



OPEN

# A combined transcriptomic approach to identify candidates for an anti-tick vaccine blocking *B. afzelii* transmission

Jos J. A. Trentelman<sup>1✉</sup>, Radek Sima<sup>2</sup>, Nicolas Krezdorn<sup>3</sup>, Julen Tomás-Cortázar<sup>4</sup>, Diego Barriales<sup>4</sup>, Katsuhisa Takumi<sup>5</sup>, Joe M. Butler<sup>1</sup>, Hein Sprong<sup>5</sup>, Michelle J. Klouwens<sup>1</sup>, Veronika Urbanova<sup>2</sup>, Sazzad Mahmood<sup>2,6</sup>, Peter Winter<sup>3</sup>, Petr Kopacek<sup>2</sup>, Juan Anguita<sup>4,7</sup>, Ondrej Hajdusek<sup>2</sup> & Joppe W. Hovius<sup>1</sup>

*Ixodes ricinus* is the vector for *Borrelia afzelii*, the predominant cause of Lyme borreliosis in Europe, whereas *Ixodes scapularis* is the vector for *Borrelia burgdorferi* in the USA. Transcription of several *I. scapularis* genes changes in the presence of *B. burgdorferi* and contributes to successful infection. To what extent *B. afzelii* influences gene expression in *I. ricinus* salivary glands is largely unknown. Therefore, we measured expression of uninfected vs. infected tick salivary gland genes during tick feeding using Massive Analysis of cDNA Ends (MACE) and RNAseq, quantifying 26,179 unique transcripts. While tick feeding was the main differentiator, *B. afzelii* infection significantly affected expression of hundreds of transcripts, including 465 transcripts after 24 h of tick feeding. Validation of the top-20 *B. afzelii*-upregulated transcripts at 24 h of tick feeding in ten biological genetic distinct replicates showed that expression varied extensively. Three transcripts could be validated, a basic tail protein, a lipocalin and an ixodegrin, and might be involved in *B. afzelii* transmission. However, vaccination with recombinant forms of these proteins only marginally altered *B. afzelii* infection in *I. ricinus*-challenged mice for one of the proteins. Collectively, our data show that identification of tick salivary genes upregulated in the presence of pathogens could serve to identify potential pathogen-blocking vaccine candidates.

*Ixodes* ticks are small parasitic arthropods that feed on the blood of vertebrate hosts. They are three host-ticks; their lifecycle consists of four life stages, egg, larva, nymph and adult, where the latter three each parasitizes different hosts. Ticks need to feed on blood of their hosts to obtain the nutrients and energy to develop into their next life stage or for successful reproduction. They do so by penetrating the skin of their host with their hypostome and, depending on the life stage, stay attached for 3–10 days to complete their blood meal. This feeding behavior presents a large window of opportunity for tick-borne pathogens to be transmitted to the host. Ticks are therefore only second to mosquitoes as the most important arthropod vectors for human disease. In contrast to the USA where *Ixodes scapularis* is the tick species most notorious for human disease<sup>1</sup>, in Europe, *Ixodes ricinus* is the tick that most affects human health<sup>2</sup>. *I. ricinus* is a vector for viruses, bacteria and protozoan parasites, and as such can cause a wide range of diseases, including tick-borne encephalitis, relapsing fever, anaplasmosis, babesiosis and most notably Lyme borreliosis.

Lyme borreliosis, also referred to as Lyme disease, is the most prevalent *I. ricinus*-borne disease; in Europe alone, over 65,000 cases of Lyme borreliosis are reported every year and some expect it to be 2–3 times higher due to underreporting<sup>3</sup>. Lyme borreliosis is caused by bacteria belonging to *Borrelia burgdorferi* sensu lato (s.l.) group and in Europe, *Borrelia afzelii* has the highest incidence rate. In humans, it is associated with (chronic)

<sup>1</sup>Center for Experimental and Molecular Medicine, Amsterdam Infection and Immunity, Amsterdam UMC, Location Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. <sup>2</sup>Biology Centre, Institute of Parasitology, Czech Academy of Sciences, Ceske Budejovice, Czech Republic. <sup>3</sup>GenXPro GmbH, Frankfurt Innovation Center Biotechnology, Frankfurt am Main, Germany. <sup>4</sup>CIC bioGUNE-Basque Research & Technology Alliance, 48160 Derio, Spain. <sup>5</sup>National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands. <sup>6</sup>Faculty of Science, University of South Bohemia, Ceske Budejovice, Czech Republic. <sup>7</sup>Ikerbasque, Basque Foundation for Science, 48012 Bilbao, Spain. ✉email: j.j.trentelman@amsterdamumc.nl

|                                                          | R1 Reads    | R2 Reads    | Total Reads |
|----------------------------------------------------------|-------------|-------------|-------------|
| Uninfected <i>I. ricinus</i> nymphs                      | 37,960,637  | 37,960,637  | 75,921,274  |
| <i>B. afzelii</i> -infected <i>I. ricinus</i> nymphs     | 40,784,777  | 40,784,777  | 81,569,554  |
| <i>B. afzelii</i> -infected <i>I. ricinus</i> whole body | 44,465,827  | 44,465,827  | 88,931,654  |
| Uninfected <i>I. ricinus</i> whole body                  | 41,339,909  | 41,339,909  | 82,688,620  |
| Total                                                    | 164,555,551 | 164,555,551 | 329,111,102 |

**Table 1.** Summary of RNA sequencing reads after cleaning.

cutaneous manifestations of Lyme borreliosis<sup>4</sup>. *B. afzelii* is acquired by the larval tick during its first blood meal, it can survive in the tick to later life stages and can be transmitted with each following blood meal. However, given their smaller size nymphal ticks are less easy to be identified (visually and sensationally) than adult ticks and they are therefore considered to be the most clinically relevant life stage with regard to human disease<sup>5</sup>. It is commonly accepted that *B. burgdorferi* s.l. transmission starts approximately 16–36 h after attachment of the tick, transmission of *B. afzelii* starts earlier than *B. burgdorferi* sensu stricto (s.s.)<sup>6,7</sup>. In spite of this, it has been shown in a mouse model that *B. afzelii*-infected ticks need to feed for longer than 24 h to establish infection. *B. afzelii* is presumably transmitted through the saliva of the feeding tick, although alternative routes of infection have been proposed<sup>8</sup>.

The saliva of the tick is crucial for the long period of attachment and the successful completion of the blood meal. It contains proteins that interfere with host defense mechanisms through for instance immunosuppressive, anticomplement or antihemostatic roles. Indeed, animals repeatedly infested with ticks have antibodies against tick saliva and display so-called tick-immunity; ticks are less able to feed and/or are rejected<sup>9–11</sup>. As the host defense mechanisms are also essential to prevent and contain infection, these tick salivary gland proteins (TSGPs) greatly increase the odds of successful infection of the vertebrate host by *B. burgdorferi* s.l.-infected ticks, as it has been shown most notably for *I. scapularis*<sup>12–16</sup>. As a consequence, anti-tick immunity also protects against *B. burgdorferi* s.l. infection via tick bites and it has been shown that this anti-tick immunity can be transferred by serum<sup>17–20</sup>. These observations show the potential of anti-tick vaccines, by targeting tick proteins, specifically TSGPs, one could prevent tick feeding and/or pathogen transmission. Neutralization of specific TSGPs by antibodies indeed reduced *B. burgdorferi* s.s. infection *in vivo*<sup>18,19</sup>. As with all biological processes, the expression, translation and secretion of TSGPs is a dynamic process. The expression of these TSGPs is highly upregulated during the tick feeding process<sup>21–24</sup>, but it is also known that infection with *B. burgdorferi* s.s. induces alterations in gene expression that contribute to the successful infection of the host<sup>13–15,25–29</sup>. Based on their properties, TSGPs can be divided in large multi-gene families that have distinct functionalities, as reviewed before<sup>30</sup>. These multi-gene families are thought to be the result of gene duplication early in evolution<sup>31</sup>.

We used a combined transcriptomic approach to gain insight into the transcriptional changes within the salivary glands of *I. ricinus* during the complex interplay between the tick, the host and the pathogen. The respective strengths of both Massive Analysis of cDNA Ends (MACE) and RNAseq were combined to identify tick transcripts and subsequent processes influenced by *B. afzelii*. Gene expression in salivary glands of *I. ricinus* nymphs in different stages of feeding (unfed, 24 h and fully fed) and in different states of infection with regards to *B. afzelii* (infected and uninfected) were analyzed and characterized to identify specific TSGPs upregulated in *B. afzelii*-infected salivary glands. TSGPs upregulated in *B. afzelii*-infected salivary glands and those biologically validated were tested as *B. afzelii*-blocking anti-tick vaccines.

