1. Molecular data collection: further details

1.1. DNA extraction and sequencing of bat-associated Bartonella *strains*

Genomic DNA was extracted from 129 bat-associated *Bartonella* cultures using a simple heat extraction protocol (incubation at 95°C for 10 min) and diluted 1:10 in extraction buffer (Qiagen, Valencia, CA). Amplification of targeted genetic loci (Table S1) used published primers and protocols (Bai et al., 2015, 2013; Buffet et al., 2013; McKee et al., 2017). Amplification of *groEL* was unsuccessful for many strains with the available primers (Zeaiter et al., 2002), so this locus was not sequenced for any bat-associated strains and was only available from MLSA and genomic data. Positive PCR amplicons were purified using the Qiagen QIAquick PCR Purification Kit and sequenced in both directions with the same primers on an Applied Biosystems Model 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA). Reads were then assembled in Lasergene v14 (DNASTAR, Madison, WI). Repeated amplification or sequencing was performed for some missing genes, but for 28 strains there was one or more sequence that could not be obtained: *ftsZ* (2), *nuoG* (2), *ribC* (22), or *rpoB* (3).

1.2. Sequence alignment and data cleaning

For all bat-associated and reference *Bartonella* strains, sequences from each genetic locus were aligned separately with MAFFT v7.187 (Katoh and Standley, 2013). Ends of alignments and poorly aligned sites were trimmed with Gblocks v0.91b (Castresana, 2000), and final alignments were manually checked for ambiguous base pairs and edited. The final alignment lengths and coverage across taxa are listed in Table S1, with an average of 78% coverage across the nine loci. We concatenated all loci using Phyutility v2.2 (Smith and Dunn, 2008) to produce a full supermatrix of 8345 base pairs (including gap sites) for later analyses. Preliminary phylogenetic analysis of bat-associated strains determined that seven showed evidence of homologous recombination with another bat-associated strain (even after repeated amplification and sequencing) and one showed highly discordant phylogenetic positions across sequenced loci, so these strains were removed from the database.

1.3. Molecular data validation

Previous analyses have shown that the protein-coding loci (*ftsZ*, *gltA*, *groEL*, *nuoG*, *ribC*, *rpoB*) are under purifying selection with low ratios of non-synonymous to synonymous substitutions (Bai et al., 2015; Buffet et al., 2013). The 16S rRNA locus is known for being highly conservative within a bacterial genus (Kosoy et al., 2018; La Scola et al., 2003). As a spacer sequence, ITS is unlikely to be under selection. We examined GC content across the full alignment for all 332 taxa using DAMBE v7.0.48 (Xia, 2018). The eubartonellae clade and *B. tamiae* exhibited a stationary GC content distribution between 0.38-0.48 while *B. apis*, *Candidatus* Tokpelaia hoelldoblerii, and *Brucella abortus* had progressively higher GC content values (Fig. S1). Previous studies have shown that the similarity in GC content between *B. tamiae* and eubartonellae can affect phylogenetic results. Specifically, nucleotide alignments show that *B. tamiae* is a sister taxon to eubartonellae while protein alignments or nucleotide alignments without the third codon position show that *B. tamiae* is a sister taxon to *B. apis* (Bisch et al., 2018; Segers et al., 2017). Since this inference was focused primarily on the eubartonellae clade and not on its putative sister taxa in arthropods, we determined that the stationary GC content distribution for eubartonellae was acceptable for phylogenetic analysis and required no correction.

To confirm the absence of homologous recombination within taxa in the database, we generated a network phylogeny in SplitsTree v4.14.8 (Huson, 2005) using the concatenated alignment and the Neighbor-Net method (Bryant and Moulton, 2003) on uncorrected pairwise distances. The network phylogeny showed a moderately tree-like structure (Fig. S2) with parallelograms connecting closely related taxa and basal splits indicative of shared evolutionary history. A pairwise homoplasy (PHI) test

(Bruen et al., 2005) for recombination implemented in SplitsTree found no statistically significant evidence for recombination (P = 1) for the concatenated alignment or each locus separately (Table S2).

Separate loci were tested for the presence of nucleotide substitution saturation by plotting uncorrected versus adjusted distances (Tamura-Nei model (Tamura and Nei, 1993)) and transitions/transversions versus adjusted distances using DAMBE and the R package *ape* (Paradis et al., 2016, 2004; R Core Team, 2020). Adjusted distances did not show substantial saturation, exhibiting a strongly linear relationship with only slightly asymptotic behavior at the farthest distances (Fig. S3). Transitions and transversions fell along a straight line (Fig. S4) and transitions largely outnumbered transversions for all loci except ITS (Xia, 2018), indicating no substantial evidence of saturation. The absence of significant saturation was confirmed for all loci (Table S3) in DAMBE using the test developed by Xia et al. (2003). Based on all the tests above, we determined that these molecular loci would be appropriate for phylogenetic analysis and accurate estimation of divergence times.

2. Phylogenetic analysis: further details

2.1. Phylogenetic model selection

The best sequence evolution model was chosen according to the Akaike information criterion (AIC) using jModelTest v2.1.6 (Darriba et al., 2012) via the CyberInfrastructure for Phylogenetic RESearch (CIPRES) Science Gateway portal v3.3 (Miller et al., 2010). The generalized time-reversible model with a proportion of invariant sites and gamma rate variation across sites (GTR+I+G) was chosen for all loci except *ssrA*, which best fit the Tamura-Nei model (TN+I+G) (Table S1). Since substitution models had little effect on the tree topology, we chose to analyze all loci using the GTR+I+G for consistency and to correspond with the maximum likelihood analysis, which used a GTR+I+G model. A maximum likelihood (ML) tree was generated from the concatenated alignment of nine loci using RAxML v8.2.12 (Stamatakis, 2014) on CIPRES with 1000 bootstrap iterations to estimate node support. The ML tree was used to compare topologies with the Bayesian tree and for tip-association tests.

2.2. Phylogenetic model priors and run settings

The prior distributions for substitution rate and speciation model parameters are listed in Table S4. The prior for all ancestral state transitions was a gamma distribution with shape and scale parameters set to one, and the prior for the mean rate of order and ecozone transitions were set to the CTMC approximate reference prior (Ferreira and Suchard, 2008). We ran three chains in BEAST using the model settings above for the final analysis. The chains were run for $2x10^8$ iterations, sampling parameters every $2x10^4$ iterations. We inspected posterior distributions for all model parameters to assess convergence, mixing, and high effective sample sizes (ESS > 200) using Tracer v1.7.1 (Drummond and Rambaut, 2007). We chose the maximum clade credibility (MCC) tree from the posterior tree iterations after burn-in using TreeAnnotator (Drummond and Rambaut, 2007) and the tree with the highest MCC score was used for all subsequent analyses. The final tree was visualized and edited in FigTree v1.4.4. Molecular clocks for the nine genetic loci were summarized by the median and highest posterior density (HPD) of their distributions. The divergence date of the most recent common ancestor of mammal-infecting *Bartonella* (eubartonellae, excluding *B. tamiae* and *B. apis*) was summarized from the MCC tree by the median and HPD.

2.3. Testing alternative models in BEAST

To increase confidence in the robustness of our conclusions with respect to phylogenetic model choice, we performed additional runs in BEAST using alternative models and subsets of sequence data. The amount of data and the complexity of models led to long computational runtimes (up to 7 days) that reached the limit permitted on CIPRES (Miller et al., 2010). For this reason, we did not pursue a

formal model selection approach through estimation of marginal likelihoods (Baele et al., 2013, 2012) and instead chose to run a non-exhaustive series of models using combinations of alternative model settings to assess the combined effects on the topology and divergence times on the resulting tree.

For models that used a TN+I+G model for *ssrA*, the two prior distributions for the kappa priors were chosen to be lognormal with log mean of one and a log standard deviation of 1.25 with an initial value of 2. For models that used an uncorrelated relaxed clock model with a lognormal distribution of clock rates along branches, the means for each locus were set as for the exponential distribution detailed in the main text. An additional prior was set for the standard deviation of the lognormally distributed clock rates using an exponential distribution with a mean of 0.33. All model combinations were run until parameters converged to stationary distributions as determined through visual inspection in Tracer v1.7.1. Burn-in iterations were removed and the maximum clade credibility (MCC) tree was selected using TreeAnnotator. We then compared the topology and divergence dates (particularly the estimated divergence date of eubartonellae) of the MCC trees.

Regardless of substitution, codon partitioning, clock, or tree models, we found only limited variation in the topology of the tree across all runs with no major changes in the position of large clades that would influence the results or conclusions in the main text. The divergence dates of eubartonellae (Table S4) and *a posteriori* defined clades (Table S8) varied little across runs, indicating that the molecular data, taxon sampling, and choice of prior on the 16S rRNA clock were more important to phylogenetic inference than any other model settings. The only major differences observed in the topology and divergence dates of the tree were observed when a strict clock was used. These runs predicted a younger divergence date for eubartonellae (~57 mya) and showed a different arrangement of clades A-C and the clades that contain *B. bacilliformis*, *B. rochalimae*, and *B. clarridgeiae*. All runs using strict molecular clocks had much lower likelihoods than runs using relaxed clocks, so the use of a strict clock was rejected. As long as variation in clock rates were allowed to be uncorrelated across the branches of the tree, the topology and divergence dates on the tree were stable.

We note that the exclusion of ITS sequences had little effect on tree topology and divergence dates, so this locus may have had limited phylogenetic signal. Nevertheless, we retained this locus for the final run used in the main text. Since codon partitioning and the choice of relaxed clock models had little influence on the trees, we chose not to use codon partitioning and to use exponential distributions for the uncorrelated relaxed clocks rather than lognormal distributions for the final runs in the main text to reduce the number of independent parameters that needed to be estimated.

3. *Bartonella* lineages associated with arthropods: further details

Several *Bartonella* lineages in the database were labeled as being associated primarily with arthropods for the ancestral state reconstruction analysis. *B. apis* was originally isolated from western honeybees (*Apis mellifera*) and has not been associated with any mammalian hosts (Kešnerová et al., 2016). While several strains of *B. apis* have been characterized from honeybees in North America and Europe (Kešnerová et al., 2016), we chose to associate this species with the Palearctic ecozone to reflect the hypothesized historical distribution of domesticated *Apis mellifera* in northern Africa or the Middle East (Cridland et al., 2017). *B. tamiae* was originally isolated from humans in Thailand (Kosoy et al., 2008), this likely represents an accidental association. Genetic sequences identified as *B. tamiae* or closely related to this species have been obtained from several arthropod species including bat flies and bat ticks (Bai et al., 2018; Leulmi et al., 2016) and chigger mites collected from rodents (Kabeya et al., 2010). Given its basal position relative to the mammal-associated eubartonellae clade and closer affinities with *B. apis* (Kešnerová et al., 2016), we chosen to associate *B. tamiae* primarily with arthropods for this analysis.

We labeled *B. bacilliformis* and its relatives *B. ancashensis* and *Candidatus* B. rondoniensis as being associated with arthropods instead of a particular mammalian order because a reservoir host has

not been conclusively determined for these species. *Bartonella bacilliformis* causes severe morbidity and mortality in humans, and prevalence is generally low in human populations, so humans are unlikely to be the reservoir host (Sanchez Clemente et al., 2012). Furthermore, repeated attempts to isolate *B. bacilliformis* from alternative plant or animal reservoirs have been unsuccessful (Birtles et al., 1999; Garcia-Quintanilla et al., 2019; Herrer, 1953). Despite the uncertainty about the reservoir host, *B. bacilliformis* is known to be vectored by *Lutzomyia* spp. sandflies (Battisti et al., 2015; Billeter et al., 2008; Breitschwerdt and Kordick, 2000). A recent study also reported the presence of *B. bacilliformis* in ticks collected from tapirs and peccaries in Peru (del Valle-Mendoza et al., 2018), although this finding has been disputed (Ruiz, 2019). The phylogenetically related *Candidatus* B. rondoniensis was also described from the assassin bug *Eratyrus mucronatus* in French Guiana (Laroche et al., 2017). While the host or vector of *B. ancashensis* is unknown (Mullins et al., 2015) it is part of a clade that includes *B. bacilliformis* and *Candidatus* B. rondoniensis. Given the uncertainty of the mammalian hosts for this *Bartonella* clade, we chose to associate this group primarily with arthropods since it appears to be the ancestral trait (Neuvonen et al., 2016). Future work that conclusively determines the mammalian hosts of *B. bacilliformis* and its allies is clearly necessary and could improve the inference of ancestral hosts for *Bartonella* lineages.

