2013-42745-P granted by the Spanish Ministry of Economy and Competitiveness.

827

828 **References**

- 829 Anderson, J.M., Sonenshine, D.E., Valenzuela, J.G., 2008. Exploring the mialome of
830 ticks: an annotated catalogue of midgut transcripts from the hard tick, *Dermacentor*
831 *variabilis* (Acari: Ixodidae). BMC Genomics 9, 552
- 832 Astigarraga, A., Oleaga-Pérez, A., Pérez-Sánchez, R., Encinas-Grandes, A., 1995. A
833 study of the vaccinal value of various extracts of concealed antigens and salivary gland
834 extracts against *Ornithodoros erraticus* and *Ornithodoros moubata*. Vet. Parasitol. 60,
835 133-147.
- 836 Boinas, F., Ribeiro, R., Madeira, S., Palma, M., de Carvalho, I.L., Nuncio, S., Wilson,
837 A.J., 2014. The medical and veterinary role of *Ornithodoros erraticus* complex ticks
838 (Acari: Ixodida) on the Iberian Peninsula. J. Vector Ecol. 39, 238-48.
- 839 Boldbaatar, D., Umemiya-Shirafuji, R., Liao, M., Tanaka, T., Xuan, X., Fujisaki, K.,
840 2010. Multiple vitellogenins from the *Haemaphysalis longicornis* tick are crucial for
841 ovarian development. J. Insect Physiol. 56, 1587-1598.
- 842 Brown, M.S., Goldstein, J.L., 1986. A receptor-mediated pathway for cholesterol
843 homeostasis. Science 232, 34-47.
- 844 Budachetri, K., Karim, S., 2015. An insight into the functional role of thioredoxin
845 reductase, a selenoprotein, in maintaining normal native microbiota in the Gulf Coast
846 tick (*Amblyomma maculatum*). Insect Mol. Biol. 24, 570-81.
- 847 Cabezas-Cruz, A., Alberdi, P., Valdés, J.J., Villar, M., de la Fuente, J., 2017.
848 *Anaplasma phagocytophilum* Infection Subverts Carbohydrate Metabolic Pathways in
849 the Tick Vector, *Ixodes scapularis*. Front. Cell. Infect. Microbiol. 7, 23.
- 850 Chmurzyńska, A., 2006. The multigene family of fatty acid-binding proteins (FABPs):
851 function, structure and polymorphism. J. Appl. Genet. 47, 39-48.
- 852 Chrudimská, T., Chrudimský, T., Golovchenko, M., Rudenko, N., Grubhoffer, L., 2010.
853 New defensins from hard and soft ticks: similarities, differences, and phylogenetic
854 analyses. Vet. Parasitol. 167, 298-303.
- 855 Čikoš, Š., Bukovská, A., Koppel, J., 2007. Relative quantification of mRNA:
856 comparison of methods currently used for real-time PCR data analysis. BMC Mol. Biol.
857 8, 113.
- 858 Citelli, M., Lara, F.A., da Silva Vaz, I., Jr., Oliveira, P.L., 2007. Oxidative stress
859 impairs heme detoxification in the midgut of the cattle tick, *Rhipicephalus (Boophilus)*
860 *microplus*. Mol. Biochem. Parasitol. 151, 81-88.

861 Cramaro, W.J., Revets, D., Hunewald, O.E., Sinner, R., Reye, A.L., Muller, C.P., 2015.
862 Integration of *Ixodes ricinus* genome sequencing with transcriptome and proteome
863 annotation of the naïve midgut. BMC Genomics. 16, 871.

864 Crispell, G., Budachetri, K., Karim, S., 2016. *Rickettsia parkeri* colonization in
865 *Amblyomma maculatum*: the role of superoxide dismutases. Parasit. Vectors 9, 291.

866 Cruz, C.E., Fogaca, A.C., Nakayasu, E.S., Angeli, C.B., Belmonte, R., Almeida, I.C.,
867 Miranda, A., Miranda, M.T., Tanaka, A.S., Braz, G.R., Craik, C.S., Schneider, E.,
868 Caffrey, C.R., Daffre, S., 2010. Characterization of proteinases from the midgut of
869 *Rhipicephalus (Boophilus) microplus* involved in the generation of antimicrobial
870 peptides. Parasit. Vectors 3, 63

871 Cutler, S.J., 2010. Relapsing fever – a forgotten disease revealed. J. Appl. Microbiol.
872 108, 1115-1122.

873 de la Fuente, J., Estrada-Pena, A., Venzal, J.M., Kocan, K.M., Sonenshine, D.E., 2008.
874 Overview: Ticks as vectors of pathogens that cause disease in humans and animals.
875 Front. Biosci. 13, 6938-6946.

876 de la Fuente, J., Kopáček, P., Lew-Tabor, A., Maritz-Olivier, C., 2016. Strategies for
877 new and improved vaccines against ticks and tick-borne diseases. Parasite Immunol. 38,
878 754-769.

879 Díaz-Martín, V., Manzano-Román, R., Obolo-Mvoulouga, P., Oleaga, A., Perez-
880 Sanchez, R., 2015. Development of vaccines against *Ornithodoros* soft ticks: An
881 update. Ticks Tick Borne Dis. 6, 211-220

882 Dong, X.Y., Tang, S.Q., 2010. Insulin-induced gene: a new regulator in lipid
883 metabolism. Peptides 31, 2145-50.

884 Dreher-Lesnick, S.M., Mulenga, A., Simser, J.A., Azad, A.F., 2006. Differential
885 expression of two glutathione S-transferases identified from the American dog tick,
886 *Dermacentor variabilis*. Insect Mol. Biol. 15, 445-453.

887 Finn, R.D., Coghill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter,
888 S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman,
889 A., 2016. The Pfam protein families database: towards a more sustainable future.
890 Nucleic Acids Res. 44, 279-85.

891 Fankhauser, N., Maser, P., 2005. Identification of GPI anchor attachment signals by a
892 Kohonen self-organizing map. Bioinformatics 21, 1846-1852.

893 Gallardo, M.C., Reoyo, A.T., Fernández-Pinero, J., Iglesias, I., Muñoz, M.J., Arias,
894 M.L., 2015. African swine fever: a global view of the current challenge. Porcine Health
895 Manag. 1, 21.

