RESEARCH

The diversity and evolutionary relationships of ticks and tick-borne bacteria collected in China

JunHua Tian[1,3†], Xin Hou[2†], MiHong Ge[4], HongBin Xu[5], Bin Yu[3], Jing Liu[3], RenFu Shao[6,7], Edward C. Holmes[8], ChaoLiang Lei[1*] and Mang Shi[2*]

Abstract

Background: Ticks (order Ixodida) are ectoparasites, vectors and reservoirs of many infectious agents affecting humans and domestic animals. However, the lack of information on tick genomic diversity leaves significant gaps in the understanding of the evolution of ticks and associated bacteria.

Results: We collected > 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. Metagenomic sequencing was performed on individual ticks to obtain the complete or near-complete mitochondrial (mt) genome sequences from 46 tick species, among which mitochondrial genomes of 23 species were recovered 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 genetic diversity of ticks based on an analysis of mitochondrial DNA as well as on the prevalence of tick-borne pathogens in China. It also sheds new light on the long-term evolutionary and ecological relationships between ticks and their associated bacteria.

Keywords: Ticks, Metagenomics, Bacteria, Evolution, Phylogeny, Co-divergence

Background

Ticks (Acari: Ixodidae) are hematophagous arthropods and act as vectors for various infectious pathogens. 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, > 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*, *Haemaphysalis conicinna*, *Rhipicephalus microplus* and *Rhipicephalus sanguineus*

†Junhua Tian and Xin Hou have contributed equally to this work

*Correspondence: ioir@mail.hzau.edu.cn; shim23@mail.sysu.edu.cn

[1] Hubei Key Laboratory of Resources Utilization and Sustainable Pest Management, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, Hubei Province 430070, China
[2] School of Medicine, Shenzhen Campus of Sun Yat-Sen University, Sun Yat-Sen University, Shenzhen, Guangdong Province 518107, China
Full list of author information is available at the end of the article

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 2 of 14

sensu lato [5, 6]. *Dermacentor sinicus*, *Ixodes sinensis*, *Haemaphysalis tibetensis* and *Haemaphysalis qinghaiensis* have been only reported in China, although there is a lack of genomic data for these species [7].

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 [8]. For ticks, complete mt genomes representing 66 species from 18 genera have been sequenced to date [9, 10]. 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 show great consistency in genus level classification [9]. A single genome re-arrangement event has been identified in the Ixodidae, including the genera *Rhipicephalus*, *Dermacentor*, *Amblyomma* and *Haemaphysalis*, which possess a translocation of tRNA genes ($trnL_1$, $trnL_2$, $trnC$) and an inversion of the *trnC* gene [11, 12].

Ticks have long been regarded as disease vectors, with an increasing number of human disease associations described in recent years [13–16]. Since 1982, more than 30 emerging tick-borne disease agents have been identified from at least 28 tick species, causing a variety of human infections including rickettsioses, Q fever and borrelioses. Together, these have accounted for nearly 50,000 cases per year since 2013 as reported by National Notifiable Disease Surveillance System in the USA [17, 18]. Among these, rickettsial disease and Q fever are caused by two groups of obligate intracellular endosymbiont bacteria in ticks: the spotted fever group (SFG) rickettsiae [19, 20] and *Coxiella burnetii* [21], 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 is generally carried by *Ornithodoros sonrai*, *Ornithodoros erraticus* and *Ornithodoros moubata* [22, 23]. In addition to known pathogens, ticks harbor a number of other bacterial symbionts in the genera *Rickettsia* and *Coxiella*. While their public health significance remains unclear, these microbes are highly prevalent in tick species and can be transmitted transovarially [24–27].

Despite growing interest in tick-borne pathogens, our knowledge of the genetic diversity of ticks in China and the bacteria they carry is 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 to our knowledge, we revealed the genetic diversity of both ticks and their bacterial symbionts, enabling a more systematic study of their co-evolutionary history.

Methods

Sample collection

Between 1959 and 2019 > 20,000 ticks belonging to eight genera (*Rhipicephalus, Hyalomma, Dermacentor, Amblyomma, Haemaphysalis, Ixodes, Argas* and *Carios*) were collected from different geographic regions in China (Fig. 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% (*I. simplex*, *A. testudinarium* and *Hae. qinghaiensis*) were preserved in alcohol at room temperature for 1–60 years, while others were captured alive and stored at the temperature of − 80 ℃ 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, including questing and partially engorged ticks, were first washed three times in PBS solutions before they were homogenized 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 150 bp 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 100 g) for insoluble debris removal. The supernatant was further centrifuged at 10,000 g 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 [28] with default parameters. To obtain mt genome sequences, the assembled contigs

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 3 of 14

Fig. 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), *Argas* (purple) and *Carios* (pink)

were compared against reference mt genomes in Gen-Bank using blastn [29]. 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 [30]. We used Bowtie2 [31] 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 [32]. The annotated mt genomes of these ticks were deposited in GenBank under accession numbers OM368258—OM368330 and MK344649.

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 4 of 14

To infer phylogenetic trees, we used 74 mt genomes generated in this study plus an additional 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 [33]. Ambiguously aligned regions were removed. Two rRNA genes were aligned using the MAFFT version 7.4 [34] employing the L-INS-i algorithm with all ambiguous aligned regions removed using TrimAL version 1.2rev59 [35]. 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 [36] and MrBayes version 3.2 [37], respectively. For ML analyses, the best-fit nucleotide substitution model was selected using ModelFinder [38], and tree topologies were evaluated with an ultrafast bootstrap approximation approach [39] 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%.