Finally, the host origin of *B. senegalensis* is unclear since it was isolated from the soft tick *Ornithodoros sonrai* in Senegal (Mediannikov et al., 2013). Although the ticks were found in rodent burrows, the presence of the bacteria was not confirmed in any mammals, so we chose to associate this species with arthropods. Similar to the clade that includes *B. bacilliformis*, future studies involving *B. senegalensis* and its host associations will improve our knowledge of evolution within the *Bartonella* genus.

We confirmed that none of these choices had an effect on the results by repeating stochastic character mapping using alternative assignments of traits to these tips. Bats were always inferred to be the ancestral hosts of eubartonellae. Thus, we chose to retain these trait assignments for the analysis in the main text.

4. Revision of *Bartonella* tree topology: further details

In contrast with past phylogenetic analyses of the *Bartonella* genus that used only maximum likelihood analysis of concatenated genes (Engel et al., 2011; Guy et al., 2013; Harms et al., 2017; Wagner and Dehio, 2019; Zhu et al., 2014), we showed that neither *B. bacilliformis* nor *B. australis* are the most deeply branching lineages in the genus. Instead we found that *B. bacilliformis* and its allies *B. ancashensis* (Mullins et al., 2015) and *Candidatus* B. rondoniensis (Laroche et al., 2017), constituting the clade previously named “Lineage 1”, are in fact most closely related to ruminant-associated *Bartonella* species including *B. bovis*, *B. schoenbuchensis*, and others (Fig. 1; Fig. S8); a finding supported by Wagner and Dehio (Wagner and Dehio, 2019). This ruminant clade was named clade C in our analysis and “Lineage 2” in other studies (Engel et al., 2011; Guy et al., 2013; Harms et al., 2017; Wagner and Dehio, 2019; Zhu et al., 2014). The group of species including *B. rochalimae, B. clarridgiae*, and allies, previously named “Lineage 3”, was found to be distantly related to a clade containing kangaroo-associated *B. australis* (Fournier et al., 2007) and other *Candidatus* strains from marsupials (Kaewmongkol et al., 2011a, 2011b), and two lineages associated with bats, one from Africa (Bai et al., 2015) and one from Europe (Urushadze et al., 2017). These clades (“Lineages 1-3”) were all found to be part of a strongly supported monophyletic clade (posterior probability, PP = 1) that includes the deeply branching sister group clade A associated with neotropical bats (Fig. 1). The bat-associated and marsupial-associated clades could potentially be elevated to the level of lineages equal to the others. Alternatively, unification of these lineages into a monophyletic clade would suggest a redefinition of lineages into subclades.

Broad taxon sampling also expanded "Lineage 4", a well-supported clade (PP = 1) that contains all other *Bartonella* species separate from "Lineages 1-3" and most of the diversity in the genus (Fig. 1). Specifically, we discovered four new bat-associated clades (D, G, L, N) within this lineage, with clade D as the sister group to all other "Lineage 4" clades and clade G as the sister group to a large clade of predominantly rodent-associated *Bartonella* species (enclosed within clade O). Clade L, containing strains from North American and European vespertilionid bats (Lilley et al., 2017; Stuckey et al., 2017; Urushadze et al., 2017; Veikkolainen et al., 2014) and *Candidatus* B. mayotimonensis (Lin et al., 2010), along with clade N associated with neotropical bats, are contained within a clade that includes the *B. vinsonii* species complex associated primarily with rodents (Bai et al., 2011; Kosoy, 2010; Kosoy et al., 2012, 1997; Morway et al., 2008; Rubio et al., 2014; Schulte Fischedick et al., 2016). We also recovered a monophyletic clade (PP = 1) that unites rodent-associated *B. birtlesii*, *B. doshiae*, and *B. taylorii*, similar to a previous MLSA study (Buffet et al., 2013). Subdivisions within lineage 4 could be based on radiations within distinct mammalian groups, as we have done (Fig. 1A; Tables S5-S6). This revision of the *Bartonella* tree through increased taxon sampling and characterization of bat-associated strains illustrates the diversity in this genus that remains uncharacterized.

5. Estimated clock rates for *Bartonella* genetic markers: further details

To verify that the molecular clock approach could capture variation across loci using a single strong prior distribution on the 16S rRNA gene, we analyzed clock rates for each of the nine loci. Clock rates predictably varied by gene function (Table S9). The 16S locus had a very low median clock rate at $5.2x10^{-10}$ nucleotide substitutions $site^{-1}$ $year^{-1}$ (95% HPD: $3.4\text{-}7.1x10^{-10}$) across branches. As this locus codes for a functional RNA with a conserved 3D structure, this low rate was deemed reasonable and was very close to previous estimates of 16S rRNA divergence of 1-2% per 50 million years in *E. coli* and *Buchnera* symbionts of aphids (Moran et al., 1993; Ochman et al., 1999). Protein-coding loci and the functional transfer-messenger RNA locus *ssrA* had branch rates five to nine times higher than 16S rRNA while ITS had rates 22 times higher than 16S rRNA (Table S9).

6. Biogeographic patterns of select *Bartonella* clades: further details

In support of our hypothesis that *Bartonella* clades co-diverge with their mammalian hosts, there were several instances of deep separations of host-associated *Bartonella* strains that are most compatible with an ancient origin of the *Bartonella* genus. Inoue *et al.* (2011) discovered phylogenetically distinct clades of *B. washoensis* infecting ground squirrels in North America and Asia, a result we replicated in our tree within clade E (Fig. 1; Fig. S8). The squirrels harboring these bacteria are from two separate genera, *Spermophilus* from Eurasia and *Urocitellus* from North America, that diverged 7.8 mya according to studies published on TimeTree (Kumar et al., 2017). Therefore, it is unlikely squirrels from these two genera have been in recent close contact that could lead to *Bartonella* transmission and the divergence observed in the *Bartonella* clades reflects their independent evolution in isolated hosts. Similar patterns were seen in *Bartonella* clades infecting bats. One involves the separation of two clades within bat-associated clade L (Fig. 1). One clade within this group (Fig. S8) contains strains from vespertilionid bats in Europe (Urushadze et al., 2017; Veikkolainen et al., 2014) and the other clade contains a strain from North American bats (Lilley et al., 2017) and an agent of human endocarditis, *Candidatus* B. mayotimonensis (Lin et al., 2010). *Myotis* and *Eptesicus* spp. bats in North America diverged from their congeners in Eurasia 16.2 and 15.3 mya respectively according to TimeTree. Within the large clade D harbored by Old World bats (Fig. 1; Fig. S8) there are two *Bartonella* strains infecting *Hipposideros* spp. bats, *H. larvatus* from Thailand (McKee et al., 2017) and *H. vittatus* from Kenya (Kosoy et al., 2010). While *Hipposideros* species have repeatedly moved

between Africa and Asia according to phylogenetic analysis (Foley et al., 2017), these two species have been separated for 34 million years. These divergence times between geographically isolated hosts are reflected in the estimated times for their *Bartonella* divergence times: 5.2 mya for *B. washoensis*-like strains in ground squirrels, 10.5 mya for *Candidatus* B. mayotimonensis-like strains in vespertilionid bats, and 27.6 mya for the two *Hipposideros*-associated strains. These results provide confidence in the molecular clock approach and an ancient diversification of the *Bartonella* genus, however more work is clearly needed to reconstruct historical biogeographical patterns of bartonellae and their hosts.

7. Network analysis of *Bartonella* host and ecozone transitions: further details

Based on our stochastic character mapping of host orders and ecozones onto 1000 posterior sampled trees, we built separate networks using host orders and ecozones as nodes and the median number of transitions as edges. Considering only transitions within and between mammalian orders, the ecozone network had 22 non-zero median transitions between 6 nodes, resulting in a network density of 73% considering all possible directed transitions. The host order network showed only 10 non-zero transitions between 9 nodes, for a density of 14%. The ecozone network also had higher median counts of transitions than the order network, with up to 12 observed transitions between the Palearctic and Indo-Malayan ecozones (Table S12). Rodents are a source of transitions to Carnivora, Eulipotyphla, and Lagomorpha, while bats are a source to Diprotodontia and other marsupials, Artiodactyla, and Carnivora. Many of these transitions are strongly supported within the MCC tree with posterior probability greater than 0.9 (Fig. 1A; Fig. S8A). Notable transitions from Rodentia include those to Carnivora at the ancestor to *B. rochalimae* and to *Bartonella* sp. JM-1 from *Martes melampus* within clade E (including *B. washoensis*); to Eulipotyphla at the ancestor to *B. florenciae*, to *Bartonella* sp. DB5-6 from *Sorex araneus* within clade J (including *B. birtlesii*, *B. doshiae*, and *B. taylorii*), and to *B. tribocorum* and *B. queenslandensis* strains from shrews within clade H; and to Lagomorpha at the ancestor to *B. alsatica*. Well-supported transitions from Chiroptera include to Diprotodontia and other marsupials at the ancestor to clade B (including *B. australis*). Rodents and bats showed an equal number of transitions between each other (Fig. 3A), however the sources of these transitions are equivocal with lower posterior probabilities for the ancestral host (Fig. 1A).

8. Analysis of taxonomic and geographic biases in bat *Bartonella* studies

Considering that our study was meant to draw conclusions about the evolution of the *Bartonella* genus with its mammalian hosts, there is a need to demonstrate that our choice of *Bartonella* lineages in our phylogenetic analysis are representative of the available hosts, or at least the host species that have been sampled to date. To do this, we compared the number of bat species with *Bartonella* lineages in our analysis to two previous studies of *Bartonella* diversity in bats globally (Corduneanu et al., 2018; Frank et al., 2018). The relative distribution of bat species in different families represented in the *Bartonella* studies was then compared to the distribution of bat species among families within the recognized global diversity of bats, as recently compiled by bat taxonomists (https://batnames.org/) (Simmons and Cirranello, 2020). We chose to look at the relative counts by the number of bat species, not the number of sequences associated with each species, because there are several species that have been sampled in multiple studies and are therefore overrepresented in terms of the number of sequences, particularly *Desmodus rotundus* and *Eidolon helvum*.

Examining the relative number of distinct species in a family out of the total species in each study's dataset, we see that sampling is biased in the bat *Bartonella* studies compared to bat diversity as a whole (Fig. S9). In all three *Bartonella* studies, the diverse families Molossidae, Pteropodidae, and Vespertilionidae are underrepresented while the family Phyllostomidae is overrepresented. The number of bat species represented in our study is only slightly lower than the other two *Bartonella* studies (50

vs. 56 and 76). Furthermore, the number of bat families represented in our study (11) is only slightly lower than the other studies (both have 13 families).

We also examined the geographic biases present in our study compared to these previous global analyses of bat *Bartonella* diversity (Corduneanu et al., 2018; Frank et al., 2018) and the geographic focus of other bat *Bartonella* studies. We classified 44 total *Bartonella* studies in bats into different geographic ecozones based on the country where sampling took place. This comparison showed that a greater proportion of *Bartonella* studies in bats have been performed in the Palearctic and Neotropic ecozones (Fig. S10). The global analyses presented by Corduneanu et al., Frank et al., and in the present study all reflect this geographic bias. Thus, we must acknowledge that while our choice of *Bartonella* lineages from bats may not be representative of the global diversity of bats, the chosen lineages are representative of the geographic and taxonomic diversity that has been sampled to date in *Bartonella* studies of bats.

