A Systems Level Analysis Reveals Transcriptomic and Proteomic Complexity in *Ixodes Ricinus* Midgut and Salivary Glands During Early Attachment and Feeding*\(^{\circ}\)**

Alexandra Schwarz‡‡‡, Stefan Tenzer§‡‡, Michael Hackenberg¶, Jan Erhart‡, Aslihan Gerhold-Ay¶, Johanna Mazur¶, Jörg Kuharev§, José M. C. Ribeiro**‡‡, and Michail Kotsyfakis‡‡‡§§

Although pathogens are usually transmitted within the first 24–48 h of attachment of the castor bean tick *Ixodes ricinus*, little is known about the tick’s biological responses at these earliest phases of attachment. Tick midgut and salivary glands are the main tissues involved in tick blood feeding and pathogen transmission but the limited genomic information for *I. ricinus* delays the application of high-throughput methods to study their physiology. We took advantage of the latest advances in the fields of Next Generation RNA-sequencing and Label-free Quantitative Proteomics to deliver an unprecedented, quantitative description of the gene expression dynamics in the midgut and salivary glands of this disease vector upon attachment to the vertebrate host. A total of 373 of 1510 identified proteins had higher expression in the salivary glands, but only 110 had correspondingly high transcript levels in the same tissue. Furthermore, there was midgut-specific expression of 217 genes at both the transcriptome and proteome level. Tissue-dependent transcript, but not protein, accumulation was revealed for 552 of 885 genes. Moreover, we discovered the enrichment of tick salivary glands in proteins involved in gene transcription and translation, which agrees with the secretory role of this tissue; this finding also agrees with our finding of lower tick t-RNA representation in the salivary glands when compared with the midgut. The midgut, in turn, is enriched in metabolic components and proteins that support its mechanical integrity in order to accommodate and metabolize the ingested blood. Beyond understanding the physiological events that support hematophagy by arthropod ectoparasites, we discovered more than 1500 proteins located at the interface between ticks, the vertebrate host, and the tick-borne pathogens. Thus, our work significantly improves the knowledge of the genetics underlying the transmission lifecycle of this tick species, which is an essential step for developing alternative methods to better control tick-borne diseases. *Molecular & Cellular Proteomics* 13: 10.1074/mcp.M114.039289, 2725–2735, 2014.

Similar to other hard ticks, *Ixodes ricinus* ticks feed on vertebrate host blood for several days, depending on their developmental stage. Nymphs of this tick species feed for up to 6 days, and adults for over a week, under laboratory conditions (1), making them unique among arthropod disease vectors because of their extensive length of attachment and feeding. During feeding, hard ticks overcome numerous host homeostatic mechanisms including vertebrate hemostasis and immunity. Intensive research over the past three decades has revealed that tick salivary secretion contains modulators of vertebrate coagulation (2), platelet aggregation, and complement activation, as well as substances that interfere with innate and adaptive host immunological mechanisms at the cellular and molecular level (3) in order to preserve blood flow and to prevent tick rejection. Although the effects of tick salivary secretion on host physiology are under investigation, very little is known about how the tick reacts during the first few hours of attachment. The tick has developed strategies to accommodate and metabolize the relatively large quantity of blood that it receives upon feeding. Although blood bolus volume is significantly reduced because of water excretion by the tick salivary glands, a hard tick will increase in total volume and mass by many tens of times its original size and...
weight during a meal (4). Feeding is likely to increase mechanical stress on tick midgut epithelial cells, especially because hard ticks, unlike mosquitoes and many other disease vectors, digest the blood within the intracellular environment of the gut epithelial cells (5).

The transmission of many tick-borne pathogens takes place within the first 24 to 48 h of tick attachment, or even earlier (6). Accordingly, an intervention that interrupts the very initial steps of tick feeding, blood metabolism, and pathogen transmission could be an effective measure to interfere with the disease transmission by ticks. But only very limited information is available regarding tick midgut and salivary gland proteome and transcriptomic composition in the first 24 h of tick attachment to the host (7). An integrated, systems-level view of the molecular events that mediate the establishment of a tick feeding cavity would therefore improve our understanding of tick attachment molecular mechanisms as well as the protein composition of the pathogen transmission interface. However, there is little genomics information available for most tick species, which hampers high-throughput studies of tick physiological responses (8, 9). In order to address this limitation, we herein extensively sequenced different transcriptome libraries to generate a sequence database which supported the application of a recently developed label-free data-independent acquisition based strategy (10) to study the proteome dynamics in our tissue collections. As a result of the study of the transcriptome and proteome dynamics simultaneously in the same tissue collections, we herein describe a systems-level view of gene expression in I. ricinus midgut and salivary glands over the first 24 h of tick attachment (for nymphs and adults) and up to 36 h for adults, because it is known that adult I. ricinus ticks feed slower than nymphal ticks (1). In doing so, we report the most comprehensive proteome coverage of this important class of pathogen vectors to date and begin to define the dynamic responses of these physiological compartments during the tick lifecycle.

**EXPERIMENTAL PROCEDURES**

Unless otherwise indicated, the protocols followed standard procedures (11) and all the experiments were performed at room temperature (25 ± 1 °C). All water used was of 18-megaohm quality, produced by a MilliQ apparatus (Millipore Corp., Prague, Czech Republic). Quantitative PCR and Illumina sequencing methodology is described as supplemental Experimental Procedures S1.