- 896 Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, Mackay JP. 2007. Sticky fingers:
897 zinc-fingers as protein-recognition motifs. Trends Biochem Sci. 32, 63-70.
- 898 Garver, W.S., Heidenreich, R.A., 2002. The Niemann-Pick C proteins and trafficking of
899 cholesterol through the late endosomal/lysosomal system. Curr. Mol. Med. 2, 485-505.
- 900 Gibson, A.K., Smith, Z., Fuqua, C., Clay, K., Colbourne, J.K., 2013. Why so many
901 unknown genes? Partitioning orphans from a representative transcriptome of the lone
902 star tick *Amblyomma americanum*. BMC Genomics. 14, 135.
- 903 Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I.,
904 Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N.,
905 Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K.,
906 Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data
907 without a reference genome. Nat. Biotechnol. 29, 644-652
- 908 Guerrero, F.D., Miller, R.J., Perez de Leon, A.A., 2012. Cattle tick vaccines: many
909 candidate antigens, but will a commercially viable product emerge? International J.
910 Parasitol. 42, 421-427.
- 911 Hajdusek, O., Sima, R., Perner, J., Loosova, G., Harcubova, A., Kopacek, P., 2016.
912 Tick iron and heme metabolism - New target for an anti-tick intervention. Ticks Tick
913 Borne Dis. 7, 565-72.
- 914 Han, J., Pluhackova, K., Böckmann, R.A., 2017. The multifaceted role of SNARE
915 proteins in membrane fusion. Front. Physiol. 8, 5.
- 916 Heekin, A.M., Guerrero, F.D., Bendele, K.G., Saldivar, L., Scoles, G.A., Dowd, S.E.,
917 Gondro, C., Nene, V., Djikeng, A., Brayton, K.A., 2013. Gut transcriptome of replete
918 adult female cattle ticks, *Rhipicephalus (Boophilus) microplus*, feeding upon a *Babesia*
919 *bovis*-infected bovine host. Parasitol Res. 112, 3075-90.
- 920 Hegedus, D., Erlandson, M., Gillott, C., Toprak, U., 2009. New insights into peritrophic
921 matrix synthesis, architecture, and function. Annu. Rev. Entomol. 54: 285-302.
- 922 Holden, H.M., Thoden, J.B., Timson, D.J., Reece, R.J., 2004. Galactokinase: structure,
923 function and role in type II galactosemia. Cell. Mol. Life Sci. 61, 2471-84.
- 924 Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering,
925 C., Bork, P., 2017. Fast Genome-Wide Functional Annotation through Orthology
926 Assignment by eggNOG-Mapper. Mol. Biol. Evol. 34, 2115-2122.
- 927 Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L., Daszak,
928 P., 2008. Global trends in emerging infectious diseases. Nature 451, 990-993.
- 929 Kaewhom, P., Stich, R.W., Needham, G.R., Jittapalapong, S., 2008. Molecular analysis
930 of calreticulin expressed in salivary glands of *Rhipicephalus (Boophilus) microplus*
931 indigenous to Thailand. Ann. N. Y. Acad. Sci. 1149, 53-7.

932 Kariu, T., Smith, A., Yang, X., Pal, U., 2013. A chitin deacetylase-like protein is a
933 predominant constituent of tick peritrophic membrane that influences the persistence of
934 Lyme disease pathogens within the vector. PLoS One. 8, e78376.

935 Kocan, K.M., De La Fuente, J., Blouin, E.F., Garcia-Garcia, J.C., 2004. *Anaplasma*
936 *marginale* (Rickettsiales: Anaplasmataceae): Recent advances in defining host-pathogen
937 adaptations of a tick-borne rickettsia. Parasitology 129, S285-S300.

938 Kongsuwan, K., Josh, P., Zhu, Y., Pearson, R., Gough, J., Colgrave, M.L., 2010.
939 Exploring the midgut proteome of partially fed female cattle tick (*Rhipicephalus*
940 *(Boophilus) microplus*). J. Insect. Physiol. 56:212-26.

941 Kotsyfakis, M., Schwarz, A., Erhart, J., Ribeiro, J.M., 2015. Tissue- and time-dependent
942 transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the
943 vertebrate host. Sci Rep. 5, 9103.

944 Kumar, B., Manjunathachar, H.V., Nagar, G., Ravikumar, G., de la Fuente, J.,
945 Saravanan, B.C., Ghosh, S., 2017. Functional characterization of candidate antigens of
946 *Hyalomma anatolicum* and evaluation of its cross-protective efficacy against
947 *Rhipicephalus microplus*. Vaccine. 35, 5682-5692.

948 Lai, R., Takeuchi, H., Lomas, L.O., Jonczy, J., Rigden, D.J., Rees, H.H., Turner, P.C.,
949 2004. A new type of antimicrobial protein with multiple histidines from the hard tick,
950 *Amblyomma hebraeum*. FASEB J. 18, 1447-1449.

951 Laity, J.H., Lee, B.M., Wright, P.E., 2001. Zinc finger proteins: new insights into
952 structural and functional diversity. Curr. Opin. Struct. Biol. 11, 39-46.

953 Landulfo, G.A., Patané, J.S.L., Silva, D.G.N.D., Junqueira-de-Azevedo, I.L.M.,
954 Mendonca, R.Z., Simons, S.M., Carvalho, E., Barros-Battesti, D.M., 2017. Gut
955 transcriptome analysis on females of *Ornithodoros mimon* (Acari: Argasidae) and
956 phylogenetic inference of ticks. Rev. Bras. Parasitol. Vet. 26, 185-204.

957 Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat.
958 Methods 9, 357-359.

959 Lara, F.A., Pohl, P.C., Gandara, A.C., Ferreira, J. da S., Nascimento-Silva, M.C.,
960 Bechara, G.H., Sorgine, M.H., Almeida, I.C., Vaz Ida, S. Jr, Oliveira, P.L., 2015. ATP
961 Binding Cassette Transporter Mediates Both Heme and Pesticide Detoxification in Tick
962 Midgut Cells. PLoS One. 10, e0134779.

963 Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq
964 data with or without a reference genome. BMC Bioinformatics 12, 323.

965 Liao, M., Boldbaatar, D., Gong, H., Huang, P., Umemiya, R., Harnnoi, T., Zhou, J.,
966 Tanaka, T., Suzuki, H., Xuan, X., Fujisaki, K., 2008. Functional analysis of protein
967 disulfide isomerases in blood feeding, viability and oocyte development in
968 *Haemaphysalis longicornis* ticks. Insect. Biochem. Mol. Biol. 38, 285-95.

969 Liao, M., Hatta, T., Umemiya, R., Huang, P., Jia, H., Gong, H., Zhou, J., Nishikawa, Y.,
 970 Xuan, X., Fujisaki, K., 2007. Identification of three protein disulfide isomerase
 971 members from *Haemaphysalis longicornis* tick. Insect. Biochem. Mol. Biol. 37, 641-54.

972 Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and
 973 dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550.

974 Manning, G., Plowman, G.D., Hunter, T., Sudarsanam, S., 2002. Evolution of protein
 975 kinase signaling from yeast to man. Trends Biochem. Sci. 27, 514-20.