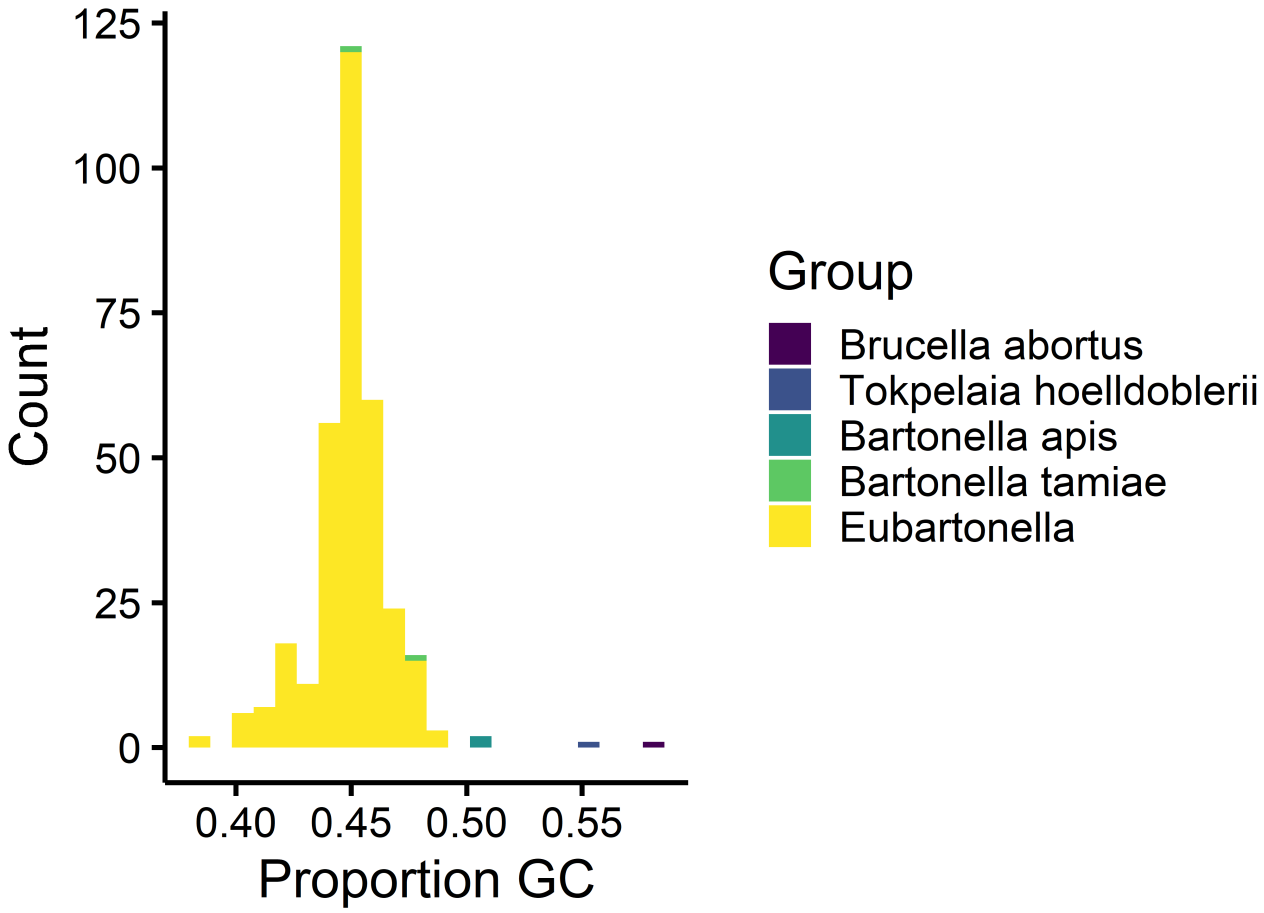

Fig. S1. Histogram of GC content across taxa. Nucleotide content was calculated across the concatenated alignment of nine loci for all 332 taxa.

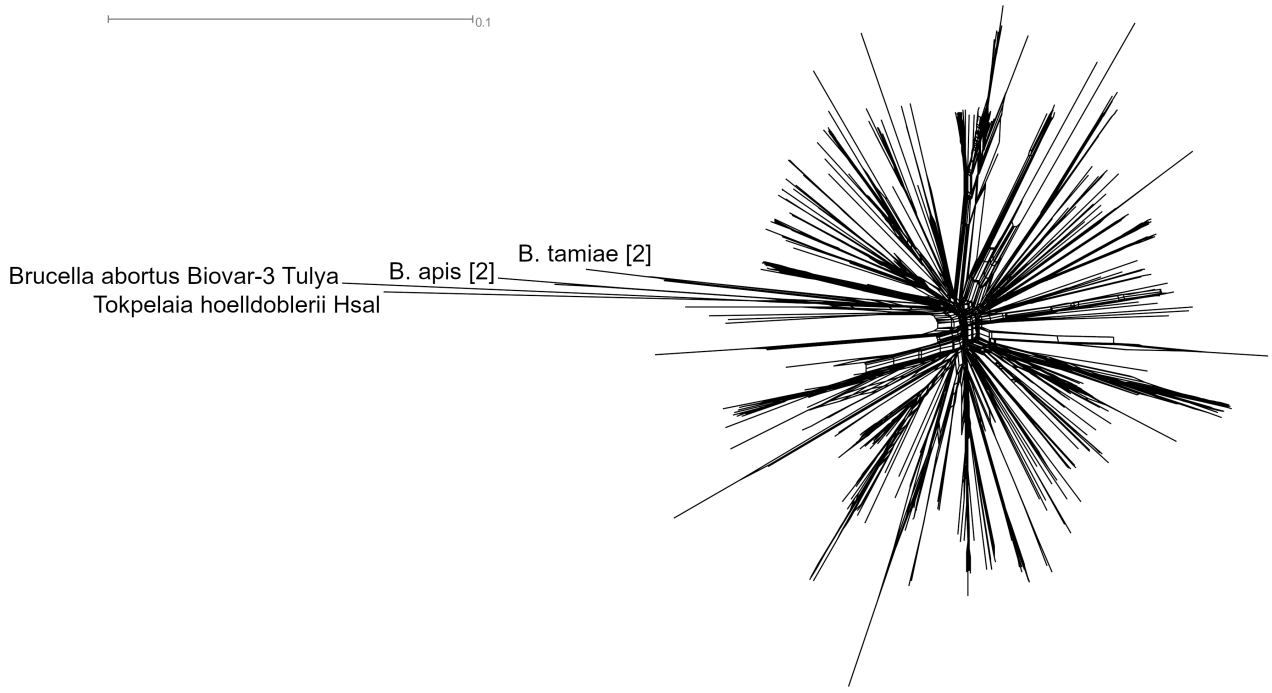

Fig. S2. Network phylogeny of *Bartonella* strains. The network was produced using the Neighbor-Net method on uncorrected pairwise distances calculated from an 8345 base pair alignment of nine genetic loci. Distances are show as the number of nucleotide substitutions per site.

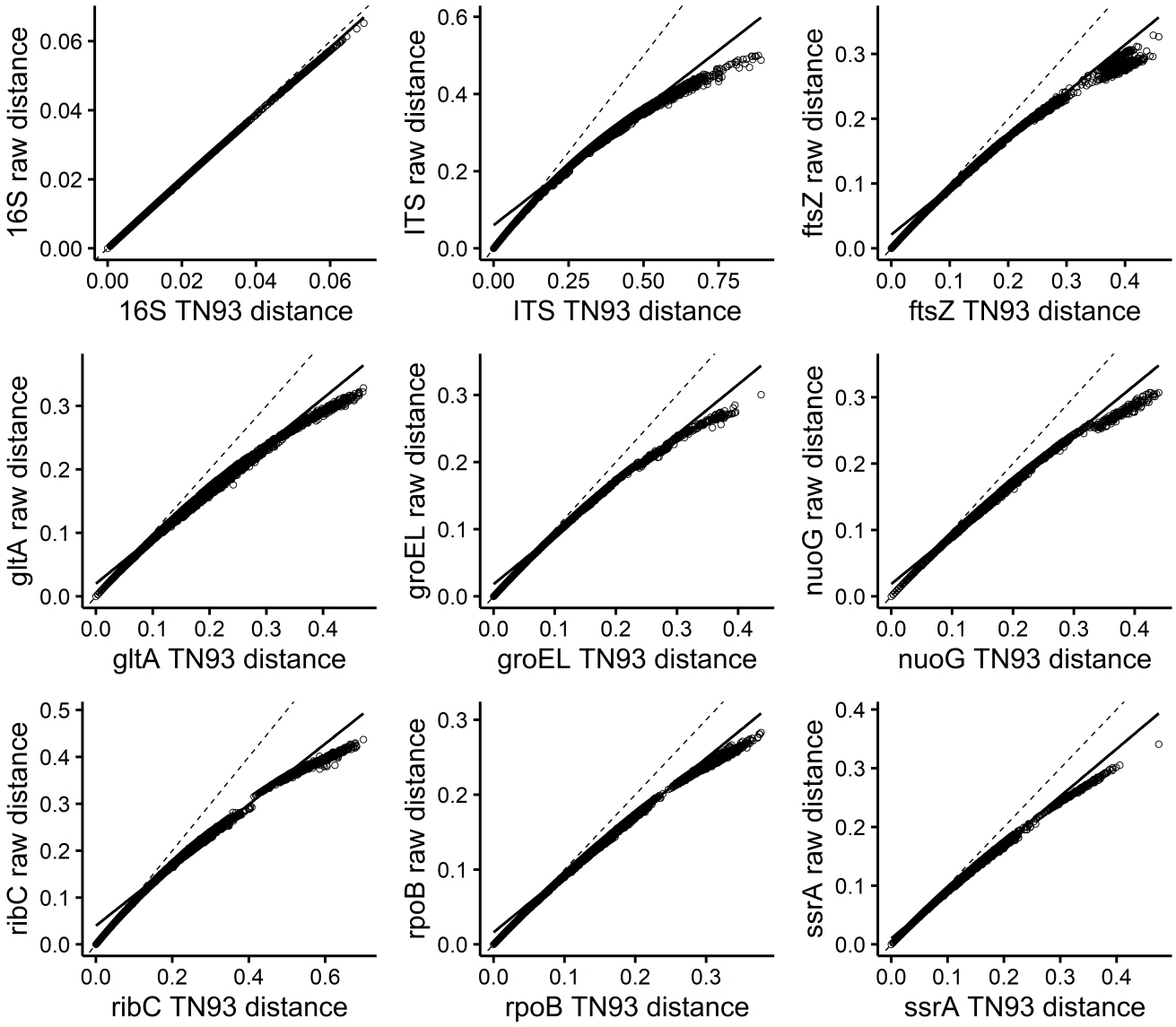

Fig. S3. Nucleotide substitution saturation across nine sequenced loci using uncorrected versus adjusted distances. Points represent pairwise distances for all taxa sequenced at each locus. Raw distances represent the uncorrected pairwise distances and adjusted distances were calculated using the Tamura-Nei model. The dashed line shows the 1:1 line for uncorrected versus adjusted distances and the solid line shows the best-fit line for linear regression.

