# scientific reports

OPEN

# Antimicrobial resistance gene lack in tick-borne pathogenic bacteria

Márton Papp[1], Adrienn Gréta Tóth[1], Gábor Valcz[2,3], László Makrai[4], Sára Ágnes Nagy[1], Róbert Farkas[5] & Norbert Solymosi[1,6✉]

Tick-borne infections, including those of bacterial origin, are significant public health issues. Antimicrobial resistance (AMR), which is one of the most pressing health challenges of our time, is driven by specific genetic determinants, primarily by the antimicrobial resistance genes (ARGs) of bacteria. In our work, we investigated the occurrence of ARGs in the genomes of tick-borne bacterial species that can cause human infections. For this purpose, we processed short/long reads of 1550 bacterial isolates of the genera *Anaplasma* (n = 20), *Bartonella* (n = 131), *Borrelia* (n = 311), *Coxiella* (n = 73), *Ehrlichia* (n = 13), *Francisella* (n = 959) and *Rickettsia* (n = 43) generated by second/third generation sequencing that have been freely accessible at the NCBI SRA repository. From *Francisella tularensis*, 98.9% of the samples contained the FTU-1 beta-lactamase gene. However, it is part of the *F. tularensis* representative genome as well. Furthermore, 16.3% of them contained additional ARGs. Only 2.2% of isolates from other genera (*Bartonella*: 2, *Coxiella*: 8, *Ehrlichia*: 1, *Rickettsia*: 2) contained any ARG. We found that the odds of ARG occurrence in *Coxiella* samples were significantly higher in isolates related to farm animals than from other sources. Our results describe a surprising lack of ARGs in these bacteria and suggest that *Coxiella* species in farm animal settings could play a role in the spread of AMR.

Diseases of bacterial origin are form a significant portion of public health issues. Vector-borne pathogens are gaining further importance as climate change, urbanization, and globalization affect the population size and geographical distribution of various pathogen vectors[1]. Furthermore, antimicrobial resistance (AMR) of bacteria is often referred to as one of the main challenges of the 21st century[2]. Several studies prove the widespread dissemination of antimicrobial resistance genes (ARGs) in various environmental or alimentary sample types[3–8]. Vector-borne infections are an important and widespread aspect of microbial diseases. Indeed, around 80% of the human population of the Earth is estimated to be at risk of one or more vector-borne pathogens (VBPs)[9]. Since many of these VBPs are bacterial, treating a set of vector-borne infections relies upon the efficacy of antibiotics. Despite the potential and relatively dynamic changes expected in the spread of both AMR and VBPs in the coming decades, the associations of these global health issues are less studied.

Besides such significant arthropod pathogen vectors as mosquitoes, sandflies, blackflies, fleas, lice, tsetse flies, or triatome bugs, soft and hard ticks also serve as vectors for pathogenic microorganisms of human and veterinary medical significance[9,10]. Considering that the spectrum of tick-borne bacterial pathogens is relatively wide, the assessment of the antimicrobial resistance status of tick-borne pathogens, such as the members of the genera *Anaplasma*, *Bartonella*, *Borrelia*, *Coxiella*, *Ehrlichia*, *Francisella* and *Rickettsia* is an anticipatory public health step[11–17].

Our study aimed to obtain detailed knowledge of the ARG spectrum of the bacteria that could potentially cause the primary tick-borne infections. Therefore, whole-genome sequencing (WGS) data of bacterial isolates belonging to the genera of *Anaplasma*, *Bartonella*, *Borrelia*, *Coxiella*, *Ehrlichia*, *Francisella*, *Rickettsia* were processed using a uniform bioinformatics pipeline[8].

[1]Centre for Bioinformatics, University of Veterinary Medicine, Budapest 1078, Hungary. [2]Translational Extracellular Vesicle Research Group, Eötvös Loránd Research Network-Semmelweis University, Budapest 1089, Hungary. [3]Department of Image Analysis, 3DHISTECH Ltd., Budapest 1141, Hungary. [4]Department of Microbiology and Infectious Diseases, University of Veterinary Medicine, Budapest 1143, Hungary. [5]Department of Parasitology and Zoology, University of Veterinary Medicine, Budapest 1078, Hungary. [6]Department of Physics of Complex Systems, Eötvös Loránd University, Budapest 1117, Hungary. ✉email: solymosi.norbert@gmail.com

## Materials and methods

**Samples.** The available samples and corresponding sequencing data for each genus were retrieved and downloaded from the SRA database on the following days: *Anaplasma* 2022 Aug 16, *Bartonella* 2022 Aug 19, *Borrelia* 2022 Aug 21, *Coxiella* 2022 Aug 27, *Ehrlichia* 2022 Aug 12, *Francisella* 2022 Aug 12, *Rickettsia* 2022 Aug 12. Of the downloaded samples, the number of WGS samples per genus was as follows: *Anaplasma*: 40, *Bartonella*: 251, *Borrelia*: 707, *Coxiella*: 155, *Ehrlichia*: 16, *Francisella*: 1088, *Rickettsia*: 128.

