

# The Diversity and Evolutionary Relationships of Ticks and Tick-borne Bacteria in China

**Junhua Tian**

Huazhong Agriculture University

**Xin Hou**

Sun Yat-Sen University

**Mihong Ge**

Wuhan Academy of Agricultural Sciences

**Hongbin Xu**

Jiangxi Province Center for Disease Control and Prevention

**Bin Yu**

Wuhan Centers for Disease Prevention and Control

**Jing Liu**

Wuhan Centers for Disease Prevention and Control

**Renfu Shao**

University of the Sunshine Coast

**Edward C. Holmes**

The University of Sydney

**Chaoliang Lei**

Huazhong Agriculture University

**Mang Shi** (✉ [shim23@mail.sysu.edu.cn](mailto:shim23@mail.sysu.edu.cn))

Sun Yat-Sen University <https://orcid.org/0000-0002-6154-4437>

---

## Research Article

**Keywords:** Ticks, metagenomics, bacteria, evolution, co-divergence

**Posted Date:** December 20th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1123049/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Ticks (order Ixodida) are ectoparasites of vertebrates that transmit diverse pathogens to humans and domestic animals. However, information on the genomic diversity of ticks in China is currently limited to a small number of common species, leaving substantial knowledge gap in the evolution of both ticks and their associated bacterial.

## Results

We collected more than 20,000 contemporary and historical (up to 60 years of preservation) tick samples representing a wide range of tick biodiversity across diverse geographic regions in China, including 18 common species, nine rare species, and two undetermined species. Metagenomic sequencing was performed on individual ticks to obtain the complete or near-complete mitochondrial (mt) genome sequences from 46 tick species, among which 30 species were revealed for the first time. These new mt genomes data greatly expanded the diversity of many tick groups and revealed five cryptic species. Utilizing the same metagenomic sequence data we identified divergent and abundant bacteria in *Haemaphysalis*, *Ixodes*, *Dermacentor* and *Carios* ticks, including nine species of pathogenetic bacteria and potentially new species within the genus *Borrelia*. We also used these data to explore the evolutionary relationship between ticks and their associated bacteria, revealing a pattern of long-term co-divergence relationship between ticks and *Rickettsia* and *Coxiella* bacteria.

## Conclusions

In sum, our study provides important new information on the prevalence of tick-borne pathogens in China and sheds new light on the long-term evolutionary and ecological relationships between ticks and their associated bacteria.

## Introduction

Ticks (Acari: Ixodidae) are blood-feeding ectoparasites of vertebrates that are notorious for their role in the transmission of a variety of infectious diseases. Modern ticks are classified into three families: Argasidae (soft ticks), Ixodidae (hard ticks) and Nuttalliellidae. While there is only one extant species in the Nuttalliellidae, considered the closest extant relative to the ancestral tick lineage, more than 700 and 200 recognized species have been identified within the Ixodidae and Argasidae, respectively. [1–3].

China covers a large geographic area and possesses a variety of ecosystems. To date, at least 125 tick species from nine genera have been reported across 34 provinces in China, representing 13.9% of tick species identified globally [4]. The most frequently reported species in China are *Haemaphysalis*.

*longicornis*, *Dermacentor. silvarum*, *Ixodes. persulcatus*, *Ha. conicinna*, *Rhipicephalus. microplus*, and *Rh. sanguineus* [5, 6]. *De. sinicus*, *Ix. sinensis*, *Ha. tibetensis* and *Ha. qinghaiensis* have been only reported in China, although there is a lack of genomic data for these species.

Mitochondrial (mt) genomes are widely used in molecular systematics because their relatively high rate of evolutionary change provides greater phylogenetic resolution at the genus or family levels [7]. For ticks, complete mt genomes representing 66 species from 18 genera have been sequenced to date [8, 9]. As in most arthropods, tick mt genomes are circular, 14-16 kb in length and contain 37 genes, including 13 protein coding genes, 22 tRNA genes and two rRNA genes. Phylogenetic analyses based on protein coding and rRNA gene sequences shows great consistency in genus level classification [8]. A single genome re-arrangement event has been identified in the Ixodidae, including the genera *Rhipicephalus*, *Dermacentor*, *Amblyomma* and *Haemaphysalis*, that possess a translocation of tRNA genes (*trnL*<sub>1</sub>, *trnL*<sub>2</sub>, *trnC*) and an inversion of the *trnC* gene [10, 11].

Ticks has long been regarded as disease vectors, with an increasing number of human disease associations described in recent years [12–15]. Since 1982, more than 30 emerging tick-borne disease agents have been identified from 28 tick species, causing a variety of human infections including rickettsial disease, Q fever and Borrelioses. Together, these account for nearly 50,000 cases per year since 2013 as reported by National Notifiable Disease Surveillance System in the USA [16, 17]. Among these, rickettsial disease and Q fever are caused by two groups of obligate intracellular endosymbiont bacteria in ticks: the spotted fever group (SFG) of the genus *Rickettsia* [18, 19] and *Coxiella burnetii* [20], harbored by hard and soft ticks, respectively. In contrast, Borrelioses, represented by Lyme disease and relapsing fever, are mainly caused by bacteria in the genus *Borrelia* (phylum Spirochaetes). Lyme borreliosis (LB) species (*B. burgdorferi*) are generally carried by *Ixodes* ticks, whereas the relapsing fever *Borrelia* group are generally carried by *Ornithodoros. sonrai*, *Or. erraticus* and *Or. moubata* [21, 22]. In addition, ticks harbor a number of other bacteria in the genera *Rickettsia*, *Coxiella*, and *Borrelia*. While their public health significance remains unclear, these microbes are highly prevalent in tick species and can be transmitted transovarially [23–26].

Despite growing interest in tick-borne pathogens, our knowledge of the genetic diversity of ticks in China and the bacteria they carry are limited to a small number of common species. Herein, we collected and sequenced 46 tick species representing the biodiversity of ticks across China. For the first time we revealed the genetic diversity of both ticks and their bacterial symbionts, enabling a more systematic study of their co-evolutionary history.

## Results

### Morphological identification and characterization of ticks

Between 1959 and 2019 more than 20,000 ticks were collected from a wide range of hosts (e.g. cattle, goats, camels, hedgehogs) across China (Figure 1, Table 1, Additional file 1: Table S1). Species

identification of these ticks were carried out based on morphological characters, such as the number of punctuations, size and shape of mouthparts, and genital aperture [27–29]. This revealed a total of at least 46 species comprising two families (Ixodidae and Argasidae) and eight genera: *Haemaphysalis* (19 species), *Ixodes* (10 species), *Dermacentor* (6 species), *Rhipicephalus* (four species), *Hyalomma* (two species), *Amblyomma* (three species), *Argas* (one species) and *Carios* (one species). Among these were the six most common tick species in China [6]: *Rh. microplus*, *Rh. sanguineus*, *Ix. persulcatus*, *Ha. longicornis*, *De. silvarum* and *Hy. asiaticum*. In addition, other common species were collected and analyzed, including *Ix. sinensis* (the vector of *B. burgdorferi* [30]), *Ix. ovatus*, *De. steini*, *Ha. yeni*, *Ha. concinna*, *Hy. scupense*, *Rh. turanicus*, *Rh. haemaphysaloides* and *Ar. persicus*, all of which were sampled from at least two provinces. In comparison, other species were more locally defined, such as *Ha. tibetensis* from Tibet, *Ha. qinghaiensis* and *Ha. danieli* from Qinghai, *Hy. asiaticum*, *Ha. punctata*, *De. marginatus*, and *De. niveus* from Xinjiang, *Ix. acutitarsus* from Hubei, and *Ha. lagrangei* and *Ha. mageshimaensis* from Hainan. In addition, we obtained a number of every rare species, particularly *Am. javanense*, *Ix. simplex*, *Ix. nuttallianus*, *Ix. crenulatus*, *Ix. kuntzi*, *Ha. kitaokai*, and *Ca. vespertilionis*, some of which were collected from wildlife animals including bats, pangolins and flying squirrels (Pteromyini). Other samples were collected by drag-flagging methods or were historical samples preserved in ethanol for more than 60 years. For example, the oldest sample from our data set (*Am. javanense*) was collected in the 1960s from a wild Chinese pangolin (*M. pentadactyla*). In the case of two samples (i.e. C20, A29) we could only identify them to the genus level, in *Amblyomma* and *Ixodes*, respectively. Since we were unable to identify them at the species level using morphological characteristics, they were tentatively assigned as potential new species.

Table 1

Sampling locations, animal hosts and bacterial pathogens of tick species in China at the genus level. The number of libraries denotes that how many libraries are included in each tick genus while bracketed numbers denote the numbers of mt genomes successfully obtained in each group.