**Ethics Statement**—All animal experiments were carried out in accordance with the Animal Protection Law of the Czech Republic (§17, Act No. 246/1992 Sb) and with the approval of the Akademie Věd České Republiky (approval no. 161/2010).

**Ticks, Tissue Dissections, and Total RNA/protein Isolations—**I. ricinus feeding experiments, using 1080 nymphs and 420 adult females and males, were performed as previously described (1). Because we were interested in the discovery of transcripts expressed shortly after tick attachment and early during tick blood feeding, groups of either 30 nymphal and/or 10 female ticks were removed from the animals after three hours of tick attachment to the animals, followed by subsequent tick collections every three hours until nymphs had attached to the experimental animals for 24 h and adult females for 36 h. A total of 990 nymphal and 370 adult female salivary glands and midguts (including tissues of unfed nymphs and female ticks) were dissected in sterile DEPC/PBS buffer and transferred into lysis buffer RA1 of the NucleoSpin RNA XS or RNA II kits (Macherey-Nagel, Düren, Germany); all dissected tissues were stored at −70 °C.

RNA was extracted from ten pooled tissue samples representing the tick attachment periods 0–12 h, 12–24 h, or 24–36 h (three female salivary gland samples, two nymphal salivary gland samples, three female midgut samples and two nymphal midgut samples) according to the NucleoSpin RNA XS and RNA II kits manufacturer’s instructions. Total RNA concentration of all the samples was measured using the NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). All the RNA samples were precipitated with ethanol using a final concentration of 0.3 M Na-acetate, pH 5.2. The precipitated RNA was overlaid with 70% ethanol and stored at −70 °C. The same tissue samples were also subjected to proteomics analysis. Accordingly, the flow-through containing nymphal/female salivary gland or midgut proteins from the different samples after the total RNA isolation was collected according to the manufacturer’s instructions of the NucleoSpin RNA XS and RNA II kits and the samples pooled similarly as for the RNA samples resulting in ten samples. The protein concentrations of all the samples were measured using the Protein Assay Dye Reagent Concentrate (Bio-Rad, Prague, Czech Republic) and all the samples were stored at −70 °C until use.

**Library Construction—**Four samples mentioned above for 0–12 h and 12–24 h from nymphal salivary glands (SG) and midguts (MG), and six samples for 0–12 h, 12–24 h and 24–36 h for adult SG and MG were used to make Illumina Truseq libraries which were sequenced using a Illumina Hiseq 2000 machine and the raw data deposited at the Sequence Read Archives (SRA) of the National Center for Biotechnology Information (NCBI) under accessions SRR641305, SRR641306, SRR641307, SRR641308, SRR641309, SRR641327, SRR641328, SRR641329, SRR641330, and SRR641331.

**Read Assembly and Bioinformatics**—The reads indicated above, plus reads deriving from our previous work (SRA accessions SRR592662, SRR592663, SRR592664, SRR592665, SRR592674, SRR592675, SRR592676, and SRR592677) (12) were reassembled using Abyss (13, 14) andSOAPdenovo trans (15) with various k values and the assembly of the assemblies done with 15 iterations of a blastn and cap3 pipeline (16), generating 198,504 contigs larger than 150 nt, with an average size of 662 nt and L50 = 501. Coding sequences (CDS) were extracted based on larger open reading frames (ORF) containing a signal peptide and by contigs that matched protein sequences derived from the Swissprot, invertebrate Refseq and GenBank-extracted protein sequences having “acari” in their organism name. The 25,808 nonredundant set of CDS were mapped to a hyperlinked excel spreadsheet that charts these sequences to several database comparisons as well as the number of reads deriving from each of the 10 libraries (16, 17) . These 25,808 CDS supported the subsequent quantitative proteome analysis.

1 The abbreviations used are: SG, salivary glands; A, adult ticks; CDS, coding sequences; cs, cytoskeletal; det, detoxification; extmat, extracellular matrix and adhesion; Gb, giga base pairs; IEP, isoelectric point; imm, immunity related; met, metabolism; MG, midgut; N, nymphal ticks; ne, nuclear export; nr, nuclear regulation; ORF, open reading frames; pe, protein export machinery; pm, protein modification machinery; pr, proteasome machinery; ps, protein synthesis machinery; s, secreted proteins; SRA, Sequence Read Archives; st, signal transduction; tf, transcription factors; tm, transcription machinery; tr, transporters; uc/uk, unknowns; UPLC, ultra performance liquid chromatography.
Protein Digestion—Aliquots (25 µg) of protein extract were lyophilized to remove isopropanol and redissolved in lysis buffer (7 M urea, 2 M thiourea, and 2% CHAPS). Two biological replicates of each experimental sample were prepared. Subsequently, the proteins were digested using a modified FASP method (18). Briefly, the redissolved proteins were loaded on the filter and detergents were removed by washing three times with buffer containing 8 M urea. The proteins were then reduced using DTT, alkylated using iodoacetamide, and the excess reagent quenched by addition of additional DTT and washed through the filters. Buffer was exchanged by washing with 50 mM NH₄HCO₃ and the proteins were digested overnight with trypsin (Trypsin Gold, Promega, Madison, WI) at an enzyme to protein ratio of 1:50. After overnight digestion, the peptides were recovered by centrifugation and two additional washes using 50 mM NH₄HCO₃. Flow-throughs were combined, lyophilized, and redissolved in 20 µl 0.1% formic acid by sonication. The resulting tryptic digest solutions were diluted with aqueous 0.1% v/v formic acid to a peptide concentration of 200 ng/µl and spiked with 25 fmol/microliters of enolase 1 (Saccharomyces cerevisiae) tryptic digest standard (Waters Corporation, Manchester, UK).

