Microbiome analysis of *Ixodes scapularis* ticks from New York and Connecticut

Rafal Tokarza,*, Teresa Tagliafierroa, Stephen Sameroffb, D. Moses Cucura, Alexandra Oleynik, Xiaoyu Chea, Komal Jain, W. Ian Lipkina

**a** Center for Infection and Immunity, Mailman School of Public Health, Columbia University, NY, United States
**b** Division of Vector Control, Suffolk County Department of Public Works, Yaphank, NY, 11980, United States

**ARTICLE INFO**

**Keywords:**
Ticks
Metagenomics
High-throughput sequencing
Borrelia
Bartonella
*Ixodes scapularis*

**ABSTRACT**

We employed high throughput sequencing to survey the microbiomes of *Ixodes scapularis* collected in New York and Connecticut. We examined 197 individual *I. scapularis* adults and pools from 132 adults and 197 nymphs. We detected *Borrelia burgdorferi* sensu stricto in 56.3% of individual ticks, *Anaplasma phagocytophilum* in 10.6%, *Borrelia miyamotoi* in 5%, *Babesia microti* in 7.6%, and Powassan virus in 3.6%. We did not detect *Borrelia mayonii*, *Ehrlichia muris eauclairensis*, Bartonella spp. or pathogenic *Babesia* species other than *B. microti*. The most abundant bacterium (65%), and only rickettsial species identified, was the endosymbiont *Rickettsia buchneri*. A filarial nematode was found in 13.7% of adult ticks. Fourteen viruses were detected including South Bay virus (22%) and blacklegged tick phleboviruses 1 and 2 (73%). This study provides insight into the microbial diversity of *I. scapularis* in New York State and Connecticut.

1. Introduction

The blacklegged tick, *Ixodes scapularis*, is among the most clinically important tick species in the United States (US). *I. scapularis* transmits the greatest diversity of pathogens of any tick within the US, including agents of Lyme disease (*Borrelia burgdorferi* sensu stricto and *Borrelia mayonii*), babesiosis (*Babesia microti*), anaplasmosis (*Anaplasma phagocytophilum*), *Borrelia miyamotoi* disease (*B. miyamotoi*), ehrlichiosis (*Ehrlichia muris eauclairensis*) and tick-borne encephalitis (Powassan virus) (Spelman et al., 1979; Burgdorfer et al., 1982; Pancholi et al., 1995; Telford et al., 1997; Tokarz et al., 2010; Pritt et al., 2011, 2016a, 2016b) Combined, these agents account for > 90% of all reported tick-borne diseases (Nelson et al., 2015; Connally et al., 2016). The promiscuous host selection of *I. scapularis*, coupled with the recent discovery and characterization of *B. mayonii* and *E. muris eauclairensis*, suggests that *I. scapularis* may harbor other yet undiscovered human pathogens. The fact that *I. scapularis* is among the most frequent ticks encountered by humans advocates the need to catalogue all *I. scapularis*-associated microbes in order to identify the full spectrum of agents that can be transmitted by this tick.

Over the past decade, the employment of high-throughput sequencing (HTS) has substantially enhanced the detection of tick-borne agents. The use of HTS for tick microbiome analyses has also facilitated the discovery of novel tick borne-agents (Tokarz et al., 2014; Bouquet et al., 2017; Cross et al., 2018; Tokarz et al., 2018). In our previous work, we employed HTS to examine the virome of *I. scapularis* (Tokarz et al., 2014, 2018). In this study, we employed a metagenomic approach to survey the complete *I. scapularis* microbiome to gain insight into the capacity of *I. scapularis* to serve as a vector for additional agents of clinical significance.

2. Materials and methods

2.1. Tick extractions

All ticks were collected by dragging from sites in New York and Connecticut (Fig. 1). Ticks were extracted individually or in pools. All individual ticks were collected in 2016 and 2017 for this study. The homogenates from pooled ticks were obtained from ticks collected in 2015 and 2016 that were also previously used for virome analyses (Tokarz et al., 2018). To remove environmental contaminants, prior to extraction each individual tick was washed with 1 ml of 3% hydrogen peroxide, followed by 3 washes with 1 ml of nuclease free water. Ticks were homogenized in 100 μl of viral transport media, and the entire

* Corresponding author at: Center for Infection and Immunity, Mailman School of Public Health, Columbia University, 722 West 168th Street, Room 1701, New York, NY, 10032, United States.
E-mail address: rt2249@columbia.edu (R. Tokarz).