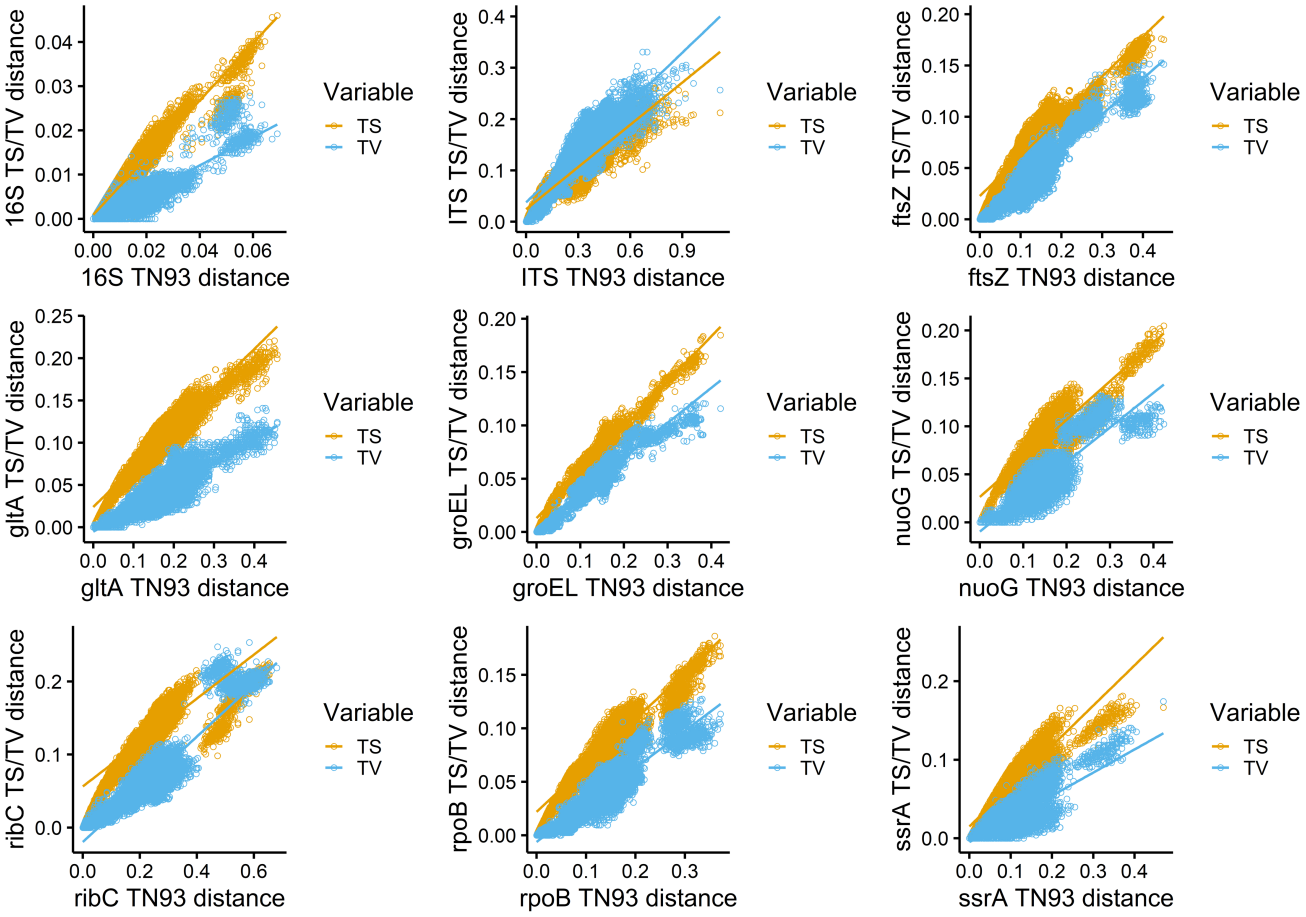

Fig. S4. Nucleotide substitution saturation across nine sequenced loci using adjusted distances versus transitions and transversions. Points represent pairwise distances for all taxa sequenced at each locus. Adjusted distances were calculated using the Tamura-Nei model. Transitions (TS) are colored orange and transversions (TV) are colored blue. The solid lines show the best-fit lines for linear regression for transitions and transversions.

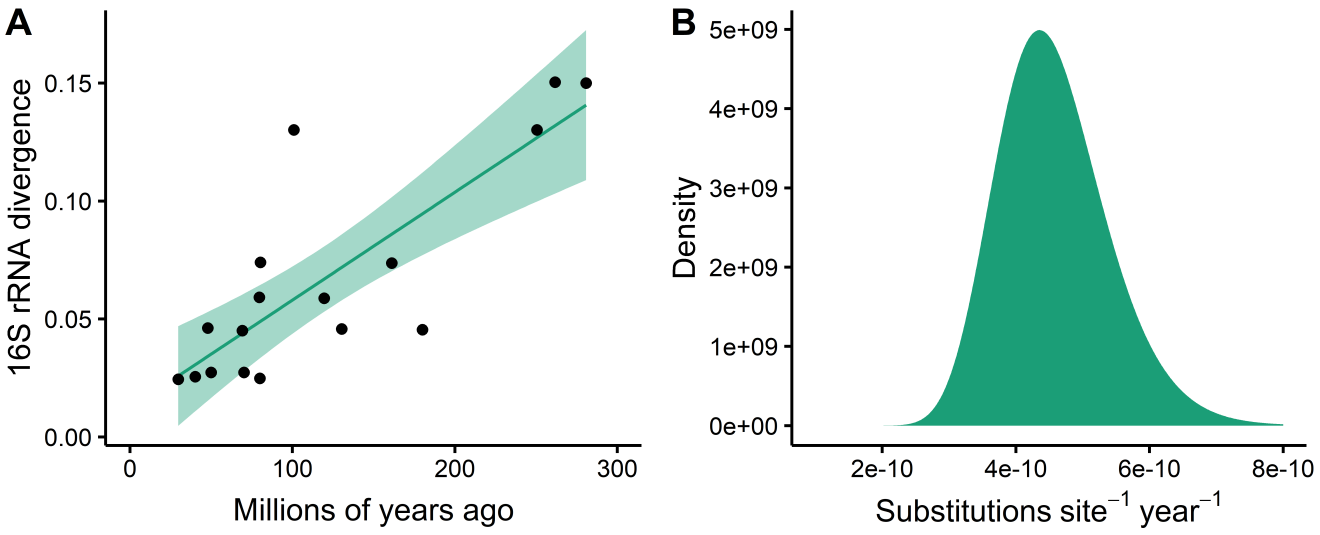

Fig. S5. Estimated molecular clock for 16S ribosomal RNA (rRNA). (A) Linear regression of 16S rRNA divergence and host divergence times for bacterial symbionts of arthropods from Kuo and Ochman (2009). (B) A lognormal distribution for the 16S rRNA molecular clock estimated by moment matching to the normal distribution of the fitted mean and standard error of the regression.

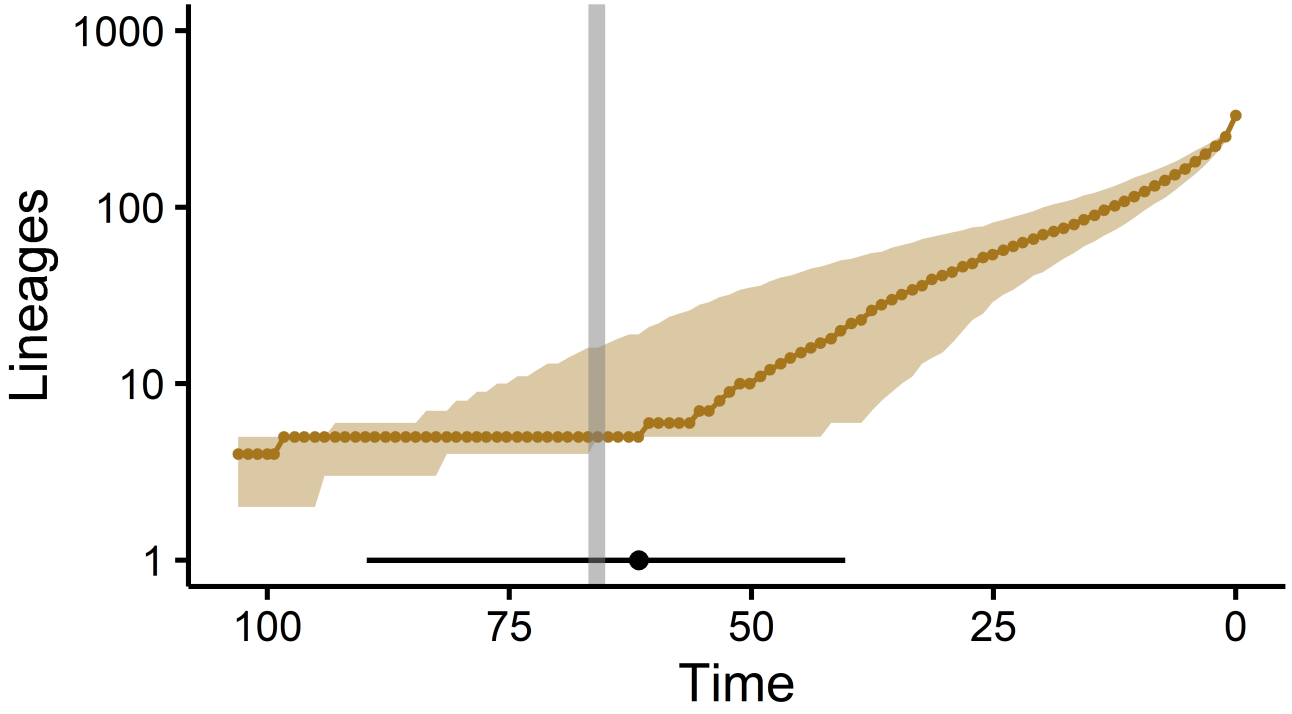

Fig. S6. Number of *Bartonella* lineages through time. Circles show the median number of lineages and shading representing the 95% HPD interval. The diversification date of eubartonellae is shown as a circle at the bottom of the figure with a line for the 95% HPD interval. Time is shown in millions of years. The Cretaceous-Paleogene extinction event is drawn as a gray line at 66 million years ago.

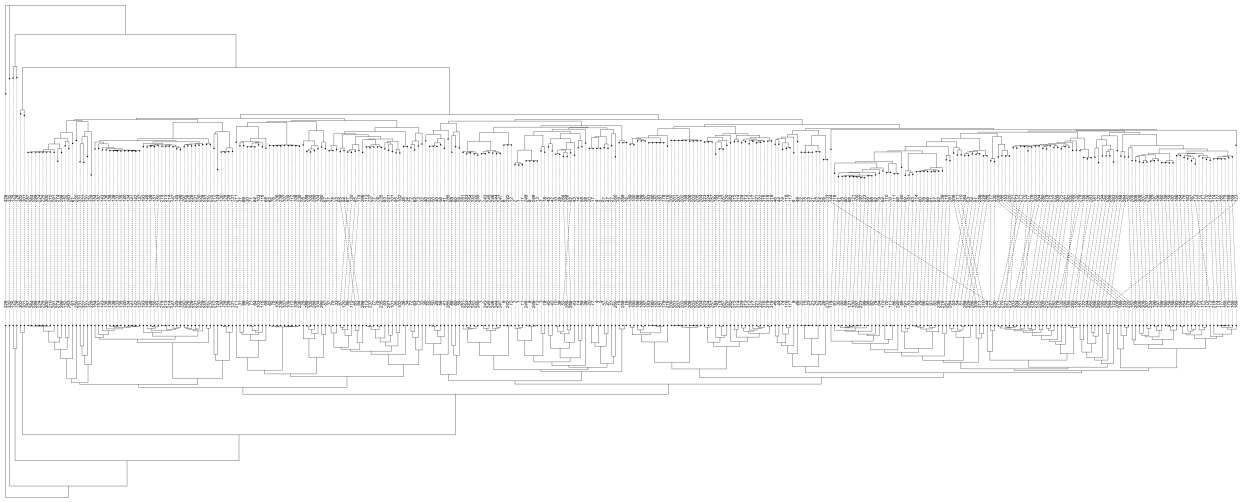

Fig. S7. Comparison of Bayesian and maximum likelihood trees. The Bayesian tree (top) used separate sequence evolution models for each of the nine partitioned loci. The maximum likelihood tree (bottom) used concatenated sequences of all nine loci.

Fig. S8. Timed maximum clade credibility tree of *Bartonella* lineages including ancestral reconstruction of (A) host orders and (B) ecozones. Posterior probabilities (PP) for nodes are indicated by the size of circles. Branch lengths are in millions of years. Branches are colored according to their most probable (PP > 0.5) host order or ecozone, with host or ecozone probability shown by the color of circles at each node.