Mass spectrometric analysis of tryptic peptides was performed using the nanoAcquity system (Waters Corporation) equipped with a BEH C18 1.7 µm, 75 µm x150 mm analytical reversed-phase column (Waters Corporation) in direct injection mode as described previously (19). A 0.2 µl sample (40 ng of total protein) was injected per technical replicate (three technical replicates were performed for each of the two biological replicates of the ten different experimental conditions; a total of sixty LC-MS experiments). Mobile phase A was water containing 0.1% v/v formic acid, whereas mobile phase B was ACN containing 0.1% v/v formic acid. Peptides were separated with a gradient of 3–40% mobile phase B over 90 min at a flow rate of 300 nL/min, followed by a 10 min column rinse with 90% mobile phase B. The columns were re-equilibrated at initial conditions for 15 min. The analytical column temperature was maintained at 55 °C. The lock mass compound, [Glu1]-fibrinopeptide B (100 fmol/µl), was delivered by the auxiliary pump of the LC system at 300 nL/min to the reference sprayer of the NanoLockSpray source of the mass spectrometer.

For each protein, we made a global comparison of the null model with the model containing the variables developmental stage, tissue, and feeding time, we applied a multivariable linear mixed effects model with log-transformed ppm values for each protein. The calculated p values are the adjusted FDRs using the Benjamini-Hochberg procedure (23). In particular, we compared nymphs with adults, salivary glands with midguts, and for the time points we made two comparisons: first, the 24 h of feeding time point was compared with 12 h of feeding time point, and afterward the 36 h of feeding time point was compared with the 12 h of feeding time point (in the case of adult ticks only). Additionally, for each protein we calculated a linear model without covariates, a so-called null model. For each protein, we made a global comparison of the null model with the model containing the variables developmental stage, tissue, and feeding time using ANOVA. The changes in protein amounts were selected as statistically significant by the Bonferroni corrected p values with a local significance level of 3.13 e-5. The calculations were performed using R 2.15.2.

Data Access—Public disclosure of the sequences: This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession number [DDBJ/EMBL/GenBank: GANP00000000]. The version described in this paper is the first version, [DDBJ/EMBL/GenBank:GANP01000000] (Subid: SUB334716, Bioproject: PRJNA217984]). All the identified protein and transcript annotations as well as their quantity dynamics are publicly available in supplemental Link S1: http://exon.niaid.nih.gov/Transcriptome/Ixodesricinus/ Supplemental-link1.xlsx.
RESULTS

Improving the Transcriptome Coverage—We have previously obtained sequencing data from four normalized cDNA samples originating only from the salivary glands of adult and nymphal ticks feeding on the animal host until repletion (1). Herein we focused on gene expression at the very initial steps of tick attachment and moreover we aimed to study the tick physiological response in both the tick midgut and the salivary glands. Thus, our initial transcriptome coverage was very low because of the different experimental design of our previous project (1). Accordingly, we improved the coverage of our transcriptomes by the generation of an additional 315 million Illumina sequence reads of these four cDNA samples (29.1 Gb of new sequence reads). In addition, non-normalized RNAs/cDNAs originating from our tissue collections were obtained and sample-specific Illumina libraries were constructed and extensively sequenced (ten Illumina libraries from our ten experimental samples, 270 million Illumina reads in total; 35.4 Gb of total raw sequence reads). The raw sequence reads from our previous work (1), together with the 64.5 Gb of new raw sequence reads, were then assembled and 25,808 Coding sequences were extracted (please see Experimental Procedures) which supported the successful application of a recently developed data-independent acquisition-based nanoUPLC-MS protocol (10) for the proteomic analysis of our samples.

The Proteome of Tick Salivary Glands and Midgut is Highly Complex—Using a novel nanoUPLC-MS approach for small quantities of protein samples (20 microgramms) (10) we identified 17,769 peptides representing 1510 tick-derived proteins in tick salivary glands and midgut samples taken from the nymph and adult specimens at 12, 24, and 36 h (36h only from adults) postattachment (supplemental Link S1). Of these proteins, 1291 were common to all tissue/age combinations analyzed (Fig. 1A) when a stringent identification threshold was used (proteins containing a minimum of two peptides, FDR < 1%); the majority of proteins (75%) were identified by at least four peptides, and 40% by nine or more peptides (Fig. 1B). The dynamic range of the identified proteins spanned four
orders of magnitude (Fig. 1C) with excellent reproducibility between technical (Fig. 1C) and biological (Fig. 1D) replicates, indicating high reproducibility of sample preparation and robustness of the nanoUPLC-MS workflow. The detected proteins covered isoelectric points (IEPs) ranging from 3.31 to 12.48 and molecular weights from 5.7 kDa to 637 kDa.