Identification and characterization of bacterial pathogens in ticks

The initial taxonomy profiling was performed by metaphlan2 [40]. Furthermore, 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 [31] and generating consensus sequences from the mapped reads. The clean reads were subsequently mapped to *groEL* gene sequences to estimated gene abundance as the number of mapped reads per million total reads (RPM, RPM of a gene = Number of reads mapped to a gene $\times 10^6$/Total number of mapped reads from given library). 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 [41] to test whether bacteria pathogens (*Rickettsia* and *Coxiella*) form close co-phylogenetic relationship with their tick hosts. This analysis considered tick 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 [42] 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 [37] 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 [43]. 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-tick co-divergence using ParaFit as implemented in the COPYCAT software package version 2.0 [44], comparing distance matrices derived from the bacteria and tick host phylogenies. Significance testing was based on 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 [45].

The influence of geographic and tick genetic distance on pathogen genetic diversity

Bacterial and tick genetic distance matrices were derived from pairwise genetic comparisons using MEGA version 5.2 [33]. Geographic distances (Euclidean distance) were calculated using spatial coordinates of the samples derived from information on their geographic location. We used Mantel correlation analysis [46] 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 10,000 permutations. To access which of the two factors—geographic or tick genetic distances—best explained total variation in the bacteria genetic distance matrices, we performed a multiple linear regression analysis [47] on these distance matrices. The statistical significance of each regression was evaluated by performing 1000 permutations. All statistical analyses were performed using

Tian *et al. Parasites & Vectors* (2022) 15:352

Page 5 of 14

the Ecodist package implemented in R 3.4.4 [48], and all statistical results were considered significant at a *P*-value of 0.05.

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 (Fig. 1, Table 1, Additional file 1: Table S1). Species identification of hard ticks was carried out based on morphological characters, such as palps, basis capituli, cornuae, auriculae, dentition formula, punctuations, coxal spurs, amongst others [49–51]. Identification of soft ticks was conducted using taxonomic keys proposed by Hoogstraal [54], Teng [50] and Sun et al. [52–54]. This revealed a total of at least 46 species comprising two families (Ixodidae and Argasidae) and eight genera: *Haemaphysalis* (19), *Ixodes* (10), *Dermacentor* (six), *Rhipicephalus* (four), *Hyalomma* (two), *Amblyomma* (three), *Argas* (one) and *Carios* (one). Among these genera are the six most common tick species in China [6]: *R. microplus*, *R. sanguineus* s.l., *I. persulcatus*, *Hae. longicornis*, *D. silvarum* and *Hyalomma asiaticum*. In addition, other common species were collected and analyzed, including *I. sinensis* (vector of *B. burgdorferi* [55]), *Ixodes ovatus*, *D. steini*, *Haemaphysalis yeni*, *Hae. concinna*, *Hyalomma scupense*, *Rhipicephalus turanicus*, *Rhipicephalus haemaphysaloides* and *Argas persicus* from at least two provinces. In comparison, other species were more locally defined, such as *Hae. tibetensis* from Tibet, *Hae. qinghaiensis* and *Haemaphysalis danieli*

from Qinghai, *Hy. asiaticum*, *Haemaphysalis punctata*, *Dermacentor marginatus* and *Dermacentor niveus* from Xinjiang, *Ixodes acutitarsus* from Hubei and *Haemaphysalis lagrangei* and *Haemaphysalis mageshimaensis* from Hainan. In addition, we obtained a number of very rare species, particularly *Amblyomma javanense*, *Ixodes simplex*, *Ixodes nuttallianus*, *Ixodes crenulatus*, *Ixodes kuntzi*, *Haemaphysali kitaokai* and *Carios 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 > 60 years. For example, the oldest sample from our data set (*A. javanense*) was collected in the 1960s from a wild Chinese pangolin (*Manis pentadactyla*). In the case of two samples (i.e. C20, A29), identification could only take place 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.

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 23 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% (*Hae.*

Table 1 Sampling locations, animal hosts and bacterial pathogens of tick species in China at the genus level

Species(genus)	Location	Host	No. of libraries	Pathogen
Rhipicephalus	Hubei, Jiangxi, Yunnan, Beijing, Zhejiang, Xinjiang, Hainan	Cattle, dog, hedgehog	12 (11)	*Rickettsia, Coxiella,*
Hyalomma	Xinjiang, Hubei	Cattle, goat, camel	4 (3)	–
Dermacentor	Hubei, Jiangxi, Inner Mongolia, Beijing, Hebei, Zhejiang, Xinjiang	Cattle, goat, rabbit, hedgehog, wild boar	18 (15)	*Rickettsia, Coxiella,* Rickesiales unclassified
Amblyomma	Hubei, Zhejiang, Hainan	Wild boar, Malaysia pangolin, Chinese pangolin	6 (2)	*Rickettsia*
Haemaphysalis	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)	*Rickettisa, Coxiella,*
Ixodes	Hubei, Jiangxi, Jiangsu, Zhejiang, Hainan, Inner Mongolia, Jilin, Tibet, Heilongjiang	Cattle, goat, Eurasian badger, hog badger, rabbit, bat, pteromyini, tupaia	22 (16)	*Rickettisa, Coxiella, Wolbachia*
Carios	Hubei, Henan	–	2 (2)	*Coxiella, Rickettisa, Borrelia*
Argas	Heilongjiang, Xinjiang	Chicken	2 (2)	*Rickettsia, Coxiella*
Total	–	–	96 (74)	–