Species(genus)	Location	Host	No. of libraries	Pathogen
<i>Rhipicephalus</i>	Hubei, Jiangxi, Yunnan, Beijing, Zhejiang, Xinjiang, Hainan	Cattle, dog, hedgehog	12 (11)	<i>Rickettsia</i> , <i>Coxiella</i> ,
<i>Hyalomma</i>	Xinjiang, Hubei	Cattle, goat, camel	4 (3)	—
<i>Dermacentor</i>	Hubei, Jiangxi, Inner Mongolia, Beijing, Hebei, Zhejiang, Xinjiang	Cattle, goat, rabbit, hedgehog, wild boar	18 (15)	<i>Rickettsia</i> , <i>Coxiella</i> , <i>Rickettsiales</i> unclassified
<i>Amblyomma</i>	Hubei, Zhejiang, Hainan	Wild boar, Malaysia pangolin, Chinese pangolin	6 (2)	<i>Rickettsia</i>
<i>Haemaphysalis</i>	Hubei, Jiangxi, Fujian, Zhejiang, Hainan, Yunnan, Beijing, Shaanxi, Heilongjiang, Liaoning, Qinghai, Tibet, Xinjiang, Inner Mongolia	Cattle, yak, pheasant, dog, hedgehog, muntjac, goat, wild boar, hog badger	30 (23)	<i>Rickettsia</i> , <i>Coxiella</i> ,
<i>Ixodes</i>	Hubei, Jiangxi, Jiangsu, Zhejiang, Hainan, Inner Mongolia, Jilin, Tibet, Heilongjiang	Cattle, goat, Eurasian badger, hog badger, rabbit, bat, pteromyini, tupaia	22 (16)	<i>Rickettsia</i> , <i>Coxiella</i> , <i>Wolbachia</i>
<i>Carios</i>	Hubei, Henan	—	2 (2)	<i>Coxiella</i> , <i>Rickettsia</i> , <i>Borrelia</i>
	Heilongjiang, Xinjiang	Chicken	2 (2)	<i>Rickettsia</i> , <i>Coxiella</i>
Total	—	—	96 (74)	—

## Mitochondrial Genomes Of 74 Ticks Of 46 Species

We sequenced the total DNA of 96 individual or mixed tick samples, which generated an average of 7.87 Gb of clean reads for *de novo* assembly and annotation. Complete or near-complete mt genome sequences were successfully obtained from 74 of the 96 libraries, including 30 species whose mt genomes were reported for the first time. The length of the newly identified mt genomes ranged from 14,428 bp to 15,307 bp, and the AT content varied from 72.29% (*Ixodes* sp. A29) to 81.06% (*Ha. danieli* Z14), similar to previously identified mt genomes of ticks [8]. Furthermore, the structure, composition, as well as the arrangement of genes largely followed their closest relatives within the same genus [8, 31].

The only differences were observed in the length and composition of non-coding regions, some of which contain more than one tandem repeat region (Additional files 2, 3, 4: Figure S1-S2, Table S2). For example, the mt genome of *De. marginatus* E48 has an extra copy of the non-coding region so that its length (i.e. 15,307 bp) has surpassed that of *Ix. tasmani* (NC 041086.1, 15227 bp) [8] to become the longest tick mt genome identified to date (Additional file 2: Figure S1). Furthermore, we found inconsistencies in the control region within some individual samples. For example, cloning of sequencing of PCR products spanning the control region between *trnQ* and *trnF* reveals various copy numbers of short repeat sequences within the same (single-tick, *De. marginatus* E1) sample (Additional file 3: Figure S2C).

## Molecular Identification And Genetic Diversity Of Ticks

Both maximum likelihood (ML) and Bayesian phylogenetic trees were estimated based on sequences of 13 protein coding genes and two rRNA genes derived from 136 tick mt genomes, including 74 generated in this study and 62 reference mt genomes from GenBank. The ML and Bayesian methods resulted in highly similar tree topologies that placed the diversity of Chinese ticks within a global context with high resolution (Figure 2, Additional file 5, 6, 7, 8, 9, 10, 11, 12: Figure S3-S10). Importantly, the newly added genomes greatly expanded the diversity of many groups, particularly the genera *Haemaphysalis*, *Ixodes* and *Dermacentor* (Figure 2). In addition to the new species identified in this study, 28 tick species previously only known through morphological characteristics or incomplete mt genomes (e.g. *Ix. kuntzi*, *Ix. acutitarsus*, *Ha. mageshimaensis* and *Ha. colasbelcourii*) were also included (Figure 2, Additional file 5, 6: Figure S3-4). Furthermore, at least five potential cryptic species were identified – *Rh. sanguineus*, *De. steini*, *De. marginatus*, and *Ix. ovatus* – adding to the previously reported cryptic species identified in *Rh. microplus* [6, 32]. Each contained at least two divergent (70.93% ~ 94.21% identity) phylogenetic clusters while sharing the same morphological characteristics (Figure 2). Conversely, *De. sinicus*, *De. nuttalli*, and *De. silvarum* shared a very close relationship (> 98.46% identity) even though these were separate species based on morphological characteristics. Interestingly, *De. nuttalli*, and *De. silvarum* cannot be distinguished based on mt genome phylogeny, although they had quite distinctive Trochanter I dorsal spur (Additional file 13: Figure S11)

## Discovery and characterization of bacterial endosymbionts and pathogens in ticks

Based on DNA sequencing, we searched for members of the tick-borne bacterial groups that are known to harbor human pathogens: the order Rickettsiales, genus *Coxiella* and genus *Borrelia*. Overall, 56% (54/96) of tick libraries were positive for these bacterial groups, among which *Coxiella* had the highest prevalence (40/96, 42%), followed by *Rickettsia* (26/96, 27%), *Wolbachia* (1/96, 1%), and *Borrelia* (1/96, 1%) (Figure 3A, Additional file 14: Table S3).

Most of the species in the order Rickettsiales belonged to the genus *Rickettsia*, within which 14 bacterial species were identified from all tick genera included in this study (with the exception of *Hyalomma*; Figure 3B), including a number of human pathogens. For example, *R. raoultii*, which causes human tick-borne lymphadenitis [33, 34], was identified from *De. marginatus* and *De. niveus* in Jinghe, Xinjiang province, a region where *R. raoultii* have previously been reported [35–37]. Within Xinjiang (Bole and Zhaosu) we identified *R. sibirica* and *R. africa* circulating in *De. sinicus* and *Ix. vespertilionis*, which are responsible for a range of tick-borne diseases, including Siberian tick typhus (STT) in Asia and African tick bite fever (ATBF) in Africa [19]. In addition, we discovered *R. heilongjiangensis*, the newly reported agent of Far-Eastern spotted fever (FESF) [38]. This bacterium was previously found in *De. sylvarum* ticks from Heilongjiang, and herein it was associated with the tick species *Ha. campanulata* and *Ha. cornigera* in Hubei and Jiangxi provinces located in central China. Other pathogenic *Rickettsia* species included *R. tamurae*, *R. monacensis* and *R. helvetica* identified from *Am. testudinarium*, *Ix. sinensis* and *Ix. kuntzi*.

Notably, we found a potentially novel species within Spotted Fever Group that was relatively divergent (< 99.46% genetic identity in six genes) from the rest of the bacteria in this group (Additional file 15: Table S4). Since the species were identified from *Ha. megaspinasa*, we tentatively named it *Rickettsia endosymbiont of Haemaphysalis megaspinasa*. Furthermore, we identified four genetically divergent *Rickettsia* species that occupied basal phylogenetic positions. Among these, *Rickettsia endosymbiont of Ixodes persulcatus* H5 and N2 clustered with *R. canadensis* (97.78% identity), *Rickettsia endosymbiont of Argas persicus* H1 fell with *R. bellii* and *Rickettsia sp.* MEAM1 (*Bemisia tabaci*) (92.16% and 90.96% identity), and *Rickettsia endosymbiont of Ixodes vespertilionis* A54b, *Rickettsia endosymbiont of Carios vespertilionis* X1 formed a monophyletic group with *Rickettsia endosymbiont of Culicoides newsteadi* despite a high level of divergence (86.08% and 85.97% identity, respectively) (Additional file 15, 16, 17, 18: Table S4, Figure S12-S14).