Received 18 January 2019; Received in revised form 21 March 2019; Accepted 10 April 2019

1877-959X/ © 2019 Elsevier GmbH. All rights reserved.

Please cite this article as: Rafal Tokarz, et al., Ticks and Tick-borne Diseases, https://doi.org/10.1016/j.ttbdis.2019.04.011
volume was used for extraction of total nucleic acids (TNA) on the EasyMag platform (Biomerieux). TNA were eluted in 40 μl.

2.2. Library preparation

TNA (11 μl) was reverse transcribed using SuperScript III (ThermoFisher Scientific), treated with RNase H, followed by second-strand synthesis with Klenow fragment. Double-stranded DNA was sheared to a 200-bp average fragment length. Sheared DNA was purified and used for Illumina library construction using the KAPA Hyper Prep kit (KK8504, Kapa Biosystems). Libraries were quantified using an Agilent Bioanalyzer 2100 and sequencing was carried out on the Illumina HiSeq 4000 platform (Illumina, San Diego, CA, USA). Libraries were sequenced on 9 lanes (Table 1). Seven lanes included individual tick samples (16 to 32 tick samples per lane). Two lanes contained tick pools, including a lane with 8 pools of nymphs (20–25 nymphs per pool) and another lane that consisted of 10 pools of adult ticks. Six pools were from I. scapularis; we also sequenced 3 pools of A. americanum and 1 D. variabilis for microbiome comparison of the three tick species.

2.3. Bioinformatic and statistical analysis

The demultiplexed fastq files were trimmed using cutadapt software (v 1.8.3) (Martin, 2011) followed by generation of quality reports using FastQC software (v 0.11.5) (Andrews, 2010). The quality reports were used to determine filtering criteria based on the average quality scores of the reads, presence of indeterminate nucleotides and homopolymeric content of the reads. The reads were quality filtered and end-trimmed with PRINSEQ software (v 0.20.3) (Schmieder and Edwards, 2011). Host background levels were determined by mapping filtered reads against tick reference database using Bowtie2 mapper (v 2.2.9) (Langmead and Salzberg, 2012). The host-subtracted reads were de-novo assembled using MIRA (4.0) (Cheverux et al., 1999) and

---

Table 1

<table>
<thead>
<tr>
<th>Sample type</th>
<th>Site</th>
<th>Number of Raw Reads</th>
<th>Read Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Tick Pools (N = 8)</td>
<td>Fifty Foot Cliff Preserve, CT</td>
<td>368,789,299</td>
<td>5,876,299 – 28,104,634</td>
</tr>
<tr>
<td>2. Tick Pools (N = 10)</td>
<td>Suffolk County, NY</td>
<td>341,244,749</td>
<td>16,404,599 – 63,921,410</td>
</tr>
<tr>
<td>3. Individual ticks</td>
<td>NYC, NY</td>
<td>330,404,206</td>
<td>15,696,114 – 26,332,782</td>
</tr>
<tr>
<td>4. Individual ticks</td>
<td>Mianus State Park, CT</td>
<td>359,207,128</td>
<td>5,703,401 – 29,673,420</td>
</tr>
<tr>
<td>5. Individual ticks</td>
<td>Manorville, NY</td>
<td>331,736,276</td>
<td>3,958,492 – 21,425,762</td>
</tr>
<tr>
<td>8. Individual ticks</td>
<td>Caumsett State Park, NY</td>
<td>328,089,508</td>
<td>6,228,648 – 19,948,223</td>
</tr>
</tbody>
</table>