For further analysis, the raw sequencing data from Illumina, Nanopore or PacBio platforms were considered only. Moreover, the analysis was restricted to samples from species of potential tick-borne spread and potentially associated human diseases. Due to the limited resources on the pathogenic potential of each bacterial species examined in the analysis, we have considered them as potentially pathogenic if information was found that they were isolated from humans. De novo assembly was performed on the sequencing data of these samples, followed by the prediction of the ARG content of the generated contigs. Only samples with an available representative genome at the NCBI database were included with the condition that sequencing coverage reached at least 95% of the genome. After the selection steps, a total of 1550 samples (*Anaplasma*: 20, *Bartonella*: 131, *Borrelia*: 311, *Coxiella*: 73, *Ehrlichia*: 13, *Francisella*: 959, *Rickettsia*: 43) were left for analysis (BioSample IDs of the samples at each genus considered for the analysis can be found in Supplementary File 1 and 8).

**Bioinformatic analysis.** For Illumina sequenced data, the quality-based filtering and trimming of the raw short reads was performed by TrimGalore (v.0.6.6, https://github.com/FelixKrueger/TrimGalore), setting 20 as the quality threshold. Only reads longer than 50 bp were retained. Nanopore sequenced reads were adapter trimmed and quality-based filtered by Porechop (v0.2.4, https://github.com/rrwick/Porechop) and Nanofilt (v2.6.0)[18], respectively. The long reads from the PacBio sequencing platform were prepared for assembly using the pbclip (https://github.com/fenderglass/pbclip) tool.

The Illumina reads were assembled to contigs by MEGAHIT (v1.2.9)[19] using default settings. The assembly of the long read sequencing data was performed with Flye (v2.9-b1779)[20], and 3x round polishing was applied with Racon (v1.4.3)[21] on the resulting contigs. In the case of hybrid sequencing, the sequenced reads were assembled by MaSuRCA (v4.0.9)[22].

All possible open reading frames (ORFs) were predicted by Prodigal (v2.6.3)[23] on each contig. The protein-translated ORFs were searched for ARG sequences against the Comprehensive Antibiotic Resistance Database (CARD, v.3.2.3)[24,25] by the Resistance Gene Identifier (RGI, v5.2.0), running with Diamond[26] as the aligner. To capture ARGs with high-confidence rates for the following steps of the analysis, ORFs classified as perfect or strict matches were further filtered by 90% identity and 90% coverage thresholds.

ARGs predicted by RGI were further screened for potential associated mobile genetic elements to asses their potential for participation in Horizontal Gene Transfer (HGT) events. Contigs, throughout which ARGs were identified, were screened for the probability of plasmid origin by PlasFlow (v.1.1)[27]. The prediction of integrative mobile genetic elements (iMGE) on the contigs harboring ARGs were assessed by the MobileElementFinder (v1.0.3) and its database (MGEdb, v1.0.2)[28]. An ARG was considered to be associated with an iMGE if it was within the distance of the median for the longest composite transposon stored for each bacterial species in the MGEdb database (distance threshold: 10098 bp). Bacteriophage sequences were predicted with VirSorter2 (v2.2.3)[29] and were filtered to dsDNAphage and ssDNA sequences only. ARGs within phage sequences at their whole length were considered to be associated with bacteriophages.

To mitigate the effect of the database and tool used for the ARG prediction on our results, we have further analysed them with two additional pipelines. AMRFinderPlus (v.3.10.36)[30] is an alignment-based tool similar to RGI. We used AMRFinderPlus with a database (National Database of Antibiotic Resistant Organisms, NDARO, version 2022-05-26.1) without the SPSVERBc1 option to retain as much comparability as possible between its results and those from RGI and CARD. Similarly, the determined genes were filtered by 90% identity and 90% coverage thresholds. KmerResistance[31] aligns raw reads to a user-defined database with the aid of the KMA aligner[32] while also handling potential contamination. We have searched only the Illumina sequencing data from the samples considered in our analysis (for the samples included in the KmerResistance analysis, see Supplementary File 1) with KmerResistance (v2.0) against the CARD (v3.2.3) and NDARO (v2022-05-26.1) databases. Therefore, the number of samples analysed with KmerResistance as well was 1492. From the CARD database, only genes under the protein homolog models were used for the KmerResistance analysis. Results were filtered by 90% template coverage and 90% query identity thresholds. As the AMRFinderPlus analysis was run without the SPSVERBc1 option, virulence genes and genes confering resistance to metal- and biocide compounds only were filtered from the KmerResistance results when run against the NDARO database. Data management and the figure generation were performed in the R-environment (v4.1.2)[33].

The association of the origin of the samples and the ARG positivity was analysed based on the metadata available for each biosample at the SRA database. The origin of samples was determined by the sample_origin and sample_host metadata columns. Based on this metadata information, the samples were manually sorted into the following categories: companion animal, environment, farm animal, human, vector, wild animal and NA when no information was available. Also, an NA category was used in the case of indicated hosts of frequently used laboratory animal species, as in those cases, we couldn't determine if the host and the sample were experimentally manipulated. Additionally, if any other metadata indicated the laboratory origin of the strain, the NA category was used for the sample origin.

## Results

The ARG content of available WGS samples of human-pathogenic bacteria with possible tick-borne origin from the SRA database was analysed. After selecting samples with raw data of sufficient quality (for details, see Materials and Methods), 1550 samples remained for the ARG analysis. Samples represent species from 7 bacterial genera (*Anaplasma*, *Bartonella*, *Borrelia*, *Coxiella*, *Ehrlichia*, *Francisella* and *Rickettsia*), originating from at least 33 different countries. Even though most samples were supplemented with metadata of their country of origin, there were still 142 samples with this information missing. Countries with at least one sample included in our analysis are shown in Fig. 1. Predicted ARG content was analysed with RGI and CARD, a robust pipeline often employed to survey antimicrobial resistance genes in different environments[3–7]. However, as resistance gene presence was rather scarce in the samples analysed in this study, we have included two additional methods for resistance gene prediction to augment our results. Even though we utilized these additional methodologies, we base our findings on the widely used and accepted RGI pipeline. Further information about the samples can be found in Supplementary File 1.