976 Mans, B.J., de Castro, M.H., Pienaar, R., de Klerk, D., Gaven, P., Genu, S., Latif, A.A.
 977 Ancestral reconstruction of tick lineages. Ticks Tick Borne Dis. 7, 509-35.

978 Manzano-Román, R., Encinas-Grandes, A., Pérez-Sánchez, R., 2006. Antigens from the
 979 midgut membranes of *Ornithodoros erraticus* induce lethal anti-tick immune responses
 980 in pigs and mice. Vet. Parasitol. 135, 65-79.

981 Manzano-Román, R., Garcia-Varas, S., Encinas-Grandes, A., Perez-Sanchez, R., 2007.
 982 Purification and characterization of a 45-kDa concealed antigen from the midgut
 983 membranes of *Ornithodoros erraticus* that induces lethal anti-tick immune responses in
 984 pigs. Vet. Parasitol. 145, 314-325.

985 Manzano-Román, R., Diaz-Martin, V., Oleaga, A., Siles-Lucas, M., Perez-Sanchez, R.,
 986 2012. Subolesin/akirina orthologs from *Ornithodoros* spp. soft ticks: cloning, RNAi
 987 gene silencing and protective effect of the recombinant proteins. Vet. Parasitol. 185,
 988 248-259.

989 Maritz-Olivier, C., van Zyl, W., Stutzer, C., 2012. A systematic, functional genomics,
 990 and reverse vaccinology approach to the identification of vaccine candidates in the
 991 cattle tick, *Rhipicephalus microplus*. Ticks Tick Borne Dis. 3, 179-187.

992 May, A.P., Whiteheart, S.W., Weis, W.I., 2001. Unraveling the mechanism of the
 993 vesicle transport ATPase NSF, the N-ethylmaleimide-sensitive factor. J. Biol. Chem.
 994 276, 21991-21994.

995 Mendlovic, F., Conconi, M., 2010. Calreticulin: a Multifaceted Protein. Nature
 996 Education 4, 1.

997 Moreti, R., Perrella, N.N., Lopes, A.R., 2013. Carbohydrate digestion in ticks and a
 998 digestive alpha-L-fucosidase. J. Insect Physiol. 59, 1069-1075

999 Narasimhan, S., Fikrig, E., 2015. Tick microbiome: the force within. Trends Parasitol.
 1000 31, 315-323.

1001 Oleaga, A., Obolo-Mvoulouga, P., Manzano-Román, R., Pérez-Sánchez, R., 2015.
 1002 Midgut proteome of an argasid tick, *Ornithodoros erraticus*: a comparison between
 1003 unfed and engorged females. Parasit. Vectors 8, 525.

1004 Oleaga, A., Obolo-Mvoulouga, P., Manzano-Román, R., Pérez-Sánchez, R., 2017a.
1005 Functional annotation and analysis of the *Ornithodoros moubata* midgut genes
1006 differentially expressed after blood feeding. Ticks Tick Borne Dis. 8, 693-708.

1007 Oleaga, A., Obolo-Mvoulouga, P., Manzano-Román, R., Pérez-Sánchez, R., 2017b. A
1008 proteomic insight into the midgut proteome of *Ornithodoros moubata* females reveals
1009 novel information on blood digestion in argasid ticks. Parasit. Vectors 10, 366.

1010 Oleaga-Perez, A., Perez-Sanchez, R., Encinas-Grandes, A., 1990. Distribution and
1011 biology of *Ornithodoros erraticus* in parts of Spain affected by African swine fever.
1012 Vet. Record 126, 32-37.

1013 Olkkonen, V.M., 2013. OSBP-related proteins: liganding by glycerophospholipids
1014 opens new insight into their function. Molecules 18, 13666-79.

1015 Perner, J., Provazník, J., Schrenková, J., Urbanová, V., Ribeiro, J.M., Kopáček, P.,
1016 2016a. RNA-seq analyses of the midgut from blood- and serum-fed *Ixodes ricinus*
1017 ticks. Sci. Rep. 6, 36695.

1018 Perner, J., Sobotka, R., Sima, R., Konvickova, J., Sojka, D., Oliveira, P.L., Hajdusek,
1019 O., Kopacek, P., 2016b. Acquisition of exogenous haem is essential for tick
1020 reproduction. Elife 7;5. pii: e12318.

1021 Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0:
1022 discriminating signal peptides from transmembrane regions. Nat.Methods 8, 785-786.

1023 Popara, M., Villar, M., Mateos-Hernández, L., de Mera, I.G., Marina, A., del Valle, M.,
1024 Almazán, C., Domingos, A., de la Fuente, J., 2013. Lesser protein degradation
1025 machinery correlates with higher BM86 tick vaccine efficacy in *Rhipicephalus*
1026 *annulatus* when compared to *Rhipicephalus microplus*. Vaccine 31, 4728-4735.

1027 Rachinsky, A., Guerrero, F.D., Scoles, G.A., 2008. Proteomic profiling of
1028 *Rhipicephalus (Boophilus) microplus* midgut responses to infection with *Babesia bovis*.
1029 Vet. Parasitol. 152,294-313.

1030 Rahal, A., Kumar, A., Singh, V., Yadav, B., Tiwari, R., Chakraborty, S., Dhama, K.,
1031 2014. Oxidative stress, prooxidants, and antioxidants: the interplay. Biomed. Res. Int.
1032 2014, 761264.

1033 Rebaudet, S., Parola, P., 2006. Epidemiology of relapsing fever borreliosis in Europe.
1034 FEMS Immunol. Med. Microbiol. 48, 11-15

1035 Reddy, V.S., Shlykov, M.A., Castillo, R., Sun, E.I., Saier, M.H. Jr, 2012. The major
1036 facilitator superfamily (MFS) revisited. FEBS J. 279, 2022-35.

1037 Ronzone, E., Wesolowski, J., Bauler, L.D., Bhardwaj, A., Hackstadt, T., Paumet, F.,
1038 2014. An α -helical core encodes the dual functions of the chlamydial protein IncA. J.
1039 Biol. Chem. 289, 33469-80.

1040 Sánchez-Vizcaíno, J.M., Mur, L., Gomez-Villamandos, J.C., Carrasco, L., 2015. An
 1041 update on the epidemiology and pathology of African swine fever. *J. Comp. Pathol.*
 1042 152, 9-21.

1043 Schwarz, A., Tenzer, S., Hackenberg, M., Erhart, J., Gerhold-Ay, A., Mazur, J.,
 1044 Kuharev, J., Ribeiro, J.M., Kotsyfakis, M., 2014. A systems level analysis reveals
 1045 transcriptomic and proteomic complexity in *Ixodes ricinus* midgut and salivary glands
 1046 during early attachment and feeding. *Mol. Cell. Proteomics* 13:2725-35.