Transcriptomic and Proteomic Data are Divergent but Correlated—We subjected all ten different tick tissue collections to parallel extraction of the total transcriptome and proteome; thus, next, we compared the relative amounts of proteins with their respective transcript levels in the same tissue, as reflected by the total amount of Illumina reads per transcript within each sample. But before further proceeding downstream, we ensured with Real-Time quantitative PCR that Illumina reads counting reflects the transcript accumulation in the tissue collections of our study (please see supplemental Data S1 and the related supplemental Data S2).

Transcript levels displayed a slightly wider dynamic range than proteins of about five orders of magnitude (supplemental Data S3). Notably, the relative abundance of proteins (as determined by the TOP3 approach) and the normalized Illumina reads were not significantly correlated for most of the analyzed samples (supplemental Data S3), consistent with other studies comparing protein and transcript levels in specialized cells such as thrombocytes (26). Although absolute levels correlated only poorly, relative changes in expression (expressed as the log2 ratio of the relative abundance of protein or transcript levels between two samples) were correlated when comparing salivary gland and midgut tissues in adults ticks (R² = 0.452) (Fig. 2A) and nymphs (R² = 0.331)
Systems Biology Analysis of *Ixodes ricinus* Attachment

(Fig. 2B) (averaged over all timepoints and biological replicates). Correlations between developmental stage-dependent changes (adult *versus* nymph) were weaker, but still significant in both salivary glands ($R^2 = 0.161$) (Fig. 2C) and midgut ($R^2 = 0.059$) (Fig. 2D). No significant correlations were observed in time-dependent (24h *versus* 12h postattachment) changes in proteome and transcriptome levels in any tissue analyzed (Fig. 2E and Fig. 2F), indicating that proteomic and transcriptomic dynamics are likely to have different kinetics. The proteome is highly conserved upon tick attachment.

Our initial analyses of relative changes in expression did not take variations between technical and biological replicates into account. In order to identify significant differences between samples at the proteomic level, we constructed a statistical model to analyze the dynamics of all the identified proteins and we considered in the statistical model all the varying experimental parameters and the variations in the proteome analysis methodology (*i.e.* protein identification consistency and variability in estimated protein amounts between technical and biological replicates). Most importantly, the statistical model also corrected for multiple hypothesis testing. Although the midgut and salivary glands are divergent in morphology and physiological function, their individual proteomic composition was highly conserved, with 885 of 1510 protein identifications (58.6% of total) not significantly changing in quantity between midgut and salivary glands (Fig. 3A). Less than 2.5% of proteins significantly altered in quantity solely as a function of feeding time or developmental stage (37 out of 1510; Fig. 3B and Supplemental Link 1). Moreover, the 20 most abundant proteins (of arthropod origin) were conserved across all experimental samples, suggesting a tightly conserved proteome in the tick salivary glands and midgut upon tick attachment to the host (ssal data supplement).

The Proteome is Highly Stable upon Tick Attachment but Transcrip
tomic Dynamics are Highly Variable—The dynamics of the 885 proteins that did not show any statistically significant difference between the different experimental conditions are represented as a heatmap in supplemental Link S2. Although slight fluctuations are occasionally observed between different experimental conditions, no statistically significant differences were detected. We next estimated the corresponding transcriptome dynamics in the same tissues for the genes that encode for these 885 proteins using Illumina sequencing read counting from our sample-specific non-normalized RNA/cDNA libraries; the results are represented as a heat map in supplemental Link S3. Comparison of the transcriptomic and proteomic dynamics shows discordance between the two sets of data and that there is tissue-specific gene activation at the transcriptional level which is not reflected at the proteome level. Accordingly, 416 of 885 genes (47%; supplemental Link S4) showed at least three times more total Illumina reads in the salivary glands compared with the midgut. Only 333 genes (37.6%; supplemental Link S6) had less than three times more Illumina reads in the two tissues and were more consistent with their protein expression levels (summarized in Fig. 3C). A more detailed analysis of the distribution of the 885 genes in different functional classes in relation to their transcriptomic and proteomic dynamics is presented in the corresponding paragraph of the supplemental Data S1 (Supplemental Results) and the related supplemental Data S4. Proteomic and transcriptomic dynamics are fully concordant in polypeptides with increased expression in the midgut, but this is not always the case for the polypeptides that are overexpressed in salivary glands.

Of 625 proteins identified as differing in quantity across different experimental conditions (Supplemental Link S7), the majority (59.7%; 373 out of 625) showed higher protein expression in the salivary glands (Fig. 3B; supplemental Link S7) and approximately one third (34.7%; 217 out of 625) showed higher protein expression in the midgut. Only 5.6% of proteins (35 out of 625) were differentially expressed between other experimental conditions (*i.e.* developmental stage or feeding time). Less than one third (29.5%; 110 out of 373) of polypeptide products overexpressed in tick salivary glands (Fig. 3D; supplemental Link S8) also had at least three times more total Illumina reads in the salivary glands (when compared with the midgut; see supplemental Link S9 for the related transcriptomic dynamics of these 110 genes). Most of these genes (83 out of 110; 75.5%) encode for secreted proteins and 11 genes encode for proteases (10%; Fig. 3F). Of note, only 72 out of these 110 genes had more than 15 Illumina reads in the midgut samples.