In addition to *Rickettsia*, we identified two potential new species within the order Rickettsiales. One, *Wolbachia endosymbiont of Ixodes vespertilionis* A54, clustered with *W. pipientis* strain FL2016 (95.37%) and *Wolbachia endosymbiont of Drosophila melanogaster* (95.16%) within the genus *Wolbachia*. The identification of *Wolbachia* in ticks has been reported in recent years [39, 40]. The other - *Rickettsiales endosymbiont of Dermacentor* - clustered with an unclassified Rickettsiales bacterium Ac37b identified from *Am. cajennense* in Brazil (86.62% identity). Together they may represent a new genus or even family within the Rickettsiales (Additional file 15, 16, 18: Table S4, Figure S12-S13).

Bacteria of the genus *Coxiella* had the highest prevalence among the tick species examined (74%, 40/54). The newly discovered *Coxiella* species in the present study are highly diverse and greatly expand the genetic diversity within this group (Figure 3C). Indeed, new genetic lineages were defined based on our phylogenetic analysis, including *Coxiella endosymbiont of Dermacentor marginatus*, *Coxiella endosymbiont of Haemaphysalis concinna* and *Coxiella endosymbiont of Ixodes ovatus*, most of which were divergent from existing members of *Coxiella* and generally associated with specific tick genera (Figure 3C, Additional file 19: Figure S15). In contrast, we discovered a *C. burnetii*, the causative agent of Q fever [41], from a *De. sinicus* tick sampled from Xinjiang province, which had high abundance (29.92

RPM) and was closely related with the “Dugway 5J108-111” strain sampled from the United States and transmitted by *De. andersoni* [42] (Figure 3C, Additional file 19: Figure S15). In addition to *C. burnetii*, we identified a single species of *Borrelia* from a soft tick *Ca. vesperitilionis* in Henan province. Based on the phylogenetic analyses, the newly identified bacteria, named *Borrelia henanensis* X1, fell within a clade “RF” that contains pathogens causing tick-borne relapsing fever (Figure 3D, Additional file 20: Figure S16) [43, 44].

## Ecological and evolutionary patterns in ticks and their associated bacterial symbionts

We used Mantel tests to examine whether the tick host and/or geographic factors shape the genetic diversity of the bacteria they carry. For both *Rickettsia* and *Coxiella*, our results revealed positive and significant ( $P < 0.0005$ ) correlations between tick and bacteria genetic distance matrices. However, no such significant correlation was found between bacterial genetic distance and geographic distance. Similar results were obtained using (i) partial Mantel analyses, in which we tested the effect between two factors while controlling for the third, and (ii) multiple linear regression analyses in which we tested the effect between three matrices (Table 2, Additional file 21: Table S5). These results suggested that bacterial genetic diversity was primarily shaped by host genetic distance, with geographic distribution having little or no impact. The strong impact of host on bacterial genetic diversity was also reflected in the phylogenetic analysis in which we observed a significant clustering of bacterial genetic diversity at the host general level (*Rickettsia*: association index (AI) = 2.760,  $P < 0.001$ ; *Coxiella*: AI=0.969,  $P < 0.001$ ).

Table 2

Results of the Mantel test and Partial Mantel test comparing two factors (host genetic distance and geographic distance) that predict the structure of genetic diversity in bacterial pathogens.

	Model	r value (P value)
<i>Rickettsia</i>	Host <sup>a</sup>	0.5868 (0.0004)
	Host   geography <sup>b</sup>	0.5952 (0.0009)
	Geography <sup>a</sup>	0.0028 (0.3916)
	Geography   host <sup>b</sup>	-0.1225 (0.8633)
<i>Coxiella</i>	Host <sup>a</sup>	0.4939 (0.0001)
	Host   geography <sup>b</sup>	0.4939 (0.0001)
	Geography <sup>a</sup>	0.0013 (0.4499)
	Geography   host <sup>b</sup>	0.0053 (0.4378)
<sup>a</sup> Mantel test.		
<sup>b</sup> Partial Mantel test.		
Indicate that the first factor excludes the effect of the second.		

We next examined whether the phylogeny of the ticks and their bacterial symbionts exhibited a pattern of bacterial-host co-divergence over evolutionary time. We first tested hypothesis of co-divergence using an event-based framework, based on which we reconciled the phylogenies of ticks and their associated bacteria (i.e. *Rickettsia* and *Coxiella*, respectively) by accounting for four processes: co-divergence, duplication, host switching and loss [45]. This revealed significantly fewer non-co-divergence events (i.e. duplication, host switching and loss) than expected by chance alone. We similarly examined the co-divergence hypothesis using a distance method, in which we evaluated the overall phylogenetic congruence by comparing the tick and bacterial symbionts patristic distance [46]. This confirmed the significant overall similarity (ParafitGlobal,  $P = 0.0021$  and  $0.0003$ , respectively, for *Rickettsia* and *Coxiella*, at 9999 permutations) between the tick and bacterial symbionts phylogenies (Figure 4). Collectively, these results suggest that the symbiotic bacteria from genera *Rickettsia* and *Coxiella* have co-diverged with their tick hosts for many millions of years.

## Discussion

We collected more than 20,000 ticks and determined the mitochondrial genomes of at least 46 species representing the diversity of both common and rare ticks in China. Our sampling mostly covered human residential areas as well as some biodiversity hotspots, namely Shennongjia forest (Hubei province), the Tibet plateau (Tibet and Qinghai provinces), Hulun Beir prairie (Inner Mongolia). While the majority of ticks were sampled from domestic animals and were commonplace [6], those identified from wildlife or directly from the environment yielded more unique diversity. Hence, there may be many more species of ticks in China that have not yet been identified by current disease or human-centered sampling schemes. Indeed, a number of tick species, such as *Ix. kuntzi* and *Ha. colasbelcourii* identified from Taiwan and Laos [47, 48], have not been sequenced and characterized previously. Furthermore, there is a general lack of genomic surveillance of ticks in reptiles and amphibians, such that substantial evolutionary gaps remain in studying the long-term tick-host relationship.

The results of the morphological and molecular species identification performed here were largely similar, suggesting that mitochondrial sequencing can reliably identify tick species. However, there were several inconsistencies, mainly reflected in the presence of five cryptic species complexes, within which morphologically identical ticks can be separated by high levels of intra-specific genetic diversity (e.g. 73.40%, 70.93%) or that had paraphyletic phylogenetic structure (e.g. *De. marginatus* and *Rh. sanguineus*) (Figure 2). Although we treat all these as cryptic species, we could not completely exclude that there may be some additional morphological characteristics that can distinguish these ticks and were unrecognized during our species identification process. In addition to cryptic species, inconsistencies between morphological and mitochondrial taxonomy were also reflected in the observation that several morphologically distinct ticks (*De. nuttali*, *De. silvarum*, *De. sinicus*) had very similar mt sequences (i.e. > 98.46% genetic identity). This suggests a relatively recent speciation event, although this needs to be further confirmed with nuclear genes.

Our study also revealed the high prevalence and large diversity of bacterial endosymbionts within the ticks examined, many of which fell with groups/species that contain human pathogens, including members of Spotted Fever Group (nine species, genus *Rickettsia*), *C. burnetii*, as well as *B. henanensis*. Given that our study only covers 219 individuals from 46 species of ticks, the prevalence rate for pathogenetic bacteria within these ticks is relatively high. This broadly agrees with previous studies that revealed a 41.2% - 68.5% prevalence for pathogenetic bacteria [49, 50]. Furthermore, the abundance of some of the bacterial pathogens was high (1.42 - 162.57 RPM), further facilitating inter-host transmission. Interestingly, we did not detect other pathogenic genera within order Rickettsiales that are frequently carried by ticks, such as *Anaplasma* and *Ehrlichia*, although this may simply reflect a limited sample size.

Our results greatly enrich the diversity of both ticks and their associated bacteria, revealing that both endosymbiotic bacteria groups, namely those of the *Rickettsia* and *Coxiella* genera, had close association with ticks. This is also reflected in both the strong host structure on the bacteria phylogeny and the high resemblance between bacteria and host phylogenies indicative of long-term a co-divergence. Previous studies have suggested a lack of co-divergence between *Rickettsia* [51] and *Coxiella*-like endosymbionts

[26]. However, these were mainly based on bacteria detected from PCR assays that are highly sensitive and may easily include both endosymbionts as well as bacteria transmitted through co-feeding. In addition, some previous studies of co-divergence are based on the rRNA gene which is too conserved to distinguish closely related species [52]. In contrast, the unbiased metagenomic sequencing performed here only detects those bacteria at relatively high abundance, such that these data are more reflective of the presence of endosymbiotic bacteria. Clearly, more data is required to fully resolve the co-evolutionary history between ticks and their endosymbiont bacteria.