1047 Schulz, M.H., Zerbino, D.R., Vingron, M., Birney, E., 2012. Oases: robust de novo
 1048 RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28,
 1049 1086-1092.

1050 Sedgwick, S.G., Smerdon, S.J., 1999. The ankyrin repeat: a diversity of interactions on
 1051 a common structural framework. *Trends Biochem. Sci.* 24, 311–316.

1052 Smith, T.F., 2008. Diversity of WD-repeat proteins. *Subcell. Biochem.* 48, 20-30.

1053 Šmit, R., Postma, M.J., 2016. Vaccines for tick-borne diseases and cost-effectiveness of
 1054 vaccination: a public health challenge to reduce the diseases' burden. *Expert. Rev.*
 1055 *Vaccines*. 15, 5-7.

1056 Sojka, D., Franta, Z., Horn, M., Caffrey, C.R., Mareš, M., Kopáček, P., 2013. New
 1057 insights into the machinery of blood digestion by ticks. *Trends Parasitol.* 29, 276-85.

1058 Sojka, D., Pytelkova, J., Perner, J., Horn, M., Konvickova, J., Schrenkova, J., Mares,
 1059 M., Kopacek, P., 2016. Multienzyme degradation of host serum albumin in ticks. *Ticks*
 1060 *Tick Borne Dis.* 7, 604-613.

1061 Sonenshine, D.E., Anderson, J.M., 2014. Mouthparts and digestive system: Anatomy
 1062 and Molecular Biology of feeding and digestion. In: Sonenshine, D.E., Roe, R.M.,
 1063 editors. *Biology of Ticks*. Oxford University Press, p.122-162

1064 Sonenshine, D.E., Roe, R.M. (Ed.), 2014. *Biology of Ticks*, Vol I. Oxford University
 1065 Press.

1066 Stutzer, C., van Zyl, W.A., Olivier, N.A., Richards, S., Maritz-Olivier, C., 2013. Gene
 1067 expression profiling of adult female tissues in feeding *Rhipicephalus microplus* cattle
 1068 ticks. *Int. J. Parasitol.* 43, 541-554.

1069 Tautz, D., Domazet-Lošo, T., 2011. The evolutionary origin of orphan genes. *Nat. Rev*
 1070 *Genet.* 12, 692-702.

1071 Umemiya-Shirafuji, R., Tanaka, T., Boldbaatar, D., Tanaka, T., Fujisaki, K., 2012. Akt
 1072 is an essential player in regulating cell/organ growth at the adult stage in the hard tick
 1073 *Haemaphysalis longicornis*. *Insect. Biochem. Mol. Biol.* 42, 164-73.

1074 Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., Leunissen, J.A., 2007.
1075 Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* 35, 71-74.

1076 van Zyl, W.A., Stutzer, C., Olivier, N.A., Maritz-Olivier, C., 2015. Comparative
1077 microarray analyses of adult female midgut tissues from feeding *Rhipicephalus* species.
1078 *Ticks Tick Borne Dis.* 6, 84-90.

1079 Villar, M., Popara, M., Ayllón, N., Fernández de Mera, I.G., Mateos-Hernández, L.,
1080 Galindo, R.C., Manrique, M., Tobes, R., de la Fuente, J., 2014. A systems biology
1081 approach to the characterization of stress response in *Dermacentor reticulatus* tick
1082 unfed larvae. *PLoS One* 9, e89564.

1083 Willadsen, P., 2008. Anti-tick vaccines. In: Bowman, A.S., Nuttall, P. (Ed.). *Ticks:*
1084 *Biology, Disease and Control.* Cambridge University Press, pp. 424-446.

1085 Xu, X.L., Cheng, T.Y., Yang, H., Liao, Z.H., 2015. De novo assembly and analysis of
1086 midgut transcriptome of *Haemaphysalis flava* and identification of genes involved in
1087 blood digestion, feeding and defending from pathogens. *Infect. Genet. Evol.* 38, 62-72.

1088 Zhu, J., Khalil, S.M., Mitchell, R.D., Bissinger, B.W., Egekwu, N., Sonenshine, D.E.,
1089 Roe, R.M., 2016. Mevalonate-farnesal biosynthesis in ticks: Comparative synganglion
1090 transcriptomics and a new perspective. *PLoS One* 11, e0141084.

1091

1092 **Figure legends.**

1093 Figure 1. *Ornithodoros erraticus* midgut transcriptome. Classification of the annotatable
1094 transcript by their molecular function. Only the categories represented by more than 25
1095 genes have been included in the pie charts. The number of transcripts is indicated inside
1096 the parentheses.

1097 Figure 2. *Ornithodoros erraticus* midgut transcriptome. Classification of the annotatable
1098 transcript by their biological process. Only the categories represented by more than 25
1099 genes have been included in the pie charts. The number of transcripts is indicated inside
1100 the parentheses.

1101 Figure 3. Classification by Protein Class of the genes identified in the midgut
1102 transcriptome of *Ornithodoros erraticus*. Bars and numbers inside represent number of
1103 genes in each protein class. Blue: genes not differentially expressed upon feeding (NS);
1104 red: genes downregulated; green: genes upregulated.

1105 Figure 4: Scatter plot of the gene expression counts between fed and unfed. The log₂
1106 fold-change is plotted against the mean expression of each transcript in the two
1107 conditions. Each point represents one transcript: black dots are transcripts with similar
1108 expression before and after feeding; red and blue dots are downregulated and
1109 upregulated transcripts, respectively. E-value is the score of the BLASTX searches on
1110 selected transcript.

1111 Figure 5. Number of up- and downregulated sequences involved in active biological
1112 pathways.

1113 Figure 6. Classification of differential expressed genes (DEGs) in the midgut
1114 transcriptome of fed *versus* unfed *Ornithodoros erraticus* females by molecular
1115 function (A) and by biological process (B).

1116 Figure 7. Correlation between values of fold-change at 48 h post feeding *versus* unfed,
1117 obtained by RT-qPCR and by RNA-seq for the six selected genes shown in the table
1118 below the graph.

1119 Figure 8. RT-qPCR analysis of six differentially expressed genes upregulated after
1120 feeding in the midgut of *O. erraticus* female ticks. Analysis was performed in two
1121 independent cDNA sets obtained at different time-points after blood ingestion. The
1122 fold-change in the expression level for each gene and time-point was calculated with
1123 respect to its level in unfed ticks. Car: organic cation/carnitine transporter; Int: integrin
1124 alpha 9; HDP: hydroxyprostaglandin dehydrogenase; Chi: chitinase; NPC1: Niemann–
1125 Pick type C1 domain-containing protein; Glu1: glucose transporter type 1. Values
1126 represent the mean ± SD of two biological replicates (n = 2).