To our surprise, one quarter (24.9%; 93 out of 373; Fig. 3D) of genes with higher polypeptide expression in the salivary glands had at least three times more total transcript reads in the midgut samples (when compared with the salivary gland samples) (supplemental Link S10). The distribution of these 93 genes in the different functional classes is shown in Fig. 3F. Of note, 24 of these 93 genes (supplemental Link S10) were transcriptionally up-regulated in the midgut as feeding progressed (at least three times greater transcript abundance in midgut samples in adult ticks feeding for 36 h compared with 12 h). The remaining 170 (out of 373) genes (45.6%; Fig. 3D) had higher expression in the tick salivary glands at the proteome level, but Illumina reads remained relatively unchanged in the different tissue collections (supplemental Link S11). The distribution of these 170 genes in the different functional classes is shown in Fig. 3G. The proteomic and transcriptomic dynamics were highly concordant for the 217 genes displaying higher polypeptide expression in the tick midgut (Fig. 3B), with at least three times more total Illumina reads in the tick midgut samples compared with tick salivary glands for 194 out of 217 genes (89.4%; supplemental Link S12). The distribution of these 217 genes in the different functional classes is shown in Fig. 3H.
A complex gene expression pattern operates in *I. ricinus* midgut and salivary glands upon establishing a feeding cavity. A, 885 proteins were identified as having statistically unchanged protein levels in our experiments, whereas 625 proteins were identified as having statistically changed protein levels. B, Of the 625 genes that encode for proteins with statistically changed levels between experimental conditions, 217 showed higher protein expression in the midgut, 373 showed higher protein expression in the salivary glands, and only 35 proteins showed a difference of another type. C, Of the 885 genes that encode for proteins with stable expression, only 333 showed less than three times difference in total Illumina reads between all the midgut and the salivary gland samples. 136 out of the 885 genes showed at least three times greater numbers of total Illumina reads in the salivary glands and 416 genes showed at least three times greater numbers of total Illumina reads in the midgut. D, Of the 373 genes that show higher expression of polypeptide products in the salivary glands, only 110 showed at least three times greater numbers of total Illumina reads in the same tissue; 170 genes showed less than three times difference in total Illumina reads between the midgut and the salivary glands, and 93 genes showed at least three times greater numbers of total Illumina reads.
Different Genesets are not only Differentially Regulated but also Differentially Enriched for Genes Involved in Critical Cellular Functions—We have shown herein that a complex gene expression program operates in the tick midgut and salivary glands upon initiation of feeding by nymphal and adult ticks. The gene expression dynamics are further represented and visualized in Fig. 4A as a synthesis of three colors (red, blue, and green; please see the figure legend for further explanation). Most of the 110 “blue” genes (salivary gland expression at the proteome and transcriptome level) encode for secreted proteins and proteases, in agreement with the secretory nature of the tick salivary glands. These genes are differentially expressed in adults and nymphs at the level of transcription (but not at the proteome level, supplemental Link S9), suggesting that diversification in protein composition of nymphal and adult salivary secretion might take place later than 24 h of tick feeding (because protein expression changes between the two tissues were not observed over the time course of our study). Most of the 136 “light cyan” genes (no difference at the proteomic level, salivary gland expression at the transcriptomic level) belong to the same functional categories as the “blue” genes; thus we hypothesize that the difference in transcript accumulation in the salivary glands for these genes was not yet reflected as a difference in peptide accumulation in the same tissue, perhaps because of translational repression, secretion of the protein products, or time-dependent regulation of protein translation. Of interest, many “dark cyan” genes encode for proteins that play a role in protein synthesis, suggesting a dominance of translational machinery components in the salivary glands compared with the midgut; this finding potentially reflects that the major secretory role of this tissue in tick feeding physiology is mediated by an increased protein synthesis activity compared with the midgut. Therefore, we further tested this experimental finding by preparing small RNA sequencing libraries from the same tissue collections used in our study. As shown in Fig. 4B, indeed higher accumulation of tick tRNAs was observed in the midgut of adult ticks, verifying that the higher representation of protein synthesis components in the adult salivary glands coincides with the lower representation of tick tRNAs in the same tissue (and when compared with the midgut).

A large proportion of the 93 “magenta” genes (salivary gland expression at proteomic level and midgut expression at transcriptomic level) are transcriptionally induced in the midgut of adult ticks. Many “magenta” genes encode for components of the transcriptional machinery, suggesting a higher protein expression of transcriptional machinery components in the salivary glands compared with the midgut. Most of the 217 “red” genes (midgut expression at both transcriptomic and proteomic level) encode mainly for metabolic and secreted proteins, as well as for extracellular matrix proteins, signal transduction proteins, and detoxification proteins. This functional profile is consistent with the major physiological role of the tick midgut in “accommodating” and metabolizing ingested food. Most of the 416 “yellow” genes (midgut expression at the transcriptomic level and no difference at the proteomic level) encode for metabolic, signal transduction, and cytoskeletal proteins; given the importance of these cell functions in the tick midgut physiology (and the transcriptional pattern of these genes as shown in supplemental Link S4), we can speculate that many of the “yellow” genes might become “red” after 36 h of feeding. Finally, most of the 333 “green” genes (no difference at the transcriptomic or proteomic level) encode for ribosomal and housekeeping/metabolic proteins, which are expected to be relatively consistent (at both the transcript and protein level) in all of our experimental samples.