### Resistance genes predicted by RGI and CARD.

Even though we could include a large number of samples in our analysis, a surprisingly low number of resistance genes were predicted from the ORFs. Furthermore, the ARGs were distributed among only a handful of samples. The percentage of ARG positive samples was 0% for the *Anaplasma* and *Borrelia* genera, 1.5% for the *Bartonella* genus, 11.0% for the *Coxiella* genus, 7.7% for the *Ehrlichia* genus and 4.7% for the *Rickettsia* genus. The exception was the *Francisella* genus, where almost all samples carried at least one ARG (98.9%). The resistance genes predicted for each sample are shown in Supplementary File 2. Detected ARGs for each species are summarized in Table 1.

Although there is a high number of ARG-positive samples from the *Francisella* genus, a large part of them (791 samples) was predicted to contain only one ARG. Furthermore, all of these samples had the *FTU-1* gene as the one predicted resistance gene. *FTU-1* was very prevalent in the analysed *Francisella* samples as all positive ones harboured this gene (948 samples out of the total 959). This gene is a class A β-lactamase gene found in almost every *Francisella tularensis* genomes[34], including the reference genome of the species (based on the analysis with the same pipeline of the *F. tularensis* reference genome assembly available at NCBI: https://www.ncbi.nlm.nih.gov/assembly/GCF_000008985.1/). As a consequence of the high prevalence of this gene, we have additionally calculated the ratio for the *Francisella* samples containing ARGs without taking *FTU-1* gene into consideration. The percentage of positive samples without this gene is 16.3%, which is still the highest ratio found in our analysis. The number of ARG positive and negative samples can be found in Table 2. Supplementary File 2 shows all the resistance genes predicted by RGI for each positive biosample included in our analysis.

### Mobility analysis.

Contigs, where ARGs were predicted by RGI, were further analysed to determine if these genes could be associated with any mobile genetic elements. An ARG was considered potentially mobile if the contig it was found on was predicted to be of plasmid origin, if it was within a predicted phage sequence or if it was found within 10098 bp distance from the closest detected iMGE (for details on the mobile element determination and the selection of the cut-off value see Materials and Methods). The number of potentially mobile and
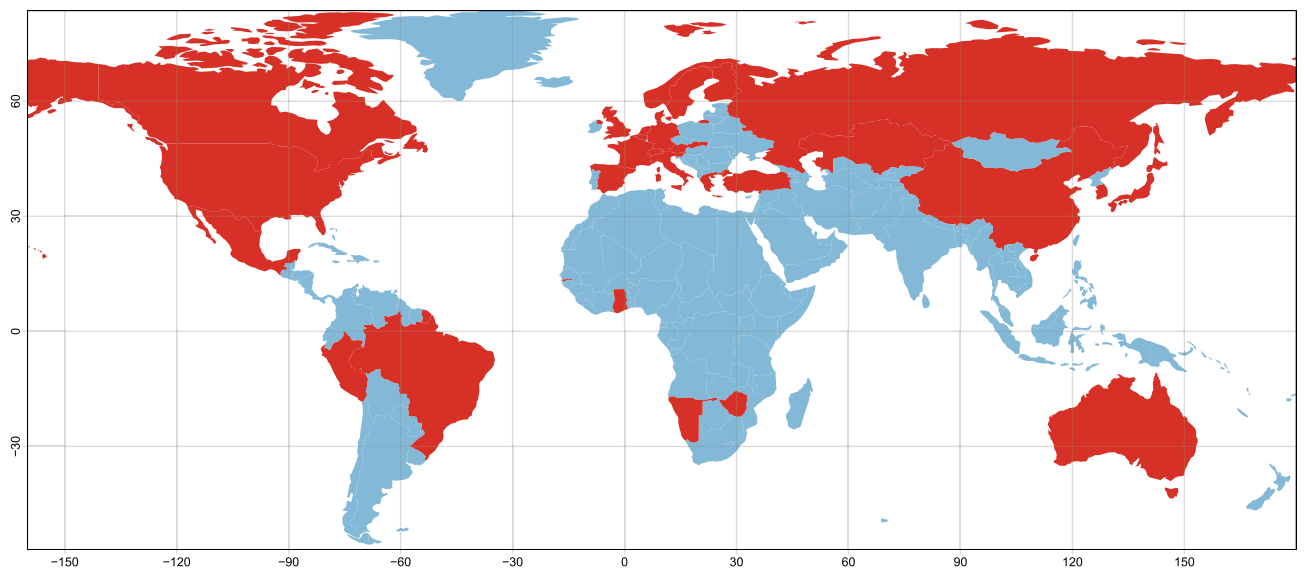


**Figure 1.** World map indicating the countries with at least one sample included in the analysis (red area). The countries with samples are the following: Australia, Austria, Belgium, Brazil, Canada, China, Denmark, Finland, France, Gambia, Germany, Ghana, Greece, Italy, Japan, Kazakhstan, Liechtenstein, Mexico, Namibia, Netherlands, Norway, Peru, Russia, Slovakia, Slovenia, South Korea, Spain, Sweden, Switzerland, Turkey, UK, USA, Zimbabwe. There were 142 samples without a designation on their country of origin.