1127 Supplementary Fig 1. Fold-change distribution of differential expressed genes (DEGs)
1128 in the midgut from fed *O. erraticus* females as compared to unfed females.

1129 Supplementary Fig. 2. Electrophoresis in 1.5% agarose gel showing the RT-qPCR
1130 products for the six target genes and the two reference genes. Car: organic
1131 cation/carnitine transporter (121 bp); Int: integrin alpha 9 (105 bp); HDP:
1132 hydroxyprostaglandin dehydrogenase (103 bp); Chi: chitinase (146 bp); NPC1:
1133 Niemann–Pick type C1 domain-containing protein (149 bp); Glu1: glucose transporter
1134 type 1 (145 bp). Actin (116 bp) and 16S rRNA (222 bp) are the reference genes. M:
1135 DNA markers.

1136

1137 **Supplementary Tables**

1138 Supplementary Table 1. Primer sequences (5' to 3') and reaction conditions used for
1139 quantitative real time PCR (RT-qPCR).

1140 Supplementary Table 2. Midgut transcriptome of *Ornithodoros erraticus* females. UF-
1141 1, UF-2: midgut transcriptome from Unfed ticks. 48h-1, 48h-2: midgut transcriptome
1142 from Fed ticks. Differently expressed genes (DEGs) in FED ticks as compared to the
1143 UNFED ticks are those having a fold change > 2 (upregulated) or <-2 (downregulated),
1144 and a padj < 0.05.

1145 Supplementary Table 3. Pfam domain families found in the whole transcriptome and in
1146 the up- and downregulated genes in the *Ornithodoros erraticus* midgut at 48 hours
1147 post-feeding.

1148 Supplementary Table 4. Biological pathways in which the genes that are differentially
1149 expressed in the midgut from fed females, as compared to unfed females, are involved.

Figure 1.

Figure 2.

Figure 3.

Figure 4.

Figure 5.

A) Molecular function

B) Biological Process

Figure 6.

Transcript	Protein	E value
ci 000076487	Organic cation/carnitine transporter (Car)	0.0
ci 000090662	Integrin alpha-9 (Int)	1e-143
ci 000089943	15-hydroxyprostaglandin dehydrogenase [NAD(+)] (HPD)	7e-36
ci 000016961	Chitinase (Chi)	6e-86
ci 000086679	Niemann-Pick type C1 domain-containing protein (NPC1)	8e-72
ci 000078671	Glucose transporter type 1 (Glu1)	0.0

Figure 7.

Figure 8.

Table 2. Selected upregulated genes putatively involved in protein digestion. FPKM values are the mean of the two biological replicates analyzed from each physiological condition (Unfed: UF-1 and UF-2; Fed: 48h-1 and 48h-2). For transcript names, sequences, FPKM data and other information see GenBank GFWV000000000 (BioProject: PRJNA401392) and Supplementary Table 2. (P) Protein identified in the intestinal proteome of *O. erraticus* (Oleaga et al., 2015).

Protein	Entry	Contig name	FoldChange	FPKM Unfed	FPKM Fed	Molecular function
Carboxypeptidase	A1IHK5	ci 000029082	3.40	144.27	496.03	serine-type carboxypeptidase
		ci 000023303	3.09	16.77	55.59	
Cathepsins B	B7P3P0 A4GTA7 B7SP43	ci 000022293	2.67	4948.81	13244.03	cysteine-type peptidase
		ci 000028552	3.09	907.37	2812.42	
		ci 000034486	2.86	658.21	1891.37	
Cathepsin C	A8D0K2	ci 000070677	3.27	35.14	119.17	cysteine-type peptidase
Cathepsins D	A2I3V1 Q45HJ6	ci 000042359	4.98	73.45	373.74	aspartic-type endopeptidase
		ci 000016629	4.80	3952.97	19053.88	
Cathepsin L1	E5SRN7	ci 000009126	4.23	33.04	145.13	cysteine-type peptidase
Cathepsin L-like proteinase	G9F9V4	ci 000021500	3.81	24.11	97.71	cysteine-type peptidase
Cysteine protease and A protease inhibitor	B7P7I7	ci 000090378	3.96	206.18	822.52	cysteine-type peptidase
Metalloproteases (P)	B4NL40 Q09JT3 Q09JP9 Q09JP5	ci 000079690	2.67	23.60	65.90	metalloendopeptidase
		ci 000012653	8.09	27.24	243.69	
		ci 000093578	6.06	25.71	163.47	
		ci 000004523	3.84	3.67	17.82	
Midgut cysteine proteinase-2	Q86GZ5	ci 000031613	12.84	16.78	233.75	cysteine-type peptidase
Midgut serine proteinase-3	Q8T4N2	ci 000039478	3.11	8.39	29.12	serine-type endopeptidase
Pro-cathepsin B2	Q8I7B2	ci 000009150	5.56	77.12	442.23	cysteine-type peptidase
		ci 000043642	2.88	11.02	34.58	
Serine carboxypeptidase	Q6NY68 B7PJ32	ci 000077749	2.07	10269.91	21320.69	serine-type carboxypeptidase
		ci 000020742	3.00	342.00	1031.43	
Serine protease (P)	Q6L7Z5 B7P183 Q6U8A8	ci 000009706	8.58	5.77	60.03	serine-type endopeptidase
		ci 000071404	4.84	1.05	11.09	
		ci 000085239	4.46	1230.86	5511.84	
		ci 000079695	2.55	17.82	48.02	
		ci 000026727	5.03	4.19	26.53	
		ci 000028094	19.51	102.29	2044.10	
		ci 000034856	8.08	2.09	26.23	
Tick legumain (P)	A4PF00	ci 000010690	6.15	23.61	153.05	cysteine protease

Table 3. Selected upregulated genes putatively involved in carbohydrate metabolism and transport. FPKM values are the mean of the two biological replicates analyzed from each physiological condition (Unfed: UF-1 and UF-2; Fed: 48h-1 and 48h-2). For transcript names, sequences, RPKM data and other information see GenBank GFWV000000000 (BioProject: PRJNA401392) and Supplementary Table 2. (P) Protein identified in the intestinal proteome of *O. erraticus* (Oleaga et al., 2015).