**DISCUSSION**

Herein we describe the first in-depth combined analysis of the proteome and transcriptome in the midgut and salivary glands of *I. ricinus* nymphs and adults upon attachment to the animal host. To our knowledge, these data represent the most comprehensive coverage of the proteome of an arthropod disease vector and provides the first time-resolved data describing changes in the proteome of salivary glands and midgut tissues shortly after tick attachment. Notably, this coverage was achieved without any protein and/or peptide prefractionation prior to LC-MS, indicating that the unbiased nanoUPLC-MS approach is highly sensitive. Over 1500 genes have been simultaneously profiled in the same sample for tissue-specific and developmental stage-specific expression analysis as a function of time at both the proteomic and transcriptomic level, albeit in an organism of unknown genome and with very low genetic information available. Of special note, our study was based only on transcriptome sequencing information and did not require any genomic information, demonstrating that proteomic analysis is doable in model organisms with limited genomic information, provided that sufficient transcriptome coverage of the studied samples is available. Accordingly, we herein demonstrate the feasibility in the midgut. E. Distribution in the different functional classes of the 110 genes which show higher poly peptide expression and at least three times greater transcript accumulation in the salivary glands. F. Distribution in the different functional classes of the 93 genes which show higher poly peptide expression in the salivary glands, but at least three times higher transcript accumulation in the midgut. G. Distribution in the different functional classes of the 170 genes which show higher poly peptide expression in the salivary glands, but less than three times difference in total Illumina reads between the midgut and the salivary glands. H. Distribution in the different functional classes of the 217 genes that show higher poly peptide expression in the midgut. The numerical data that supports the pie-charts presented in Figs. 3E–H are provided in the supplemental Table S1 supplemental data.
of similar approaches in various other research model organisms with an unknown genome.

We show that only a small proportion of genes are expressed either exclusively in the midgut or exclusively in the salivary glands at both transcriptomic and proteomic level (the “blue” and the “red” genes), but there is a good correlation between relative changes at proteomic and transcriptomic levels between these tissues in both nymphs and adult ticks (Fig. 2). Accordingly, whenever possible, the coupled gene expression profiling at both the transcriptome and the proteome level is suggested for the specific biological system, based on the herein presented data. The proteome composition is relatively conserved in both tissues during this early phase of tick feeding; however, we show that at the transcriptomic level that changes in gene expression occur in the salivary glands and the midgut in response to tick feeding, even at this early phase postattachment to the host, presumably because of the diverse physiological roles of the two tissues. In the salivary glands, tick attachment to the host is accompanied by a higher expression of transcriptional and translational machinery components (at the proteomic level) compared with the midgut (“dark cyan” and “magenta” genes); this major difference might account for the observed “phase” or “delay” when comparing gene activation (at the transcriptomic and proteomic level) between the tick midgut and salivary glands. Of course, this does not exclude other mechanisms to account for the observed differences in gene and protein expression dynamics, such as the tissue-specific presynthesis, secretion of proteins, or tissue-specific posttranscriptional and post-translational regulatory mechanisms.

Tick salivary secretion plays an important role in tick lifecycle (3) and herein we demonstrate that protein synthesis machinery components are overrepresented in tick salivary glands to support the specific function (protein secretion) of tick salivary glands.

Based on our data we propose that although the tick midgut and salivary glands share a conserved proteome early in I. ricinus attachment, the completion of the attachment to the host is accompanied by tissue-specific changes in gene expression, which may further diversify the proteome composition in both tissues later in tick feeding. In the time points of