| Species | ARG |
|---|---|
| *Bartonella henselae* | PC1 *β*-lactamase (*blaZ*) |
| *Bartonella quintana* | *APH(3')-IIIa*, *SAT-4* |
| *Coxiella burnetii* | *AcrE*, *AcrS*, *APH(3")-Ib*, *APH(6)-Id*, *ArmR*, *bacA*, *bcr-1*, *CARB-16*, *cpxA*, *emrA*, *emrB*, *emrK*, *emrY*, *ErmC*, *Escherichia coli AcrAB-TolC* with *MarR* mutations conferring resistance to ciprofloxacin and tetracycline, *E. coli EF-Tu* mutants conferring resistance to pulvomycin, *E. coli emrE*, *E. coli mdfA*, *E. coli soxR* with mutation conferring antibiotic resistance, *E. coli soxS* with mutation conferring antibiotic resistance, *E. coli UhpT* with mutation conferring resistance to fosfomycin, *evgA*, *FosA*, *gadW*, *gadX*, *marA*, *mdtE*, *mdtG*, *mdtP*, *MexI*, *OXA-488*, *Pseudomonas aeruginosa emrE*, *P. aeruginosa soxR*, *TEM-116*, *tet(44)*, *tet(H)*, *TolC*, *ugd* |
| *Ehrlichia ruminantium* | *APH(3")-Ib*, *APH(6)-Id* |
| *Francisella tularensis* subsp. *holarctica* | *AAC(6')-Ii*, *aad(6)*, *acrB*, *acrD*, *AcrE*, *AcrF*, *AcrS*, *APH(3')-IIIa*, *bacA*, *baeR*, *baeS*, *cpxA*, *efmA*, *emrA*, *emrB*, *emrK*, *emrY*, *Enterococcus faecium* chloramphenicol acetyltransferase, *eptA*, *ErmB*, *E. coli acrA*, *E. coli AcrAB-TolC* with *AcrR* mutation conferring resistance to ciprofloxacin, tetracycline, and ceftazidime, *E. coli AcrAB-TolC* with *MarR* mutations conferring resistance to ciprofloxacin and tetracycline, *E. coli ampC β*-lactamase, *E. coli EF-Tu* mutants conferring resistance to pulvomycin, *E. coli emrE*, *E. coli gyrA* conferring resistance to fluoroquinolones, *E. coli mdfA*, *E. coli soxR* with mutation conferring antibiotic resistance, *E. coli soxS* with mutation conferring antibiotic resistance, *evgA*, *FTU-1*, *gadW*, *gadX*, *kdpE*, *Klebsiella pneumoniae KpnH*, *marA*, *mdtA*, *mdtB*, *mdtC*, *mdtE*, *mdtF*, *mdtG*, *mdtH*, *mdtM*, *mdtN*, *mdtO*, *mdtP*, *msbA*, *msrC*, *PmrF*, *SAT-4*, *tet(L)*, *tet(M)*, *TolC*, *ugd*, *vanA*, *vanH* gene in *vanA* cluster, *vanR* gene in *vanA* cluster, *vanS* gene in *vanA* cluster, *vanX* gene in *vanA* cluster, *vanY* gene in *vanA* cluster, *vanZ* gene in *vanA* cluster, *YojI* |
| *Francisella tularensis* | *FTU-1* |
| *Francisella tularensis* subsp. *mediasiatica* | *FTU-1* |
| *Francisella tularensis* subsp. *novicida* | *FTU-1* |
| *Francisella tularensis* subsp. *tularensis* | *FTU-1*, *TEM-1*, *TEM-116* |
| *Rickettsia conorii* | *OXA-192* |
| *Rickettsia japonica* | *APH(3')-Ia* |

**Table 1.** Resistance genes detected for each species by RGI and CARD. In case of *Francisella tularensis*, the samples were listed at the subspecies level, where that information was available, otherwise, they were summarized at the species level.

| Genus | Number of negative samples | Number of positive samples | Number of all samples |
|---|---|---|---|
| *Anaplasma* | 20 | 0 | 20 |
| *Bartonella* | 129 | 2 | 131 |
| *Borrelia* | 311 | 0 | 311 |
| *Coxiella* | 65 | 8 | 73 |
| *Ehrlichia* | 12 | 1 | 13 |
| *Francisella* | 11 | 948 | 959 |
| *Francisella* (without *FTU-1*) | 803 | 156 | 959 |
| *Rickettsia* | 41 | 2 | 43 |

**Table 2.** ARG positive and negative samples by the genus of origin.

| Genus | Number of mobile ARGs | Number of non-mobile ARGs |
|---|---|---|
| *Anaplasma* | 0 | 0 |
| *Bartonella* | 1 | 2 |
| *Borrelia* | 0 | 0 |
| *Coxiella* | 24 | 31 |
| *Ehrlichia* | 2 | 0 |
| *Francisella* | 327 | 1359 |
| *Francisella* (without *FTU-1*) | 303 | 434 |
| *Rickettsia* | 2 | 0 |

**Table 3.** Number of mobile and immobile ARGs for each genus analysed.

immobile ARGs can be found for each genus in Table 3. Further details of mobile elements associated with each predicted ARG are presented in Supplementary File 2.