## Conclusion

In sum, our analysis of more than 20,000 ticks collected over broad geographic range across China provides insights into diversity and evolutionary history of ticks and their associated bacteria symbionts/pathogens. Our data reveal that the genetic diversity of ticks in China goes beyond a few common species and includes rare and under-explored species, for which more diversity in wildlife hosts that remains to be discovered. Importantly, despite their low occurrence, these uncommon tick species can harbor diverse pathogens, some of which could pose a potential threat to human health.

## Materials And Methods

### Sample collection

Between 1959 and 2019 more than 20,000 ticks belonging to eight genera (*Rhipicephalus*, *Hyalomma*, *Dermacentor*, *Amblyomma*, *Haemaphysalis*, *Ixodes*, *Argas* and *Carios*) were collected from different geographic regions in China (Figure 1, Additional file 1: Table S1). Most ticks were directly picked from infested wild and domestic animals, although some were captured using a drag-flagging method. Upon capture, each tick was sexed and morphologically identified to species by trained field biologists and further confirmed by sequencing mitochondrial genome sequences. Among these samples, 42% (*Ix. simplex*, *Am. testudinarium* and *Ha. qinghaiensis*) were preserved in alcohol at room temperature for 1-60 years, while others were captured alive and stored at -80°C until DNA extraction.

### Dna Extraction And Sequencing

For most of the samples we used a direct total DNA extraction and sequencing approach. For each sample, the ticks were first homogenized in PBS solution using the Mixer mill MM400 (Restsch). The homogenates were then subject to DNA extraction using the QIAamp DNA Mini Kit (Qiagen) and Mollusc DNA Kit D3373 (OMEGA) following the manufacturer's recommendations. Library preparations were performed by BGI · Tech and/or Novogene and then 150bp pair-end sequencing on an Illumina HiSeq 4000 platform.

For two of the libraries (A3 and A4), the mitochondria were first purified from cell homogenates before DNA extraction. Library preparation and sequencing were performed as described above. To purify the

mitochondria, cell homogenates were first subjected to low speed centrifugation (10 min at 100g) for insoluble debris removal. The supernatant was further centrifuged at 10000g for 5 min to obtain the mitochondria. The total sequenced data generated varied from 0.2 to 30.1 Gbp. For some of the libraries (A41, C7, D23 and C23) where mitochondrial genome received partial coverage, we increased the sequencing depth to 4.5 to 12.61 Gbp to obtain better coverage.

## Mitochondrial Genome Assembly And Phylogenetic Analyses

Clean sequence reads were assembled *de novo* using MEGAHIT version 1.2.6 [53] with default parameters. To obtain mt genome sequences, the assembled contigs were compared against reference mt genomes in GenBank using blastn [54]. We set the *E*-value to 1E-10 to maintain high sensitivity and a low false-positive rate. Contigs with unassembled overlaps were merged to form longer mt contigs using the SeqMan program implemented in the Lasergene software package version 7.1 [55]. We used Bowtie2 [56] for confirmation and/or extension of the mt genomes. The protein-coding genes and rRNA genes of the assembled mt genomes were annotated using the Spin program implemented in the Staden package [57]. The annotated mt genomes of these ticks were deposited in GenBank under accession numbers XXX-XXX.

To infer phylogenetic trees, we used 74 mt genomes generated in this study plus an additional of 62 reference tick mt genomes obtained from GenBank. Individual alignments were performed on each of the 13 mt protein coding genes (*ATP6*, *ATP8*, *COX1*, *COX2*, *COX3*, *CYTB*, *NAD1*, *NAD2*, *NAD3*, *NAD4*, *NAD4L*, *NAD5*, *NAD6*) and two rRNA genes (*12S rRNA* and *16S rRNA*). Protein coding genes were aligned based on codons using the ClustalW implemented in MEGA version 5.2 [58]. Ambiguously aligned regions were removed. Two rRNA genes were aligned using the MAFFT version 7.4 [59] employing the L-INS-i algorithm with all ambiguous aligned regions removed using TrimAL version 1.2rev59 [60]. Individual gene alignments were then concatenated to form super-alignments for subsequent phylogenetic analysis. These comprised: an (i) “all gene” data set containing 13 protein coding genes and two rRNA genes, and (ii) “a protein coding gene” data set that only contained the 13 protein coding genes. All data sets were analyzed using both maximum likelihood (ML) and Bayesian algorithms implemented in IQ-TREE version 1.6.12 [61] and MrBayes version 3.2 [62], respectively. For ML analyses, the best-fit nucleotide substitution model was selected using ModelFinder [63] and tree topologies were evaluated with an ultrafast bootstrap approximation approach [64] with 1000 replicates. For Bayesian tree inference, we used the substitution model as described above with 1,000,000 generations in two runs and each was sampled every 500 generations with a burn-in of 25%.

## Discovery And Characterization Of Bacterial Pathogens In Ticks

For each library, we searched for the existence of marker genes of Rickettsiales (*groEL*, *gltA*), *Borrelia* (*groEL*, *bamA*, *flaB*) and *Coxiella* (*groEL*, *rpoB*). The assembled contigs were compared against the database of reference marker genes downloaded from GenBank using blastx, with an *E*-value cut-off set at 1E-10. For libraries with incomplete marker gene coverage, partial gene sequences were obtained by first mapping reads to a reference marker gene sequence of bacterial pathogens using Bowtie2 [56] and generating consensus sequences from the mapped reads. Based on these genes, we then compared the bacterial pathogens identified in this study with those previously described by estimating phylogenetic relationships using the ML and Bayesian methods described above.

## The Co-phylogenetic Relationship Between Ticks And Their Bacterial Pathogens

We used the BaTS (Bayesian tip-association significance testing) program [65] to test whether bacteria pathogens (*Rickettsia* and *Coxiella*) form close co-phylogenetic relationship with their tick hosts. This analysis considered host and pathogen phylogenies at the genus level: that is, *Rhipicephalus*, *Hyalomma*, *Dermacentor*, *Amblyomma*, *Haemaphysalis*, *Ixodes*, *Argas* and *Carios*. Specifically, we estimated the Association Index [66] and compared it with a null distribution generated using 1000 replicates of state randomization (i.e. tick genera) across a credible set of pathogen trees generated by MrBayes version 3.2 [62] as described above.

To examine the extent of bacteria-tick co-divergence, we performed event-based co-phylogenetic reconstructions using the Jane program, version 4.0 [45]. The 'cost' scheme for analyses in Jane was set as follows: co-divergence = 0, duplication = 1, host switch = 1, loss = 1, failure to diverge = 1. The number of generations and the population size were both set to 100. The significance of co-divergence was derived by comparing the estimated costs to null distributions calculated from 100 randomizations of host tip mapping. In addition, we performed a distance-based analysis to test the hypothesis of bacterial-host co-divergence using ParaFit as implemented in the COPYCAT software package version 2.0 [46], comparing distance matrices derived from the bacteria and host phylogenies. Significance testing was based of 9999 randomizations of the association matrices. Additionally, to visualize the association between bacteria and their tick hosts, a tanglegram was generated by matching each bacterial species to their associated ticks using TreeMap version 3.0beta [67].

## The influence of geographic and host genetic distance on pathogen genetic diversity

Bacterial and host genetic distance matrices were derived from pairwise genetic comparisons using MEGA version 5.2 [58]. Geographic distances (Euclidean distance) were calculated using spatial coordinates of the samples derived from information on their geographic location. We used Mantel correlation analysis [68] to test the extent of the correlation between these matrices. Both a simple

Mantel's test and Partial Mantel's test were performed, and the correlation was evaluated using 10000 permutations. To access which of the two factors – geographic or host genetic distances – best explained total variation in the bacteria genetic distance matrices we performed a multiple linear regression analysis [69] on these distance matrices. The statistical significance of each regression was evaluated by performing 1000 permutations. All statistical analyses were performed using the Ecodist package implemented in R 3.4.4 [70], and all statistical results were considered significant at a *P*-value of 0.05.

## Declarations

## Acknowledgements

We thank Rongman Xu and Yi Sun for sample collection and morphologically identification of tick species.

## Authors' contributions

M.S and C.L.L designed this study. J.H.T contributed samples and experiments. X.H. performed data analyses. M.G, H.B.X, B.Y and J.L provided help during sample collection. X.H and M.S wrote the manuscript. R.F.S and E.C.H provided suggestion for manuscript. All authors reviewed and approved the final version of the manuscript.

## Funding

This work was funded by Shenzhen Science and Technology Program (KQTD20200820145822023), Guangdong Province "Pearl River Talent Plan" Innovation and Entrepreneurship Team Project (2019ZT08Y464), "Medical youth top talent project of Hubei" Fellowship and Science and Technology Project of Jiangxi (20142BBG70097), and Australian Research Council Australia Laureate Fellowship (FL170100022).