**Fig. 4.** Differential expression in the tick tissues of genes that play a role in diverse physiological procedures. **A,** Color synthesis representation of the transcriptomic and proteomic dynamics of 1475 genes. The number of genes in each piece of the pie-chart is represented by the number inside each piece of the pie-chart. The color code explanation is provided in the inset of the figure along with the functional classes to which the majority of the genes of each piece of the pie-chart belong (in abbreviation format: cs-cytoskeletal; det-detoxification; extmat-extracellular matrix; met-metabolism; pm-protein metabolism; ps-protein synthesis; s-secreted; st-signal transduction; tm-transcription machinery). Red represents the genes that show midgut-specific expression at both the transcriptomic and proteomic level; blue the genes that show salivary gland-specific expression at both the transcriptomic and proteomic level; green the genes that do not show tissue specificity in their expression (at any level); dark cyan (blue + green = cyan) represents the 170 genes that show salivary gland-specific expression at the proteomic level but no tissue-specificity at the transcriptomic level; light cyan represents the 136 genes that show salivary gland-specific expression at the transcriptomic level, but not at the proteomic level; magenta (blue + red = magenta) represents the genes that show accumulation of their polypeptide products in the salivary glands, but at least three times greater transcript accumulation in the midgut; yellow (green + red = yellow) represents the genes that show at least three times greater transcript accumulation in the midgut, but no specificity in their expression at the proteomic level. In the same figure, apart from the total number of genes belonging to each piece of the pie-chart, we also show the functional classes to which the majority of the genes of each piece of the pie-chart belong. **B,** Fold-increased tRNA relevant abundance in Ixodes ricinus midgut when compared with the salivary glands. The X axis represents the small RNA read length that corresponds to tRNA.
our study, and at the transcription level, the salivary gland secretion starts already being diversified between adult and nymphal ticks and, moreover, the midgut proteome is getting enriched in proteins essential for its mechanical integrity and the fulfillment of its metabolic role. As well as contributing to knowledge about the gene and protein expression programs that are activated upon tick attachment to the vertebrate host, our work can serve as the basis for application development. *I. ricinus* transmits various pathogens such as Lyme disease, Rickettsia, Babesia spp., (27, 28) as well as tick-borne encephalitis viruses (29). Its transmission lifecycle is active in Western, Central, and Eastern Europe, and extends into Northern and Southern Europe (30), where this tick serves as one of the most important arthropod disease vectors. The conserved composition and the relatively stable dynamics of the proteome between midgut and salivary glands in both adult and nymphal ticks (in the timepoints of our study) combined with the fact that our study provides the first high coverage proteomic dataset on *I. ricinus* and identifies over 1500 proteins expressed at the critical interface of tick-borne pathogen transmission during early tick attachment and feeding open a new field toward the development of novel applications to better control tick-borne diseases. Given that the herein identified proteins are expressed at the very earliest stages of tick infestation in the two main tissues involved in tick hematophagy suggests a potential to be used for antitick (or transmission blocking) vaccine development or the development of tick exposure diagnostics. The two tick compartments (salivary glands and midgut) differ regarding their enrichment in protein components that play a role in diverse cellular functions which support their diverse physiological roles (for example increased protein synthesis in the salivary glands, increased metabolic activity in the midgut). Although our work provides essentially novel information toward the development of novel methods to better control tick-borne diseases, the development of these methods is a subject of future research and beyond the scope of this work.

Acknowledgments—We thank the editorial assistance from Nextgenediting (www.nextgenediting.com) and the anonymous reviewers for their critical and constructive comments. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

* This work was supported by the Grant Agency of the Czech Republic (GACR grant P502/12/2409 to MK), by the Academy of Sciences of the Czech Republic (grant no. Z60220518 to Biology Center), and by the 7th Framework Program of the European Union (EU FP7; Marie Curie Reintegration grant PIRG07-GA-2010-268177 to MK). Library constructions and Illumina sequencing of all the samples were supported by the European Union, 7th Framework Program, project TRANSCAN (FP7-INFRASTRUCTURES-2008-228403 to MK). This publication reflects only the authors’ views and the European Union is not liable for any use that maybe made of the information contained herein. This work was supported by the Alexander von Humboldt Foundation (Feodor Lynen Research Fellowship to AS) and by the Forschungszentrum Immunologie (FZI) and the Naturwissenschaftlich-Medizinische Forschungszentrum (NMFZ) of the Johannes Gutenberg University Mainz (support to ST). This work was supported by the Academy of Sciences of the Czech Republic (Jan Evangelista Purkyne Fellowship to MK) and from the European Molecular Biology Organization (EMBO Short Term Fellowship, ASTF number 482–2012 to MK). This work was supported by the Intramural Research Program of the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health where JMCR is an employee. Because JMCR is a government employee and this is a government work, the work is in the public domain in the United States. Notwithstanding any other agreements, the National Institutes of Health reserves the right to provide the work to PubMedCentral for display and use by the public, and PubMed-Central may tag or modify the work consistent with its customary practices. Rights can be established outside the United States subject to a government use license.

†† These authors contributed equally to this work.

Author contributions: The work presented here was a collaboration between all authors. MK defined the research theme. AS, ST, MH, JK, JMCR, and MK designed the methods and experiments, performed the laboratory and bioinformatics experiments, analyzed the data, interpreted the results, and wrote the paper. AGA and JM designed and performed the statistical analysis of the proteomic data; JE carried out tick feeding experiments. All authors have contributed to, seen, and approved the manuscript.

Competing Interest: The authors have declared that no competing interests exist.

REFERENCES

1. Schwarz, A., von Reumont, B. M., Erhart, J., Chagas, A. C., Ribeiro, J. M., and Kotsyfakis, M. (2013) *De novo Ixodes ricinus salivary gland transcriptome analysis using two next-generation sequencing methodologies*. *Faseb J* **27**, 4745–4756

2. Omeljan, J., Calvo, E., Pedra, J. H., Franciscetti, I. M., and Kotsyfakis, M. (2013) *Tick salivary secretion as a source of anthemostatics*. *J. Proteomics* **75**, 3842–3854

3. Kazimirova, M., and Stibanianova, I. (2013) *Tick salivary compounds: their role in modulation of host defences and pathogen transmission*. *Front. Cell. Infect. Microbiol.* **3**, 43

4. Toutouni, L. N., Gern, L., and Aeschlimann, A. (1995) Biology of *Ixodes (Pholeoixodes) hexagonus* under laboratory conditions. Part II. Effect of mating on feeding and fecundity of females. *Exp. Appl. Acarol.* **19**, 233–245