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

Tian *et al. Parasites & Vectors* (2022) 15:352

Page 6 of 14

danieli Z14), similar to previously identified mt genomes of ticks [9]. Furthermore, the structure, composition and arrangement of genes largely followed their closest relatives within the same genus [9, 56]. 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 file 1: Fig. S1–S2, Table S2). For example, the mt genome of *D. marginatus* E48 has an extra copy of the non-coding region so that its length (i.e. 15,307 bp) has surpassed that of *Ixodes tasmani* (NC 041,086.1, 15,227 bp) [9] to become the longest tick mt genome identified to date (Additional file 1: Fig. 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, *D. marginatus* E1) sample (Additional file 1: Fig. 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 mt genomes of tick, 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 (Fig. 2, Additional file 1: Fig. S3–S6). Importantly, the newly added genomes greatly expanded the diversity of many groups, particularly the genera *Haemaphysalis, Ixodes* and *Dermacentor* (Fig. 2). In addition to the new species identified in this study, 23 tick species previously only known through morphological characteristics or incomplete mt genomes (e.g. *I. kuntzi*, *I. acutitarsus*, *Hae. mageshimaensis* and *Haemaphysalis colasbelcouri*) were also included (Fig. 2, Additional file 1: Fig. S3–S6). Furthermore, at least five potential cryptic species were identified—*R. sanguineus* s.l., *D. steini*, *D. marginatus* and *I. ovatus*—adding to the previously reported cryptic species identified in *R. microplus* [6, 57]. Each contained at least two divergent (70.93–94.21% identity) phylogenetic clusters while sharing the same morphological characteristics based on palps, basis capituli, shape and ornamentation on scutum, spurs on coxae I–IV, syncoxae and ala, etc., although more morphological features need to be examined to confirm this observation (Fig. 2). Conversely, *D. sinicus*, *Dermacentor nuttalli* and *D. silvarum* shared a very close relationship (>98.46% identity) even though these were separate species based on morphological characteristics. Interestingly, *D. nuttalli* and *D. silvarum* cannot be distinguished based on mt genome phylogeny,

although they had quite distinctive trochanter I dorsal spur (Additional file 1: Fig. S7).

Discovery and characterization of bacterial endosymbionts and pathogens in ticks

We first used metaphlan2 [40] for bacterial taxonomic profiling, which revealed the presence of > 32 genera including *Acinetobacter*, *Pseudomonas*, *Helicobacter* and *Escherichia* (Additional file 1: Fig. S8). Among these, we identified of tick endosymbiotic bacteria or bacteria that are known to harbor human pathogens: namely, the order Rickettsiales, genus *Coxiella* and genus *Borrelia*. These discoveries were further confirmed and characterized with analyzing the marker genes (Fig. 3A). 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%) (Fig. 3A, Additional file 1: 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*; Fig. 3B), including a number of human pathogens. For example, *Rickettsia raoultii,* which causes human tick-borne lymphadenitis [58, 59], was identified from *D. marginatus* and *D. niveus* in Jinghe, Xinjiang province, a region where *R. raoultii* have previously been reported [60–62]. Within Xinjiang (Jinghe and Yining), we identified *Rickettsia sibirica* and *Rickettsia africae* circulating in *D. sinicus* and *Ixodes 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 [20]. In addition, we discovered *Rickettsia heilongjiangensis*, the newly reported agent of Far-Eastern spotted fever (FESF) [63]. This bacterium was previously found in *D. silvarum* ticks from Heilongjiang, and herein it was associated with the tick species *Haemaphysalis campanulata* and *Haemaphysalis cornigera* in Hubei and Jiangxi provinces located in central China. Other pathogenic *Rickettsia* species included *Rickettsia tamurae, Rickettsia monacensis* and *Rickettsia helvetica* identified from *Amblyomma testudinarium*, *I. sinensis* and *I. kuntzi*.

We found a potentially novel species within Spotted Fever Group that was relatively divergent (< 99.46% genetic identity in six genes) from the other bacteria in this group (Additional file 1: Table S4). Since the species were identified from *Haemaphysalis megaspinosa*, we tentatively named it *Rickettsia* endosymbiont of *Haemaphysalis megaspinosa.* Furthermore, we identified four genetically divergent *Rickettsia* species that occupied basal phylogenetic positions. Among these, *Rickettsia* endosymbiont *of Ixodes persulcatus* H5 and N2 clustered

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 7 of 14

Fig. 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 are colored accordingly along with number of sequencing libraries

with *Rickettsia canadensis* (97.78% identity), *Rickettsia* endosymbiont of *Argas persicus* H1 fell with *Rickettsia 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 1: Table S4, Fig. S9–S11).

In addition to *Rickettsia*, we identified two potential new species within the order Rickettsiales. One,

Wolbachia endosymbiont of *Ixodes vespertilionis* A54, clustered with *Wolbachia pipientis* strain FL2016 (95.37%) and *Wolbahcia* endosymbiont of *Drosophila melanogaster* (95.16%) within the genus *Wolbachia*. The identification of *Wolbachia* in ticks has been reported in recent years [64, 65]. The other—*Rickettsiales* endosymbiont of *Dermacentor*—clustered with an unclassified Rickettsiales bacterium Ac37b identified from *Amblyomma cajennense* in Brazil (86.62% identity). Together they may represent a new genus or even family within the Rickettsiales (Additional file 1: Table S4, Fig. S9–S10).