* ticks originated from multiple sites.
* each row represents one lane on Illumina 4000.
* nymphs.
* only 132 were I. scapularis, in 6 pools.
MEGAHIT (v 1.0.4) (Li et al., 2015) assemblers. The contigs and unique singletons were annotated by homology search with Megablast against the GenBank nucleotide database. The sequences that showed poor or no homology at the nucleotide level were screened with BLASTX against the viral GenBank protein database. Sequences from viral BLASTX analysis were subjected to second round of BLASTX homology search against the complete GenBank protein database to correct for biased e-values and taxonomic mis-assignments. The abundance of the bacterial and invertebrate reads was determined from the top MEGABLAST hits. The read abundance was further normalized as counts per ten thousand filtered reads. Prevalence data for each agent was determined from the normalized reads obtained from individual ticks. In order to exclude non-tick borne bacterial sequences that can originate from environmental contaminants, we compared and subtracted reads that were also detected in negative control samples. Negative controls consisted of non-tick samples that went through the entire sequencing protocol, including extraction, library preparation and HTS.

The association of each identified agent with either male or female ticks was determined by comparing the prevalence of each agent in male ticks versus female ticks using Fisher’s exact test, deriving a two-tailed p-value. Multiple comparisons were adjusted using Hochberg’s step-up procedure controlling the family-wise error rate (FWER) at α = 0.05 level (Hochberg, 1988). We also examined the association for every pair of agents using PAIRS (PAIRS, 2008) in which the observed C-score (Stone and Roberts, 1990) was compared to the expected C-score of the null model that represented complete randomness.

2.4. PCR

We designed a quantitative PCR assay for detection of R. buchneri targeting the ompB gene. For primer design, we used sequence data generated from this study, aligned with the ompB sequence of the reference strain (accession number EF432951.2). The assay consisted of fwd primer GATTACGGCAATCGAAGCTA, reverse primer, AGCATCGCAATCGAAGCTA, and probe GCCGCAGGTACAGTTATCACGG. We also designed an assay targeting the 16S rRNA gene of the genus Bartonella. To identify optimal sequences for primer design, we aligned sequences from 7 Bartonella species (B. henselae (accession number AJ223778), B. vinsonii (NR_104902), B. phocense (NR_115254), B. queenslandensis (EU111755), B. senegalesis (NR_125574), B. acomydis (NR_113288), B. grahamii (ABS29507)) to agents with high homology to Bartonella (Rhizobium pelarcatum (NR_116790)), agents identified in our HTS data (Burkholderia and Mycobacteriaceae sp.), and known I. scapularis-transmitted pathogens (B. burgdorferi (L39081), A. phagocytophilum (AB196720)). Primer sites were chosen that displayed the highest specificity to Bartonella.

3. Results and discussion

We sequenced TNA from 197 individual adult I. scapularis (105 males and 92 females) on 7 lanes of Illumina HiSeq 4000 platform. Sequencing yielded between 240 and 359 million reads per lane, with a minimum of 3.96 million and a maximum of 29.7 million reads per individual tick sample (Table 1). For confirmation of the presence (or absence) of agents identified in individual ticks, we examined 17 tick pools on two lanes that generated 368 million and 341 million reads, respectively, with a range of 5.8 and 63.9 million reads per single pool.

3.1. Borrelia

We identified two species of Borrelia in the HTS data. Borrelia burgdorferi s.s. was detected in 111 (56.3%) individual ticks, and Borrelia miyamotoi in 10 (5.07%) (Table 2). Although several other Borrelia species have been detected in ticks throughout the US, we did not find evidence of any of these species in our samples (Marconi et al., 1995; Postic et al., 2007; Rudenko et al., 2009a, 2009b, Margos et al., 2010; Pritt et al., 2016a, 2016b, Margos et al., 2017). There was no significant difference in the prevalence of Borrelia in male or female ticks. Our results are consistent with published surveillance studies of I. scapularis by our group and others that reported significantly higher infection rates of B. burgdorferi s.s. relative to B. miyamotoi (Tokarz et al., 2017; Wroblewski et al., 2017; Cross et al., 2018). At the five sites analyzed in our study, the prevalence of B. burgdorferi s.s. ranged from 40% to 62.5%.

Two human pathogens transmitted by I. scapularis, B. mayonii and Ehrlichia muris eauclairensis, were not detected in individual ticks or tick pools. Although these agents have been detected in up to 3% of I. scapularis in the Upper Midwest, they have not yet been found outside of this region (Pritt et al., 2011; Johnson et al., 2015; Murphy et al., 2017). Our results confirm that B. mayonii and E. muris eauclairensis have not gained a significant foothold in the I. scapularis populations in the Northeast (Stromdahl et al., 2014).