5. Sojka, D., Franta, Z., Horn, M., Caffrey, C. R., Mares, M., and Kopacek, P. (2013) *New insights into the machinery of blood digestion by ticks*. *Trends Parasitol* **29**, 276–285

6. Hynote, E. D., Mervine, P. C., and Stricker, R. B. (2012) Clinical evidence for rapid transmission of Lyme disease following a tickbite. *Diagn. Microbiol. Infect. Dis.* **72**, 188–192

7. Narasimhan, S., Deponte, K., Marcanzono, N., Liang, X., Royce, T. E., Nelson, K. F., Booth, C. J., Koski, B., Anderson, J. F., Kantor, F., and Fikrig, E. (2007) Immunity against *Ixodes scapularis* salivary proteins expressed within 24 hours of attachment thwarts tick feeding and impairs Borrelia transmission. *PloS One* **2**, e4651

8. McNally, K. L., Mitzel, D. N., Anderson, J. M., Ribeiro, J. M., Valenzuela, J. G., Myers, T. G., Godinez, A., Wolfinbarger, J. B., Best, S. M., and Bloom, M. E. (2012) Differential salivary gland transcript expression profile in *Ixodes scapularis* nymphs upon feeding or flavivirus infection. *Ticks Tick Borne Dis.* **3**, 18–26

9. Stutzner, C., van Zyl, W. A., Olivier, N. A., Richards, S., and Martz-Olivier, C. (2013) Gene expression profiling of adult female tissues in feeding *Rhi-
18. Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal de novo salticid gaster gland transcriptome analysis using two next-generation sequencing methodologies. *Faseb J.* 27, 4745–4756

19. Birol, I., Jackman, S. D., Nielsen, C. B., Qian, J. Q., Varhol, R., Stazyk, G., Morin, R. D., Zhao, Y., Hirst, M., Schein, J. E., Horsman, D. E., Connors, J. M., Gascoyne, R. D., Marra, M. A., and Jones, S. J. (2009) De novo transcriptome assembly with ABySS. *Bioinformatics* 25, 2872–2877

20. Geromanos, S. J., Vissers, J. P., Silva, J. C., Dorschel, C. A., Li, G. Z., Gorenstein, M. V., Bateman, R. H., and Langridge, J. I. (2009) The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics* 9, 1683–1695

21. Silva, J. C., Denny, R., Dorschel, C. A., Gorenstein, M., Kass, I. J., Li, G. Z., McKenna, T. T., Nold, M. J., Richardson, K., Young, P., and Geromanos, S. (2009) Quantitative proteomic analysis by accurate mass retention time pairs. *Anal. Chem.* 77, 2187–2200

22. Giles, K., Pringle, S. D., Worthington, K. R., Little, D., Wildgoose, J. L., and Bateman, R. H. (2004) Applications of a laboratory-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom.* 18, 2401–2414

23. Patzig, J., Jahn, O., Tenzer, S., Wichert, S. P., de Monasterio-Schrader, P., Rosfa, S., Kuharev, J., Yan, K., Bornmüller, U., Bremer, J., Aguzzi, A., Orfaniotou, F., Hesse, D., Schwab, M. H., Mobius, W., Nave, K. A., and Werner, H. B. (2011) Quantitative and integrative proteome analysis of peripheral nerve myelin identifies novel myelin proteins and candidate neuropathy loci. *J. Neurosci.* 31, 16369–16386

24. Bradshaw, R. A., Burlingame, A. L., Carr, S., and Aebersold, R. (2006) Reporting protein identification data: the next generation of guidelines. *Mol. Cell. Proteomics* 5, 787–788

25. Hochberg, Y., and Benjamini, Y. (1990) More powerful procedures for multiple significance testing. *Stat. Med.* 9, 811–818

26. Burkhart, J. M., Vaudel, M., Gambaryan, S., Radau, S., Walter, U., Martens, L., Geiger, J., Sickmann, A., and Zahedi, R. P. (2012) The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways. *Blood* 120, e73–e82

27. Florin-Christensen, M., and Schnittger, L. (2009) Piroplasmids and ticks: a long-lasting intimate relationship. *Front. Biosci.* 14, 3064–3073

28. Stanek, G. (2009) *Pandora’s Box: pathogens in Central Europe*. Wien Klin. Wochenschr. 121, 673–683

29. Mansfield, K. L., Johnson, N., Phipps, L. P., Stephenson, J. R., Fooks, A. R., and Solomon, T. (2009) Tick-borne encephalitis virus - a review of an emerging zoonosis. *J. Gen. Virol.* 90, 1781–1794

30. Medlock, J. M., Hansford, K. M., Bormane, A., Derdakova, M., Estrada-Pena, A., George, J. C., Golovilova, I., Jaenson, T. G., Jensen, J. K., Jensen, P. M., Kazimirnova, M., Oteo, J. A., Papa, A., Pfister, K., Plantard, O., Randolph, S. E., Rizzoli, A., Santos-Silva, M. M., Spong, H., Vial, L., Hendrickx, G., Zeller, H., and Van Bortel, W. (2013) Driving forces for changes in geographical distribution of *Ixodes ricinus* ticks in Europe. *Parasit. Vectors* 6, 1