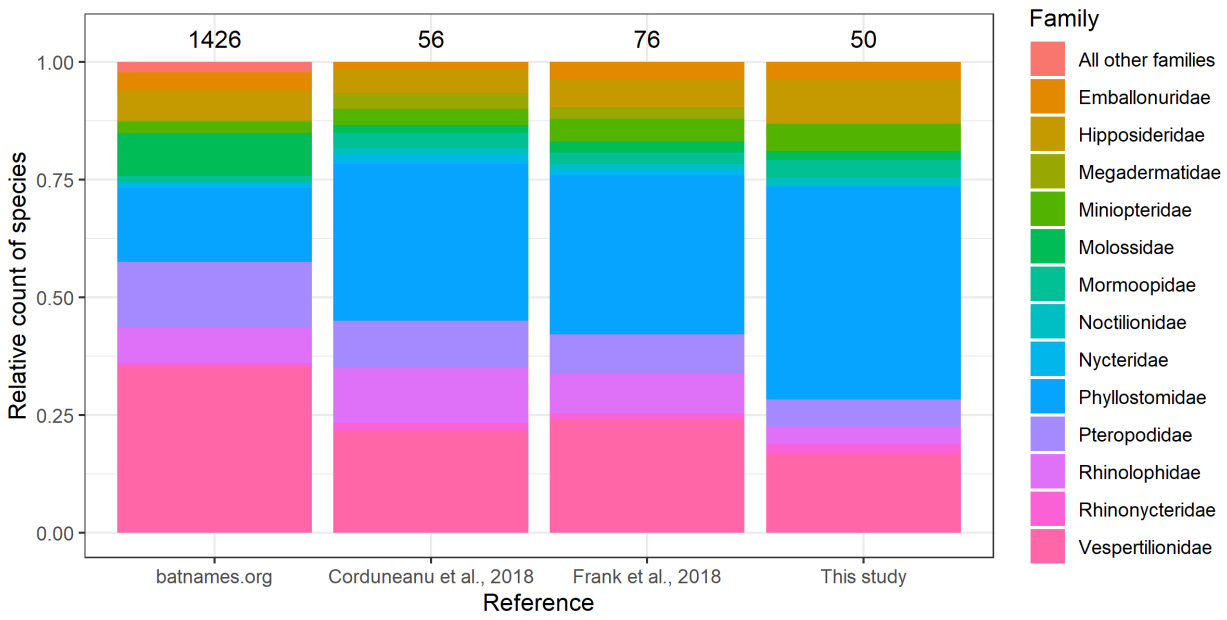

Fig. S9. Relative count of bat species represented in each dataset. The total number of species in each reference is listed above the bars.

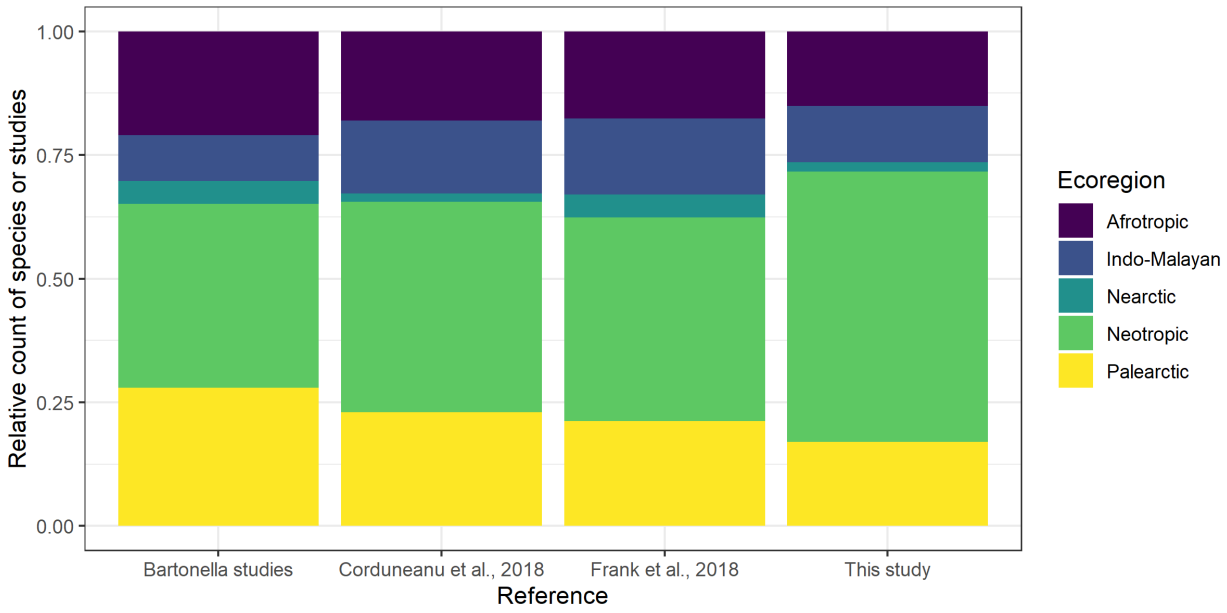

Fig. S10. Relative counts of bat species by geographic ecozone. Relative counts for *Bartonella* studies in bats are based on the number of studies in an ecozone relative to the total number of *Bartonella* studies in bats (44).

Table S1. Features of sequenced genetic loci. The number of taxa with sequences for each locus and the coverage of taxa out of 332 are listed, as well as the number of sites (base pairs) included in the final database. The best DNA substitution model was chosen based on the Akaike information criterion (AIC) in jModelTest. The proportion of invariant sites and the substitution rate gamma shape parameter were estimated from the best model. GTR, generalized time-reversible; TN, Tamura-Nei; G, gamma distributed rate variation; I, proportion of invariant sites.

Locus	Name	Taxa	Coverage	Sites	AIC best model	Invariant sites	Gamma
16S	16S ribosomal RNA	289	0.87	1511	GTR+I+G	0.77	0.33
ITS	16S-23S internal transcribed spacer	251	0.76	1833	GTR+I+G	0.11	1.1
ftsZ	cell division protein	327	0.98	885	GTR+I+G	0.46	0.64
gltA	citrate synthase	332	1	348	GTR+I+G	0.32	0.71
groEL	heat-shock chaperonin protein	116	0.35	1632	GTR+I+G	0.47	0.81
nuoG	NADH dehydrogenase gamma subunit	227	0.68	342	GTR+I+G	0.48	0.74
ribC	riboflavin synthase	256	0.77	561	GTR+I+G	0.2	0.86
rpoB	DNA-directed RNA polymerase beta subunit	322	0.97	849	GTR+I+G	0.48	0.68
ssrA	transfer-messenger RNA	220	0.66	384	TN+I+G	0.31	0.6

Table S2. Results of pairwise homoplasy index (PHI) tests for homologous recombination. The parameter k represents the number of informative sites within a window of 100 base pairs. P-values greater than 0.05 indicate that the observed PHI was outside of the expected distribution of PHI for the tree, thereby failing to reject the null hypothesis of no recombination.

Locus	k	Expected mean PHI	Expected variance PHI	Observed PHI	P-value
16S	8	0.42	2.6×10^{-4}	0.423	0.57
ITS	50	0.61	3.6×10^{-5}	0.78	1
ftsZ	43	1.81	2.5×10^{-4}	1.9	1
gltA	57	1.24	4.9×10^{-4}	1.21	0.067
groEL	39	1.1	3.9×10^{-5}	1.13	1
nuoG	45	1.62	9.5×10^{-4}	1.64	0.71
ribC	69	1.64	2.5×10^{-4}	1.7	1
rpoB	43	1.93	2.9×10^{-4}	1.97	0.98
ssrA	45	0.63	2.2×10^{-4}	0.7	1
Concatenated	40	0.95	6.5×10^{-6}	1.2	1

Table S3. Results of tests for substitution saturation. The index of substitution saturation (Iss) was calculated for each locus across a series of subtrees randomly pruned to a number of taxa based on 100 iterations. The critical index (Iss.c) is the value at which the sequences will begin to fail to recover the true tree and was calculated for each locus across the series of sampled taxa. If Iss is smaller than Iss.c and the P-value is less than 0.05, then we conclude that the sequences have not experienced severe substitution saturation and can be used for phylogenetic reconstruction. The Iss and Iss.c values shown here assume a symmetrical tree topology and use the proportion of invariant sites from Table S1.

Locus	Taxa	Iss	Iss.c	T	DF	P-value
16S	4	0.15	0.8	8.1	24	0
	8	0.15	0.78	7.1		0
	16	0.16	0.59	4.7		0.0001
	32	0.16	0.78	6.5		0
ITS	4	0.014	1.3	76.8	21	0
	8	0.014	1.6	82.8		0
	16	0.017	0.61	26.9		0
	32	0.019	2.1	85.3		0
ftsZ	4	0.33	0.82	6.1	47	0
	8	0.33	0.82	5.6		0
	16	0.33	0.58	2.7		0.0086
	32	0.34	0.85	5.6		0
gltA	4	0.27	0.78	12.2	121	0
	8	0.25	0.74	11.4		0
	16	0.26	0.63	8.6		0
	32	0.27	0.7	10.2		0
groEL	4	0.27	0.8	17.4	265	0
	8	0.28	0.75	15.8		0
	16	0.28	0.72	14.5		0
	32	0.28	0.7	14		0
nuoG	4	0.3	0.78	10.7	117	0
	8	0.29	0.73	9.6		0
	16	0.29	0.65	7.9		0
	32	0.3	0.69	8.3		0
ribC	4	0.27	0.79	11.3	106	0
	8	0.26	0.76	10.3		0
	16	0.26	0.61	7.2		0
	32	0.27	0.74	9.6		0
rpoB	4	0.31	0.78	12.1	163	0

Locus	Taxa	Iss	Iss.c	T	DF	P-value
	8	0.3	0.74	10.8		0
	16	0.31	0.68	9.2		0
	32	0.32	0.68	9.2		0
ssrA	4	0.11	1.4	15.3	13	0
	8	0.11	1.9	18.5		0
	16	0.11	0.66	5.9		0.0001
	32	0.11	2.6	26.2		0

Table S4. Prior distributions for phylogenetic analysis in BEAST.

Parameter	Distribution	Initial value
A-C substitutions	gamma(0.05, 10)	1
A-G substitutions	gamma(0.05, 20)	1
A-T substitutions	gamma(0.05, 10)	1
C-G substitutions	gamma(0.05, 10)	1
G-T substitutions	gamma(0.05, 10)	1
Base frequencies	uniform(0, 1)	0.25
Gamma shape parameter	exponential(0.5)	0.5
Proportion of invariant sites	uniform(0, 1)	0.5
Birth-death birth rate	uniform(0, 1E5)	0.01
Birth-death relative death rate	uniform(0, 1)	0.5
Proportion of taxa sampled	beta(1, 1)	0.01
16S rRNA UCED clock rate	lognormal(-21.5, 0.18)	4.6×10^{-10}
Host state transition rates	gamma(1, 1)	1
Ecozone state transition rates	gamma(1, 1)	1

Table S5. Robustness of mammal-infecting eubartonellae divergence date to model choice. RelTime divergence dates were estimated in MEGA using uniform prior distributions based on the confidence intervals of 15 host divergence dates listed in Table S7 and a maximum likelihood tree based on a concatenated alignment of all nine loci. BEAST divergence dates were estimated using separate prior distributions for all nine genetic loci separately with a strong prior distribution on the 16S rRNA locus and diffuse continuous-time Markov chain priors on the remaining loci. Separate BEAST runs using alternative sequence evolution, tree, and clock models were run until parameters converged. Intervals in parentheses show either the 95% highest posterior density interval for BEAST analyses or the 95% maximum likelihood confidence interval for RelTime. The primary model used in the main text is in bold. All runs were performed with all nine loci except those marked with a dagger, which were run with ITS. Codon partitioning was added to the last two runs in the table (marked with a double dagger). GTR, generalized time-reversible; TN, Tamura-Nei; I, proportion of invariant sites; G, gamma distributed rate variation; BD, birth-death; BDI, birth-death with incomplete sampling.