### Comparing species and environment of origin between ARG positive and negative samples.

Understanding the ecological environment where the samples included originated from is useful to gain a deeper insight into the potential factors underlying our results. The background of the samples, either positive or negative considering their ARG content, was assessed based on the origin of samples and species of the isolates. Samples were enrolled into origin categories based on their metadata (for details, see Materials and Methods). The ARG positive and negative sample distribution among the categories of origin are summarized in Table 4.

Interestingly, only samples originating from farm animals were positive to ARGs in the case of *Coxiella burnetii*, while such high dependence wasn't revealed for any other genus. To compare the association of positive samples to farm animal origins at *Coxiella* genus and all the remaining samples (excluding *Francisella*), Fisher's exact tests were performed on the contingency tables from Fig. 2. A significant difference was revealed in the case of *Coxiella* (OR: Inf, 95% CI 3.6–Inf, $p<0.001$), however, such differences could not have been found in case of the rest of the samples (OR: Inf, 95% CI 0.01–Inf, $p > 0.9999$).

The species of origin of ARG positive and negative samples are also of a key importance to understand the presence of genotypic resistance among microorganisms from different ecological environments. The number of ARG positive and ARG negative samples from each species can be found in Supplementary File 3.

| Genus | Companion animal | Environment | Farm animal | Human | Vector | Wild animal | NA |
|---|---|---|---|---|---|---|---|
| *Anaplasma* | 0/2 | 0/0 | 0/9 | 0/6 | 0/1 | 0/0 | 0/2 |
| *Bartonella* | 1/11 | 0/0 | 0/0 | 1/86 | 0/0 | 0/2 | 0/30 |
| *Borrelia* | 0/0 | 0/0 | 0/0 | 0/47 | 0/254 | 0/1 | 0/9 |
| *Coxiella* | 0/0 | 0/1 | 8/19 | 0/6 | 0/15 | 0/1 | 0/23 |
| *Ehrlichia* | 0/0 | 0/0 | 0/0 | 0/7 | 1/1 | 0/0 | 0/4 |
| *Francisella* | 4/0 | 8/0 | 0/0 | 112/1 | 37/1 | 314/2 | 473/7 |
| *Francisella* (without *FTU-1*) | 0/4 | 0/8 | 0/0 | 2/111 | 0/38 | 0/316 | 154/326 |
| *Rickettsia* | 0/0 | 0/0 | 0/0 | 1/16 | 1/20 | 0/0 | 0/5 |

**Table 4.** Number of ARG positive/negative samples in different origin categories.
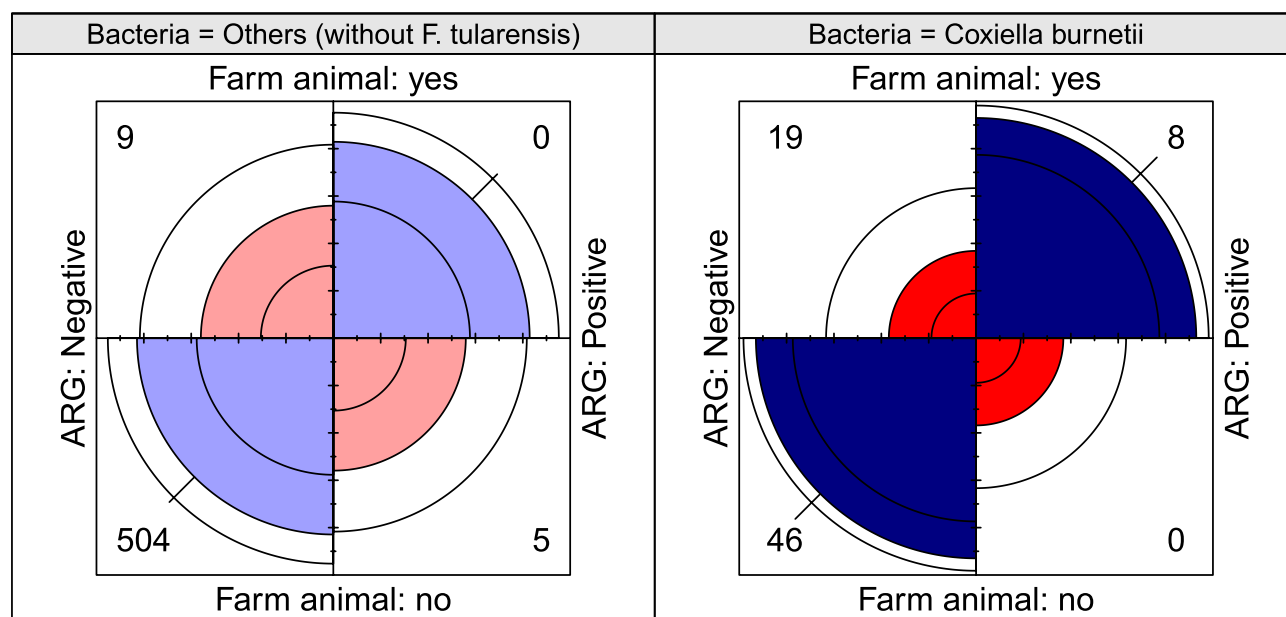


**Figure 2.** Fourfold displays showing the difference between ARG positivity and farm animal origin, in the case of *Coxiella burnetii* (right-hand side) and the rest of the samples (without *Francisella* genus, left-hand side). The plot is visualising the contingency tables for these categories[35]. Numbers in quadrants correspond to the number of observed cases in that category. Quarter circles are proportional to the ratio of the counts associated with the diagonal frequencies after standardisation. A significant imbalance is apparent in the case of *Coxiella* based on the 95% confidence bands around the quadrants, as the Fisher's exact tests showed ($p < 0.001$) also (see the Results section for details).