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 8 of 14

Fig. 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. bernetii* is highlighted by a black arrow. Within the *Borrelia* phylogeny, "RF" denotes the Relapsing fever group, "REB" denotes the Reptile and echnida-associated *Borrelia*, while "LB" indicates the Lyme borreliosis group [69]

Bacteria of the genus *Coxiella* had the highest prevalence among the tick species examined (40/96, 42%). The newly discovered *Coxiella* species in the present study are highly diverse and greatly expand the genetic diversity within this group (Fig. 3C). Indeed, new genetic lineages were defined based on our phylogenetic analysis, including the *Coxiella* endosymbiont *of Dermacentor marginatus*, the *Coxiella* endosymbiont *of Haemaphysalis concinna* and the *Coxiella* endosymbiont *of Ixodes ovatus*, most of which were divergent from existing members of

Coxiella and generally associated with specific tick genera (Fig. 3C, Additional file 1: Fig. S12). *Coxiella burnetii*, the causative agent of Q fever [66], has been reported from *D. sinicus* tick sampled from Xinjiang province and was closely related with the "Dugway 5J108-111" strain sampled from the US [67] (Fig. 3C, Additional file 1: Fig. S11). In addition to *C. burnetii*, we identified a single species of *Borrelia* from a soft tick *Carios vesperitilionis* in Henan province. Based on the phylogenetic analyses, the newly identified bacteria, named *Borrelia henanensis* X1,

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 9 of 14

fell within a clade "RF" that contains pathogens causing tick-borne relapsing fever (Fig. 3D, Additional file 1: Fig. S13) [68, 69].

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 1: Table S5). These results suggested that bacterial genetic diversity was primary shaped by tick genetic distance, with geographic distribution having little or no impact. The strong impact of tick on bacterial genetic diversity was also reflected in the phylogenetic analysis in which we observed a significant clustering of bacterial genetic diversity at the tick general level [*Rickettsia*: association index (AI) = 2.760, $P < 0.001$; *Coxiella*: AI = 0.969, $P < 0.001$].

We next examined whether the phylogeny of the ticks and their bacterial symbionts exhibited a pattern of bacterial-tick 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 [43]. 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 [44]. 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 (Fig. 4). Collectively, these results suggest that the symbiotic bacteria from genera *Rickettsia* and *Coxiella* have co-diverged with their tick hosts for at least 264 million years.

Discussion

We collected > 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) and Hulun Beir prairie (Inner Mongolia). While most 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 *I. kuntzi* and *Hae. colasbelcouri* identified from Taiwan and Laos [70, 71], 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-bacteria relationship.

Our study inferred the evolutionary history of ticks based on the entire sets of mitochondrial genes, which revealed a well-supported phylogenetic tree that resolves the inter-species relationships of ticks with high resolution (Fig. 2). Mitochondrial genes are an important tool in tick molecular systematics because they evolve faster than most nuclear genes and are therefore better suited to address evolutionary questions at lower taxonomic levels [8, 53, 54], and their small size and the abundance of mitochondria in cells make them easy to analyze. Indeed, the sequencing depth in this study (0.2–30.1 Gbp per library) is sufficient for mitochondrial genes but inadequate for nuclear genes.

The results of the morphological and molecular species identification performed here were largely similar, suggesting that sequencing can reliably identify tick

Table 2 Results of the Mantel test and partial Mantel test comparing two factors (tick genetic distance and geographic distance) that predict the structure of genetic diversity in bacterial pathogens

	Model	*r* value (*P* value)
Rickettsia	Tick[a]	0.5868 (0.0004)
	Tick[a] | geography[b]	0.5952 (0.0009)
	Geography[a]	0.0028 (0.3916)
	Geography | tick[b]	− 0.1225 (0.8633)
Coxiella	Tick[a]	0.4939 (0.0001)
	Tick[a] | geography[b]	0.4939 (0.0001)
	Geography[a]	0.0013 (0.4499)
	Geography | tick[b]	0.0053 (0.4378)

[a] Mantel test

[b] Partial Mantel test

\| Indicates that the first factor excludes the effect of the second

	P value (Parafit)	*P* value (Jane4)
Rickettsia	0.0021	< 0.01
Coxiella	0.0003	< 0.01

Fig. 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 colors correspond to different tick groups as shown by figure legend at the right bottom

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 had paraphyletic phylogenetic structure (e.g. *D. marginatus* and *R. sanguineus* s.l.) (Fig. 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 (*D. nuttali*, *D. silvarum*, *D. 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 level 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 [72, 73]. Furthermore, the abundance of some of the bacterial pathogens was high (1.42—162.57 RPM), further facilitating inter-tick 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 tick structure on the bacteria phylogeny and the high resemblance between bacteria and tick phylogenies indicative of long-term co-divergence. Previous studies have suggested a lack of co-divergence between *Rickettsia* [74] and *Coxiella*-like endosymbionts [27]. 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 [75]. 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

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 11 of 14

bacteria. Clearly, more data are required to fully resolve the co-evolutionary history between ticks and their endosymbiont bacteria.

Conclusions

In sum, our analysis of > 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.