## Availability of data and materials

The tick mt genome sequences reported in this study are available on NCBI, under accession number XXX-XXX. Sequence alignments underlying analyses and phylogenetic trees are available from figshare (<https://doi.org/10.6084/m9.figshare.XXXXXX>). Raw sequencing data are available at the NCBI SRA database as BioProject PRJNAXXXXXX (<https://www.ncbi.nlm.nih.gov/sra/PRJNAXXXXXX>)

## Ethics approval and consent to participate

This study was conducted under the approved project by the XXXXX.

## Consent for publication

All authors have read and agreed to publish the manuscript.

## Competing interests

The authors declare no competing interests.

## References

1. Penalver E, Arillo A, Delclos X, Peris D, Grimaldi DA, Anderson SR, Nascimbene PC, Perez-de la Fuente R: **parasitised feathered dinosaurs as revealed by Cretaceous amber assemblages**. *Nat Commun* 2017, **8**(1):1924.
2. Nicholson WL, Sonenshine DE, Noden BH, Brown RN: **Ticks (Ixodida)**. In: *Medical and Veterinary Entomology*. 2019: 603-672.
3. Mans BJ, de Klerk D Fau - Pienaar R, Pienaar R Fau - Latif AA, Latif AA: **Nuttalliella namaqua: a living fossil and closest relative to the ancestral tick lineage: implications for the evolution of blood-feeding in ticks**. *PLoS One* 2011, **6**(1932-6203 (Electronic)).
4. Zhang YK, Zhang XY, Liu JZ: **Ticks (Acari: Ixodoidea) in China: Geographical distribution, host diversity, and specificity**. *Arch Insect Biochem Physiol* 2019, **102**(3):e21544.
5. Zhang G, Zheng D, Tian Y, Li S: **A dataset of distribution and diversity of ticks in China**. *Sci Data* 2019, **6**(1):105.
6. Jia N, Wang J, Shi W, Du L, Sun Y, Zhan W, Jiang JF, Wang Q, Zhang B, Ji P *et al*: **Large-Scale Comparative Analyses of Tick Genomes Elucidate Their Genetic Diversity and Vector Capacities**. *Cell* 2020, **182**(5):1328-1340 e1313.
7. Shao R, Barker SC: **Mitochondrial genomes of parasitic arthropods: implications for studies of population genetics and evolution**. *Parasitology* 2007, **134**(Pt 2):153-167.
8. Wang T, Zhang S, Pei T, Yu Z, Liu J: **Tick mitochondrial genomes: structural characteristics and phylogenetic implications**. *Parasites & vectors* 2019, **12**(1):451.
9. Kelava S, Mans BJ, Shao R, Moustafa MAM, Matsuno K, Takano A, Kawabata H, Sato K, Fujita H, Ze C *et al*: **Phylogenies from mitochondrial genomes of 120 species of ticks: Insights into the evolution of the families of ticks and of the genus Amblyomma**. *Ticks Tick Borne Dis* 2021, **12**(1):101577.
10. Shao R, Barker SC, Mitani H, Takahashi M, Fukunaga M: **Molecular mechanisms for the variation of mitochondrial gene content and gene arrangement among chigger mites of the genus Leptotrombidium (Acari: Acariformes)**. *J Mol Evol* 2006, **63**(2):251-261.

11. Cameron SL, Johnson KP, Whiting MF: **The mitochondrial genome of the screamer louse *Bothriometopus* (phthiraptera: ischnocera): effects of extensive gene rearrangements on the evolution of the genome.** *J Mol Evol* 2007, **65**(5):589-604.
12. Yu XJ, Liang Mf Fau - Zhang S-Y, Zhang Sy Fau - Liu Y, Liu Y Fau - Li J-D, Li Jd Fau - Sun Y-L, Sun Yl Fau - Zhang L, Zhang L Fau - Zhang Q-F, Zhang Qf Fau - Popov VL, Popov VI Fau - Li C, Li C Fau - Qu J *et al*: **Fever with thrombocytopenia associated with a novel bunyavirus in China.** *N Engl J Med* 2011, **364**(1533-4406 (Electronic)).
13. Jia N, Zheng Y-C, Jiang J-F, Ma L, Cao W-CJNEJoM: **Human infection with *Candidatus Rickettsia tarasevichiae*.** *N Engl J Med* 2013, **369**(12):1178-1180.
14. Jia N, Liu HB, Ni XB, Bell-Sakyi L, Zheng YC, Song JL, Li J, Jiang BG, Wang Q, Sun Y *et al*: **Emergence of human infection with Jingmen tick virus in China: A retrospective study.** *EBioMedicine* 2019, **43**:317-324.
15. Li H, Li XM, Du J, Zhang XA, Cui N, Yang ZD, Xue XJ, Zhang PH, Cao WC, Liu W: ***Candidatus Rickettsia xinyangensis* as Cause of Spotted Fever Group Rickettsiosis, Xinyang, China, 2015.** *Emerg Infect Dis* 2020, **26**(5):985-988.
16. Yu Z, Wang H, Wang T, Sun W, Yang X, Liu J: **Tick-borne pathogens and the vector potential of ticks in China.** *Parasit Vectors* 2015, **8**:24.
17. Rosenberg R, Lindsey NP, Fischer M, Gregory CJ, Hinckley AF, Mead PS, Paz-Bailey G, Waterman SH, Drexler NA, Kersh GJ *et al*: **Vital Signs: Trends in Reported Vectorborne Disease Cases - United States and Territories, 2004-2016.** *MMWR Morb Mortal Wkly Rep* 2018(1545-861X (Electronic)).
18. Walker DH, Ismail N: **Emerging and re-emerging rickettsioses: endothelial cell infection and early disease events.** *Nat Rev Microbiol* 2008, **6**(5):375-386.
19. Parola P, Paddock CD, Socolovschi C, Labruna MB, Mediannikov O, Kernif T, Abdad MY, Stenos J, Bitam I, Fournier PE *et al*: **Update on tick-borne rickettsioses around the world: a geographic approach.** *Clin Microbiol Rev* 2013, **26**(4):657-702.
20. Njeru J, Henning K, Pletz MW, Heller R, Neubauer H: **Q fever is an old and neglected zoonotic disease in Kenya: a systematic review.** *BMC Public Health* 2016, **16**:297.
21. Trape JF, Diatta G, Arnathau C, Bitam I, Sarih M, Belghyti D, Bouattour A, Elguero E, Vial L, Mane Y *et al*: **The epidemiology and geographic distribution of relapsing fever borreliosis in West and North Africa, with a review of the *Ornithodoros erraticus* complex (Acari: Ixodida).** *PLoS One* 2013, **8**(11):e78473.
22. Kernif T, Leulmi H, Raoult D, Parola P: **Emerging Tick-Borne Bacterial Pathogens.** *Microbiol Spectr* 2016, **4**(3).
23. Moran NA, McCutcheon Jp Fau - Nakabachi A, Nakabachi A: **Genomics and evolution of heritable bacterial symbionts.** 2008(0066-4197 (Print)).
24. Matsuura Y, Kikuchi Y Fau - Meng XY, Meng Xy Fau - Koga R, Koga R Fau - Fukatsu T, Fukatsu T: **Novel clade of alphaproteobacterial endosymbionts associated with stinkbugs and other arthropods.** 2012(1098-5336 (Electronic)).