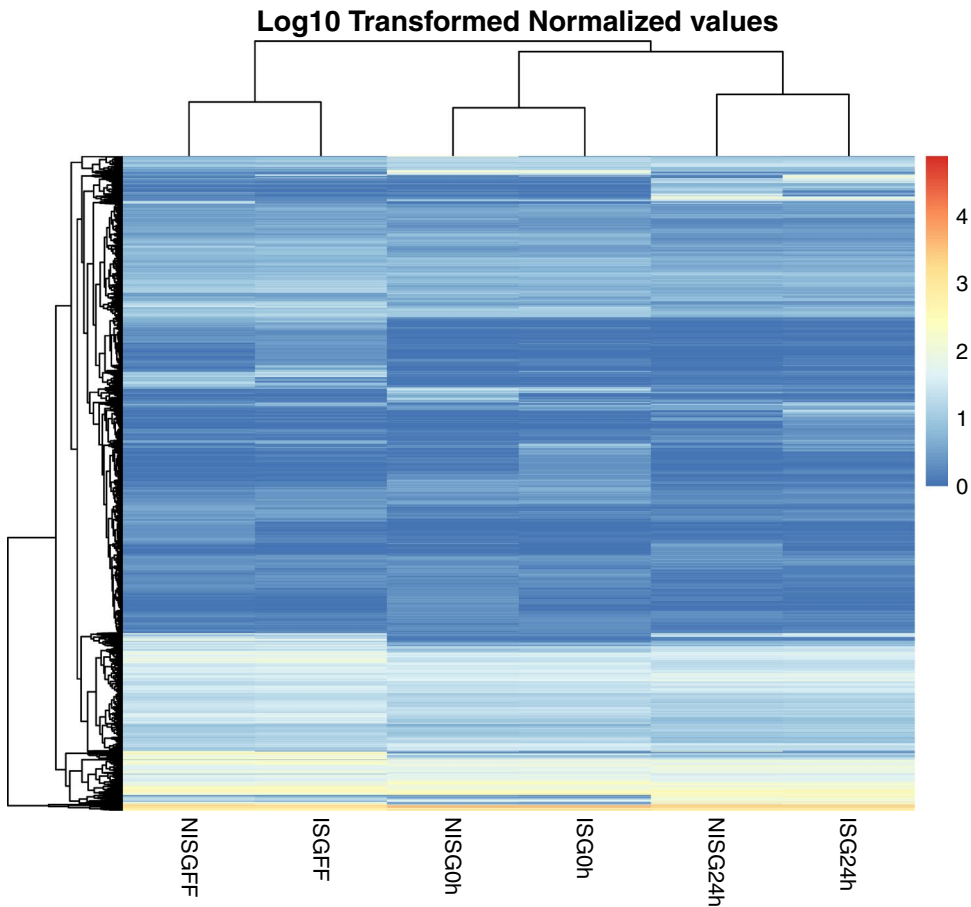
Results

**RNA sequencing.** In order to obtain long sequences to serve as our own framework for the annotation of the ensuing MACE analyses, RNA was prepared simultaneously for the construction of both the MACE and RNA sequencing libraries. Salivary gland and whole body RNA was isolated from *B. afzelii* CB43-infected *I. ricinus* nymphs and uninfected *I. ricinus* nymphs from the same parental lineage fed for 0, 24 or 72 h. RNA was pooled for all time points of *B. afzelii*-infected salivary glands, uninfected salivary glands, *B. afzelii*-infected whole body and uninfected whole body tick samples to obtain four cDNA libraries for RNAseq. The resulting cDNA libraries were used for paired-end sequencing and resulted in a total of 329,111,102 reads (Table 1) to be used for analysis, after elimination of duplicates and quality trimming. From these reads, 32,897 high quality contigs could be assembled. These formed our Master Reference exome (Master Reference) for the MACE analyses and represent an unprecedented source of *I. ricinus* sequence information.

**MACE analysis.** MACE was chosen as a quantitative tool as sequencing of the polyA captured cDNA molecules will result in one short read per molecule. In contrast, with RNAseq, one cDNA molecule will result in multiple reads. As such, MACE is excellent for detailed quantification of gene expression and has proven to be able to identify even low expressed genes<sup>32,33</sup>. RNA from salivary glands extracted at each time point (unfed, 24 h fed and fully fed) of *B. afzelii* CB43-infected (ISG) or uninfected (NISG) were used to prepare a total of six cDNA libraries for MACE. A total of 74.651.134 sequencing reads were processed (Table 2) and mapped against our Master Reference or assembled de novo and subsequently annotated against SwissProt, TrEMBL and the NCBI nucleotide database. Annotation resulted in the identification of a total of 93,096 unique transcripts of which on average 250–500 bp were covered by the MACE reads. We focused on transcripts that had more than one

| Library | Reads      |
|---------|------------|
| ISG0h   | 12,954,933 |
| ISG24h  | 15,916,749 |
| ISGFF   | 12,701,261 |
| NISG0h  | 12,449,936 |
| NISG24h | 9,553,944  |
| NISGFF  | 11,074,311 |
| Total   | 74,651,134 |

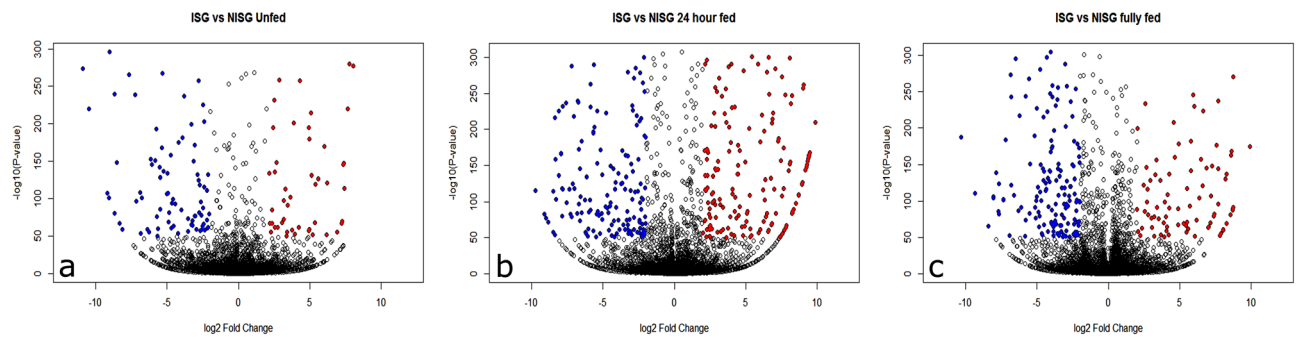
**Table 2.** Summary of MACE reads after cleaning.



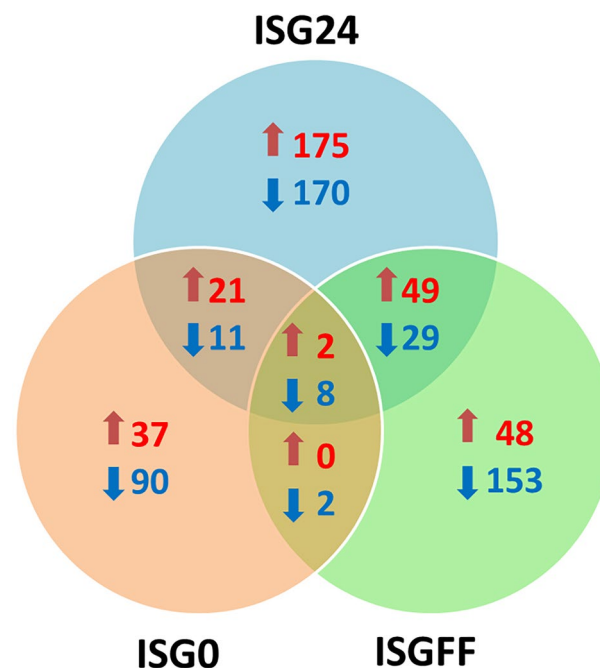
**Figure 1.** Unsupervised hierarchial cluster analyses of gene expression. Heatmap of log 10 transformed normalized reads illustrating gene expression of nymphal *I. ricinus* uninfected salivary glands (NISG) and *B. afzelii*-infected salivary glands (ISG) that were unfed (0 h), fed for 24 h (24 h) or fully engorged (FF). Each condition is represented in a single column. Gene expression is illustrated by color code, the color scale ranges from blue for low normalized reads to red for very high normalized reads.

normalized read (read per million reads) in at least one of the MACE libraries to reduce the background signals. As a result, the number of transcripts used for further analysis was 26,179 unique transcripts.

**Differential gene expression.** An unsupervised hierarchical clustering of all transcripts was performed (Fig. 1). From this analysis, it becomes clear that salivary gland gene expression was mostly affected by the stage of the feeding process. Differences in gene expression were most pronounced between early time points, 0 and 24 h fed, versus 72 h fed tick salivary glands (1790 and 1665 differentially expressed transcripts, respectively). Transcripts were considered to be differentially expressed when the change in gene expression was 4 times lower or 4 times higher, the corresponding  $p$  value  $< 1^{-50}$ . Although gene expression was largely driven by the feeding status of the ticks, *B. afzelii* infection also altered gene expression (Fig. 2); in *B. afzelii*-infected unfed salivary glands (ISG0), 60 transcripts were upregulated and 110 transcripts were downregulated. In 72 h-fed *B. afzelii*-



**Figure 2.** Volcano plot of comparison in salivary gland gene expression upon *B. afzelii* infection. (a) Relative gene expression in ISG vs NISG at 0 h. (b) Relative gene expression in ISG vs NISG at 24 h. (c) Relative gene expression in fully fed ISG vs NISG. Red dots are significantly upregulated genes, blue dots are significantly down regulated genes ( $\log_2\text{fold} < -2$  or  $> 2$  and the corresponding  $p$  value  $< 10^{-50}$ ).

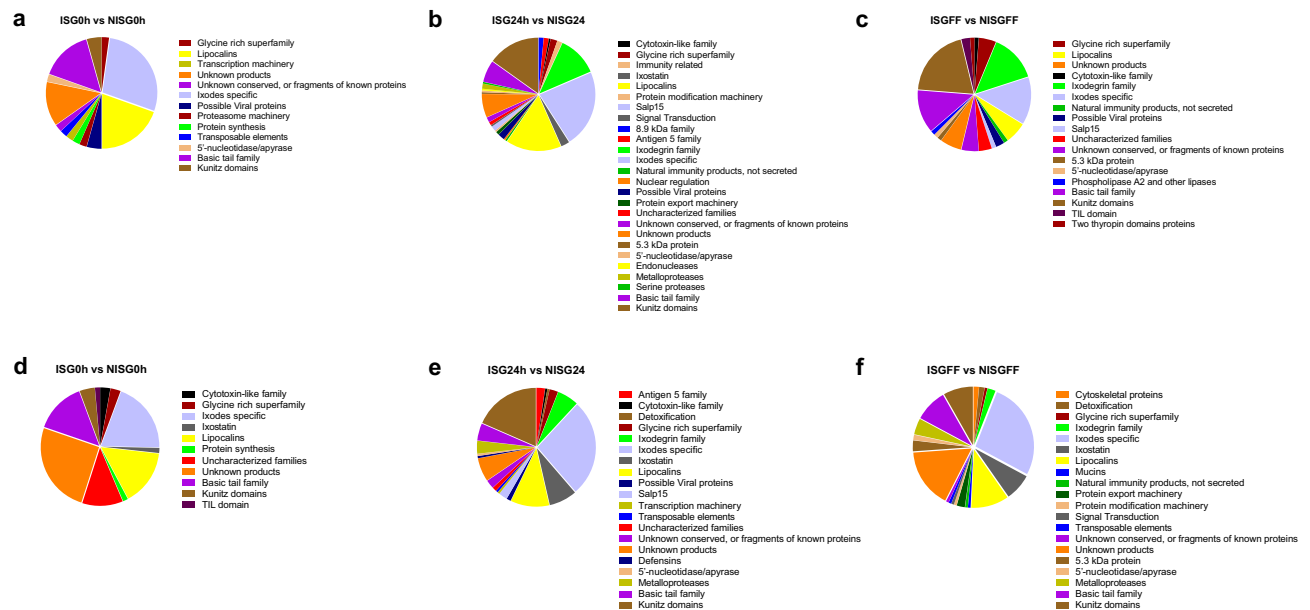


**Figure 3.** VENN diagram depicting differentially expressed tick salivary transcripts upon *B. afzelii* infection and their behaviour through time. Transcripts that are differentially up-regulated ( $> 2\log_2\text{fold}$  change in normalized reads,  $p < 10^{-50}$ ) are depicted in red. Transcripts that are differentially down-regulated ( $> -2\log_2\text{fold}$  change in normalized reads,  $p < 10^{-50}$ ) are depicted in blue.

infected salivary glands (ISGFF) 99 transcripts were upregulated, while 192 were downregulated. Interestingly, most transcripts were differentially expressed upon 24 h feeding in *B. afzelii*-infected salivary glands (ISG24); 247 transcripts were upregulated and 218 were downregulated. Only a fraction of the genes were upregulated or downregulated at all time points (Fig. 3). Overall, *B. afzelii* infection influenced the expression of 795 unique salivary gland transcripts ( $> 2\log_2\text{fold}$  change or  $< -2\log_2\text{fold}$  change,  $p < 10^{-50}$ ); 332 unique transcripts were up-regulated in one or more time points, whereas 463 unique transcripts were down-regulated in one or more time points. Interestingly, most transcripts that were affected by infection in a single time point only, were differentially expressed at 24 h of tick feeding (345 genes; 175 upregulated, 170 downregulated).