**Comparison between different methods.**    To understand the effect of the applied methods on our observations, we have complemented the analysis by including two additional methods. A similar approach to RGI is the NCBIs AMRFinderPlus pipeline which uses the NDARO database for ARG annotation. This tool utilizes a separate (the so-called plus database) for virulence-, metal- and biocide resistance genes. Even though CARD stores information regarding disinfectants, we have not used the plus genes in our analysis with AMRFinderPlus to keep as much consistency between the databases as possible. Furthermore, to compare the results from a pipeline with a fundamentally different approach, KmerResistance was also employed on the raw data with both the CARD and NDARO databases. For consistency to the analysis with AMRFinderPlus, plus genes were removed from the results of the KmerResistance analysis with the NDARO database. (Further details on the analysis pipelines are presented in the Materials and Methods section.).

Due to the inherent differences between databases and methods, their results are not directly comparable in a study like ours. However, to still augment the confidence of our findings, we compared the ARG positive samples among the four methods. Venn diagrams of ARG-positive samples from the so-called methodologies for each genus incorporated in the analysis can be found in Supplementary Fig. 1–8 (Supplementary File 4). Detailed ARG predictions from each method can be found in Supplementary Files 5-7. Results for *Anaplasma*, *Borrelia*, *Coxiella*, *Ehrlichia* and *Rickettsia* genera were in line by each employed method, with differences affecting only a few samples (Supplementary Fig. 1, 3-5 and 8 in Supplementary File 4). In the case of the genus *Bartonella*, KmerResistance with CARD and NDARO databases has predicted an additional 13 and 11 ARG positive samples compared to RGI, respectively (Supplementary Fig. 2 in Supplementary File 4). The short read-based approach predicted more positive samples in the case *Francisella* genus as well (when FTU-1 gene wasn't considered), where KmerResistance with CARD found 60 and KmerResistance with NDARO identified 34 ARG-positives more (Supplementary Fig. 7 in Supplementary File 4).

## Discussion

We have revealed a lack of ARGs in the available and high-quality WGS samples of human-pathogenic and potentially tick-borne bacteria. *Francisella* showed the highest ratio of samples with at least one ARG (16.3%), however, ARG-positivity was generally very low in the rest of the analysed genera (only 2.2 %). This finding is surprising at first, considering the vast amount of antibiotic pollution and ARG prevalence of the natural and artificial environments[3–7,36–38]. The effect of natural environments on the spread of antimicrobial resistance (AMR) is well represented by the fact that wild animals were found to harbour different ARGs[39,40], even though it might correlate with anthropogenic influence[41,42]. Furthermore, it is well known that antibiotic resistance is not only dependent on human activity as its origins root long before the onset of clinical antibiotic usage[43–45]. Considering this widespread occurrence of resistance, we believe that the potential factors influencing our findings deserve a closer look.

Our analysis has included data generated by second and third-generation sequencing technologies. Even though these methods have proven useful in numerous research areas, including antimicrobial resistance, they also have certain limitations. For example, if some genomic elements (e.g. ARGs) are absent from the sample, as in our case, only with sufficient sequencing depth can one conclude that true biological variation is the reason for the missingness[46]. To prevent such artefacts in our results, we have limited our analysis to samples with at least 95% coverage to their respective representative genome. Furthermore, we have used a 90% coverage and identity threshold for filtering the raw ARG predictions for limiting potential false positive calls. Another potential source of bias in genome-based resistance gene surveillance studies is that the results can only depend on the currently available resistance gene resources and tools, that might have many differences among them[47]. To mitigate these effects on our results, the samples included in our analysis were analysed with two additional tools and one additional database. NCBI's AMRFinderPlus tool with the NDARO database and the short read-based KmerResistance tool with both the CARD and NDARO databases were employed, besides the RGI with CARD database approach. These results indicate that RGI was highly concordant with the additional methods used for the analysis, even though KmerResistance has predicted more positive samples for the *Bartonella* and *Francisella* genera. It can be assumed that short-read technologies are more sensitive, which could explain why more positive samples were found with KmerResistance; however, these methods have not been systematically compared yet. Furthermore, it is worth noting that many of the analysed genera are intracellular bacteria, which might enhance the risk of contamination during sample collection and isolation procedures. Even tough it might not be possible to exactly determine if a gene is originating from contamination, it would further highlight the lack of ARGs in the analysed bacteria.

Phenotypic resistance to various antibiotics has been described in the bacteria we have studied, but to the best of our knowledge, multi-drug resistant strains have not yet been found. Interestingly, the first choice for treatment of infections caused by the the majority of bacteria included in our study is doxycycline[14,48–56]. For *F. tularensis* aminoglycosides and fluoroquinolones are also commonly recommended[57]. Resistance to fluoroquinolone antibiotics has been identified in *Bartonella*, *Borrelia* and *Ehrlichia* species[50,55,56]. Nevertheless, fluoroquinolone resistance in *Ehrlichia* spp. is probably due to intrinsic effects[58]. In *Ehrlichia*, resistance to macrolides, aminoglycosides and certain β-lactams has also been detected[50,59], which may be consistent with the aminoglycoside resistance genes we found in *E. ruminantium* (Table 1). In contrast to the lack of ARGs in *Borrelia* species involved in our study, β-lactam and rifampicin resistance have already been described for them[54,55] (Table 1). However, it should be noted that using the two KmerResistance-based approaches, we did find an aminoglycoside resistance gene and an efflux pump in this group (*APH(3')-Ia*, *E. coli emrE*, Supplementary Files 6, 7). *C. burnetii* and *F. tularensis* samples were found to encode a wide range of resistance genes in our analysis (Table 1). While phenotypically manifested ciprofloxacin and chloramphenicol resistance have already been described in *C. burnetii*[52], furthermore, resistance to penicillin has also been described[60], several strains appear