25. Toft C, Andersson SG: **Evolutionary microbial genomics: insights into bacterial host adaptation.** *Nat Rev Genet* 2010, **11**(7):465-475.
26. Duron O, Binetruy F, Noel V, Cremaschi J, McCoy KD, Arnathau C, Plantard O, Goolsby J, Perez de Leon AA, Heylen DJA *et al*: **Evolutionary changes in symbiont community structure in ticks.** *Mol Ecol* 2017, **26**(11):2905-2921.
27. Hoogstraal H: **African Ixodoidea. Ticks of the Sudan (with special reference to Equatoria Province and with preliminary reviews of the genera Boophilus, Margaropus, and Hyalomma).** In: *Apha Meeting & Exposition: 1956*.
28. Walker AJTAH, Production: **The Genus Rhipicephalus (Acari, Ixodidae): A Guide to the Brown Ticks of the World; Jane B. Walker, James E. Keirans and Ivan G. Horak.** 2000, **32**(6):417-418.
29. Walker AR, Bouattour A, Camicas JL, Estrada-Pena A, Preston PM: **Ticks of domestic animals in Africa: a guide to identification of species;** 2014.
30. Ye X, Sun Y, Ju W, Wang X, Cao W, Wu MJP, Vectors: **Vector competence of the tick Ixodes sinensis (Acari: Ixodidae) for Rickettsia monacensis.** 2014, **7**(1):512.
31. Shao R, Barker SC, Mitani H, Aoki Y, Fukunaga M: **Evolution of duplicate control regions in the mitochondrial genomes of metazoa: a case study with Australasian Ixodes ticks.** *Mol Biol Evol* 2005, **22**(3):620-629.
32. Li LH, Zhang Y, Wang JZ, Li XS, Yin SQ, Zhu D, Xue JB, Li SG: **High genetic diversity in hard ticks from a China-Myanmar border county.** *Parasit Vectors* 2018, **11**(1):469.
33. Oteo JA, Portillo A: **Tick-borne rickettsioses in Europe.** 2012(1877-9603 (Electronic)).
34. Seo Mg Fau - Kwon O-D, Kwon Od Fau - Kwak D, Kwak D: **High Prevalence of Rickettsia raoultii and Associated Pathogens in Canine Ticks, South Korea.** 2020(1080-6059 (Electronic)).
35. Tian ZC, Liu Gy Fau - Shen H, Shen H Fau - Xie J-R, Xie Jr Fau - Luo J, Luo J Fau - Tian M-Y, Tian MY: **First report on the occurrence of Rickettsia slovaca and Rickettsia raoultii in Dermacentor silvarum in China.** 2012(1756-3305 (Electronic)).
36. Dong Z, Yang Y, Wang Q, Xie S, Zhao S, Tan W, Yuan W, Wang Y: **A case with neurological abnormalities caused by Rickettsia raoultii in northwestern China.** *BMC Infect Dis* 2019, **19**(1):796.
37. Zhao S, Yang M, Jiang M, Yan B, Zhao S, Yuan W, Wang B, Hornok S, Wang Y: **Rickettsia raoultii and Rickettsia sibirica in ticks from the long-tailed ground squirrel near the China-Kazakhstan border.** *Exp Appl Acarol* 2019, **77**(3):425-433.
38. Mediannikov OY, Sidelnikov Y Fau - Ivanov L, Ivanov L Fau - Mokretsova E, Mokretsova E Fau - Fournier P-E, Fournier Pe Fau - Tarasevich I, Tarasevich I Fau - Raoult D, Raoult D: **Acute tick-borne rickettsiosis caused by Rickettsia heilongjiangensis in Russian Far East.** *Emerg Infect Dis* 2004, **10**(1080-6040 (Print)):810-817.
39. Adegoke A, Kumar D, Bobo C, Rashid MI, Durrani AZ, Sajid MS, Karim S: **Tick-Borne Pathogens Shape the Native Microbiome Within Tick Vectors.** *Microorganisms* 2020, **8**(9).

40. Chao LL, Castillo CT, Shih CM: **Molecular detection and genetic identification of Wolbachia endosymbiont in Rhipicephalus sanguineus (Acari: Ixodidae) ticks of Taiwan.** *Exp Appl Acarol* 2021, **83**(1):115-130.
41. Hemsley CM, O'Neill PA, Essex-Lopresti A, Norville IH, Atkins TP, Titball RW: **Extensive genome analysis of Coxiella burnetii reveals limited evolution within genomic groups.** *BMC Genomics* 2019, **20**(1):441.
42. Beare PA, Unsworth N, Andoh M, Voth DE, Omsland A, Gilk SD, Williams KP, Sobral BW, Kupko JJ, Porcella SF *et al.*: **Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus Coxiella.** *Infection and immunity* 2009, **77**(2):642-656.
43. Elbir H, Larsson P, Upreti M, Normark J, Bergström S: **Genome Sequence of the Relapsing Fever Borreliosis Species Borrelia hispanica.** *Genome announcements* 2014, **2**(1).
44. Margos G, Gofton A, Wibberg D, Dangel A, Marosevic D, Loh S-M, Oskam C, Fingerle V: **The genus Borrelia reloaded.** *PloS one* 2018, **13**(12):e0208432.
45. Conow C, Fielder D, Ovadia Y, Libeskind-Hadas R: **Jane: a new tool for the cophylogeny reconstruction problem.** *Algorithms for molecular biology : AMB* 2010, **5**:16.
46. Meier-Kolthoff JP, Auch AF, Huson DH, Göker M: **COPYCAT: cophylogenetic analysis tool.** *Bioinformatics (Oxford, England)* 2007, **23**(7):898-900.
47. Hoogstraal H, Kohls GM: **Description, Hosts, and Ecology of Ixodes Kuntzi N. Sp., Kuntz's Taiwan Flying Squirrel Tick (Ixodoidea: Ixodidae)1.** *Journal of Medical Entomology* 1965, **2**(3):209-214.
48. Vongphayloth K, Brey PT, Robbins RG, Sutherland IW: **First survey of the hard tick (Acari: Ixodidae) fauna of Nakai District, Khammouane Province, Laos, and an updated checklist of the ticks of Laos.** *Systematic and Applied Acarology* 2016, **21**(2).
49. Wang Q, Pan YS, Jiang BG, Ye RZ, Chang QC, Shao HZ, Cui XM, Xu DL, Li LF, Wei W *et al.*: **Prevalence of Multiple Tick-Borne Pathogens in Various Tick Vectors in Northeastern China.** *Vector Borne Zoonotic Dis* 2021, **21**(3):162-171.
50. Kim TY, Kwak YS, Kim JY, Nam SH, Lee IY, Mduma S, Keyyu J, Fyumagwa R, Yong TS: **Prevalence of Tick-Borne Pathogens from Ticks Collected from Cattle and Wild Animals in Tanzania in 2012.** *Korean J Parasitol* 2018, **56**(3):305-308.
51. Kang Y-J, Diao X-N, Zhao G-Y, Chen M-H, Xiong Y, Shi M, Fu W-M, Guo Y-J, Pan B, Chen X-P: **Extensive diversity of Rickettsiales bacteria in two species of ticks from China and the evolution of the Rickettsiales.** *BMC evolutionary biology* 2014, **14**(1):1-12.
52. Liu W, Li L, Khan MA, Zhu F: **Popular molecular markers in bacteria.** *Molecular Genetics, Microbiology and Virology* 2012, **27**(3):103-107.
53. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W: **MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph.** *Bioinformatics (Oxford, England)* 2015, **31**(10):1674-1676.

54. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *Journal of molecular biology* 1990, **215**(3):403-410.
55. Clewley JP: **Macintosh sequence analysis software. DNASTar's LaserGene.** *Molecular biotechnology* 1995, **3**(3):221-224.
56. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nature methods* 2012, **9**(4):357-359.
57. Bonfield JK, Smith Kf, Staden R: **A new DNA sequence assembly program.** *Nucleic acids research* 1995, **23**(24):4992-4999.
58. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: **MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Molecular biology and evolution* 2011, **28**(10):2731-2739.
59. Katoh K, Standley DM: **MAFFT multiple sequence alignment software version 7: improvements in performance and usability.** *Molecular biology and evolution* 2013, **30**(4):772-780.
60. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T: **trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses.** *Bioinformatics (Oxford, England)* 2009, **25**(15):1972-1973.
61. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ: **IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies.** *Molecular biology and evolution* 2015, **32**(1):268-274.
62. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics (Oxford, England)* 2003, **19**(12):1572-1574.
63. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS: **ModelFinder: fast model selection for accurate phylogenetic estimates.** *Nature methods* 2017, **14**(6):587-589.
64. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS: **UFBoot2: Improving the Ultrafast Bootstrap Approximation.** *Molecular biology and evolution* 2018, **35**(2):518-522.
65. Parker J, Rambaut A, Pybus OG: **Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty.** *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2008, **8**(3):239-246.
66. Wang TH, Donaldson YK, Brettell RP, Bell JE, Simmonds P: **Identification of shared populations of human immunodeficiency virus type 1 infecting microglia and tissue macrophages outside the central nervous system.** *Journal of virology* 2001, **75**(23):11686-11699.
67. Jackson AP, Charleston MA: **A cophylogenetic perspective of RNA-virus evolution.** *Molecular biology and evolution* 2004, **21**(1):45-57.
68. Mantel N: **The detection of disease clustering and a generalized regression approach.** *Cancer research* 1967, **27**(2):209-220.
69. Lichstein JW: **Multiple regression on distance matrices: a multivariate spatial analysis tool.** *Plant Ecology* 2006, **188**(2):117-131.