Thus, although differential gene expression in *I. ricinus* salivary glands was mostly driven by tick feeding, *B. afzelii* also influenced gene expression in *I. ricinus* salivary glands, and mostly at 24 h of tick feeding.

**Characterization of *B. afzelii*-induced differentially expressed tick salivary gland genes.** To provide more insight into the possible biological functions of the differentially expressed tick salivary gland transcripts, these were assigned to known tick protein families. To this end, the corresponding contigs were aligned (blastx) to contigs of a previously described *I. ricinus* bioproject<sup>23</sup>, in which genes were eloquently assigned to different families of tick proteins. Our contigs that had a match with contigs from the previously described



**Figure 4.** Distribution of differentially expressed transcripts over tick protein families for each time point. (a, b, c) Distribution of up-regulated transcripts ( $> 2 \log_2$  fold change,  $p < 1^{-50}$ ) over tick protein families in *B. afzelii*-infected salivary glands of unfed (ISG0h vs NISG0h), 24 h (ISG24h vs NISG24h) and 72 h (ISGFF vs NISGFF) fed nymphs respectively. (d, e, f) Distribution of down-regulated transcripts ( $< -2 \log_2$  fold change,  $p < 1^{-50}$ ) over tick protein families in *B. afzelii*-infected salivary glands of unfed (ISG0h vs NISG0h), 24 h (ISG24h vs NISG24h) and 72 h (ISGFF vs NISGFF) fed nymphs respectively.

bioproject, with an Expect value below 0.00001, were assigned to the respective tick protein family<sup>23</sup>. Using this strategy, 81% of the differentially expressed transcripts could be annotated to a tick protein family.

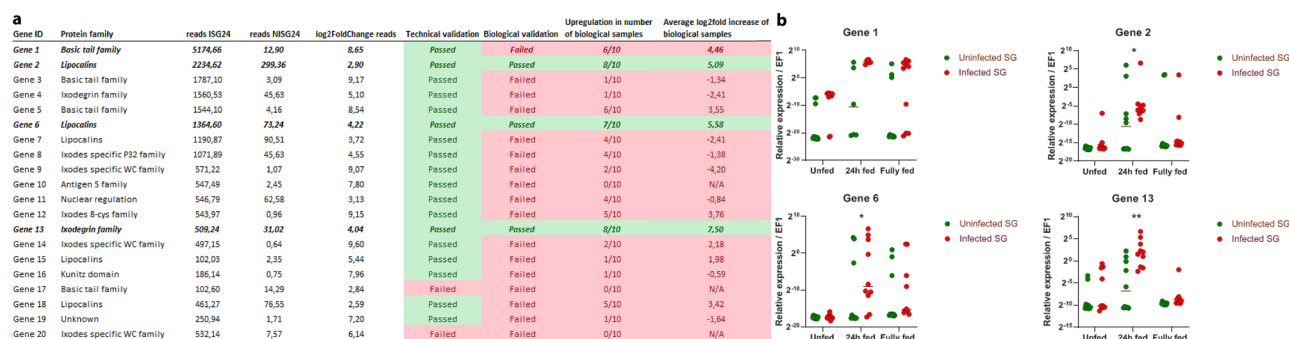
The functional annotation was limited to the main classes, only the classes of enzymes, antimicrobial peptides and protease inhibitor domains were divided into subclasses. Transcripts belonging to the glycine-rich superfamily, lipocalins, *Ixodes* specific family, and kunitz domain inhibitor family accounted for most of the transcripts upregulated by *B. afzelii* infection at any given time point (Fig. 4). Some tick protein families were only upregulated in ISG24h; most notably those related to immunity (1.46% of the upregulated transcripts at 24 h), ixostatin (2.44%), signal transduction related transcripts (0.49%), 8,9 kDa family (1.46%), antigen 5 family (1.46%), protein export machinery (0.98%), protein modification machinery (0.49%), metalloproteases (1.46%) and serine proteases (0.49%). Other families were upregulated at both ISG24h and ISGFF, those time points at which the tick is feeding and transmission of *B. afzelii* is taking place. Among these upregulated transcripts, members of the ixodegrin family (11.71% and 13.75% respectively) and Salp15 family (1.46% and 1.25% respectively) members were observed. In addition, although a few transcripts were upregulated in ISG0h (0.35% of upregulated transcripts), a marked increase of upregulated transcripts belonging to the kunitz domain inhibitor family were observed in ISG24h (15.12%) and ISGFF (20%) as well. Most of the transcripts upregulated at ISG24h belonged to the kunitz domain inhibitor (15.12%), ixodegrins (11.71%), *Ixodes* specific (22.44%) and lipocalin (16.10%) families.

Regarding the transcripts downregulated in infected salivary glands, the families affected at all time points were the glycine rich superfamily, *Ixodes* specific, lipocalins, kunitz domain families and transcripts that are considered as unknown products (with no homology to known sequences). Downregulated only in ISG24 were transcripts belonging to the Antigen 5 family (2.38% of the downregulated transcripts), Salp15 (2.38%), defensins (0.6%) and transcription machinery (0.6%).

Next to transcripts that were present in both uninfected and infected salivary glands, some transcripts could exclusively be detected in infected salivary glands, of which those only expressed in ISG24h are depicted in Supplemental Fig. 1. These transcripts were associated with the TIL—(Trypsin Inhibitor like cysteine rich) domain (4.17% of the transcripts only expressed in ISG24h compared to NISG24h), lipocalin (12.5%), Salp15 (4.17%), ixodegrin (16.67%), *Ixodes* specific (16.67%), ixostatin (8.33%) and, particularly, Kunitz domain families (33.33%).

Overall, *B. afzelii* was shown to affect *I. ricinus* salivary gland expression of transcripts encoding proteins belonging to multiple tick proteins families. Interestingly, we observed unique expression, as well as up-regulation and down-regulation, of transcripts within certain tick protein families, most notably *Ixodes* specific, lipocalins, basic tail protein, ixodegrin, kunitz domain inhibitor and ixostatin tick protein families.

**Selection of vaccine candidates; technical and biological validation.** Tick salivary transcripts upregulated upon infection with *B. afzelii* might be important for transmission of *B. afzelii* and/or subsequent successful infection of the vertebrate host. Therefore, proteins encoded by transcripts that were highly upregulated in ISG24 were considered as potential candidates for a *Borrelia* transmission blocking vaccine. Significantly



**Figure 5.** Technical and biological validation of top 20 vaccine candidates. **(a)** Top 20 genes highly upregulated in infected tick salivary glands at 24 h after onset feeding, ( $>2 \log_2$  fold change ISG24h vs NISG24h,  $p < 1 \times 10^{-50}$ ) were considered potential *Borrelia* transmission blocking vaccine candidates. Top 20 was ranked based on expression levels in infected salivary glands at 24 h as determined by MACE. Of the 10 biologically distinct tick pools used for biological validation, the number of pools that showed upregulation of the respective transcript are indicated as well as the average log2fold difference in all 10 tick pools. **(b)** Gene expression profiles of biologically validated Genes 1, 2, 6 and 13 in the salivary glands of 10 biologically distinct tick pools as determined by RT-qPCR. Elongation factor 1 alpha was used as a reference gene. Lines indicated median expression values. Significantly upregulated transcripts are indicated by \* (Friedman test paired analysis, Dunn's multiple comparison  $p < 0.05$ ).

upregulated genes ( $>2 \log_2$ ,  $p < 1 \times 10^{-50}$ ) were ranked based on expression levels determined by MACE and the 20 most abundantly expressed transcripts were selected for technical and biological validation. Primers were designed based on the nucleotide sequence identified by MACE and qRT-PCR was performed on the cDNA used for MACE (technical validation) or cDNA from tick pools derived from 10 genetically distinct ticks (biological validation). Technical validation showed that expression levels determined by qRT-PCR could confirm the MACE results for nearly all transcripts (Fig. 5), underscoring the robustness and accuracy of our approach. However, biological validation using cDNA from 10 genetically distinct tick pools showed marked variability in gene expression of the selected transcripts. Of the 20 selected transcripts, 3 genes were significantly upregulated in *B. afzelii*-infected tick salivary glands in most of the 10 genetically distinct tick pools; Gene 2, Gene 6 and Gene 13 (Fig. 5). In silico analysis showed that Gene 6 and Gene 2 are in fact highly similar; their sequence analysis showed 86% similarity at the amino acid level and Gene 6 appears to have a deletion compared to Gene 2. All 3 significantly upregulated transcripts encode a signal sequence and are likely to encode secreted proteins. Although not significantly upregulated at 24 h after the onset of feeding in the biological validation, Gene 1 was considered to be an interesting candidate. Gene 1 was only detected in 6 out of 10 tick pools, but in these tick pools Gene 1 was highly expressed upon infection at 24 h (Fig. 5). In addition, Gene 1 also encoded a signal sequence and showed a high degree of homology to basic tail proteins, although there were no conserved domains that might indicate possible functions of the encoded protein. Gene 2 and 6 were putative lipocalins and contain predicted histamine binding domains. Gene 13 was classified as a putative ixodegrin, containing a prokineticin domain and was part of the colipase-like superfamily. As Gene 2 and 6 were highly similar, Gene 1, 2 and 13 were selected for cloning and recombinant protein production in *E. coli*. For the selected targets the amino acid sequence, the predicted protein model, conserved domains and other characteristics are shown in Table 3.