Method	Sequence evolution model	Tree model	Clock model	Divergence date
RelTime	GTR+G	Concatenated maximum likelihood	Relative rates	66.3 (63.5-69.1)
BEAST	All loci GTR+I+G	Coalescent, constant size	Strict lognormal	57.6 (38.1-82.5)
BEAST	All loci GTR+I+G	BD	Strict lognormal	57.2 (37.2-81.7)
BEAST	All loci GTR+I+G	BDI	Strict lognormal	56.9 (35.6-80)
BEAST	All loci GTR+I+G	Coalescent, constant size	Relaxed lognormal	69.6 (45.1-103.1)
BEAST	All loci GTR+I+G	BD	Relaxed lognormal	65.4 (42-96.7)
BEAST	All loci GTR+I+G	BDI	Relaxed lognormal	63.5 (42.3-94.2)
BEAST	ssrA TN+I+G, other loci GTR+I+G	BDI	Relaxed lognormal	63.4 (40.9-97.1)
BEAST	All loci GTR+I+G	BDI	Relaxed exponential	**61.6 (40.3-89.7)**
BEAST	ssrA TN+I+G, other loci GTR+I+G	BDI	Relaxed exponential	64.5 (40.8-101.5)
BEAST	All loci GTR+I+G	BDI	Relaxed lognormal	63.4 (41.8-95.8)†
BEAST	ssrA TN+I+G, other loci GTR+I+G	BDI	Relaxed lognormal	64 (40.6-101.3)†
BEAST	All loci GTR+I+G	BDI	Relaxed exponential	59.9 (39.5-90.3)†
BEAST	ssrA TN+I+G, other loci GTR+I+G	BDI	Relaxed exponential	59.5 (37.2-84.1)†
BEAST	All loci GTR+I+G	BDI	Relaxed lognormal	67.2 (41.3-97.4)‡
BEAST	All loci GTR+I+G	BDI	Relaxed lognormal	65.3 (41.5-97.1)‡

Table S6. Summary of *Bartonella* clades and host associations. Host clades above or below the order level associated with each *Bartonella* clade and any named *Bartonella* species or *Candidatus*-level species are listed. Clades A, D, G, L, and N are novel bat-associated clades described in this study. Clade O contains the predominantly rodent-associated clades H-N. Host clades are detailed in Table S6.

Bartonella clade	Tips in clade	Host order(s)	Host clade	*Bartonella* species in clade
A	51	Chiroptera	Noctilionoidea	*Candidatus* B. rolaini
B	4	Dasyuromorphia Diprotodontia Peramelemorphia	Marsupialia	*B. australis*
				Candidatus B. antechini
				Candidatus B. bandicootii
				Candidatus B. woyliei
C	32	Artiodactyla	Pecora	*B. bovis*
				B. capreoli
				B. chomelii
				B. dromedarii
				B. melophagi
				B. schoenbuchensis
				Candidatus B. davousti
D	55	Chiroptera	Yinpterochiroptera	*B. naantaliensis*
E	19	Rodentia	Sciuridae	*B. jaculi*
				B. heixiaziensis
				B. washoensis
F	10	Carnivora	Felidae	*B. henselae*
				B. koehlerae
G	15	Chiroptera	Vespertilionoidea	
H	29	Rodentia	Murinae	*B. elizabethae*
				B. fuyuanensis
				B. grahamii
				B. queenslandensis
				B. rattimassiliensis
				B. mastomydis
				B. tribocorum
I	4	Rodentia	Gerbillinae	*B. pachyuromydis*
J	22	Rodentia	Arvicolinae	*B. birtlesii*
				B. doshiae
				B. taylorii
K	3	Rodentia	Neotominae	*B. vinsonii*

Bartonella clade	Tips in clade	Host order(s)	Host clade	*Bartonella* species in clade
L	7	Chiroptera	Myotis	*Candidatus* B. mayotimonensis
M	2	Rodentia	Sigmodontinae	
N	30	Chiroptera	Phyllostomidae	
O	88	Rodentia	Muroidea	

Table S7. Description of *Bartonella* clades and host associations. MRCA, most recent common ancestor.

Bartonella clade	Host clade	Description
A	Noctilionoidea	MRCA for families Noctilionidae, Mormoopidae, and Phyllostomidae in order Chiroptera
B	Marsupialia	MRCA for orders Dasyuromorphia, Diprotodontia, and Peramelemorphia in infraclass Marsupialia
C	Pecora	MRCA for families Bovidae and Cervidae in order Artiodactyla
D	Yinpterochiroptera	MRCA for families Rhinolophidae, Hipposideridae, and Pteropodidae in order Chiroptera
E	Sciuridae	MRCA for genera *Sciurus*, *Tamiasciurus*, *Glaucomys*, *Eutamias*, *Urocitellus*, *Spermophilus*, *Otospermophilus*, and *Cynomys* in family Sciuridae
F	Felidae	MRCA for genera *Panthera*, *Lynx*, *Felis*, *Puma*, and *Acinonyx* in family Felidae
G	Vespertilionoidea	MRCA for families Vespertilionidae and Molossidae in order Chiroptera
H	Murinae	MRCA for genera *Bandicota*, *Rattus*, *Niviventer*, *Melomys*, *Uromys*, and *Mus* in subfamily Murinae
I	Gerbillinae	MRCA for genera *Pachyuromys*, *Meriones*, and *Sekeetamys* in subfamily Gerbillinae
J	Arvicolinae	MRCA for subfamily Arvicolinae
K	Neotominae	MRCA for subfamily Neotominae
L	*Myotis*	MRCA for species *Myotis blythii* and *Myotis lucifugus* in family Vespertilionidae

M	Sigmodontinae	MRCA for genera *Hylaeamys*, *Akodon*, and *Sigmodon* in subfamily Sigmodontinae
N	Phyllostomidae	MRCA for family Phyllostomidae
O	Muroidea	MRCA for superfamily Muroidea (excluding *genus* Typhlomys)

Table S8. Comparison of divergence date estimates for 15 *Bartonella* clades with divergence dates of the associated hosts within each clade collated from TimeTree. The number of published studies used to estimate host divergence dates is listed. Both host and *Bartonella* clade divergence dates are in units of millions of years. Intervals in parentheses show either the 95% highest posterior density interval for *Bartonella* clade dates or the 95% confidence interval for host clade dates. Intervals in brackets show the ranges. Details regarding *Bartonella* clades are found in Tables S5-S6. The positive correlation between *Bartonella* and host clades is depicted in Fig. 2.

Bartonella clade	TimeTree studies	Host clade date	*Bartonella* clade date
A	19	43 (41-46) [36.7-60.4]	45.9 (29.6-68.4) [24.1-110.1]
B	17	62 (58-67) [49.8-82]	35.1 (21.9-53.8) [16-84.7]
C	10	27.3 (23.1-31.5) [20.8-38.7]	15.4 (9.4-23.1) [6.9-36.2]
D	21	58 (56-61) [46-71.2]	49.6 (32.3-72.5) [25.2-110.1]
E	11	35 (29-40) [17.8-47.6]	18.8 (10.7-28.9) [8.7-52.9]
F	12	15.2 (12.3-18.1) [9.6-26.3]	8.4 (4.8-13) [3.3-27.8]
G	15	49 (45-52) [36-60.4]	36.9 (22.9-55.1) [18.7-84.3]
H	84	20.9 (18.3-23.4) [8.8-53.6]	20.8 (13.1-30.6) [9.7-47.5]
I	6	18.4 (10.3-26.4) [11-28.4]	16.4 (9.1-25.1) [6.9-38.5]
J	3	18.6 [15.2-20.9]	25.2 (15.5-37.1) [11.7-47.2]
K	8	19.3 (12.1-26.4) [8.6-32]	11.3 (6.1-18.6) [4.2-35.4]
L	6	18.1 (9.3-27) [10.8-32.8]	10.5 (5.7-17.2) [4.7-31.9]
M	5	19.8 (10-29.5) [11.6-29.7]	7.4 (2.7-15.2) [1.6-30.7]
N	16	31 (29-33) [25-35.3]	23.9 (15.5-35.2) [11.9-54.4]
O	16	45 (42-49) [35.9-60.1]	40.4 (26.4-59) [21.1-95]

Table S9. Posterior median estimates of clock rates across genetic loci. Numbers in parentheses show the 95% highest posterior density (HPD) interval. UCED, uncorrelated exponential distribution. Median UCED clock rate represents the molecular clock rate ($x10^{-9}$ substitutions $site^{-1}$ $year^{-1}$). Median branch clock rate represents the molecular clock rate estimate weighted by branch lengths ($x10^{-9}$ substitutions $site^{-1}$ $year^{-1}$).

Locus	Median UCED clock rate	Median branch clock rate
16S	0.47 (0.31-0.63)	0.52 (0.34-0.71)
ITS	9.2 (5.9-13.3)	11.5 (7.5-16.6)
ftsZ	3.1 (2-4.6)	3.8 (2.4-5.4)
gltA	3.8 (2.3-5.5)	3.7 (2.4-5.3)
groEL	2.5 (1.6-3.7)	2.4 (1.5-3.5)
nuoG	3.3 (2-4.8)	3.3 (2.1-4.8)
ribC	4 (2.4-5.6)	3.9 (2.5-5.6)
rpoB	5 (3.1-7.1)	4.8 (3-6.7)
ssrA	2.9 (1.8-4.2)	2.8 (1.7-4.1)

Table S10. Tip-association tests of host trait clustering on trees. Observed credible intervals were drawn from 1000 posterior sampled trees. Null distributions were produced from 100 resampling steps for each sampled tree. ML, maximum likelihood; AI, association index; PS, parsimony score.

	Posterior sampled trees		Single ML tree	
Trait	Order	Ecozone	Order	Ecozone
States	12	7	12	7
Observed AI	1.4 (1.39-1.43)	6.13 (5.87-6.25)	2.2	6.1
Null AI	25.5 (23.1-27.4)	28.1 (25.8-30)	20.9 (19-22.8)	23 (21.2-24.9)
Observed PS	24 (24-24)	61.9 (61-62)	54	102
Null PS	153.5 (147-160)	178.4 (172.5-186.1)	172.2 (148-205)	193.8 (175-216)

Table S11. Posterior probability of host and ecozone states for the mammal-infecting eubartonellae ancestor. For the MCC tree, 100 stochastic character mapping simulations were run on 100 randomly resampled trees and for the posterior sampled trees, 100 stochastic simulations were run on 10 randomly chosen trees with 10 random resampling iterations of tips. The distribution of the posterior probability of the ancestral state over 100 trees is summarized by the median and the interquartile range (in parentheses).

	Sampled tips	Probability host is Chiroptera	Probability ecozone is Palearctic
MCC tree	87	0.99 (0.95-1)	0.75 (0.7-0.8)
	32	0.95 (0.91-0.98)	0.63 (0.56-0.69)
	21	0.93 (0.88-0.96)	0.64 (0.56-0.69)
Posterior sampled trees	87	0.99 (0.95-1)	0.77 (0.7-0.82)
	32	0.92 (0.87-0.95)	0.67 (0.55-0.74)
	21	0.93 (0.9-0.97)	0.63 (0.57-0.72)

Table S12. Results of stochastic character mapping of host orders and ecozones on 1000 posterior sampled trees. The posterior distribution of the number of transitions is given as the median and the 95% HPD interval (in parentheses).

Network	Transition	Count
Order	Arthropoda → Chiroptera	1 (0-1)
	Arthropoda → Outgroup	1 (0-1)
	Carnivora → Arthropoda	1 (0-1)
	Chiroptera → Arthropoda	1 (0-2)
	Chiroptera → Artiodactyla	1 (0-1)
	Chiroptera → Carnivora	1 (0-2)
	Chiroptera → Diprotodontia	1 (0-2)
	Chiroptera → Rodentia	2 (1-4)
	Diprotodontia → Dasyuromorphia	1 (0-1)
	Diprotodontia → Peramelemorphia	1 (0-1)
	Rodentia → Carnivora	3 (2-5)
	Rodentia → Chiroptera	2 (1-3)
	Rodentia → Eulipotyphla	4 (3-4)
	Rodentia → Lagomorpha	1 (0-1)
	All order transitions	26 (24-30)
Ecozone	Afrotropic → Indo-Malayan	2 (0-5)
	Afrotropic → Nearctic	3 (1-5)
	Afrotropic → Palearctic	3 (0-7)
	Australasia → Indo-Malayan	1 (0-2)
	Australasia → Palearctic	1 (0-3)
	Indo-Malayan → Afrotropic	2 (0-4)
	Indo-Malayan → Australasia	1 (0-2)
	Indo-Malayan → Nearctic	1 (0-3)
	Indo-Malayan → Neotropic	2 (0-3)
	Indo-Malayan → Palearctic	3 (1-5)
	Nearctic → Afrotropic	1 (0-4)
	Nearctic → Indo-Malayan	1 (0-3)
	Nearctic → Neotropic	1 (0-3)
	Nearctic → Palearctic	4 (2-7)
	Neotropic → Afrotropic	1 (0-3)
	Neotropic → Indo-Malayan	1 (0-3)
	Neotropic → Palearctic	2 (0-5)

Network	Transition	Count
	Outgroup → Indo-Malayan	1 (0-2)
	Palearctic → Afrotropic	8 (4-12)
	Palearctic → Australasia	3 (1-6)
	Palearctic → Indo-Malayan	12 (9-16)
	Palearctic → Nearctic	11 (8-14)
	Palearctic → Neotropic	6 (3-9)
	Palearctic → Outgroup	1 (0-3)
	All ecozone transitions	82 (71-92)

Table S13. Node properties of state transition networks. Measures are based on median counts of stochastic character mapping simulations on 1000 posterior sampled trees. The networks exclude transitions between states and the outgroup (*Brucella abortus*) or transitions between mammalian orders and arthropods.