Abbreviations

Mt: Mitochondrial; N(A)D1: NADH dehydrogenase subunit 1; N(A)D2: NADH dehydrogenase subunit 2; N(A)D3: NADH dehydrogenase subunit 3; N(A)D4: NADH dehydrogenase subunit 4; N(A)D4L: NADH dehydrogenase subunit 4L; N(A)D5: NADH dehydrogenase subunit 5; N(A)D6: NADH dehydrogenase subunit 6; COX1: Cytochrome oxidase I; COX2: Cytochrome oxidase II; COX3: Cytochrome oxidase III; ATP8: ATPase8; ATP6: ATPase8; 16 s rRNA: Large ribosomal subunit; 12 s rRNA: Small ribosomal subunit; CYTB: Cytochrome b; trnL1: TRNA-Leu1; $trnL_2$: TRNA-Leu2; trnC: TRNA-Cys; trnQ: TRNA-Gln; trnF: TRNA-Phe; groEL: Chaperonin GroEL; gltA: Citrate synthase; atpD: ATP synthase CF1 delta chain; coxB: Cytochrome c oxidase subunit II; ftsZ: Cell division protein FtsZ; sucA: 2-Oxoglutarate dehydrogenase E1 component; bamA: Outer membrane protein assembly factor BamA; flab: Flagellin FlaB; rpoB: RNA polymerase beta subunit; SFG: Spotted fever group; LB: Lyme borreliosis; STT: Siberian tick typhus; ATBF: Asia and African tick bite fever; FESF: Far-Eastern spotted fever; RF: Relapsing fever; ML: Maximum likelihood; BaTS: Bayesian tip-association significance testing; RPM: Reads per million total reads; AI: Association index; PBS: Phosphate buffered solution.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13071-022-05485-3.

Additional file 1: Table S1. Detailed information on the 96 tick libraries examined here. **Figure S1.** Gene arrangements in tick mitochondrial genomes across eight genera. There are 23 linear maps in 19 species of *Haemaphysalis*, nine in *Amblyomma* species, two in *Bothriocroton* species, three in two *Hyalomma* species, 13 in 10 *Rhipicephalus* species, 18 in 16 Ixodes species, one of Nuttalliellidae species and 26 linear maps of Argasid ticks. Each mitochondrial genome has 13 protein-coding genes, two ribosomal genes, 22 tRNA genes and misc-features varying with tick species. Mitochondrial gene arrangement and direction in *Haemaphysalis*, *Dermacentor*, *Amblyomma*, *Hyalomma*, *Rhipicephalus* and *Bothriocroton* are almost identical and those in Ixodes, Nuttalliellidae and Argasidae are almost identical. Protein-coding genes are denoted by yellow arrows, rRNA are denoted by red arrows, tRNA are denoted by pink arrows, and control regions are denoted by gray arrows. The direction of arrows indicated the direction of protein translation. Abbreviations are as follows: ND1 = NADH dehydrogenase subunit 1, ND2 = NADH dehydrogenase subunit 2, ND3 = NADH dehydrogenase subunit 3, ND4 = NADH dehydrogenase subunit 4, ND4L = NADH dehydrogenase subunit 4L, ND5 = NADH dehydrogenase subunit 5, ND6 = NADH dehydrogenase subunit 6, COX1 = cytochrome oxidase I, COX2 = cytochrome oxidase II, COX3 = cytochrome oxidase III, ATP8 = ATPase8, ATP6 = ATPase8, 16 s rRNA = large ribosomal subunit, 12 s rRNA = small ribosomal subunit,