3.2. Anaplasma

A. phagocytophilum, the agent of human granulocytic anaplasmosis, was present in 21 (10.6%) ticks. The majority (90%) of the A. phagocytophilum-positive ticks were infected with the Ap-ha strain. We did not detect other Anaplasma species.

3.3. Babesia

Babesia was the lone protozoan identified in our tick samples. We detected Babesia sequences in 32 (16.2%) ticks. Two species were present. Fifteen ticks (7.6%) were infected with B. microti, the primary species implicated in babesiosis in the United States (CDC, 2018). Another 15 ticks were infected with B. odocoi, a species pathogenic in white-tailed deer and other animals, but not connected to human disease (Spindler et al., 1958; Gallatin et al., 2003; Holman et al., 2003). Two additional ticks were co-infected with both B. microti and B. odocoi. Two other species, B. duncani and B. divergens have also been implicated in babesiosis in the United States (Persing et al., 1995; Herwaldt et al., 1996; Beattie et al., 2002; Yabsley and Shook, 2013; Herc et al., 2018; Scott and Scott, 2018). B. duncani has been found on the West Coast, although the vector(s), prevalence and true geographic distribution of this species are unknown. Similarly, limited data exists for B. divergens, although this species has been linked with I. dentatus, a tick that feeds primarily on rabbits and birds (Goehert and Telford, 2003). The absence of B. duncani and B. divergens sequences in our data suggests that I. scapularis is not a vector for these species, at least in the Northeastern part of the US.

Table 2

<table>
<thead>
<tr>
<th>Agent</th>
<th>Prevalence (# of positive ticks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td></td>
</tr>
<tr>
<td>Anaplasma phagocytophilum</td>
<td>10.6% (21)</td>
</tr>
<tr>
<td>Borrelia burgdorferi s.s.</td>
<td>56.3% (111)</td>
</tr>
<tr>
<td>Borrelia miyamotoi</td>
<td>5.07% (10)</td>
</tr>
<tr>
<td>Rickettsia buchneri</td>
<td>65.0% (128)</td>
</tr>
<tr>
<td>Invertebrate</td>
<td></td>
</tr>
<tr>
<td>Babesia microti</td>
<td>8.6% (17)</td>
</tr>
<tr>
<td>Babesia odocoi</td>
<td>8.6% (17)</td>
</tr>
<tr>
<td>Ixodes scapularis nematode</td>
<td>13.7% (27)</td>
</tr>
<tr>
<td>Virus</td>
<td></td>
</tr>
<tr>
<td>Blacklegged tick phlebovirosis</td>
<td>73.10% (144)</td>
</tr>
<tr>
<td>BLTV-associated virus 1</td>
<td>5.10% (10)</td>
</tr>
<tr>
<td>BLTV-associated virus 2</td>
<td>1% (2)</td>
</tr>
<tr>
<td>Powassan virus</td>
<td>3.60% (7)</td>
</tr>
<tr>
<td>South Bay virus</td>
<td>21.80% (43)</td>
</tr>
<tr>
<td>Suffolk virus</td>
<td>9.60% (19)</td>
</tr>
</tbody>
</table>

* 2 ticks co-infected with Babesia microti and Babesia odocoi.
3.4. Rickettsia

A Rickettsia species was first identified in I. scapularis in 1991, and has subsequently been referred to as Rickettsia cooleyi, Rickettsia midichlorii and the recently adopted name, Rickettsia buchneri (Magnarelli et al., 1991; Billings et al., 1998; Benson et al., 2004; Moreno et al., 2006; Troughton and Levin, 2007; Kurtti et al., 2015). Like many other tick-associated Rickettsia, R. buchneri is suspected to be an endosymbiont. The endosymbiotic Rickettsia of ticks are transovarially transmitted, and presumed not to establish infection in other invertebrate hosts (Cheng et al., 2013). I. scapularis has not been implicated as a host of other species of Rickettsia. R. buchneri was the lone Rickettsia identified in our study. We detected R. buchneri in 128 (65%) of adult ticks (Table 2). Congruent with reports that infection is more common in female ticks, we detected R. buchneri in 89% of females versus 44% of males (Cross et al., 2018). This female tick-specific association of Rickettsia was the only significant sex-specific association among all identified agents. When assessed by quantitative PCR, the Rickettsia burden was not significantly different between male and female ticks (data not shown). To determine the prevalence of R. buchneri in immature stages of I. scapularis, we tested TNA of 25 nymphs by quantitative PCR, all 25 were Rickettsia-positive, suggesting that the loss in Rickettsia occurs at the time of the molt from nymphs to adults.