Network	State	Degree	Weighted degree	Out-degree	Weighted out-degree	Betweenness
Order	Artiodactyla	1	1	0	0	0
	Carnivora	2	4	0	0	0
	Chiroptera	5	7	4	5	4
	Dasyuromorphia	1	1	0	0	0
	Diprotodontia	3	3	2	2	4
	Eulipotyphla	1	4	0	0	0
	Lagomorpha	1	1	0	0	0
	Peramelemorphia	1	1	0	0	0
	Rodentia	5	12	4	10	2
Ecozone	Afrotropic	7	20	3	8	0.33
	Australasia	4	6	2	2	0
	Indo-Malayan	10	26	5	9	3.67
	Nearctic	7	22	4	7	0.33
	Neotropic	6	13	3	4	0
	Palearctic	10	53	5	40	3.67

References

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., Alekseyenko, A. V, 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. 29, 2157–2167. https://doi.org/10.1093/molbev/mss084

Baele, G., Li, W.L.S., Drummond, A.J., Suchard, M.A., Lemey, P., 2013. Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. Mol. Biol. Evol. 30, 239–243. https://doi.org/10.1093/molbev/mss243

Bai, Y., Calisher, C.H., Kosoy, M.Y., Root, J.J., Doty, J.B., 2011. Persistent infection or successive reinfection of deer mice with *Bartonella vinsonii* subsp. *arupensis*. Appl. Environ. Microbiol. 77, 1728–1731. https://doi.org/10.1128/AEM.02203-10

Bai, Y., Hayman, D.T.S., McKee, C.D., Kosoy, M.Y., 2015. Classification of *Bartonella* strains associated with straw-colored fruit bats (*Eidolon helvum*) across Africa using a multi-locus sequence typing platform. PLOS Negl. Trop. Dis. 9, e0003478. https://doi.org/10.1371/journal.pntd.0003478

Bai, Y., Malania, L., Alvarez Castillo, D., Moran, D., Boonmar, S., Chanlun, A., Suksawat, F., Maruyama, S., Knobel, D., Kosoy, M.Y., 2013. Global distribution of *Bartonella* infections in domestic bovine and characterization of *Bartonella bovis* strains using multi-locus sequence typing. PLOS One 8, e80894. https://doi.org/10.1371/journal.pone.0080894

Bai, Y., Osinubi, M.O. V, Osikowicz, L., McKee, C., Vora, N.M., Rizzo, M.R., Recuenco, S., Davis, L., Niezgoda, M., Ehimiyein, A.M., Kia, G.S.N., Oyemakinde, A., Adeniyi, O.S., Gbadegesin, Y.H., Saliman, O.A., Ogunniyi, A., Ogunkoya, A.B., Kosoy, M.Y., 2018. Human exposure to novel *Bartonella* species from contact with fruit bats. Emerg. Infect. Dis. 24, 2317–2323. https://doi.org/10.3201/eid2412.181204

Battisti, J.M., Lawyer, P.G., Minnick, M.F., 2015. Colonization of *Lutzomyia verrucarum* and *Lutzomyia longipalpis* sand flies (Diptera: Psychodidae) by *Bartonella bacilliformis*, the etiologic agent of Carrión's disease. PLOS Negl. Trop. Dis. 9, e0004128. https://doi.org/10.1371/journal.pntd.0004128

Billeter, S.A., Levy, M.G., Chomel, B.B., Breitschwerdt, E.B., 2008. Vector transmission of *Bartonella* species with emphasis on the potential for tick transmission. Med. Vet. Entomol. 22, 1–15. https://doi.org/10.1111/j.1365-2915.2008.00713.x

Birtles, R.J., Canales, J., Ventosilla, P., Alvarez, E., Guerra, H., Llanos-Cuentas, A., Raoult, D., Doshi, N., Harrison, T.G., 1999. Survey of *Bartonella* species infecting intradomicillary animals in the Huayllacallán Valley, Ancash, Peru, a region endemic for human bartonellosis. Am. J. Trop. Med. Hyg. 60, 799–805. https://doi.org/10.4269/ajtmh.1999.60.799

Bisch, G., Neuvonen, M.-M., Pierce, N.E., Russell, J.A., Koga, R., Sanders, J.G., Łukasik, P., Andersson, S.G.E., 2018. Genome evolution of Bartonellaceae symbionts of ants at the opposite ends of the trophic scale. Genome Biol. Evol. 10, 1687–1704. https://doi.org/10.1093/gbe/evy126

Breitschwerdt, E.B., Kordick, D.L., 2000. *Bartonella* infection in animals: carriership, reservoir potential, pathogenicity, and zoonotic potential for human infection. Clin. Microbiol. Rev. 13, 428–438. https://doi.org/10.1128/CMR.13.3.428-438.2000.Updated

Bruen, T.C., Philippe, H., Bryant, D., 2005. A simple and robust statistical test for detecting the presence of recombination. Genetics 172, 2665–2681. https://doi.org/10.1534/genetics.105.048975

Bryant, D., Moulton, V., 2003. Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. Mol. Biol. Evol. 21, 255–265. https://doi.org/10.1093/molbev/msh018

Buffet, J.-P., Pisanu, B., Brisse, S., Roussel, S., Félix, B., Halos, L., Chapuis, J.-L., Vayssier-Taussat, M., 2013. Deciphering *Bartonella* diversity, recombination, and host specificity in a rodent community. PLOS One 8, e68956. https://doi.org/10.1371/journal.pone.0068956

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in

phylogenetic analysis. Mol. Biol. Evol. 17, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334
Corduneanu, A., Sándor, A.D., Ionică, A.M., Hornok, S., Leitner, N., Bagó, Z., Stefke, K., Fuehrer, H., Mihalca, A.D., 2018. *Bartonella* DNA in heart tissues of bats in central and eastern Europe and a review of phylogenetic relations of bat-associated bartonellae. Parasit. Vectors 11, 489. https://doi.org/10.1186/s13071-018-3070-7
Cridland, J.M., Tsutsui, N.D., Ramírez, S.R., 2017. The complex demographic history and evolutionary origin of the western honey bee, *Apis mellifera*. Genome Biol. Evol. 9, 457–472. https://doi.org/10.1093/gbe/evx009
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772–772. https://doi.org/10.1038/nmeth.2109
del Valle-Mendoza, J., Rojas-Jaimes, J., Vásquez-Achaya, F., Aguilar-Luis, M.A., Correa-Nuñez, G., Silva-Caso, W., Lescano, A.G., Song, X., Liu, Q., Li, D., 2018. Molecular identification of *Bartonella bacilliformis* in ticks collected from two species of wild mammals in Madre de Dios: Peru. BMC Res. Notes 11, 405. https://doi.org/10.1186/s13104-018-3518-z
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214. https://doi.org/10.1186/1471-2148-7-214
Engel, P., Salzburger, W., Liesch, M., Chang, C.C., Maruyama, S., Lanz, C., Calteau, A., Lajus, A., Médigue, C., Schuster, S.C., Dehio, C., 2011. Parallel evolution of a type IV secretion system in radiating lineages of the host-restricted bacterial pathogen *Bartonella*. PLOS Genet. 7, e1001296. https://doi.org/10.1371/journal.pgen.1001296
Ferreira, M.A.R., Suchard, M.A., 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. Can. J. Stat. 36, 355–368. https://doi.org/10.1002/cjs.5550360302
Foley, N.M., Goodman, S.M., Whelan, C. V, Puechmaille, S.J., Teeling, E., 2017. Towards navigating the Minotaur's labyrinth: cryptic diversity and taxonomic revision within the speciose genus *Hipposideros* (Hipposideridae). Acta Chiropterologica 19, 1–18. https://doi.org/10.3161/15081109ACC2017.19.1.001
Fournier, P.-E., Taylor, C., Rolain, J.-M., Barrassi, L., Smith, G., Raoult, D., 2007. *Bartonella australis* sp. nov. from kangaroos, Australia. Emerg. Infect. Dis. 13, 1961–1963. https://doi.org/10.3201/eid1312.060559
Frank, H.K., Boyd, S.D., Hadly, E.A., 2018. Global fingerprint of humans on the distribution of *Bartonella* bacteria in mammals. PLOS Negl. Trop. Dis. 12, e0006865. https://doi.org/10.1371/journal.pntd.0006865
Garcia-Quintanilla, M., Dichter, A.A., Guerra, H., Kempf, V.A.J., 2019. Carrion's disease: more than a neglected disease. Parasit. Vectors 12, 141. https://doi.org/10.1186/s13071-019-3390-2
Guy, L., Nystedt, B., Toft, C., Zaremba-Niedzwiedzka, K., Berglund, E.C., Granberg, F., Näslund, K., Eriksson, A.-S., Andersson, S.G.E., 2013. A gene transfer agent and a dynamic repertoire of secretion systems hold the keys to the explosive radiation of the emerging pathogen *Bartonella*. PLOS Genet. 9, e1003393. https://doi.org/10.1371/journal.pgen.1003393
Harms, A., Segers, F.H.I.D., Quebatte, M., Mistl, C., Manfredi, P., Körner, J., Chomel, B.B., Kosoy, M., Maruyama, S., Engel, P., Dehio, C., 2017. Evolutionary dynamics of pathoadaptation revealed by three independent acquisitions of the VirB/D4 type IV secretion system in *Bartonella*. Genome Biol. Evol. 9, 761–776. https://doi.org/10.1093/gbe/evx042
Herrer, A., 1953. Carrion's disease. I. Studies on plants claimed to be reservoirs of *Bartonella bacilliformis*. Am. J. Trop. Med. Hyg. 2, 637–643. https://doi.org/10.4269/ajtmh.1953.2.637
Huson, D.H., 2005. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267. https://doi.org/10.1093/molbev/msj030
Inoue, K., Kabeya, H., Hagiya, K., Kosoy, M.Y., Une, Y., Yoshikawa, Y., Maruyama, S., 2011. Multi-locus sequence analysis reveals host specific association between *Bartonella washoensis* and