## Figures

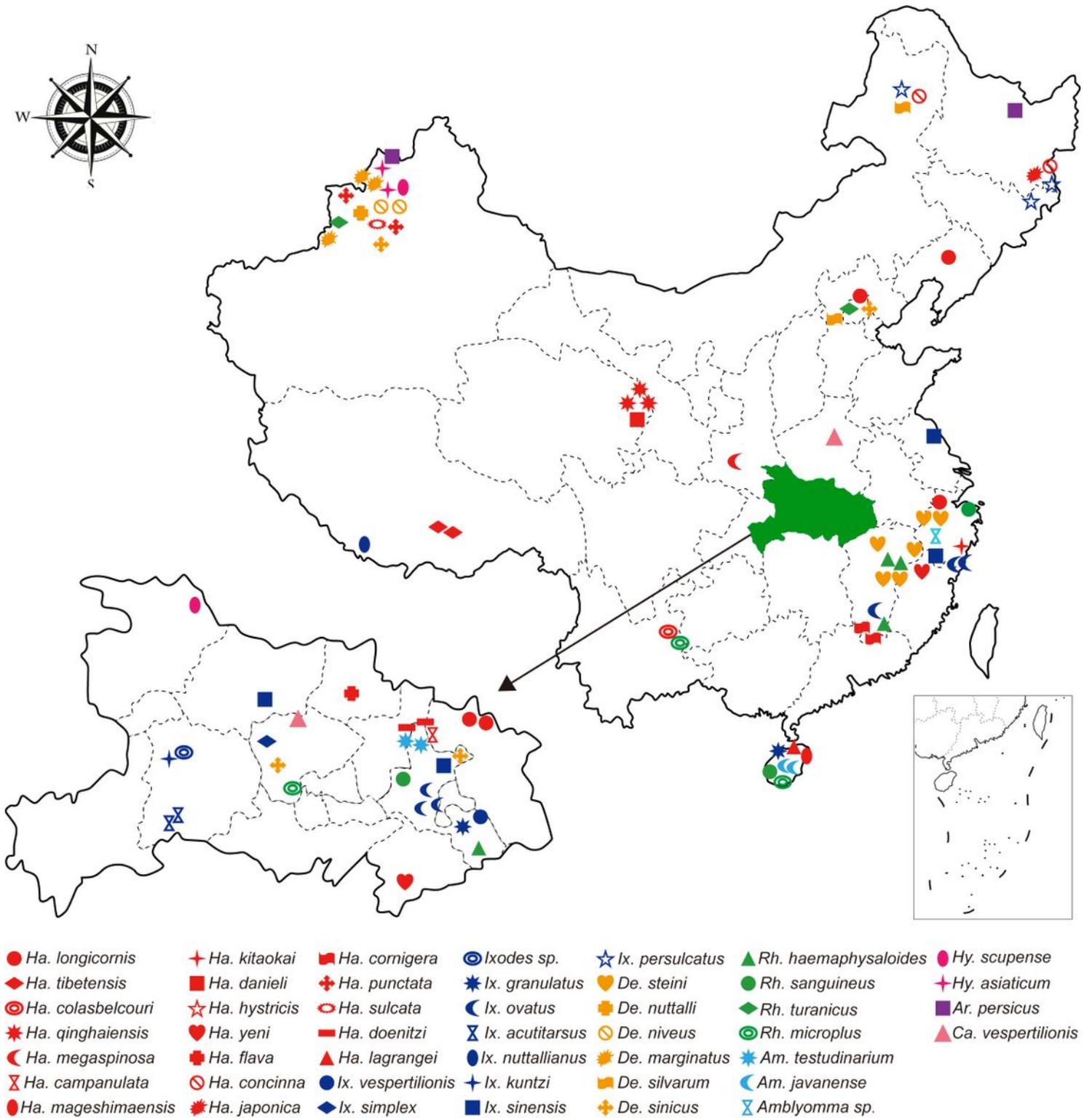
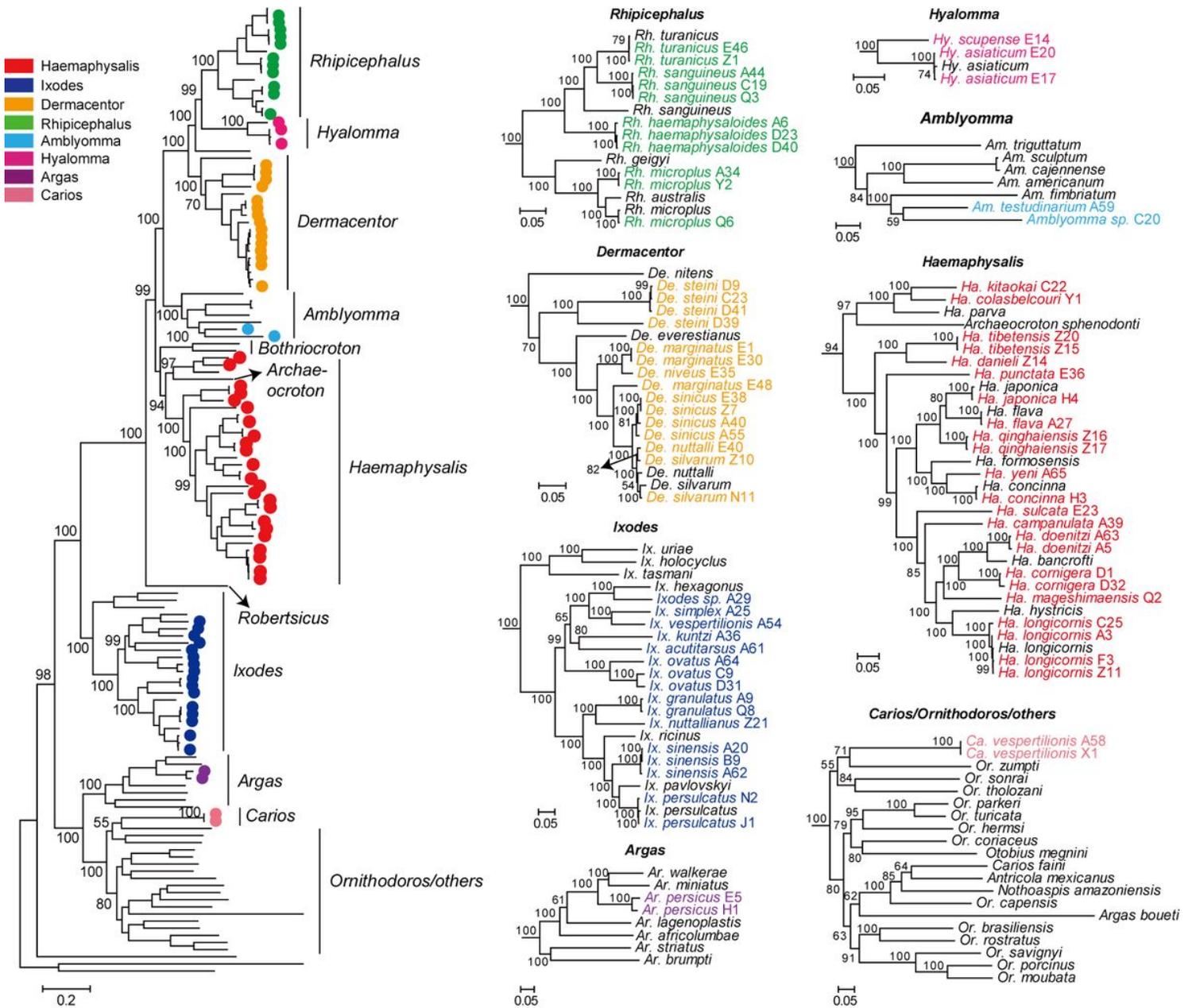