3.5. Bartonella

Bartonella species have been promoted as agents of tick-borne disease (Eskow et al., 2001; Telford and Wormser, 2010). Reports of Bartonella DNA in ticks, coupled with detection of this bacteria in patients with Lyme disease, have led to the proposal that Bartonella is a tick-transmitted pathogen and an upsurge in Bartonella testing of patients with a suspected tick-borne illness (Eskow et al., 2001; Adelson et al., 2004). Table 3 summarizes our evaluation of Bartonella 16S rRNA sequences using primers identified from the literature.

Table 3
Comparison of Bartonella 16S rRNA assays.

<table>
<thead>
<tr>
<th>Primer source</th>
<th>Primer sequence 5’ to 3’</th>
<th># of ticks tested</th>
<th># of positive ticks</th>
<th># of Bartonella positives PCR products*</th>
<th>BlastN (%) homology</th>
</tr>
</thead>
<tbody>
<tr>
<td>This study</td>
<td>Fwd- AATTTGGTCAATACCGTACGCCCT</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Rev- CAGCCTCTCCACCTCAAGATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adelson et al</td>
<td>Fwd- GGAATCTCCGCTCTCAAGGCTTG</td>
<td>45</td>
<td>10</td>
<td>0</td>
<td>Phyllobacterium sp (100%), Ensifer sp (100%)</td>
</tr>
<tr>
<td></td>
<td>Rev- GGCAATCCCGGAGATGGTTTGAGATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* bacteria were identified through BlastN homology searches.

Fig. 2. Phylogenetic characterization of a putative nematode identified in this study. Relationships were inferred through the alignment of a 1562 nucleotide fragment of the 18s rRNA gene of related filarial and non-filarial nematodes using CLUSTALW in Geneious 10.2.4. The sequence obtained in this study is indicated with an asterisk. The phylogenetic tree was constructed in MEGA 7.0.26 and the robustness of each node was determined using 1000 bootstrap replicates.
et al., 2004; Holden et al., 2006).

We did not detect *Bartonella* sequences in individual ticks or in tick-pools. We speculate that differences between our findings and published reports implicating *I. scapularis* as potential vectors of *Bartonella* may reflect several factors. First, the presence of microbial DNA within a tick does not conclusively establish the tick’s competency as a vector, or the viability of the microbe. The DNA may instead be a remnant of a previous blood meal, and its concentration sufficient to be amplified and detected by sensitive molecular assays such as PCR. Rodents, a typical source of blood meals for *I. scapularis*, are frequent hosts of *Bartonella* spp and soil bacteria. To address this confound, we designed primer pairs with the highest degree of specificity used to screen targeted a conserved fragment of the *Bartonella* 16S rRNA gene. However, both primers were also homologous to the corresponding 16S rRNA sequence from a wide range of soil bacteria and may have amplified trace amounts of soil bacterial contaminants in tick homogenates. This has also been suggested previously (Telford and Wormser, 2010). In addition, we found that the resulting 240 nucleotide fragment is not sufficient to adequately discriminate between a PCR product obtained from Bartonella spp and soil bacteria. To address this confound, we designed alternative 16S rRNA Bartonella-specific assays that would exclude soil bacteria. We tested multiple primer pairs, with the highest degree of specificity achieved with the primer pair shown in Table 3. We then tested TNA from 45 ticks from our study and did not identify any Bartonella-positive ticks. However, when we used the same samples with primers employed by Adelson et al, 10 (13%) were positive. When we sequenced these PCR products, all were 100% identical to soil bacteria and not *Bartonella* (Table 3). Our results do not exclude that a portion of Lyme disease patients may indeed have a *Bartonella* infection. Nonetheless, we postulate that these infections are not due to *I. scapularis* transmission but may be acquired through alternative means of exposure, such as contact with infected domestic cats, or flea bites (Klotz et al., 2011). We also acknowledge that our results are limited to *I. scapularis* within the New York metropolitan area and do not exclude a role for other tick species in *Bartonella* transmission (Holden et al., 2006).