Protein	Entry	Contig name	Fold Change	FPKM Unfed	FPKM Fed	Biological process
N-acetylglucosaminyltransferase I	A0A087UN81	ci 000015469	3.31	2.62	12.36	carbohydrate metabolism
Fructose 1,6-bisphosphatase (P)	A0A0D3QZ08	ci 000080606	3.16	4227.46	13417.10	
Glucosidase II	B7QCA7	ci 000010167	3.93	4.72	22.24	
Glucosylceramidase	B7PNW7	ci 000090332	12.64	108.08	1387.74	
		ci 000022520	12.01	17.86	237.50	
Putative beta-N-acetylhexosaminidase	I0YYH5	ci 000041332	3.62	74.50	274.72	
		ci 000012376	2.44	68.73	170.26	
Ribulose-phosphate 3-epimerase	F4WL89	ci 000080496	3.60	14.69	56.49	
		ci 000080493	12.18	2.10	39.57	
Uncharacterized proteins	G3MK74 F7A9H0 G3SN00 K7J396 G3MIF9	ci 000013713	2.29	280.12	647.50	
		ci 000016489	8.43	150.05	1282.42	
		ci 000090171	2.73	761.52	2087.12	
		ci 000090295	2.15	659.75	1425.18	
		ci 000005147	4.57	17.31	86.43	
PFKPB3 ¹	E5S5V1	ci 000079163	3.11	314.27	984.57	fructose metabolism
Alpha-mannosidase (P)	D6W117 B4P0P8 Q58EM3	ci 000056017	12.50	1.05	25.98	mannose metabolism
		ci 000008990	4.66	2.10	15.24	
		ci 000066951	2.24	28.87	66.87	
		ci 000023616	2.23	31.46	72.35	
		ci 000014030	2.01	389.14	786.87	
		ci 000000525	4.94	192.50	959.45	
Beta-galactosidase	B7PW84	ci 000013098	8.11	39.36	331.96	galactose metabolism
		ci 000086301	6.60	5637.90	37323.49	
Galactokinase	B7PKS9	ci 000028628	5.71	2.62	21.64	
		ci 000078330	2.59	7449.22	19359.33	
Galactose-1-phosphate uridylyltransferase	C4WU13	ci 000017643	2.09	136.92	288.05	
Uncharacterized protein	G3MIF9	ci 000039039	3.52	8.39	34.00	
Malate dehydrogenase (P)	Q7F9T0	ci 000078112	3.04	3705.14	11290.10	malate metabolism
Antibacterial peptide Cp6	A9XG02	ci 000031217	3.55	190.96	683.98	chitin metabolism
		ci 000019673	3.22	171.10	559.28	
Chitin deacetylase 4	A8W489	ci 000010020	4.82	6.83	38.92	
Chitinase 5	I4DN67	ci 000015968	67.46	0.00	72.18	
		ci 000067165	10.42	1.05	21.99	
Chitinase. putative	B7QD78	ci 000039438	2.35	7.35	19.07	

Endochitinase (EC 3.2.1.14)	P36362	ci 000033288	4.84	7.34	40.34	
		ci 000015542	3.08	334.60	1041.61	
GM25314	B4HFG8	ci 000004272	4.00	20.46	85.97	
Mucin peritrophin salivary protein	Q09JI0	ci 000077537	5.16	5.24	32.48	
		ci 000086014	2.37	461.45	1103.35	
Salivary mucin with chitin-binding domain	Q09JK5	ci 000019340	2.31	369.01	901.81	
		ci 000042173	2.14	163.82	361.64	
TIL domain containing protein	A0A131ZDX3	ci 000009248	2.61	54.00	145.15	
		ci 000032686	3.70	162.09	606.28	
Uncharacterized protein	G3MSZ2	ci 000005147	4.57	17.31	86.43	
Sugar transporter	B7QM62	ci 000005071	6.07	7.34	63.23	carbohydrate transport
GDP-fucose transporter	F1L4R9	ci 000008831	3.11	24.65	80.03	fucose transport
		ci 000061957	4.52	4.20	23.63	
		ci 000002913	2.36	90.77	217.09	
Glucose transport proteins	B7PGT7 E2C2M2	ci 000078671	3.93	104.40	418.29	
		ci 000017359	3.84	19.91	82.40	glucose transport
		ci 000034723	2.52	5.76	33.63	
		ci 000077873	5.46	138.99	771.84	
Sodium-dependent glucose transporter	B7PS38 B7PKI7	ci 000015995	5.36	307.36	1658.27	
		ci 000032460	4.57	72.37	337.58	

¹ 6-phosphofructo-2-kinase/fructose-2, 6-biphosphatase 3

Table 4. Selected upregulated genes putatively involved in lipid metabolism and transport. FPKM values are the mean of the two biological replicates analyzed from each physiological condition (Unfed: UF-1 and UF-2; Fed: 48h-1 and 48h-2). For transcript names, sequences, FPKM data and other information see GenBank GFWV000000000 (BioProject: PRJNA401392) and Supplementary Table 2.

Protein	Entry	Contig name	Fold Change	FPKM Unfed	FPKM Fed	Biological process
Acyl-CoA synthetase	G1MQ54 G3X0M2	ci 000010868	5.09	20.46	110.27	lipid biosynthesis
		ci 000030406	4.93	4.20	26.65	
CBN-LBP-7 protein	G0MRL4	ci 000056109	18.52	2.09	59.01	lipid transport
Cellular retinoic acid/retinol binding protein	Q8WR15	ci 000043737	10.72	10.49	124.79	lipid transport
Dfg10 protein	B7QG88	ci 000003995	2.57	268.09	693.96	lipid metabolism
		ci 000020263	2.37	496.77	1180.58	
Fatty-acid binding protein	E2IH93	ci 000020395	13.78	2216.29	30779.08	lipid transport
GF11181	B3MGZ2	ci 000085504	2.50	37.24	95.56	phospholipid biosynthesis
Glycerol-3-phosphate acyltransferase	B7PBT7 B7PKN6	ci 000003487	2.37	293.69	701.02	lipid metabolism
		ci 000089670	2.86	821.97	2355.81	
Glyceronephosphate O-acyltransferase	F1QAU7	ci 000015590	5.13	4.72	29.37	lipid metabolism
Glycerophosphoryl diester phosphodiesterase	B7PNT8	ci 000081902	2.23	399.15	898.46	lipid metabolism
Hemelipoglycoprotein 2	B5ABL8	ci 000022643	10.33	25.70	281.94	lipid transport
Insulin-induced gene protein	G3N9F3	ci 000077129	3.13	199.31	628.86	cholesterol metabolism
LDL receptor-related protein	B7PJ49 B7PY42	ci 000012221	2.69	96.00	310.10	lipid transport
		ci 000089577	2.34	226.51	537.26	
Niemann-Pick type C1 domain-containing protein	B7Q350 B7PPF5 Q6QVL7	ci 000038753	4.70	5.77	32.28	cholesterol transport
		ci 000080152	2.79	2977.97	8323.93	
		ci 000023386	2.25	23.58	55.62	
		ci 000079217	2.17	261.72	569.98	
		ci 000086679	3.51	3258.71	11445.34	
		ci 000042898	2.30	19.92	48.10	
		ci 000031899	6.45	22.02	150.77	
Oxysterol-binding protein	G3MGP9 G3MHI6 E2B528 E5S9Y9	ci 000068735	2.15	36.16	80.54	lipid transport
		ci 000006034	2.20	1883.04	4149.23	
		ci 000021961	2.14	80.78	174.80	
		ci 000083954	3.89	58.20	232.42	
Phospholipase A2	B2D2A9	ci 000005000	2.07	78.12	165.57	cholesterol biosynthesis
Uncharacterized proteins	G3MLU0 G3MG23 E9IWS1 G3MJ85 H2MZC2 G3MJ85	ci 000020772	2.16	538.15	1165.66	lipid metabolism
		ci 000072619	6.07	3.14	25.30	sphingolipid metabolism
		ci 000008489	3.00	23.08	72.10	phospholipid biosynthesis
		ci 000017152	2.52	725.13	1834.59	phospholipid biosynthesis
		ci 000019266	4.73	20.45	102.30	phospholipid biosynthesis
		ci 000085487	2.31	177.27	413.21	phospholipid biosynthesis
Saposin	B7Q6Z1	ci 000070562	2.61	7.87	23.23	sphingolipid metabolism
Sphingolipid transporter 1	H9GAL5	ci 000076948	2.83	144.21	414.62	lipid transport