---

CYTB = cytochrome b. **Figure S2.** Sequence alignment of special tandem repeat regions in misc-features of *Dermacentor* mt genome. The first hypervariable region located between tRNA-Gln and NAD1 (A), and the second hypervariable region located tRNA-Gln between tRNA-Phe (B), and its cloning of sequencing of PCR products spanning the control region implied various copy numbers of short repeat sequences within the same *D. marginatus* E1 sample (C). **Table S2.** Special tandem repeat regions in misc-features of *Dermacentor* mt genome. **Figure S3.** Detailed phylogenetic tree of ticks based on all 13 protein-coding genes and two rRNA genes inferred using both ML and Bayesian methods with two mites as the outgroup. Highly similar tree topologies were obtained. The trees are midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. For clarity, the bootstrap values are shown on the left with posterior probability on the right for each node. Mt genomes identified in this study are marked colored font according to different tick genera; the name of each tick genus is shown beside. **Figure S4.** Detailed phylogenetic tree of ticks based on only 13 protein-coding genes estimated using both ML and Bayesian methods with two mites as the outgroup. Highly similar tree topologies were obtained. The trees are midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. For each node, the bootstrap values are shown on the left with posterior probability on the right. Mt genomes identified in this study are marked colored font according to different tick genera; the name of each tick genus is shown beside. **Figure S5.** Detailed phylogenetic tree of ticks based on mt 12S rRNA gene estimated using ML methods. The tree is midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. Mt 12S rRNA gene identified in this study are marked colored font according to different tick genera, the name of each tick genus is shown beside. **Figure S6.** Detailed phylogenetic tree of ticks based on mt 16S rRNA gene estimated using ML methods. The tree is midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. Mt 16S rRNA gene identified in this study are marked colored font according to different tick genera, the name of each tick genus is shown beside. **Figure S7.** Differential diagnosis for *D. nuttalli* and *D. silvarum*. Adult trochanter I dorsal spur of *D. nuttalli* (A) is short, broad and blunt apically, while that of *D. silvarum* (B) is slightly long, pointed at the apex. **Figure S8.** Relative abundance of bacteria and fungi at the level of genus based on metaphlan2 results. **Table S3.** Prevalence of tick-borne bacteria in ticks at the level of genus. A total of 43 tick associated bacterial species were identified from 54 libraries and 7 tick genera (with the exception of *Hyalomma*). Bracketed numbers denote numbers of bacteria identified in each bacterial group at the level of genera. **Table S4.** The genetic similarity of bacterial strains identified in this study with its closest reference sequence. **Figure S9.** ML phylogenetic tree of the order Rickettsiales based on the groEL gene. The tree is midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. The strains identified here are marked by colored font according to different tick genera, and reference sequences are represented by black font with corresponding accession number nearby. Each bacterial group at genus level is denoted by gray font and a vertical line, while bacterial group at subfamily level is denoted by black font and a square bracket. **Figure S10.** ML phylogenetic tree of the order Rickettsiales based on the gltA gene. The tree is midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. The strains identified here are marked by colored font according to different tick genera, and reference sequences are represented by black font with corresponding accession number nearby. Each bacterial group at genus level is denoted by gray font and vertical line, while bacterial group at subfamily level is denoted by black font and square bracket. **Figure S11.** ML phylogenetic tree of the genus *Rickettsia* based on the six conserved housekeeping genes, including the atpD, coxB, ftsZ, gltA, groEL and sucA genes. Individual genes were first aligned and then concatenated to form super-alignment for phylogenetic analysis. The tree is midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. The strains identified here are marked by colored font according to different tick genera, and reference sequences are represented by black font with corresponding accession number nearby. The Rickettsiae collected here could further assigned into three groups: Spotted fever group (SFG), Typhus group (TG) and Belli group (BG). **Figure S12.** ML

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 12 of 14

phylogenetic tree of the genus *Coxiella* based on the groEL gene. The tree is midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. The strains identified here are marked by colored font according to different tick genera, and reference sequences are represented by black font with corresponding accession number nearby. A significant clustering of *Coxiella* genetic diversity at the host general level was observed, and the position of *C. burnetii* was highlighted by an arrow. **Figure S13.** ML phylogenetic tree of the genus *Borrelia* based on the flaB gene. The tree is midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. The trees are midpoint-rooted, and the scale bar represents the number of nucleotide substitutions per site. The *B. henanensis* strain X1 identified here is marked in pink font according to its tick genus (*Carios*), and reference sequences are represented by black font with corresponding accession number nearby. The *Borrelia* collected here could be assigned into three groups: The Relapsing fever group (RFG), Reptile and echidna-associated *Borrelia* (REB) and Lyme borreliosis group (LB). **Table S5.** Multiple regression of bacteria genetic distance against tick genetic distance and geographic distance shows that bacterial genetic diversity was mainly structured by tick genetic distance rather than geographical distribution.

Availability of data and materials
The tick mt genome sequences reported in this study are available on NCBI, under accession number OM368288-OM368330, MK344649. Sequence alignments underlying analyses, phylogenetic trees and related code are available from figshare (https://figshare.com/articles/dataset/The_diversity_and_evolutionary_relationships_of_ticks_and_tick-borne_bacteria_in_China/19354481). Raw sequencing data are available at the NCBI SRA database as BioProject PRJNA816261 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA816261).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Hubei Key Laboratory of Resources Utilization and Sustainable Pest Management, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, Hubei Province 430070, China. [2]School of Medicine, Shenzhen Campus of Sun Yat-Sen University, Sun Yat-Sen University, Shenzhen, Guangdong Province 518107, China. [3]Wuhan Centers for Disease Control and Prevention, Wuhan, Hubei Province 430015, China. [4]Wuhan Academy of Agricultural Sciences, Wuhan, Hubei Province 430345, China. [5]Jiangxi Province Center for Disease Control and Prevention, Nanchang, Jiangxi Province 330029, China. [6]School of Science, Technology and Engineering, University of the Sunshine Coast, Sippy Downs, QLD 4558, Australia. [7]GeneCology Research Centre, University of the Sunshine Coast, Sippy Downs, QLD 4558, Australia. [8]Sydney Institute for Infectious Diseases, School of Life & Environmental Sciences and School of Medical Sciences, The University of Sydney, Camperdown, NSW 2006, Australia.