**Transmission and vaccination studies.** Preliminary RNAi studies, with successful knock down of Gene1, 2, 6 and 13, in *B. afzelii*-infected nymphs fed on a small number of mice ( $n = 3$ ), did not show a significant reduction of tick feeding or *B. afzelii* infection (Supplemental Fig. 2). This result indicated that the absence of transcripts by itself was not enough to affect *B. afzelii* transmission. We next focused on vaccination studies where antibody-antigen interactions and complexes can lead to multiple effector mechanism that can block transmission. To this end, mice were vaccinated with recombinant proteins of Gene 1, 2, 13 or a combination of these antigens and subsequently challenged with *B. afzelii*-infected nymphal ticks. Vaccination was shown to be successful; antigen-specific total IgG levels could be detected after vaccination (Supplemental Fig. 3), although antibody levels against recombinant Gene 2 were significantly lower as compared to the other antigens (Mann-Whitney test,  $p < 0.05$ ). Vaccination with recombinant Gene 1 significantly reduced the number of infected mice tissues as determined by qPCR and although the number of mouse tissues infected as determined by culture was also lower, this effect was not significant nor was there a difference in the cumulative number of mice that were infected (Chi-square,  $p < 0.05$ ) (Table 4). For all other experimental groups, including the cocktail vaccination, no significant differences were observed in the spirochetal loads of the tissues nor in the number of infected mice (Supplemental Fig. 3).

## Discussion

To our knowledge, this is the first time that the relationship between *B. afzelii* and nymphal *I. ricinus* on the total transcript level of salivary glands is studied. In the current study, two different gene quantification tools have been combined to provide an unprecedented insight into the transcriptome of *I. ricinus* salivary glands. RNAseq

|         | Amino acid sequence                                                                                              | Predicted protein structure                                                       | Protein family     | Domain            | Signal peptide | N - glycosylation sites | O - glycosylation sites | Transmembrane helices    | GPI anchor | MHC I binding                                     | MHC II binding                                    | Linear B cell epitopes              |
|---------|------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|--------------------|-------------------|----------------|-------------------------|-------------------------|--------------------------|------------|---------------------------------------------------|---------------------------------------------------|-------------------------------------|
| Gene 1  | MMGLSTGVSLVSLGSAADCCNGTSPS<br>QNDKSGCPNMGSESTYDCHYDQVCEY<br>VNWGAGLQNGVCHTSTTTPAPASGDTF<br>VYTRPKQKXKPKPKSGKSKKE |  | Basic tail protein | None              | Yes, AA1:22    | 1                       | 4                       | 1 AA1:22, signal peptide | No         | 6 high binding, 15 low binding peptides (mouse)   | 7 high binding, 23 low binding peptides (mouse)   | 5 predicted epitopes (average 16AA) |
| Gene 2  | MMGLSTGVSLVSLGSAADCCNGTSPS<br>QNDKSGCPNMGSESTYDCHYDQVCEY<br>VNWGAGLQNGVCHTSTTTPAPASGDTF<br>VYTRPKQKXKPKPKSGKSKKE |  | Lipocalin          | Histamine binding | Yes, AA1:20    | 7                       | 12                      | 0                        | No         | 44 high binding, 112 low binding peptides (mouse) | 14 high binding, 118 low binding peptides (mouse) | 10 predicted epitopes (average 9AA) |
| Gene 13 | MMGLSTGVSLVSLGSAADCCNGTSPS<br>QNDKSGCPNMGSESTYDCHYDQVCEY<br>VNWGAGLQNGVCHTSTTTPAPASGDTF<br>VYTRPKQKXKPKPKSGKSKKE |  | Isodegrin          | Prokinectin       | Yes, AA1:23    | 1                       | 3                       | 1 AA1:28, signal peptide | No         | 8 high binding, 31 low binding peptides (mouse)   | 24 low binding proteins (mouse)                   | 2 predicted epitopes (average 46AA) |

**Table 3.** In silico analysis of validated transcripts selected for vaccination studies. Amino acid sequences encoded on the transcripts Gene 1, 2 and 13 as determined by the ExPASy Translate tool. Protein structures were predicted using Phyre2 web portal and although confidence in the predicted model was low for Gene 1 (32% of residues modelled at > 90% confidence, 55% of the sequence is predicted disordered), confidence in the predicted model was good for Gene 2 and 13 (73% and 67% of residues modelled at > 90% confidence respectively). Proteins sequences were subsequently scanned for domains with InterProScan. Signal peptide, O- and N-glycosylation sites were predicted based on amino acid sequence by SignalP 5.0 server, NetOGlyc 4.0 Server and NetNGlyc 1.0 server, respectively. HTHMM v2.0 server was used to predict transmembrane helices and GPI-SOM to predict GPI-anchor. MHC class I and II binding peptides were predicted using NetMHCpan-4.1 and NetMHCIIpan-4.0 and linear B cell epitopes with BepiPred Linear Epitope Prediction 2.0. Columns colored in red are negative and those in blue are positive, while the others are neutral.

|                                       | Skin <sup>1</sup> | Bladder <sup>1</sup> | Skin <sup>2</sup> | Bladder <sup>2</sup> | Heart <sup>2</sup> | Joint <sup>2</sup> | Cumulative |
|---------------------------------------|-------------------|----------------------|-------------------|----------------------|--------------------|--------------------|------------|
| PBS                                   | 5/6               | 4/6                  | 6/6               | 6/6                  | 6/6                | 5/6                | 6/6        |
| Recombinant Gene 1                    | 3/6               | 3/6                  | 2/6*              | 2/6*                 | 2/6*               | 2/6*               | 5/6        |
| Recombinant Gene 2                    | 4/6               | 4/6                  | 4/6               | 4/6                  | 4/6                | 4/6                | 4/6        |
| Recombinant Gene 13                   | 4/6               | 4/6                  | 4/6               | 4/6                  | 4/6                | 5/6                | 5/6        |
| Recombinant Gene 1 + Gene 2 + Gene 13 | 5/6               | 4/6                  | 4/6               | 4/6                  | 5/6                | 5/6                | 6/6        |

**Table 4.** Number of *Borrelia*-infected mice as determined for each organ. *Borrelia* infection as determined by culture<sup>1</sup> or qPCR<sup>2</sup> and shown as number of positive mice/total mice. Cumulative infection was calculated as the number of mice that were positive in at least one of the organs either by culture or qPCR. Significance was calculated compared to the PBS groups and significant differences are indicated by \* (Chi-square,  $p = 0.04$ ).

is a powerful technique to obtain accurate and qualitative sequence information of transcripts, but fragmentation of the RNA molecules and sequencing of all fragments could lead to a bias in the quantification of longer transcripts<sup>34</sup>. MACE, on the other hand, only targets sequences from the 3' end of the sequence by capturing the RNA fragment containing the pol-A tail. As a result, sequence information might be partial (i.e. not providing sequence information of the whole gene sequence), but it provides a high resolution gene expression analysis, even revealing differential expression of low-abundant transcripts, which are beyond the scope of RNAseq or microarrays<sup>35</sup>. In addition, the TrueQuant method increases the reliability of quantification by eliminating PCR bias<sup>36</sup>. By combining RNAseq and MACE, the complete sequence information provided by RNAseq results in increased mapping accuracy of MACE reads, strengthening the highly accurate quantification by MACE. RNAseq analysis was performed using pooled RNA from nymphal *I. ricinus* salivary gland and whole bodies fed for different time points (0, 24 or fully fed), with or without *B. afzelii* infection, resulting in 32,897 high-quality contigs, which is similar to or higher than the number of transcripts reported by previous RNAseq projects<sup>37–39</sup>. MACE resulted in the quantification of 26,179 transcripts selected for further analysis. Technical validation by qRT-PCR using the MACE cDNA libraries, to determine the expression profiles of the 20 most abundantly expressed *B. afzelii*-induced *I. ricinus* salivary glands transcripts, corroborated the MACE expression profiles and clearly validated our findings.

As described previously, our results confirm that the feeding process greatly affected gene expression in tick salivary glands<sup>37–39</sup>. Although the feeding process is the main differentiator of gene expression, MACE analysis showed that the expression of hundreds of transcripts is significantly affected by *B. afzelii* infection. This could have multiple underlying mechanisms that are not mutually exclusive; firstly, the transcripts could be part of the tick immune response to *Borrelia* infection. Secondly, the expression could be altered by *Borrelia* infection to increase survival in the tick. Thirdly, the transcripts could be affected by *Borrelia* infection to increase transmission through saliva and infectivity in the mammalian host. Interestingly, only a few transcripts were upregulated in *B. afzelii*-infected salivary glands of unfed ticks. This fits the general assumption that there are few to none spirochete in the salivary glands at this time point as has been observed for *B. burgdorferi* s.s. as they are located in the midgut and still have to migrate to the salivary glands upon onset of feeding<sup>40</sup>. In addition, the expression

of hardly any transcript is affected in all three time points (0, 24 and FF), making it unlikely that the identified upregulated transcripts are involved in the tick immune response against *B. afzelii*. Most of the *B. afzelii*-induced differentially expressed transcripts were observed 24 h after onset of tick feeding. This coincides with the time point that *B. burgdorferi* s.l. is thought to have found its way into the tick saliva and starts to be transmitted to the host<sup>7</sup>. Indeed, transmission experiments using the same experimental model that we have previously used, has shown that removal of *B. afzelii*-infected ticks after 24 h of tick feeding blocks successful infection of the host<sup>41–43</sup>.