Sterol O-acyltransferase	B7PS11	ci 000014830	3.19	7.86	28.05	cholesterol metabolism
Triacylglycerol lipase	B7PP52 B7PP53	ci 000089281	3.17	965.64	3073.46	cholesterol metabolism
		ci 000044799	2.29	51.36	121.67	
		ci 000020206	2.88	595.85	1720.40	
Vitellogenin	B6ZIV7 Q5EG54	ci 000054240	2.58	24.64	66.54	lipid transport
		ci 000012543	2.15	218.09	474.52	
Vitellogenin-1	E1CAX9	ci 000037347	26.72	0.53	42.95	lipid transport
		ci 000024326	21.44	7.88	196.35	
		ci 000056878	13.53	6.83	111.28	
		ci 000034412	13.35	30.41	426.90	
		ci 000022645	12.89	17.30	239.40	
		ci 000045144	11.05	42.49	485.92	
		ci 000024070	10.62	2.62	38.45	
		ci 000034427	10.04	4.19	53.15	
		ci 000062252	7.33	18.87	149.41	
		ci 000058647	4.64	4.19	64.41	

Table 5. Selected upregulated genes putatively involved in stress responses. FPKM values are the mean of the two biological replicates analyzed from each physiological condition (Unfed: UF-1 and UF-2; Fed: 48h-1 and 48h-2). For transcript names, sequences, FPKM data and other information see GenBank GFWV000000000 (BioProject: PRJNA401392) and Supplementary Table 2. (P) Protein identified in the intestinal proteome of *O. erraticus* (Oleaga et al., 2015).

Protein	Entry	Contig name	Fold Change	FPKM Unfed	FPKM Fed	Biological process
ANTIOXIDANT						
Catalase (EC 1.11.1.6) (P)	E3MMR6 H2UW17 F7B7H5	ci 000024251	4.89	61.38	307.70	hydrogen peroxide catabolic process
		ci 000086362	2.97	11685.83	34741.92	
		ci 000023954	2.24	197.98	454.83	
Glutathione peroxidase (P)	Q4PMF0 I0YYP3 F8WFK6	ci 000014039	17.86	360.96	6501.57	oxidation-reduction process
		ci 000070357	9.14	1.05	19.30	
		ci 000009744	2.06	1526.19	3157.91	
Glutathione S-transferase (P)	Q8H1P7 B7QC64 B7PG84 B7PB71 B7PR98 B7PR99 B7PRA0	ci 000026428	3.39	13.13	49.14	aromatic amino acid family metabolic process
		ci 000031927	3.25	7.34	27.18	
		ci 000089852	2.87	3417.49	9839.08	
		ci 000076478	2.39	1133.35	2715.66	
		ci 000027976	23.34	3.15	98.26	
		ci 000009670	4.84	19.94	102.12	
		ci 000080474	6.66	1347.11	9004.33	
		ci 000031669	4.11	100.17	419.46	
		ci 000013468	14.46	3.15	61.49	
Oxidoreductase yrbE	E1ZZT5	ci 000024001	3.16	18.38	63.07	oxidation-reduction process
Superoxide dismutase [Cu-Zn] (P)	I4DMD1	ci 000020892	6.02	3.67	256.69	removal of superoxide radicals
Thioredoxins (P)	Q4S0R6 O44366 F7D6R2 B5FXI6	ci 000079190	2.62	1309.20	3442.34	oxidation-reduction process
		ci 000044986	6.20	3.14	27.90	
		ci 000020415	3.76	2237.08	8635.74	
		ci 000090113	2.96	574.37	1709.60	
HEAT SHOCK RESPONSE AND PROTEIN FOLDING						
10 kDa heat shock protein, mitochondrial	F4X041	ci 000089910	3.68	1032.47	3829.60	protein folding
Calreticulin (P)	D7FZD0 D3BIF1 F7E4B5	ci 000084379	6.51	10802.62	70614.63	protein folding
		ci 000024543	5.96	142.72	864.10	
		ci 000022819	17.84	1.05	36.95	
Chaperone protein DNAj (HSP70 family)	B7QB03	ci 000006079	4.92	8.92	50.77	protein folding
DnaJ protein subfamily B (HSP70 family)	E5SN67	ci 000090121	4.89	372.32	1838.04	protein folding
Endoplasmic (HSP90 family)	F4WWD9 E5S1V7	ci 000033457	7.24	5.25	45.40	protein folding response to stress
		ci 000003052	5.15	25.19	136.38	
FK506-binding protein	B4KCX9 C1BQZ6	ci 000080248	3.93	147.91	590.11	protein folding
		ci 000023780	10.39	0.00	12.69	
		ci 000018934	2.14	1673.46	3589.66	