References
1. Penalver E, Arillo A, Delclos X, Peris D, Grimaldi DA, Anderson SR, et al. Parasitised feathered dinosaurs as revealed by Cretaceous amber assemblages. Nat Commun. 2017;8:1924.
2. Nicholson WL, Sonenshine DE, Noden BH, Brown RN. Ticks (Ixodida) medical and veterinary entomology. Amsterdam: Elsevier; 2019. p. 603–72.
3. Mans BJ, de Klerk D, Pienaar R, 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:e23675.
4. Zhang YK, Zhang XY, Liu JZ. Ticks (Acari: Ixodoidea) in China: Geographical distribution, host diversity, and specificity. Arch Insect Biochem Physiol. 2019;102:e21544.
5. Zhang G, Zheng D, Tian Y, Li S. A dataset of distribution and diversity of ticks in China. Sci Data. 2019;6:105.
6. Jia N, Wang J, Shi W, Du L, Sun Y, Zhan W, et al. Large-scale comparative analyses of tick genomes elucidate their genetic diversity and vector capacities. Cell. 2020;182:1328-1340.e13.
7. Guglielmone AA, Robbins RG, Apanaskevich DA, Petney TN, Estrada-Peña A, Horak IG. The hard ticks of the world. Dordrecht: Springer; 2014.
8. Shao R, Barker SC. Mitochondrial genomes of parasitic arthropods: implications for studies of population genetics and evolution. Parasitology. 2007;134:153–67.
9. Wang T, Zhang S, Pei T, Yu Z, Liu J. Tick mitochondrial genomes: structural characteristics and phylogenetic implications. Parasit Vectors. 2019;12:451.
10. Kelava S, Mans BJ, Shao R, Moustafa MAM, Matsuno K, Takano A, 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;12:101577.
11. 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;63:251–61.
12. 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;65:589–604.
13. Yu XJ, Liang MF, Zhang SY, Liu Y, Li JD, Sun YL, et al. Fever with thrombocytopenia associated with a novel bunyavirus in China. N Engl J Med. 2011;364:1523–32.
14. Jia N, Zheng YC, Jiang JF, Ma L, Cao WC. Human infection with Candidatus Rickettsia tarasevichiae. N Engl J Med. 2013;369:1178–80.
15. Jia N, Liu HB, Ni XB, Bell-Sakyi L, Zheng YC, Song JL, et al. Emergence of human infection with Jingmen tick virus in China: a retrospective study. EBioMedicine. 2019;43:317–24.
16. Li H, Li XM, Du J, Zhang XA, Cui N, Yang ZD, et al. Candidatus Rickettsia xinyangensis as Cause of Spotted Fever Group Rickettsiosis, Xinyang, China, 2015. Emerg Infect Dis. 2020;26:985–8.
17. 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.
18. Rosenberg R, Lindsey NP, Fischer M, Gregory CJ, Hinckley AF, Mead PS, et al. Vital Signs: trends in reported vectorborne disease cases—United States and Territories, 2004–2016. MMWR Morb Mortal Wkly Rep. 2018;67:496–501.
19. Walker DH, Ismail N. Emerging and re-emerging rickettsioses: endothelial cell infection and early disease events. Nat Rev Microbiol. 2008;6:375–86.

Tian *et al. Parasites & Vectors*   (2022) 15:352

Page 13 of 14

20. Parola P, Paddock CD, Socolovschi C, Labruna MB, Mediannikov O, Kernif T, et al. Update on tick-borne rickettsioses around the world: a geographic approach. Clin Microbiol Rev. 2013;26:657–702.

21. 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.

22. Trape JF, Diatta G, Arnathau C, Bitam I, Sarih M, Belghyti D, 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:e78473.

23. Kernif T, Leulmi H, Raoult D, Parola P. Emerging tick-borne bacterial pathogens. Microbiol Spectr. 2016. https://doi.org/10.1128/microbiolspec.EI10-0012-2016.

24. Moran NA, McCutcheon JP, Nakabachi A. Genomics and evolution of heritable bacterial symbionts. Ann Rev Genet. 2008. https://doi.org/10.1146/annurev.genet.41.110306.130119.

25. Matsuura Y, Kikuchi Y, Meng XY, Koga R, Fukatsu T. Novel clade of alphaproteobacterial endosymbionts associated with stinkbugs and other arthropods. Appl Environ Microbiol. 2012;78:4149–56.

26. Toft C, Andersson SG. Evolutionary microbial genomics: insights into bacterial host adaptation. Nat Rev Genet. 2010;11:465–75.

27. Duron O, Binetruy F, Noël V, Cremaschi J, McCoy KD, Arnathau C, et al. Evolutionary changes in symbiont community structure in ticks. Mol Ecol. 2017;26:2905–21.

28. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31:1674–6.

29. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.

30. Clewley JP. Macintosh sequence analysis software. DNAStar's LaserGene. Mol Biotechnol. 1995;3:221–4.

31. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

32 Bonfield JK, Smith KF, Staden R. A new DNA sequence assembly program. Nucleic Acids Res. 1995;23:4992–9.

33. 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. Mol Biol Evol. 2011;28:2731–9.

34. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.

35. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.

36. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.

37. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003;19:1572–4.

38. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14:587–9.

39. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35:518–22.

40. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods. 2012;9:811–4.

41. Parker J, Rambaut A, Pybus OG. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. Infect Genet Evol. 2008;8:239–46.

42. Wang TH, Donaldson YK, Brettle 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. J Virol. 2001;75:11686–99.

43. Conow C, Fielder D, Ovadia Y, Libeskind-Hadas R. Jane: a new tool for the cophylogeny reconstruction problem. Algorithms Mol Biol. 2010;5:16.

44. Meier-Kolthoff JP, Auch AF, Huson DH, Göker M. COPYCAT: cophylogenetic analysis tool. Bioinformatics. 2007;23:898–900.

45. Jackson AP, Charleston MA. A cophylogenetic perspective of RNA-virus evolution. Mol Biol Evol. 2004;21:45–57.

46. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Res. 1967;27:209–20.

47. Lichstein JW. Multiple regression on distance matrices: a multivariate spatial analysis tool. Plant Ecol. 2006;188:117–31.