As described above, it is known that certain *I. scapularis* TSGPs are upregulated upon *B. burgdorferi* infection and that some of these proteins are to be beneficial for the transmission success of the spirochete. However, this is the first study to investigate whether and to what extent *B. afzelii* influences gene expression in the salivary glands of nymphal *I. ricinus* ticks. With 465 transcripts differentially expressed at 24 h after onset feeding, the MACE analysis indicates that *B. afzelii* infection has an extensive effect on gene expression. The majority of transcripts upregulated in *Borrelia*-infected SG at 24 h belong to the kunitz domain inhibitor, ixodegrins, *Ixodes* specific and lipocalin protein families. Kunitz domain inhibitors are one of the largest families of secreted salivary gland proteins. These proteins have one or multiple kunitz domains that inhibit activity of specific proteases, most of which are involved in the coagulation pathway<sup>30,44</sup>. Ixodegrins are cysteine rich proteins that have a RGD or KGD domain, which can bind to integrins; transmembrane receptors that mediate cell–cell and cell–extracellular matrix interactions and as such can have multiple functions. Ixodegrins can block the interaction of integrins with their other ligands and block downstream processes. For instance, binding of ixodegrins to  $\alpha\text{IIb}\beta 3$  of activated platelets, prevents fibrinogen–platelet interaction and platelet aggregation<sup>45</sup>. Integrins, as they are transmembrane receptors, are also involved in immunity. For instance, macrophage 1 antigen, more recently known as complement receptor 3 (CR3), is an integrin present on polymorphonuclear leukocytes and binds fibrinogen which leads to macrophage adhesion and activation. Interestingly, CR3 interacts with uPAR, which has been shown to be important for the clearance of *Borrelia* and CR3 also binds *Borrelia* directly to polymorphonuclear leukocytes<sup>28,46–49</sup>. As the ligands of most ixodegrins are currently unknown, it might be possible that some could protect *Borrelia* from the host's immune response. *Ixodes* specific protein family comprises several smaller families and only some members of the Isac protein subfamily have been characterized; these tick proteins interfere with the complement cascade<sup>50–52</sup>. The complement cascade is an important line of defense against *B. burgdorferi* s.l.. Although sensitivity for complement-mediated killing varies between *B. burgdorferi* s.l. genospecies and *B. afzelii* is particular complement resistant, complement leads to opsonophagocytosis of *B. burgdorferi* by immune cells and in antibody-dependent complement-mediated killing<sup>53,54</sup>. Thus, it is possible that the proteins upregulated in *B. afzelii*-infected *I. ricinus* salivary glands at 24 h after feeding facilitate both *B. afzelii* transmission from the tick to the host or successful infection of the host. However, as transcripts belonging to the same protein family are both upregulated and downregulated, the characterization of the majority of proteins in each protein family is poor or non-existent, and the same TSGP can exert multiple functions, it is difficult to appreciate the exact biological role of the different families of tick proteins in *B. afzelii* transmission or infection.

One of the aims of this study was to identify possible pathogen transmission blocking anti-tick vaccine targets. Previous studies have shown that antibodies against *I. scapularis* TSGPs not only interfere with tick feeding, but antibodies induced after 24 h of tick feeding could also partially protect against *B. burgdorferi* infection<sup>11,17–19,55</sup>. In search for potential vaccine targets to block *B. afzelii* transmission by *I. ricinus*, the 20 most abundantly expressed transcripts upregulated in *B. afzelii*-infected *I. ricinus* salivary glands at 24 h after onset feeding were validated in 10 biological and genetically distinct replicates. Three transcripts—encoding 2 unique proteins—were significantly upregulated in *B. afzelii*-infected salivary glands across the 10 biological samples; Gene 2, Gene 6 and Gene 13. Despite the fact that we, in line with previously published tick transcriptome studies<sup>21,23,24</sup>, pooled salivary glands of hundreds of ticks to obtain enough RNA for both RNAseq and MACE, we were only able to biologically validate three out of the 20 selected abundantly-expressed *I. ricinus* transcripts. This suggests that there is substantial biological variation, either in transcript sequence or expression and underscores that it is critical to consider this variation, especially when selecting vaccine candidates. Ideal vaccine candidates are highly conserved and expressed in multiple biological replicates, which is the case for the three selected transcripts. Gene 2 and 6 proved to encode the same protein, a putative lipocalin with predicted histamine binding domains. Lipocalins are one of the largest and most diverse protein families in ticks. Despite their diversity in amino acid sequence, they all have a barrel structure that creates a fold and facilitates the binding of hydrophobic ligands. The targets of lipocalins are as diverse as the protein family itself; lipocalins can target inflammation, acquired immunity and the complement system. As Gene 2 and 6 have histamine binding domains, they appear to belong to the first category. Histamine release by host cells induce inflammation at the tick bite site; hence, Gene 2 and 6 could inhibit inflammation by binding histamine. Gene 13 is characterized as a putative Ixodegrin, has a prokineticin domain and is part of the colipase-like superfamily. As described above, Ixodegrins are cysteine-rich proteins and although the function for most of these proteins is unknown, there are some that act as antiplatelet inhibitors and they might affect innate immunity<sup>30,46,47,56,57</sup>. In addition, although overall not significantly upregulated in *B. afzelii*-infected salivary glands at 24 h, Gene 1 was highly expressed in the tick pools in which the transcript could be detected and therefore evaluated as a vaccine candidate. In silico analysis showed that Gene 1 showed a high degree of homology to basic tail proteins and has a very basic carboxy terminus or tail, which is one of the key features of this protein family. It is thought that the basic tail might help binding to anionic phospholipids expressed at the surface of activated platelets and mast cells and can interfere in the functioning of the subsequent host processes<sup>30,58</sup>. Indeed, several TSGPs belonging to the basic tail protein family have been described to interfere with complement or coagulation (TSLPI, Salp14, Ixonnexin and Salp-9pac)<sup>15,16,44,59–61</sup>. These processes have been proven to be important for tick feeding and *B. burgdorferi* survival in the host<sup>15,16,61</sup>. Although there are no conserved domains to directly pinpoint possible effector functions of Gene 1, the presence of a basic tail and upregulation in *B. afzelii*-infected salivary glands indicate that this TSGP could facilitate *B. afzelii* transmission or survival in the mammalian host. Thus, in silico analysis indicated that

the three selected targets could very well be involved in the manipulation of the host defense mechanisms that are essential for tick feeding and/or survival of *B. afzelii* in the host. However, a preliminary RNAi experiment, in which the four validated transcripts encoding the three targets were successfully silenced, did not reveal an essential role for the identified TSGPs in tick feeding or *B. afzelii* transmission.

In line with these findings, vaccination with recombinant forms of Gene 2 or 13 did not reduce tick feeding nor *B. afzelii* transmission to the host after challenge with *B. afzelii*-infected *I. ricinus* nymphs, compared to control mice. In silico analysis showed that all antigens are predicted to have peptides that can bind to MHC class I and class II, and are linear epitopes for B cells. Vaccination indeed did induce high antibody titers for recombinant Gene 13, confirming immunogenicity. For recombinant Gene 2, antibody levels are relatively low despite predicted immunogenicity and although the purified antigen seem to contain *E. coli* residue that could interfere with the immune response, we consider it unlikely that these trace amounts of contaminants have interfered with the antigen specific immune response (Supplemental Fig. 3). It is therefore unclear what explains the modest antibody titers. Interestingly, vaccination with recombinant Gene 1 significantly induced a robust antibody response and reduced the number of infected tissues in mice as determined by qPCR. However, vaccination with Gene 1 did not protect against infection; the cumulative number of infected mice as determined by qPCR and culture was similar for recombinant Gene 1 vaccinated and control animals. This modest effect was not observed in mice vaccinated with all three antigens. This might be explained by interference of the other two antigens with the immune response against recombinant Gene 1 upon vaccination or tick-challenge. In general, other vaccination platforms or different ways of producing the tick antigens as recombinant proteins might lead to improved vaccine efficacy. Indeed, a recent publication showed the importance of glycosylation of tick saliva proteins in tick immunity against *I. scapularis*<sup>9</sup>. Therefore, one could argue that the fact that we produced the selected tick antigens in an *E. coli* expression system, and the resulting absence of posttranslational modifications such as glycosylation, are responsible for the low observed vaccine efficacy. Whether vaccination with Gene 1 produced in an Eukaryotic expression system would increase vaccine efficiency remains to be investigated.

To conclude, in this study, using two independent next generation sequencing techniques, we clearly show that *B. afzelii* affects *I. ricinus* salivary glands gene expression during tick feeding, and that the uniquely expressed, as well as up- and downregulated tick transcripts upon *B. afzelii* infection encode proteins belonging to the same tick protein families. Four transcripts encoding three different proteins were shown to be robustly upregulated in *B. afzelii*-infected *I. ricinus* salivary glands at 24 h. Of these three proteins, only recombinant Gene 1 altered *B. afzelii* infection when tested as a transmission blocking anti-tick vaccine in the current set-up and although it did not prevent infection, it could still be an interesting antigen for further optimization, for example as part of a multivalent vaccine or produced in a different expression system. In addition, future research could focus on determining the function of these proteins in either the tick or the host.

## Material and methods

**Infection of ticks with *Borrelia afzelii*, tick feeding and RNA extraction.** *I. ricinus* ticks were obtained from the BC ASCR tick colony and were free of *Borrelia*, *Babesia*, and *Anaplasma*, as determined by PCR<sup>62,63</sup>. To obtain non-infected and *B. afzelii*-infected ticks, clean *I. ricinus* larvae—a mixture of the offspring from three individual adult females—were fed on naive or *B. afzelii* strain CB43 syringe-inoculated 6–8 weeks old C3H/HeN mice (Charles River Laboratories, Sulzfeld, Germany). Larvae were collected and allowed to molt to nymphs in a climate chamber with a humidity of about 95%, temperature 24 °C and day/night period set to 15/9 h. Infection rates for infected ticks were assessed by qPCR and ticks were used when infection rates were higher than 90%. Resulting non-infected and *B. afzelii*-infected ticks (4 to 6 weeks after molting) were fed for 0 h (220 nymphs per infection state, 440 total), 24 h (180 nymphs, 380 total) or to repletion (150, 300 total) on naive 6–8 weeks old C3H/HeN mice and dissected under a dissection microscope. Salivary glands were collected and total (small and large) RNA was extracted using a NucleoSpin miRNA kit (MACHEREY–NAGEL, Dürren, Germany) according to the manufacturer's instructions and stored at –80 °C until further use. All tick and animal experiments were approved by the BC ASCR animal ethical committee (Animal protection laws of the Czech Republic No. 246/1992 Sb., Ethics approval No. 79/2013). All experiments were performed in accordance with relevant guidelines and regulations.

**RNA sequencing.** For RNA sequencing, we created four separate RNA-Seq libraries; infected salivary glands (RNA from 550 *B. afzelii*-infected nymphs at 0, 24 h and fully fed were pooled), uninfected salivary glands (RNA from 550 uninfected nymphs at 0, 24 h and fully fed were pooled), 45 *B. afzelii*-infected whole body fully fed nymphs and 45 uninfected whole body fully fed. The RNAseq libraries were generated using the “NEBNextUltra directional RNA-Seq” (NEB, Ipswich, USA) protocol, as described by the manufacturer and based on the method previously published<sup>64</sup>. In short, mRNA was captured from 5 µg of total RNA using Oligo dT(25) beads. The purified mRNA was randomly fragmented in a Zn<sup>2+</sup> solution and first strand synthesis was performed using random hexamers. Second strand synthesis was performed using a dNTP mixture in which dTTP was exchanged with dUTP and P5-P7-Y-adapters were ligated. The second strand was eliminated prior to PCR using dUTPase. Subsequently, a PCR was performed using 14 cycles. The final products were analyzed on an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA) and product sizes ranged from 200 to 800 bp, with a major peak at 450 bp. Finally, the products were sequenced on an Illumina HiSeq2000 machine (Illumina, Inc., San Diego, CA, USA) using 2 × 100 bp. Overlapping sequencing reads were de novo assembled into GXP\_Contigs with TrinityRNAseq (Version: v2.2.0<sup>65</sup>). Further assembly output refining resulted in 32,897 high quality Contigs used as a reference database. The obtained sequences were uploaded to GenBank (Bioproject PRJNA657487).