GRP94 (HSP90 family)	A7YAU9	ci 000073405	8.95	1.05	18.87	protein folding response to stress
Grpe protein	B7PPJ5	ci 000015207	2.20	1066.93	2351.60	protein folding
Heat shock protein 90 family (P)	E1B2T4 E4WZX2 H9JT91 B4MJZ2	ci 000017317	3.63	57.71	214.47	protein folding response to stress
		ci 000077634	8.18	3485.35	28589.21	
		ci 000014902	3.79	1965.86	7468.76	
		ci 000027999	17.11	3.15	71.63	
Heat shock protein family H (HSP70 family)	E1BT08	ci 000076271	2.38	2251.97	5376.10	protein folding
Heat shock protein (HSP70 family)	Q4Z3B0	ci 000026532	7.09	1.57	19.03	protein folding
HSP70 family (P)	B7PPP4 Q3TKF8 G3MRN8	ci 000079394	2.41	288.43	701.33	protein folding
		ci 000009720	6.53	23.07	168.25	
		ci 000076959	2.42	583.36	1416.10	
		ci 000024193	7.17	6.30	53.26	
Hypoxia up-regulated protein 1 (HSP70 family)	F7ABW6	ci 000009064	2.67	1487.43	3988.69	protein folding
Molecular chaperone	F0ZEN2	ci 000044781	5.22	3.15	21.49	protein folding
Peptidyl-prolyl cis-trans isomerase (P)	B3RNT2 Q6XPZ6 B2ZWT5 C3YPY6 O44073 G3MQW7	ci 000010569	2.47	125.91	314.33	protein folding
		ci 000089853	6.85	871.91	6198.04	
		ci 000003179	6.58	5.24	42.09	
		ci 000070518	5.19	3.14	23.39	
		ci 000024469	3.90	5.25	24.13	
		ci 000006670	3.09	2475.87	7663.75	
Prefoldin, putative ptotein	B7QE91 B7Q0H9	ci 000020509	2.09	523.55	1097.14	protein folding
		ci 000013653	2.53	221.41	564.95	
Protein disulfide-isomerase (P)	E2RD86 C3YNT9 H2USA8	ci 000068708	9.24	1.05	19.37	cell redox homeostasis; protein folding response to stress
		ci 000068716	8.83	2.09	28.67	
		ci 000028655	7.40	2.10	23.21	
		ci 000089827	5.29	12479.13	66285.08	
Uncharacterized protein	G3MKV9 A5C537 B9MY68 H3FKC7 H2Z8C9	ci 000040437	2.40	15.75	40.12	protein folding
		ci 000089900	2.02	3899.76	7902.23	
		ci 000021580	6.05	283.83	1732.79	
		ci 000032755	3.45	4.19	17.70	
		ci 000089940	2.49	1876.84	4678.09	

Table 6. Selected up-regulated genes putatively involved in defense responses. FPKM values are the mean of the two biological replicates analyzed from each physiological condition (Unfed: UF-1 and UF-2; Fed: 48h-1 and 48h-2). For transcript names, sequences, FPKM data and other information see GenBank GFWV000000000 (BioProject: PRJNA401392) and Supplementary Table 2.

Protein	Entry	Contig name	FoldChange	FPKM Unfed	FPKM Fed	Biological process
Defensin	Q09JE6 A0SXP6	ci 000083012	163.42	56.20	17680.62	defense response
		ci 000051079	51.23	0.00	55.49	
Defensin B	B6UYJ6	ci 000003931	26.91	6.81	216.60	defense response
		ci 000037332	4.70	2.63	18.63	
Defensin D	B6UYJ7	ci 000083009	20.12	613.66	12485.83	defense response
		ci 000028779	8.99	7.86	81.55	
		ci 000008745	8.86	2334.29	20762.87	
		ci 000030329	7.44	395.82	3180.31	
		ci 000052240	7.24	8.39	68.71	
		ci 000083014	4.48	426.48	1928.21	
Ovatusin	B7XFT2	ci 000083008	1085.73	2.10	4041.93	defense response
		ci 000013443	137.47	115.43	16197.32	
		ci 000009157	15.38	3.16	359.31	
Persulcatusin	B7XFT1	ci 000022397	113.07	1.05	268.19	defense response
		ci 000073523	60.90	0.00	66.66	
		ci 000074271	31.68	0.00	31.49	
		ci 000073524	28.73	0.53	44.28	
Lysozyme	B3M961	ci 000007710	2.18	75.54	169.01	defense response

Table 1. Top 30 Pfam domain families represented in the transcriptome of the *Ornithodoros erraticus* midgut and among the up- or downregulated differential expressed genes (DEGs) after feeding. Complete results of Pfam analysis are shown in the Supplementary Table 3.

Pfam code	Name	Summary	No. domains in the transcriptome	No. domains upregulated	No. domains downregulated
PF00096	zf-C2H2	Zinc finger, C2H2 type	587	43	94
PF00076	RRM_1	RNA recognition motif	547	35	52
PF13894	zf-C2H2_4	C2H2-type zinc finger	425	33	70
PF13465	zf-H2C2_2	Zinc-finger double domain	416	31	64
PF12171	zf-C2H2_jaz	Zinc-finger double-stranded RNA-binding	409	33	58
PF00078	RVT_1	Reverse transcriptase	379	50	86
PF12874	zf-met	Zinc-finger of C2H2 type	341	23	54
PF00069	Pkinase	Protein kinase domain	339	51	48
PF00400	WD40	WD domain, G-beta repeat	328	56	31
PF07714	Pkinase_Tyr	Protein tyrosine kinase	311	47	43
PF00023	Ank	Ankyrin repeat	298	29	91
PF12796	Ank_2	Ankyrin repeats (3 copies)	295	31	86
PF05485	THAP	THAP domain	263	25	54
PF04156	IncA	IncA protein	252	46	34
PF13637	Ank_4	Ankyrin repeats (many copies)	252	25	69
PF13857	Ank_5	Ankyrin repeats (many copies)	251	24	65
PF13893	RRM_5	RNA recognition motif	251	29	47
PF13606	Ank_3	Ankyrin repeat	249	26	67
PF12756	zf-C2H2_2	C2H2 type zinc-finger (2 copies)	238	13	37
PF13909	zf-H2C2_5	C2H2-type zinc-finger domain	225	13	31
PF14531	Kinase-like	Kinase-like	207	34	30
PF01926	MMR_HSR1	50S ribosome-binding GTPase	201	33	29
PF13639	zf-RING_2	Ring finger domain	195	23	28
PF00005	ABC_tran	ABC transporter	193	39	33
PF07690	MFS_1	Major Facilitator Superfamily	191	67	26
PF00097	zf-C3HC4	Zinc finger, C3HC4 type	190	35	24
PF13923	zf-C3HC4_2	Zinc finger, C3HC4 type	190	25	34
PF02892	zf-BED	BED zinc finger	188	17	36
PF13920	zf-C3HC4_3	Zinc finger, C3HC4 type	183	23	36
PF05605	zf-Di19	Drought induced 19 protein (Di19), zinc-binding	181	20	23

Supplementary Fig. 1

Supplementary Fig. 2.