to show susceptibility to ampicillin[52]. Moreover, in *F. tularensis*, resistance to *β*-lactam, macrolides, linezolid and clindamycin was described previously[61–63].

However, it is important to note that the resistance gene content of the genome does not necessarily correlate with phenotypic resistance[64], and in any case, our results will be limited by the capacity of databases and tools[47,65–67]. Furthermore, there are many factors affecting the in vivo susceptibility of a microorganism, compared to what was discovered in vitro[56,68]. In addition, it should be noted that of the 156 ARG positive *F. tularensis* samples in our analysis (when the *FTU-1* gene was excluded), 153 were from BioProject PRJNA669398. In this BioProject, samples derived from the live vaccine strain of *F. tularensis*subsp. holartica (LVS) and were tested for the presence of evolutionary mutations under the pressure of doxycycline and ciprofloxacin antibiotics[69]. However, it is interesting to note that of the strains included in our study and identified as LVS based on the available metadata, 43.7% (153/350) of the strains belonging to the BioProject PRJNA669398 were found to be ARG positive (without *FTU-1*), while 0% (0/49) of the LVS strains included in other BioProjects were positive. These aspects raise the question of the reliability of the samples from this BioProject.

The general consensus is that de novo resistance develops in response to selective pressure from antibiotics or potentially other stressors (e.g. heavy metals), which is then associated with resistance mutations or the uptake of acquired ARGs via horizontal gene transfer[70–72]. Consequently, it is important to consider the ecological environment of microorganisms to investigate the underlying effects of the resistance gene deficiency we have observed. Most microorganisms we studied are obligate intracellular bacteria, strictly bound to vectors and hosts[49,56,73–79]. Exceptions are *Borrelia* species, which despite being strongly associated with hosts- and vectors are not intracellular pathogens[80], and the *Francisella* and *Coxiella* genera, where transmission pathways outside of vectors are also important[81,82]. Consequently, there is a limited medium in which selective pressure and the presence of ARGs would co-occur. A tick or possibly another arthropod vector could theoretically be subjected to more or less antibiotic exposure in its environment, for example, when it takes up fluid from plants or soil, but the extent to which this effect could represent a true selective force is questionable. In addition, the presence of available ARGs might also be limited, as it has been described that the ticks' own bacteriota limits the colonisation of microorganisms from the environment[83].

Bacteria in hosts can be subjected to more pronounced antibiotic pressure. Of course, this is not primarily in wild animals, but when bacteria are introduced into domestic animals or even humans. It is assumed that a significant proportion of human infections are treated with antibiotics, and even if this were to involve the uptake of ARGs of these strains, it is unlikely that the bacteria that have acquired resistance genes would be re-introduced into a tick. Domestic animals may be of much greater importance in this respect. In their case, antibiotic treatment is also significant, or perhaps even more significant, while a vector may also go undetected more easily. Farm animals may be particularly noteworthy in this respect. However, it is interesting that we found a strong concordance between ARG positivity and the farm animal origin of the samples only in the case of *C. burnetii* (Fig. 2). One could hypothesize that the majority of *C. burnetii* samples from farm animals were from intensive keeping conditions, whereas the opposite was true for samples from other genera. Cattle can be considered as the classic intensively kept animal in contrast to goats, sheep and horses. For *C. burnetii* almost half of the 27 samples from farm animals were derived from cattle (14), of which 5 were positive, and there were 3 additional positive samples from goats (Supplementary Files 1, 2). It is important to note, however, that one positive sample in the genus *Ehrlichia* was from a tick collected directly from a sheep (BioSample: SAMN04335506, Supplementary File 1). Nevertheless, this was classified of vector origin in our analysis as it is not known if the tick was infected a priori. It should also be noted that the *Anaplasma* genus also included 9 samples of farm animal origin, 6 of which were of cattle origin, 1 of equine origin and 2 of sheep origin, but none of these was found to be positive (Supplementary Files 1, 2). Our reasoning is, of course, limited by the availability of metadata, but it does not appear that the confounding effect of cattle, as traditionally intensively kept animals, is the cause of the phenomenon observed in the case of *Coxiella*. A difference between the *Anaplasma* and *Coxiella* genera may be due to their life cycle. Namely, the extracellular form of *Coxiella* also plays an important role in its spread, and consequently, it is not entirely bounded to vectors[81]. Of course, the ability of the environmental form of *C. burnetii* to take up genes from the environment and the extent to which these might be necessary for its survival in the presence of antibiotics is questionable. Furthermore, it is important to note that our analysis is limited by the bias that the use of public data might cause and further research is needed to determine the generalizability of these findings.