**MACE analysis.** Essentially, MACE analysis was performed as previously described (Nold-Petry et al. Mueller et al.) using the GenXPro MACE kit (GenXPro, Frankfurt am Main, Germany) and according to the manufacturer's protocol. Briefly, 1 µg of obtained large and small tick salivary gland RNA from 550 *B. afzelii* CB43-infected or 550 non-infected ticks fed for 0, 24 or approximately 72 h (fully fed) were subjected to an additional DNase treatment to remove all DNA. Quality was assessed on an Agilent 2100 Bioanalyzer and no or only negligible degradation products were observed. Next, first and second strand cDNA synthesis was performed starting from biotinylated oligo dT primers. The cDNA was fragmented randomly by ultrasonication resulting in fragments with an average size of 300 bps as determined by an Agilent 2100 Bioanalyzer. The biotinylated 3' cDNA ends were bound to a streptavidin matrix and all other fragments were eliminated through washing. To the unbound end of the fragments, a p5 "TrueQuant" sequencing adapter included in the MACE kit was ligated and a PCR was performed, using tailed Illumina p5 and p7 oligonucleotides as primers, in order to obtain a library of fragments suitable for Next Generation Sequencing on an Illumina HiSeq2000 machine. The Quality of the final library was determined using an Agilent 2100 bioanalyzer. Single end sequencing of the products produced the sequence-information of the 5' side of the bound cDNA fragment. To remove PCR-bias, all duplicate reads detected by the in house TrueQuant technology were removed from the raw datasets. In addition, low quality sequence nucleotides and poly(A)-tails were clipped off using cutadapt<sup>66</sup>. The reads were thereafter aligned to different reference sequences using Novoalign (Novocraft Technologies, Selangor, Malaysia). The main reference for the Novoalign alignment was the outcome of the RNASeq de novo assembly, described in the RNASeq section. Additionally a de novo assembly of MACE sequences that could not be mapped to sequences from the Master Reference (RNASeq) using TrinityRNASeq (Version: v2.2.0<sup>65</sup>) was performed. Subsequently, the contigs of the assemblies, "Master Reference" and "noHitAssembly" were annotated further by BLASTX to first the SwissProt and hereafter Trembl database "Arachnida" proteins<sup>67</sup>. Additional blastn analyses were performed for all Contigs against all "Ixodes" mRNA sequences available at the NCBI database, nucleotide collection from GenBank (RefSeq, TPA and PDB), *Ixodes scapularis* genome (PRJNA314100), *Ixodes ricinus* genome (PRJNA270959) and against sequences from a previous published *I. ricinus* salivary gland transcriptome<sup>23</sup> submitted to Genbank (PRJNA177622). The e-value threshold for BLASTX and BLASTN was 0.00001. Only uniquely mapped reads were accepted for quantification of the MACE tags. Finally, the expression was normalized and tested for differential gene expression between the different conditions using the DEGSeq R/Bioconductor package<sup>67</sup>. Only transcripts with at least 1 normalized read in one of the libraries were used for analysis (Supplemental file 1).

**Allocation of genes to tick protein families.** For more functional insight, the transcripts were allocated to tick protein families based on sequence homology. In short, gene sequences of the Master Reference were aligned to proteins from a previous bioproject<sup>39</sup> using blastx. Transcripts were considered to belong to a specific tick protein family if the e-value with their respective protein hit from the Bioproject Number PRJNA177622 was below 0.0001.

**Technical and biological validation.** An aliquot of total RNA from each time point analyzed by MACE was used to make cDNAs (Transcriptor High Fidelity cDNA Synthesis Kit (Roche, Basel, Switzerland)) for qRT-PCR technical validation of the MACE results. For biological validations, *Borrelia afzelii*-infected (Infection rates were assessed by qPCR and ticks were used when infection rates were higher than 90%) and uninfected nymphal *I. ricinus* ticks derived from 10 distinct egg batches laid by adult female ticks collected from the wild, were fed on mice for different time points. RNA was isolated from the salivary glands and subsequent cDNA was prepared for the individual time points. Then, gene-specific primers appropriate for unambiguous PCR confirmation of gene expression in *Borrelia*-infected nymphs at the time interval 24 h and the genes upregulated by feeding, were designed using Primer3 software (Supplemental Table 1). qRT-PCR was used to evaluate expression of the selected genes in technical and biological samples.

**In silico analysis.** The encoded protein sequence for Gene 1, 2 and 13 were determined from the transcripts nucleotide sequences using the ExPASy translate tool<sup>68</sup>. Proteins sequences were subsequently scanned for domains with InterProScan<sup>69</sup> and a predicted protein model was built using the Phyre2 web portal<sup>70</sup>. Signal peptide, O- and N-glycosylation sites were predicted based on amino acid sequence by SignalP 5.0 server<sup>71</sup>, NetOGlyc 4.0 Server<sup>72</sup> and NetNGlyc 1.0 server<sup>73</sup>, respectively. HTHMM v2.0 server<sup>74</sup> was used to predict transmembrane helices and GPI-SOM to predict GPI-anchor<sup>75</sup>. MHC class I and II binding peptides were predicted using NetMHCpan-4.1<sup>76</sup> and NetMHCIIpan-4.0<sup>76</sup> and linear B cell epitopes with BepiPred Linear Epitope Prediction 2.0<sup>77</sup>.

**Expression and purification of recombinant proteins.** Transcripts were cloned by overlapping PCR from previously designed artificial genes and cloned as NcoI-SalI fragments into the pHis-parallel 2 expression vector<sup>78</sup>. For Gene 1 forward primer (FW) CGCCATGGGAGACGATTGCAGAAACGGAAGCTAGA and reverse primer (RV) CGGTGCGACTAGTACGTTTTCCCTTCCTTAATTATTTCCTGTG was used. For Gene 2 CGCCATGGGATCTACAAAGTACTACTACCCATCCAGTG (FW) and CGGTGCGACTACACCAAGGAAAAGTGCATATTCTCGTT (RV) and for Gene 13 CGCCATGGGACAGGTACCAGTGTTCCTCCCTGG (FW) and CGGTGCGACTATTTCCTTGGCAGCAAATATGTCTG (RV) were used. Clones were induced with 1 mM Isopropyl-β-D-thiogalactoside (IPTG) for 16 h at 20 °C in *E. coli* BL21 C41(DE3). The bacterial cells were then lysed and centrifuged. The expressed insoluble proteins were extracted from the inclusion bodies with the following protocol. The pellets were thoroughly homogenized in Phosphate Buffered Saline (PBS); 2% Triton X-100 followed by an incubation at 37 °C for 30 min with shaking. The samples were ultracentrifuged at 96,000 g for

30 min and the pellets were homogenized again in PBS and incubated at 37 °C for 30 min with shaking. After a second ultracentrifugation, the pellets were homogenized in PBS; 7 M urea. The denatured proteins were dialyzed to 2 M urea overnight.

**Preliminary RNA interference study.** Silencing of the gene candidates by RNA interference (Genes 1, 2, and 13) was done as described previously (1). The *Borrelia*-infected nymphs were injected with 0.32 nl of dsRNA, rested for three days, and fed on C3H mice (5 nymphs per mouse, infection rate of ticks > 90%). The level of silencing was checked by qRT-PCR on a mix of five fully-fed nymphs per group and compared to the GFP control; expression of gene 3 was reduced by 92%, expression of gene 2 was reduced by 98%, expression of gene 13 was reduced by 99% and for gene 1 expression was reduced by 68% (primers can be found in Supplemental Table 1). The mice were screened for infection by qRT-PCR in a skin, heart, and urinary bladder, as described below.

**Vaccination-transmission studies and infection parameters.** Pathogen-free C3H/HeN mice (Charles River Laboratories) were used for the vaccination transmission experiments. Six mice per group were vaccinated with either PBS, recombinant Gene 1, Gene 2, Gene 13 or all three recombinant proteins injected subcutaneously at different sites. 20 µg of antigen was emulsified in Complete Freund's Adjuvant (Sigma-Aldrich, St. Louis, MO, USA) to 100 µl total volume for prime vaccination at day 0. For booster vaccinations at day 14 and 28, 20 µg of antigen were emulsified in Incomplete Freund's Adjuvant, 100 µl total volume. 2 weeks after the last vaccination, mice were challenged with 5 *B. afzelii*-infected *I. ricinus* nymphs (infection rate > 90%) which were allowed to feed to repletion. Before each vaccination and the tick challenge, mouse blood was collected. 3 weeks after tick infestation, mice were sacrificed and organs were collected for culture and qPCR. Half of the mouse bladder and a part of the tick bite site were cultured in BSK medium (Amsterdam UMC, AMC, The Netherlands).

Total spirochete load in mouse tissues was determined by qPCR, which targeted a fragment of the *flagellin* gene (154 bp). DNA was isolated from individual murine tissues (ear, skin, heart, and urinary bladder) using a NucleoSpin tissue kit (Macherey–Nagel) according to the manufacturer's protocol. The reaction mixture contained 12.5 µl of FastStart universal probe master (Rox) (Roche), 10 pmol of primers FlaF1A and FlaR1, 5 pmol of TaqMan probe Fla Probe1<sup>8,8</sup>, 5 µl of DNA, and PCR water up to 25 µl. Quantification of murine β-actin was performed using MmAct-F and MmAct-R primers and a MmAct-P TaqMan probe<sup>14</sup>. The following amplification program was run on a LightCycler 480 (Roche) for both targets: 95 °C for 10 min, 50 cycles at 95 °C for 15 s and 60 °C for 1 min. The spirochete burden in murine tissues was expressed as the number of spirochetes per 10<sup>5</sup> murine β-actin copies.