48 Goslee SC, Urban DL. The Ecodist package for dissimilarity-based analysis of ecological data. J Stat Softw. 2007. https://doi.org/10.18637/jss.v022.i07.

49. Lu BL. Handbook for identification of important medical animals in China. Beijing: People's Medical Publishing House; 1982.

50. Teng KF, Jiang ZJ. Economic insect fauna of China, Fasc 39 Acari: Ixodidae (in Chinese). China: Scienece Press Beijing; 1991.

51 Horak IG, Heyne H, Williams R, Gallivan GJ, Spickett AM, Bezuidenhout JD, et al. The ixodid ticks (Acari: Ixodidae) of southern Africa. Cham: Springer; 2018.

52. Sun Y, Xu RM, Wu MY. Systematics of argasid ticks (Ixodida: Argasidae) in China with a pictorial key to species. Acta Parasitol Med Entomol Sin. 2019;26:231–50.

53. Deng GF. Notes on Chinese ticks of the subgenus Argas (Acarina: Argasidae: Argas). Acta Zootaxonomica Sinica. 1983;8:255–61.

54 Hoogstraal H. Bat ticks of the genus Argas (Ixodoidea, Argasidae), 3. The subgenus Carios, a redescription of A.(C.) vespertilionis (Latreille, 1802), and variation within an Egyptian population. Ann Entomol Soc Am. 1958;51:19–26.

55. Ye X, Sun Y, Ju W, Wang X, Cao W, Wu M, et al. Vector competence of the tick Ixodes sinensis (Acari: Ixodidae) for *Rickettsia monacensis*. Parasit Vectors. 2014;7:512.

56. 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;22:620–9.

57. Li LH, Zhang Y, Wang JZ, Li XS, Yin SQ, Zhu D, et al. High genetic diversity in hard ticks from a China–Myanmar border county. Parasit Vectors. 2018;11:469.

58. Oteo JA, Portillo A. Tick-borne rickettsioses in Europe. Ticks Tick Borne Dis. 2012;3:271–8.

59. Seo MG, Kwon OD, Kwak D. High prevalence of *Rickettsia raoultii* and associated pathogens in Canine Ticks, South Korea. Emerg Infect Dis. 2020;26:2530–2.

60. Tian ZC, Liu GY, Shen H, Xie JR, Luo J, Tian MY. First report on the occurrence of *Rickettsia slovaca* and *Rickettsia raoultii* in *Dermacentor silvarum* in China. Parasit Vectors. 2012;5:19.

61. Dong Z, Yang Y, Wang Q, Xie S, Zhao S, Tan W, et al. A case with neurological abnormalities caused by *Rickettsia raoultii* in northwestern China. BMC Infect Dis. 2019;19:796.

62. Zhao S, Yang M, Jiang M, Yan B, Zhao S, Yuan W, et al. *Rickettsia raoultii* and *Rickettsia sibirica* in ticks from the long-tailed ground squirrel near the China-Kazakhstan border. Exp Appl Acarol. 2019;77:425–33.

63. Mediannikov OY, Sidelnikov Y, Ivanov L, Mokretsova E, Fournier PE, Tarasevich I, et al. Acute tick-borne rickettsiosis caused by *Rickettsia heilongjiangensis* in Russian Far East. Emerg Infect Dis. 2004;10:810–7.

64. Adegoke A, Kumar D, Bobo C, Rashid MI, Durrani AZ, Sajid MS, et al. Tick-borne pathogens shape the native microbiome within tick vectors. Microorganisms. 2020;8:1299.

65. 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;83:115–30.

66. 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;20:441.

67. Beare PA, Unsworth N, Andoh M, Voth DE, Omsland A, Gilk SD, et al. Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus Coxiella. Infect Immun. 2009;77:642–56.

68. Elbir H, Larsson P, Upreti M, Normark J, Bergström S. Genome sequence of the relapsing fever Borreliosis species *Borrelia hispanica*. Genome Announc. 2014;2:e01171-e1213.

69. Margos G, Gofton A, Wibberg D, Dangel A, Marosevic D, Loh SM, et al. The genus Borrelia reloaded. PLoS ONE. 2018;13:e0208432.

70 Hoogstraal H, Kohls GM. Description, hosts, and ecology of *Ixodes Kuntzi* N. Sp. Kuntz's Taiwan Flying Squirrel Tick (Ixodoidea: Ixodidae). J Med Entomol. 1965;2:209–14.

71. 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. Syst Appl Acarol. 2016;21:166–80.

72. Wang Q, Pan YS, Jiang BG, Ye RZ, Chang QC, Shao HZ, et al. Prevalence of multiple tick-borne pathogens in various tick vectors in Northeastern China. Vector Borne Zoonotic Dis. 2021;21:162–71.

73. Kim TY, Kwak YS, Kim JY, Nam SH, Lee IY, Mduma S, et al. Prevalence of tick-borne pathogens from ticks collected from cattle and wild animals in Tanzania in 2012. Korean J Parasitol. 2018;56:305–8.

74. Kang YJ, Diao XN, Zhao GY, Chen MH, Xiong Y, Shi M, et al. Extensive diversity of Rickettsiales bacteria in two species of ticks from China and the evolution of the Rickettsiales. BMC Evol Biol. 2014;14:167.

75. Liu W, Li L, Khan MA, Zhu F. Popular molecular markers in bacteria. Mol Gen Mikrobiol Virusol. 2012;3:14–7.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.