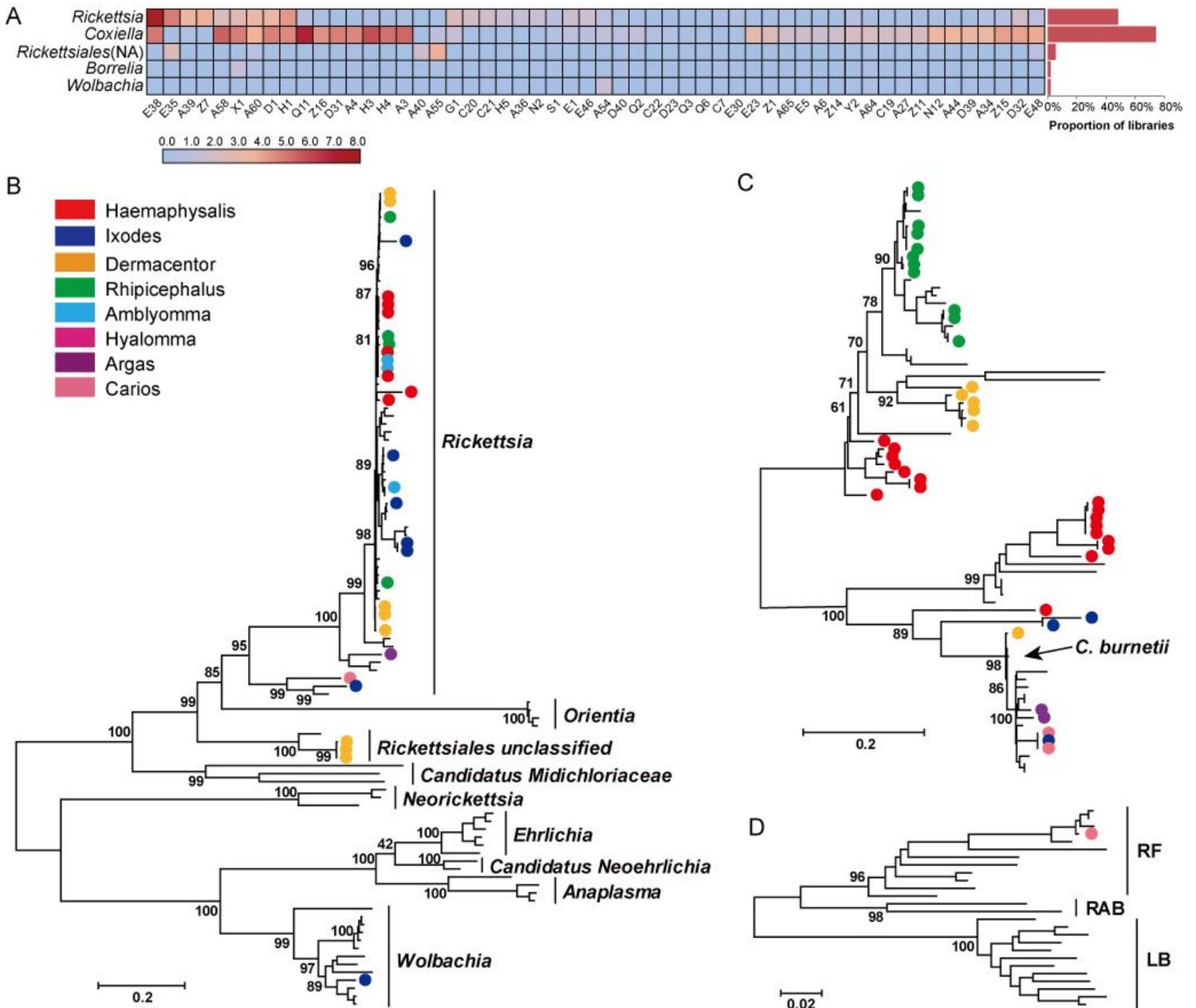
Figure 1

Sampling locations of 46 tick species collected in China. The map of Hubei province (shaded green) was magnified for clarity. We use different colors and shapes to represent 46 tick species from eight genera: *Haemaphysalis* (red), *Ixodes* (dark blue), *Dermacentor* (yellow), *Rhipicephalus* (green), *Amblyomma* (sky blue), *Hyalomma* (magenta), and *Argas* (purple), *Carios* (pink).



**Figure 2**

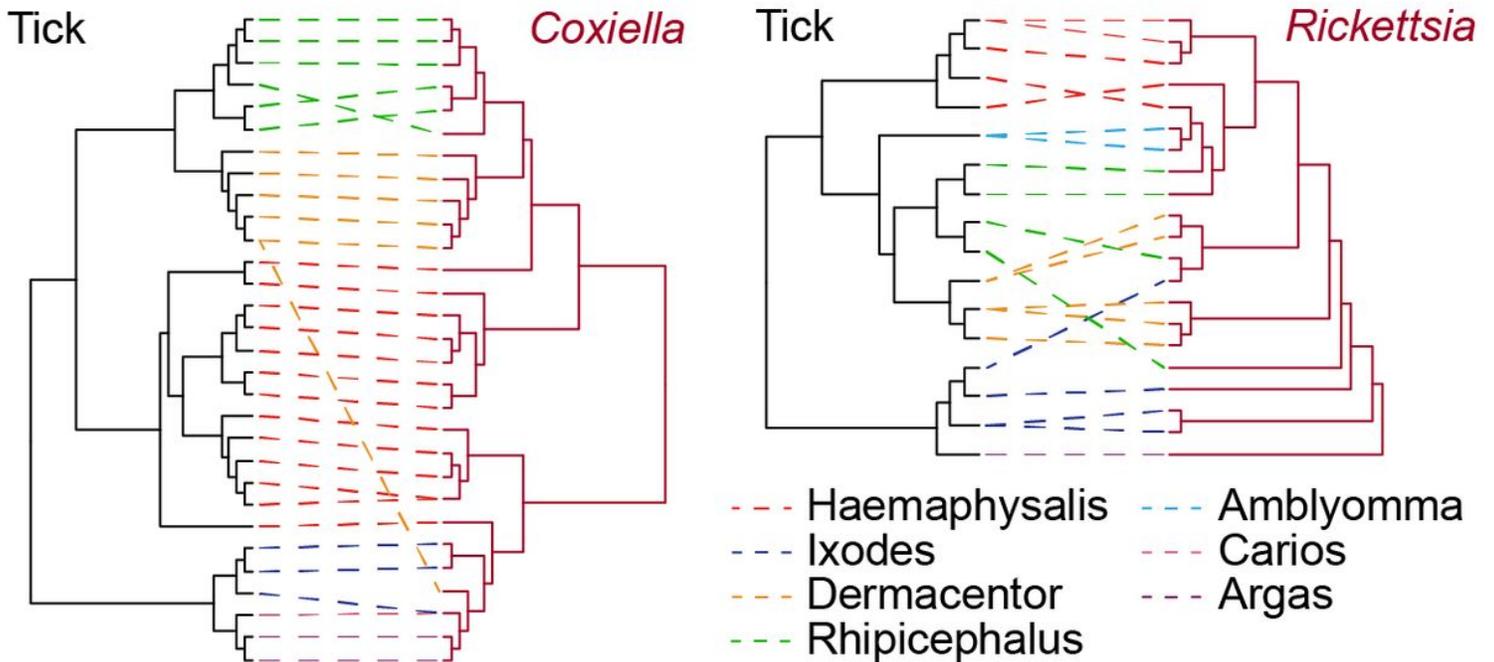
ML phylogenies of ticks based on all 13 protein coding genes and two rRNA genes. Two mite species act as the outgroup and the scale bar represents the number of nucleotide substitutions per site. For clarity, bootstrap values only shown for major nodes. The core phylogenetic tree is shown on the left, and sequences generated in this study are marked by a circle and colored according to different tick genera. The detailed subtrees of each group are shown on the right. Within each subtree, the sequences newly identified here are marked accordingly along with number of sequencing libraries.



**Figure 3**

The abundance of tick associated bacterial groups based on the *groEL* gene and the proportion of positive libraries of each group, Rickettsiales (NA) represented bacteria identified could not classified in a specific genus (A). Phylogenetic trees for bacteria from the order Rickettsiales based on *groEL* gene (B), genera *Coxiella* based on *groEL* gene (C) and *Borrelia* based on *flaB* gene(D). The trees were midpoint-rooted and the scale bar represents the number of nucleotide substitutions per site. Sequences generated in this study marked by a circle and colored according to different tick genera. Bootstrap values only shown for major nodes. Within the Rickettsiales phylogeny, different genera are denoted by vertical lines. Within the *Coxiella* phylogeny, the position of *C. burnetii* is highlighted by a black arrow. Within the *Borrelia* phylogeny, “RF” denotes the Relapsing fever group, “RAB” denotes the Reptile associated and Australian *Borrelia*, while “LB” indicates the Lyme borreliosis group [44].

	<i>P</i> value (Parafit)	<i>P</i> value (Jane4)
<i>Rickettsia</i>	0.0021	< 0.01
<i>Coxiella</i>	0.0003	< 0.01



**Figure 4**

Co-phylogenetic comparisons of *Rickettsia* and *Coxiella* bacteria phylogenies and their corresponding tick hosts. The table shows the results of the co-phylogeny analysis using Parafit and Jane4. The tanglegram shows the match between the phylogenies of the bacteria and tick hosts. The relationship between the two phylogenies is displayed to maximize topological congruence. Dotted line colours correspond to different tick groups as shown by figure legend at the right bottom.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Graphicalabstract.tif](#)
- [TicksAndPathogensSM.pdf](#)