**Antibody responses.** Total antigen-specific IgG levels were determined by ELISA. ELISA plates (Thermo Scientific) were coated with full length proteins at 0.05 µg/well in carbonate buffer (pH 9.6) and incubated overnight at 4 °C. After washing with PBS, containing 0.05% Tween 20, the plates were incubated with blocking buffer (10% of fetal calf serum (FCS, Biowest) in PBS) for 1 h. Mouse sera were added at 1:5600 dilution and incubated for 2 h at room temperature. After washing, goat anti-mouse total IgG conjugated to horseradish peroxidase (HRP) (Jackson ImmunoResearch) was added (1:1000 dilution) in blocking buffer and incubated for 1 h at room temperature. The plates were then extensively washed and incubated with KPL SureBlue substrate. The reaction was stopped with 2N H<sub>2</sub>SO<sub>4</sub>. Absorbance (450 nm) was immediately measured using a BioTek Synergy HT multi-detection microplate reader.

Received: 17 April 2020; Accepted: 17 September 2020

Published online: 18 November 2020

## References

- Eisen, R. J. & Eisen, L. The blacklegged tick, *Ixodes scapularis*: an increasing public health concern. *Trends Parasitol.* **34**, 295–309. <https://doi.org/10.1016/j.pt.2017.12.006> (2018).
- Semenza, J. C. & Suk, J. E. Vector-borne diseases and climate change: a European perspective. *FEMS Microbiol. Lett.* <https://doi.org/10.1093/femsle/fnx244> (2017).
- Hubálek, Z. Epidemiology of lyme borreliosis. *Curr. Probl. Dermatol.* **37**, 31–50. <https://doi.org/10.1159/000213069> (2009).
- van Dam, A. P. *et al.* Different genospecies of *Borrelia burgdorferi* are associated with distinct clinical manifestations of Lyme borreliosis. *Clin. Infect. Dis.* **17**, 708–717. <https://doi.org/10.1093/clinids/17.4.708> (1993).
- Embers, M. E. & Narasimhan, S. Vaccination against Lyme disease: past, present, and future. *Front. Cell. Infect. Microbiol.* **3**, 6–6. <https://doi.org/10.3389/fcimb.2013.00006> (2013).
- Cook, M. J. Lyme borreliosis: a review of data on transmission time after tick attachment. *Int. J. Gen. Med.* **8**, 1–8. <https://doi.org/10.2147/IJGM.S73791> (2014).
- Crippa, M., Rais, O. & Gern, L. Investigations on the mode and dynamics of transmission and infectivity of *Borrelia burgdorferi* sensu stricto and *Borrelia afzelii* in *Ixodes ricinus* ticks. *Vector Borne Zoonotic Dis.* **2**, 3–9. <https://doi.org/10.1089/153036602760260724> (2002).
- Pospisilova, T. *et al.* Tracking *Borrelia afzelii* from infected *Ixodes ricinus* nymphs to mice suggests a direct “gut-to-mouth” route of Lyme disease transmission. *J. bioRxiv* <https://doi.org/10.1101/316927> (2018).
- Narasimhan, S. *et al.* *Ixodes scapularis* saliva components that elicit responses associated with acquired tick-resistance. *Ticks Tick Borne Dis.* **101**, 369. <https://doi.org/10.1016/j.ttbdis.2019.101369> (2020).
- Schuijt, T. J. *et al.* Identification and characterization of *Ixodes scapularis* antigens that elicit tick immunity using yeast surface display. *PLoS ONE* **6**, e15926. <https://doi.org/10.1371/journal.pone.0015926> (2011).
- Trager, W. Acquired immunity to ticks. *J. Parasitol.* **25**, 57–81. <https://doi.org/10.2307/3272160> (1939).

12. Anguita, J. *et al.* Salp15, an *Ixodes scapularis* salivary protein, inhibits CD4(+) T cell activation. *Immunity* **16**, 849–859 (2002).
13. Dai, J. *et al.* Tick histamine release factor is critical for *Ixodes scapularis* engorgement and transmission of the Lyme disease agent. *PLoS Pathog.* **6**, e1001205. <https://doi.org/10.1371/journal.ppat.1001205> (2010).
14. Dai, J. *et al.* Antibodies against a tick protein, Salp15, protect mice from the Lyme disease agent. *Cell Host Microbe* **6**, 482–492. <https://doi.org/10.1016/j.chom.2009.10.006> (2009).
15. Schuijt, T. J. *et al.* A tick mannose-binding lectin inhibitor interferes with the vertebrate complement cascade to enhance transmission of the Lyme disease agent. *Cell Host Microbe* **10**, 136–146. <https://doi.org/10.1016/j.chom.2011.06.010> (2011).
16. Wagemakers, A. *et al.* An *Ixodes ricinus* tick salivary lectin pathway inhibitor protects *Borrelia burgdorferi* *sensu lato* from human complement (2016).
17. Brossard, M. & Girardin, P. Passive transfer of resistance in rabbits infested with adult *Ixodes ricinus* L: humoral factors influence feeding and egg laying. *Experientia* **35**, 1395–1397 (1979).
18. Narasimhan, S. *et al.* Immunity against *Ixodes scapularis* salivary proteins expressed within 24 hours of attachment thwarts tick feeding and impairs *Borrelia* transmission. *PLoS ONE* **2**, e451. <https://doi.org/10.1371/journal.pone.0000451> (2007).
19. Nazario, S. *et al.* Prevention of *Borrelia burgdorferi* transmission in guinea pigs by tick immunity. *Am. J. Trop. Med. Hygiene* **58**, 780–785 (1998).
20. Wikel, S. K., Ramachandra, R. N., Bergman, D. K., Burkot, T. R. & Piesman, J. Infestation with pathogen-free nymphs of the tick *Ixodes scapularis* induces host resistance to transmission of *Borrelia burgdorferi* by ticks. *Infect. Immun.* **65**, 335–338 (1997).
21. Chmelar, J. *et al.* Insight into the salivome of the castor bean tick, *Ixodes ricinus*. *BMC Genom.* **9**, 233 (2008).
22. Perner, J., Kropáčková, S., Kopáček, P. & Ribeiro, J. M. C. Salivome diversity of ticks revealed by RNAseq of single tick salivary glands. *PLoS Negl. Trop. Dis.* **12**, e0006410. <https://doi.org/10.1371/journal.pntd.0006410> (2018).
23. Schwarz, A. *et al.* De novo *Ixodes ricinus* salivary gland transcriptome analysis using two next-generation sequencing methodologies. *FASEB J.* **27**, 4745–4756. <https://doi.org/10.1096/fj.13-232140> (2013).
24. Kotsyfakis, M., Schwarz, A., Erhart, J. & Ribeiro, J. M. C. Tissue- and time-dependent transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the vertebrate host. *Sci. Rep.* **5**, 9103. <https://doi.org/10.1038/srep09103> (2015).
25. Anguita, J. *et al.* Salp15, an *Ixodes scapularis* salivary protein, inhibits CD4(+) T cell activation.
26. Cotte, V. *et al.* Differential expression of *Ixodes ricinus* salivary gland proteins in the presence of the *Borrelia burgdorferi sensu lato* complex. *J. Proteomics* **96**, 29–43. <https://doi.org/10.1016/j.jprot.2013.10.033> (2014).
27. Hovius, J. W. *et al.* Preferential protection of *Borrelia burgdorferi sensu stricto* by a salp 15 homologue in *Ixodes ricinus* saliva. *J. Infect. Dis.* **198**, 1189–1197. <https://doi.org/10.1086/591917> (2008).
28. Hovius, J. W., van Dam, A. P. & Fikrig, E. Tick–host–pathogen interactions in *Lyme borreliosis*. *Trends Parasitol.* **23**, 434–438. <https://doi.org/10.1016/j.pt.2007.07.001> (2007).
29. Narasimhan, S. *et al.* A tick gut protein with fibronectin III domains aids *Borrelia burgdorferi* congregation to the gut during transmission. *PLoS Pathog.* **10**, e1004278. <https://doi.org/10.1371/journal.ppat.1004278> (2014).
30. Francischetti, I. M., Sa-Nunes, A., Mans, B. J., Santos, I. M. & Ribeiro, J. M. The role of saliva in tick feeding. *Front. Biosci. (Landmark edition)* **14**, 2051–2088. <https://doi.org/10.2741/3363> (2009).
31. Mans, B. J., Featherston, J., de Castro, M. H. & Pienaar, R. Gene duplication and protein evolution in tick–host interactions. *Front. Cell. Infect. Microbiol.* **7**, 413. <https://doi.org/10.3389/fcimb.2017.00413> (2017).
32. Nold-Petry, C. A. *et al.* IL-37 requires the receptors IL-18R $\alpha$  and IL-1R8 (SIGIRR) to carry out its multifaceted anti-inflammatory program upon innate signal transduction. *Nat. Immunol.* **16**, 354–365. <https://doi.org/10.1038/ni.3103> (2015).
33. Zawada, A. M. *et al.* Massive analysis of cDNA Ends (MACE) and miRNA expression profiling identifies proatherogenic pathways in chronic kidney disease. *Epigenetics* **9**, 161–172. <https://doi.org/10.4161/epi.26931> (2014).
34. Mandelboum, S., Manber, Z., Elroy-Stein, O. & Elkon, R. Recurrent functional misinterpretation of RNA-seq data caused by sample-specific gene length bias. *PLoS Biol.* **17**, e3000481. <https://doi.org/10.1371/journal.pbio.3000481> (2019).
35. Asmann, Y. W. *et al.* 3' tag digital gene expression profiling of human brain and universal reference RNA using Illumina Genome Analyzer. *BMC Genom.* **10**, 531. <https://doi.org/10.1186/1471-2164-10-531> (2009).
36. Lenz, T. L., Eizaguirre, C., Rotter, B., Kalbe, M. & Milinski, M. Exploring local immunological adaptation of two stickleback ecotypes by experimental infection and transcriptome-wide digital gene expression analysis. *Mol. Ecol.* **22**, 774–786. <https://doi.org/10.1111/j.1365-294X.2012.05756.x> (2013).
37. Kotsyfakis, M., Schwarz, A., Erhart, J. & Ribeiro, J. M. C. Tissue- and time-dependent transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the vertebrate host. *Sci. Rep.* **5**, 1. <https://doi.org/10.1038/srep09103> (2015).
38. Schwarz, A. *et al.* A systems level analysis reveals transcriptomic and proteomic complexity in *Ixodes ricinus* midgut and salivary glands during early attachment and feeding. *Mol. Cell. Proteomics* **13**, 2725–2735. <https://doi.org/10.1074/mcp.M114.039289> (2014).
39. Schwarz, A. *et al.* De novo *Ixodes ricinus* salivary gland transcriptome analysis using two next-generation sequencing methodologies. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **27**, 4745–4756. <https://doi.org/10.1096/fj.13-232140> (2013).
40. Piesman, J. & Happ, C. M. The efficacy of co-feeding as a means of maintaining *Borrelia burgdorferi*: a North American model system. *J. Vector Ecol.* **26**, 216–220 (2001).
41. Pospisilova, T. *et al.* Tracking of *Borrelia afzelii* transmission from infected *Ixodes ricinus* nymphs to mice. *Infect. Immun.* <https://doi.org/10.1128/IAI.00896-18> (2019).
42. Crippa, M., Rais, O. & Gern, L. Investigations on the mode and dynamics of transmission and infectivity of *Borrelia burgdorferi sensu stricto* and *Borrelia afzelii* in *Ixodes ricinus* ticks. *Vector Borne Zoonotic Dis. (Larchmont, N.Y.)* **2**, 3–9. <https://doi.org/10.1089/153036602760260724> (2002).
43. Piesman, J., Mather, T. N., Sinsky, R. J. & Spielman, A. Duration of tick attachment and *Borrelia burgdorferi* transmission. *J. Clin. Microbiol.* **25**, 557–558 (1987).
44. Chmelar, J., Kotal, J., Kovarikova, A. & Kotsyfakis, M. The use of tick salivary proteins as novel therapeutics. *Front. Physiol.* **10**, 812. <https://doi.org/10.3389/fphys.2019.00812> (2019).
45. Assumpcao, T. C. F., Ribeiro, J. M. C. & Francischetti, I. M. B. Disintegrins from hematophagous sources. *Toxins (Basel)* **4**, 296–322. <https://doi.org/10.3390/toxins4050296> (2012).
46. Blasi, F. & Carmeliet, P. uPAR: a versatile signalling orchestrator. *Nat. Rev. Mol. Cell Biol.* **3**, 932–943. <https://doi.org/10.1038/nrm977> (2002).
47. Garcia, R. C., Murgia, R. & Cinco, M. Complement receptor 3 binds the *Borrelia burgdorferi* outer surface proteins OspA and OspB in an iC3b-independent manner. *Infect. Immun.* **73**, 6138. <https://doi.org/10.1128/IAI.73.9.6138-6142.2005> (2005).
48. Carreras-González, A. *et al.* Regulation of macrophage activity by surface receptors contained within *Borrelia burgdorferi*-enriched phagosomal fractions. *PLoS Pathog.* **15**, e1008163. <https://doi.org/10.1371/journal.ppat.1008163> (2019).
49. Hawley, K. L. *et al.* CD14 cooperates with complement receptor 3 to mediate MyD88-independent phagocytosis of *Borrelia burgdorferi*. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 1228–1232. <https://doi.org/10.1073/pnas.1112078109> (2012).
50. Daix, V. *et al.* *Ixodes* ticks belonging to the *Ixodes ricinus* complex encode a family of anticomplement proteins. *Insect. Mol. Biol.* **16**, 155–166. <https://doi.org/10.1111/j.1365-2583.2006.00710.x> (2007).
51. Hourcade, D. E. *et al.* Anti-complement activity of the *Ixodes scapularis* salivary protein Salp20. *Mol. Immunol.* **69**, 62–69. <https://doi.org/10.1016/j.molimm.2015.11.008> (2016).