### 3.6. Other bacteria

Previous studies have shown that the diversity of *I. scapularis* bacteriome is limited (Ross et al., 2018). In our samples, *R. buchneri*, *Borrelia*, or *Anaplasm* (when present) typically accounted for > 80% of all bacterial reads. We also identified reads for *Wolbachia*, *Mycobacteria*, and *Burkholderia*. With the exception of *Wolbachia*, the low number of reads from other genera suggest that these bacteria may represent environmental contaminants.

### 3.7. Virome

We previously identified sequences of 21 putative viruses in *I. scapularis* (Tokarz et al., 2018). Fourteen of these viruses were also present in the individual ticks from this study, along with sequences of a novel totivirus. Consistent with our previous data, the most prevalent viruses were the transovarially transmitted Blacklegged tick phleboviruses 1 and 2 (found in 73% of the ticks), *South Bay* virus (22%), and *Suffolk* virus (10%). The pathogenic Powassan virus (deer-tick lineage) was present in 7 (3.6%) ticks.

Three rare viruses, Laurel Lake virus, *Ixodes scapularis* associated virus-5 and *Ixodes scapularis* associated virus-6, were all present in a single tick (tick sample RTs-604). All three viruses have limited homology to viruses identified in metagenomic analyses of fungi. In addition, the HTS data for RTs-604 were unique as it contained > 2 million reads for *Cordyceps bronniarii*, an entomopathogenic fungus. We also identified a tick pool that was positive for these three viruses as well as *C. bronniarii*. These data suggest that all three viruses are likely of fungal origin.

### 3.8. *I. scapularis* filarial nematode

Filarial nematodes were first observed in *I. scapularis* homogenates...
34 years ago (Beaver and Burgdorfer, 1984). Recently, filarial nemato
dese sequences were identified in I. scapularis from Connecticut (by PCR) and Wisconsin (by HTS) (Namrata et al., 2014; Cross et al., 2018). We identified nematode sequences in 27 (13.7%) ticks. Phylogenetic analysis of the complete 18S rRNA sequence indicated that it belongs to a distinct filarial nematode species with the highest genetic similarity to Acanthocheilonema viteae (Fig. 2). In addition, all ticks positive for Wolbachia were also positive for the filarial nematode. Since Wolbachia are known endosymbionts of nematodes the Wolbachia reads presumably originated from the nematodes, as suggested previously (Cross et al., 2018).

3.9. Co-infections

Thirty-eight ticks (19%) were co-infected with known pathogens (A. phagocytophilum, B. burgdorferi s.s., B. miyamotoi, B. microti, and Powassan virus) (Table 4). Thirty-two ticks (16%) were co-infected with two agents, and 6 (3%) with 3 agents. We also examined associations between the bacterial, eukaryotic and viral agents identified in our study. Two positive associations were identified (B. microti with BLTV1 and BLTPV and B. burgdorferi s.s. with D. variabilis). Ticks positive for D. variabilis were also positive for the

We identiﬁed nematode sequences in 27 (13.7%) ticks. Phylogenetic analysis of the complete 18S rRNA sequence indicated that it belongs to a distinct filarial nematode species with the highest genetic similarity to Acanthocheilonema viteae (Fig. 2). In addition, all ticks positive for Wolbachia were also positive for the filarial nematode. Since Wolbachia are known endosymbionts of nematodes the Wolbachia reads presumably originated from the nematodes, as suggested previously (Cross et al., 2018).

In conjunction with recently published microbiome studies from other geographical sites, we now possess a better understanding of the spectrum of agents that can be harbored by I. scapularis, which can serve to focus research and clinical treatment.

Acknowledgments

We’d like to thank Nischay Mishra and Simon H. Williams for assistance with the manuscript.

This work was supported by a grant from the Steven & Alexandra Cohen Foundation(SACF CU15-4008).

References