squirrels. Vet. Microbiol. 148, 60–65. https://doi.org/10.1016/j.vetmic.2010.08.007
Kabeya, H., Colborn, J.M., Bai, Y., Lerdthusnee, K., Richardson, J.H., Maruyama, S., Kosoy, M.Y., 2010. Detection of *Bartonella tamiae* DNA in ectoparasites from rodents in Thailand and their sequence similarity with bacterial cultures from Thai patients. Vector-Borne Zoonotic Dis. 10, 429–434. https://doi.org/10.1089/vbz.2009.0124
Kaewmongkol, G., Kaewmongkol, S., Burmej, H., Bennett, M.D., Fleming, P.A., Adams, P.J., Wayne, A.F., Ryan, U., Irwin, P.J., Fenwick, S.G., 2011a. Diversity of *Bartonella* species detected in arthropod vectors from animals in Australia. Comp. Immunol. Microbiol. Infect. Dis. 34, 411–417. https://doi.org/10.1016/j.cimid.2011.07.002
Kaewmongkol, G., Kaewmongkol, S., Owen, H., Fleming, P.A., Adams, P.J., Ryan, U., Irwin, P.J., Fenwick, S.G., 2011b. Candidatus Bartonella antechini: A novel *Bartonella* species detected in fleas and ticks from the yellow-footed antechinus (*Antechinus flavipes*), an Australian marsupial. Vet. Microbiol. 149, 517–521. https://doi.org/10.1016/j.vetmic.2010.12.003
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010
Kešnerová, L., Moritz, R., Engel, P., 2016. *Bartonella apis* sp. nov., a honey bee gut symbiont of the class Alphaproteobacteria. Int. J. Syst. Evol. Microbiol. 66, 414–421. https://doi.org/10.1099/ijsem.0.000736
Kosoy, M., Bai, Y., Lynch, T., Kuzmin, I. V, Niezgoda, M., Franka, R., Agwanda, B., Breiman, R.F., Rupprecht, C.E., 2010. *Bartonella* spp. in bats, Kenya. Emerg. Infect. Dis. 16, 1875–1881. https://doi.org/10.3201/eid1612.100601
Kosoy, M., McKee, C., Albayrak, L., Fofanov, Y., 2018. Genotyping of *Bartonella* bacteria and their animal hosts: current status and perspectives. Parasitology 145, 543–562. https://doi.org/10.1017/S0031182017001263
Kosoy, M., Morway, C., Sheff, K.W., Bai, Y., Colborn, J., Chalcraft, L., Dowell, S.F., Peruski, L.F., Maloney, S.A., Baggett, H., Sutthirattana, S., Sidhirat, A., Maruyama, S., Kabeya, H., Chomel, B.B., Kasten, R., Popov, V., Robinson, J., Kruglov, A., Petersen, L.R., 2008. *Bartonella tamiae* sp. nov., a newly recognized pathogen isolated from three human patients from Thailand. J. Clin. Microbiol. 46, 772–775. https://doi.org/10.1128/JCM.02120-07
Kosoy, M.Y., 2010. Ecological associations between bacteria of the genus *Bartonella* and mammals. Biol. Bull. 37, 716–724. https://doi.org/10.1134/S1062359010070071
Kosoy, M.Y., Hayman, D.T.S., Chan, K.-S., 2012. *Bartonella* bacteria in nature: where does population variability end and a species start? Infect. Genet. Evol. 12, 894–904. https://doi.org/10.1016/j.meegid.2012.03.005
Kosoy, M.Y., Regnery, R.L., Tzianabos, T., Marston, E.L., Jones, D.C., Green, D., Maupin, G.O., Olson, J.G., Childs, J.E., 1997. Distribution, diversity, and host specificity of *Bartonella* in rodents from the Southeastern United States. Am. J. Trop. Med. Hyg. 57, 578–588.
Kumar, S., Stecher, G., Suleski, M., Hedges, S.B., 2017. TimeTree: a resource for timelines, timetrees, and divergence times. Mol. Biol. Evol. 34, 1812–1819. https://doi.org/10.1093/molbev/msx116
Kuo, C.-H., Ochman, H., 2009. Inferring clocks when lacking rocks: the variable rates of molecular evolution in bacteria. Biol. Direct 4, 35. https://doi.org/10.1186/1745-6150-4-35
La Scola, B., Zeaiter, Z., Khamis, A., Raoult, D., 2003. Gene-sequence-based criteria for species definition in bacteriology: the *Bartonella* paradigm. Trends Microbiol. 11, 318–321. https://doi.org/10.1016/S0966-842X(03)00143-4
Laroche, M., Berenger, J.-M., Mediannikov, O., Raoult, D., Parola, P., 2017. Detection of a potential new *Bartonella* species "*Candidatus* Bartonella rondoniensis" in human biting kissing bugs (Reduviidae; Triatominae). PLOS Negl. Trop. Dis. 11, e0005297. https://doi.org/10.1371/journal.pntd.0005297

Leulmi, H., Aouadi, A., Bitam, I., Bessas, A., Benakhla, A., Raoult, D., Parola, P., 2016. Detection of *Bartonella tamiae*, *Coxiella burnetii* and rickettsiae in arthropods and tissues from wild and domestic animals in northeastern Algeria. Parasit. Vectors 9, 27. https://doi.org/10.1186/s13071-016-1316-9

Lilley, T.M., Wilson, C.A., Bernard, R.F., Willcox, E. V, Vesterinen, E.J., Webber, Q.M.R., Kurpiers, L., Prokkola, J.M., Ejotre, I., Kurta, A., Field, K.A., Reeder, D.M., Pulliainen, A.T., 2017. Molecular detection of *Candidatus* Bartonella mayotimonensis in North American bats. Vector-Borne Zoonotic Dis. 17, 243–246. https://doi.org/10.1089/vbz.2016.2080

Lin, E.Y., Tsigrelis, C., Baddour, L.M., Lepidi, H., Rolain, J.-M., Patel, R., Raoult, D., 2010. *Candidatus* Bartonella mayotimonensis and endocarditis. Emerg. Infect. Dis. 16, 500–503. https://doi.org/10.3201/eid1603.081673

McKee, C.D., Kosoy, M.Y., Bai, Y., Osikowicz, L.M., Franka, R., Gilbert, A.T., Boonmar, S., Rupprecht, C.E., Peruski, L.F., 2017. Diversity and phylogenetic relationships among *Bartonella* strains from Thai bats. PLOS One 12, e0181696. https://doi.org/10.1371/journal.pone.0181696

Mediannikov, O., El Karkouri, K., Diatta, G., Robert, C., Fournier, P.-E., Raoult, D., 2013. Non-contiguous finished genome sequence and description of *Bartonella senegalensis* sp. nov. Stand. Genomic Sci. 8, 279–289. https://doi.org/10.4056/sigs.3807472

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, in: Proceedings of the Gateway Computing Environments Workshop (GCE). pp. 1–8.

Moran, N.A., Munson, M.A., Baumann, P., Ishikawa, H., 1993. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. Proc. R. Soc. B Biol. Sci. 253, 167–171.

Morway, C., Kosoy, M., Eisen, R., Montenieri, J., Sheff, K., Reynolds, P.J., Powers, N., 2008. A longitudinal study of *Bartonella* infection in populations of woodrats and their fleas. J. Vector Ecol. 33, 353–364. https://doi.org/10.3376/1081-1710-33.2.353

Mullins, K.E., Hang, J., Jiang, J., Leguia, M., Kasper, M.R., Ventosilla, P., Maguiña, C., Jarman, R.G., Blazes, D., Richards, A.L., 2015. Description of *Bartonella ancashensis* sp. nov., isolated from the blood of two patients with verruga peruana. Int. J. Syst. Evol. Microbiol. 65, 3339–3343. https://doi.org/10.1099/ijsem.0.000416

Neuvonen, M.-M., Tamarit, D., Näslund, K., Liebig, J., Feldhaar, H., Moran, N.A., Guy, L., Andersson, S.G.E., 2016. The genome of Rhizobiales bacteria in predatory ants reveals urease gene functions but no genes for nitrogen fixation. Sci. Rep. 6, 39197. https://doi.org/10.1038/srep39197

Ochman, H., Elwyn, S., Moran, N.A., 1999. Calibrating bacterial evolution. Proc. Natl. Acad. Sci. 96, 12638–12643. https://doi.org/10.1073/pnas.96.22.12638

Paradis, E., Blomberg, S., Bolker, B., Claude, J., Cuong, H.S., Desper, R., Didier, G., Durand, B., Dutheil, J., Gascuel, O., Heibl, C., Ives, A., Jones, B., Lawson, D., Lefort, V., Legendre, P., Lemon, J., McCloskey, R., Nylander, J., Opgen-Rhein, R., Popescu, A.-A., Royer-Carenzi, M., Schliep, K., Strimmer, K., de Vienne, D., 2016. *ape*: analyses of phylogenetics and evolution [WWW Document]. URL https://cran.r-project.org/package=ape

Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412

R Core Team, 2020. R: a language and environment for statistical computing [WWW Document]. URL http://www.r-project.org

Rubio, A. V, Ávila-Flores, R., Osikowicz, L.M., Bai, Y., Suzán, G., Kosoy, M.Y., 2014. Prevalence and genetic diversity of Bartonella strains in rodents from northwestern Mexico. Vector-Borne Zoonotic Dis. 14, 838–845. https://doi.org/10.1089/vbz.2014.1673

Ruiz, J., 2019. Dubious presence of *Bartonella bacilliformis* in ticks from Madre de Dios, Peru. BMC Res. Notes 12, 539. https://doi.org/10.1186/s13104-019-4528-1

Sanchez Clemente, N., Ugarte-Gil, C.A., Solórzano, N., Maguiña, C., Pachas, P., Blazes, D., Bailey, R.,

Mabey, D., Moore, D., 2012. *Bartonella bacilliformis*: a systematic review of the literature to guide the research agenda for elimination. PLOS Negl. Trop. Dis. 6, e1819. https://doi.org/10.1371/journal.pntd.0001819

Schulte Fischedick, F.B., Stuckey, M.J., Aguilar-Setién, A., Moreno-Sandoval, H., Galvez-Romero, G., Salas-Rojas, M., Arechiga-Ceballos, N., Overgaauw, P.A.M., Kasten, R.W., Chomel, B.B., 2016. Identification of *Bartonella* species isolated from rodents from Yucatan, Mexico, and isolation of *Bartonella vinsonii* subsp. *yucatanensis* subsp. nov. Vector-Borne Zoonotic Dis. 16, 636–642. https://doi.org/10.1089/vbz.2016.1981

Segers, F.H.I.D., Kešnerová, L., Kosoy, M., Engel, P., 2017. Genomic changes associated with the evolutionary transition of an insect gut symbiont into a blood-borne pathogen. ISME J. 1–13. https://doi.org/10.1038/ismej.2016.201

Simmons, N.B., Cirranello, A.L., 2020. Bat species of the world: a taxonomic and geographic database [WWW Document]. URL https://batnames.org/ (accessed 10.12.20).

Smith, S.A., Dunn, C.W., 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics 24, 715–716. https://doi.org/10.1093/bioinformatics/btm619

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stuckey, M.J., Boulouis, H.-J., Cliquet, F., Picard-Meyer, E., Servat, A., Aréchiga-Ceballos, N., Echevarría, J.E., Chomel, B.B., 2017. Potentially zoonotic *Bartonella* in bats from France and Spain. Emerg. Infect. Dis. 23, 539–541. https://doi.org/10.3201/eid2303.160934

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.

Urushadze, L., Bai, Y., Osikowicz, L., McKee, C., Sidamonidze, K., Putkaradze, D., Imnadze, P., Kandaurov, A., Kuzmin, I., Kosoy, M., 2017. Prevalence, diversity, and host associations of *Bartonella* strains in bats from Georgia (Caucasus). PLOS Negl. Trop. Dis. 11, e0005428. https://doi.org/10.1371/journal.pntd.0005428

Veikkolainen, V., Vesterinen, E.J., Lilley, T.M., Pulliainen, A.T., 2014. Bats as reservoir hosts of human bacterial pathogen, *Bartonella mayotimonensis*. Emerg. Infect. Dis. 20, 960–967. https://doi.org/10.3201/eid2006.130956

Wagner, A., Dehio, C., 2019. Role of distinct type-IV-secretion systems and secreted effector sets in host adaptation by pathogenic *Bartonella* species. Cell. Microbiol. 21, e13004. https://doi.org/10.1111/cmi.13004

Xia, X., 2018. DAMBE7: new and improved tools for data analysis in molecular biology and evolution. Mol. Biol. Evol. 35, 1550–1552. https://doi.org/10.1093/molbev/msy073

Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26, 1–7. https://doi.org/10.1016/S1055-7903(02)00326-3

Zeaiter, Z., Liang, Z., Raoult, D., 2002. Genetic classification and differentiation of *Bartonella* species based on comparison of partial *ftsZ* gene sequences. J. Clin. Microbiol. 40, 3641–3647. https://doi.org/10.1128/JCM.40.10.3641

Zhu, Q., Kosoy, M., Olival, K.J., Dittmar, K., 2014. Horizontal transfers and gene losses in the phospholipid pathway of *Bartonella* reveal clues about early ecological niches. Genome Biol. Evol. 6, 2156–2169. https://doi.org/10.1093/gbe/evu169