52. Valenzuela, J. G., Charlab, R., Mather, T. N. & Ribeiro, J. M. Purification, cloning, and expression of a novel salivary anticomplement protein from the tick, *Ixodes scapularis*. *J. Biol. Chem.* **275**, 18717–18723. <https://doi.org/10.1074/jbc.M001486200> (2000).
53. Coumou, J. *et al.* The role of mannose binding lectin in the immune response against *Borrelia burgdorferi* sensu lato. *Sci. Rep.* **9**, 1431. <https://doi.org/10.1038/s41598-018-37922-8> (2019).
54. Hawley, K. L., Olson, C. M. Jr., Carreras-González, A., Navasa, N. & Anguita, J. Serum C3 enhances complement receptor 3-mediated phagocytosis of *Borrelia burgdorferi*. *Int. J. Biol. Sci.* **11**, 1269–1271. <https://doi.org/10.7150/ijbs.13395> (2015).
55. Bowessidjaou, J., Brossard, M. & Aeschlimann, A. Effects and duration of resistance acquired by rabbits on feeding and egg laying in *Ixodes ricinus* L. *Experientia* **33**, 528–530. <https://doi.org/10.1007/BF01922254> (1977).
56. Hovius, J. W. R. *et al.* The urokinase receptor (uPAR) facilitates clearance of *Borrelia burgdorferi*. *PLoS Pathog.* **5**, e1000447–e1000447. <https://doi.org/10.1371/journal.ppat.1000447> (2009).
57. Tang, J. *et al.* YY-39, a tick anti-thrombosis peptide containing RGD domain. *Peptides* **68**, 99–104. <https://doi.org/10.1016/j.peptides.2014.08.008> (2015).
58. Andersen, J. F., Gudderra, N. P., Francischetti, I. M., Valenzuela, J. G. & Ribeiro, J. M. Recognition of anionic phospholipid membranes by an antihemostatic protein from a blood-feeding insect. *Biochemistry* **43**, 6987–6994. <https://doi.org/10.1021/bi049655t> (2004).
59. Das, S. *et al.* Salp25D, an *Ixodes scapularis* antioxidant, is 1 of 14 immunodominant antigens in engorged tick salivary glands. *J. Infect. Dis.* **184**, 1056–1064. <https://doi.org/10.1086/323351> (2001).
60. Narasimhan, S. *et al.* A novel family of anticoagulants from the saliva of *Ixodes scapularis*. *Insect. Mol. Biol.* **11**, 641–650. <https://doi.org/10.1046/j.1365-2583.2002.00375.x> (2002).
61. Narasimhan, S. *et al.* Disruption of *Ixodes scapularis* anticoagulation by using RNA interference. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 1141–1146. <https://doi.org/10.1073/pnas.0307669100> (2004).
62. Aase, A. *et al.* Validate or falsify: lessons learned from a microscopy method claimed to be useful for detecting *Borrelia* and *Babesia* organisms in human blood. *Infect. Dis. (Lond.)* **48**, 411–419. <https://doi.org/10.3109/23744235.2016.1144931> (2016).
63. de la Fuente, J. *et al.* Serologic and molecular characterization of *Anaplasma* species infection in farm animals and ticks from Sicily. *Vet. Parasitol.* **133**, 357–362. <https://doi.org/10.1016/j.vetpar.2005.05.063> (2005).
64. Parkhomchuk, D. *et al.* Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res.* **37**, e123–e123. <https://doi.org/10.1093/nar/gkp596> (2009).
65. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494. <https://doi.org/10.1038/nprot.2013.084> (2013).
66. Martin, M. *Cutadapt removes adapter sequences from high-throughput sequencing reads* 17(3), 2011. <https://doi.org/10.14806/ej.17.1.200> (2011).
67. Consortium, T. U. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515. <https://doi.org/10.1093/nar/gky1049> (2018).
68. Gasteiger, E. *et al.* ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–3788. <https://doi.org/10.1093/nar/gkg563> (2003).
69. Mitchell, A. L. *et al.* InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351–D360. <https://doi.org/10.1093/nar/gky1100> (2019).
70. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858. <https://doi.org/10.1038/nprot.2015.053> (2015).
71. Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423. <https://doi.org/10.1038/s41587-019-0036-z> (2019).
72. Stentoft, C. *et al.* Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* **32**, 1478–1488. <https://doi.org/10.1038/emboj.2013.79> (2013).
73. Gupta, R., Jung, E. & Brunak, S. *Prediction of N-glycosylation sites in human proteins* **46**, 203–206 (2004).
74. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes 11 Edited by F. Cohen. *J. Mol. Biol.* **305**, 567–580. <https://doi.org/10.1006/jmbi.2000.4315> (2001).
75. Fankhauser, N. & Mäser, P. Identification of GPI anchor attachment signals by a Kohonen self-organizing map. *Bioinformatics* **21**, 1846–1852. <https://doi.org/10.1093/bioinformatics/bti299> (2005).
76. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkaa379> (2020).
77. Jerspersen, M. C., Peters, B., Nielsen, M. & Marcotili, P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* **45**, W24–W29. <https://doi.org/10.1093/nar/gkx346> (2017).
78. Sheffield, P., Garrard, S. & Derewenda, Z. Overcoming expression and purification problems of RhoGDI using a family of “parallel” expression vectors. *Protein Expr. Purif.* **15**, 34–39. <https://doi.org/10.1006/prep.1998.1003> (1999).
79. Schwaiger, M., Péter, O. & Cassinotti, P. Routine diagnosis of *Borrelia burgdorferi* (sensu lato) infections using a real-time PCR assay. *Clin. Microbiol. Infect.* **7**, 461–469. <https://doi.org/10.1046/j.1198-743x.2001.00282.x> (2001).

## Acknowledgements

This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under Grant Agreement No. 602272, of which JWH was the scientific coordinator. OH, RS, VU and PK were also supported from two additional sources: Czech Science Foundation Grant No. 20-05736S and Centre for Research of Pathogenicity and Virulence of Parasites (No. CZ.02.1.01/0.0/0.0/16\_019/0000759), funded by the European Regional Development Fund (ERDF) and Ministry of Education, Youth, and Sport, Czech Republic (MEYS). CIC bioGUNE thanks the Spanish ministry of Science and Innovation (MCI) for the Severo Ochoa Excellence accreditation (SEV-2016-0644). Diego Barriales is the recipient of a MCI FPI fellowship. We thank Estibaliz Atondo, Leticia Sampedro, Jasmin Ersoz and Miguel Angel Pascual-Itoiz for their excellent technical assistance.

## Author contributions

J.W.H. was the scientific coordinator. N.K. and P.W. designed and performed RNAseq and MACE sequencing and annotation. H.S., K.T., J.M.B., N.K. and J.J.A.T. were involved in transcriptome analysis. V.U., R.S., O.H., S.M. and P.K. performed tick dissections, validations and RNAi-transmission experiments. R.S., O.H., P.K., M.J.K., J.T.C., D.B., J.A., J.J.A.T. and J.W.R. were involved in vaccination-transmission experiments. R.S., O.H., P.W., P.K., J.A., N.K., J.J.A.T. and J.W.H. designed the experiments. J.J.A.T. wrote the manuscript with input from all authors.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-76268-y>.

**Correspondence** and requests for materials should be addressed to J.J.A.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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

© The Author(s) 2020