*Francisella tularensis* is similar to *C. burnetii* in some respects. The role of vectors in its spread is not necessarily exclusive[82], and it can occur in the infected hosts in an extracellular state[84,85]. Several ARG positive samples of *F. tularensis* were found, but these were almost exclusively from an experiment in which artificial antibiotic pressure was applied to the bacteria, making these results difficult to interpret (see above). Furthermore, it is, of course, possible that the microorganisms studied do acquire resistance genes. However, these were impossible to detect because they were dropped by the bacteria. This may be because it is assumed that the tick's gut provides a fundamentally different microenvironment for the bacteria than the circulation of the host animal, including the existing selection pressure. The continuous activation of resistance mechanisms is an energy-demanding process, so without therapeutic pressure, these costs exceed the benefits. Thus, the bacteria try to get rid of them, which may explain the lack of ARGs. This effect has been described for ARGs; however, the fitness cost might depend on the resistance mechanism[70,86,87].

Our results show that the potentially tick-borne and human pathogenic microorganism strains we investigated lack antimicrobial resistance genes. Comprehending the processes underlying this phenomenon may be an important aspect in understanding the ecology of these species and, through this, in assessing the risk of phenotypic antibiotic resistance in clinical settings. However, an exception to the above statement of the tick-borne human pathogens is *Coxiella burnetii*, where resistance is a potential veterinary medical and public health threat when farm animals and their products are considered.

## Data availability

Datasets analysed in the present study are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra). For more information on the BioSample IDs and the corresponding BioProject of each analysed samples, please see the Supplementary File 8.

## References

1. Chala, B. & Hamde, F. Emerging and re-emerging vector-borne infectious diseases and the challenges for control: A review. *Front. Public Health* **9**, 715759 (2021).
2. Nolte, O. Antimicrobial resistance in the 21st century: A multifaceted challenge. *Protein Peptide Lett.* **21**, 330–335 (2014).
3. Tóth, A. G. *et al.* Antimicrobial resistance genes in raw milk for human consumption. *Sci. Rep.* **10**, 7464 (2020).
4. Tóth, A. G. *et al.* A glimpse of antimicrobial resistance gene diversity in kefir and yoghurt. *Sci. Rep.* **10**, 1–12 (2020).
5. Tóth, A. G. *et al.* Mobile antimicrobial resistance genes in probiotics. *Antibiotics* **10**, 1287 (2021).
6. Nagy, S. Á., Tóth, A. G., Papp, M., Kaplan, S. & Solymosi, N. Antimicrobial resistance determinants in silage. *Sci. Rep.* **12**, 1–10 (2022).
7. Tóth, A. G. *et al.* Canine saliva as a possible source of antimicrobial resistance genes. *Antibiotics* **11**, 1490 (2022).
8. Tóth, A. G., Judge, M. F., Nagy, S. Á., Papp, M. & Solymosi, N. A survey on antimicrobial resistance genes of frequently used probiotic bacteria, 1901 to 2022. *EuroSurveillance* **28**, 2200272 (2023).
9. Organization, W. H., UNICEF et al. Global vector control response 2017–2030. (2017).
10. Brites-Neto, J., Duarte, K. M. R. & Martins, T. F. Tick-borne infections in human and animal population worldwide. *Vet. World* **8**, 301 (2015).
11. Schneider, J. G. Human ehrlichiosis: A case study. *Clin. Lab. Sci.* **22**, 3 (2009).
12. Billeter, S. A., Cáceres, A. G., Gonzales-Hidalgo, J., Luna-Caypo, D. & Kosoy, M. Y. Molecular detection of *Bartonella* species in ticks from Peru. *J. Med. Entomol.* **48**, 1257–1260 (2011).
13. Duan, C. et al. Complete genome sequence of *rickettsia* heilongjiangensis, an emerging tick-transmitted human pathogen (2011).
14. Bakken, J. S. & Dumler, J. S. Human granulocytic anaplasmosis. *Infect. Dis. Clin.* **29**, 341–355 (2015).
15. Bush, L. M. & Vazquez-Pertejo, M. T. Tick borne illness-lyme disease. *Dis. Month* **64**, 195–212 (2018).
16. Koka, H., Sang, R., Kutima, H. L. & Musila, L. *Coxiella burnetii* detected in tick samples from pastoral communities in kenya. *BioMed Res. Int.* **2018** (2018).
17. Yeni, D. K., Büyük, F., Ashraf, A. & Shah, M. Tularemia: A re-emerging tick-borne infectious disease. *Folia Microbiologica* **66**, 1–14 (2021).
18. De Coster, W., D'hert, S., Schultz, D. T., Cruts, M. & Van Broeckhoven, C. Nanopack: Visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666–2669 (2018).
19. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
20. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).
21. Vaser, R., Sović, I., Nagarajan, N. & Šikić, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **27**, 737–746 (2017).
22. Zimin, A. V. *et al.* Hybrid assembly of the large and highly repetitive genome of aegilops tauschii, a progenitor of bread wheat, with the masurca mega-reads algorithm. *Genome Res.* **27**, 787–792 (2017).
23. Hyatt, D. *et al.* Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* **11**, 119 (2010).
24. McArthur, A. G. *et al.* The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* **57**, 3348–3357 (2013).
25. Jia, B. *et al.* CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017).
26. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
27. Krawczyk, P. S., Lipinski, L. & Dziembowski, A. PlasFlow: Predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res.* **46**, e35 (2018).
28. Johansson, M. H. *et al.* Detection of mobile genetic elements associated with antibiotic resistance in Salmonella enterica using a newly developed web tool: MobileElementFinder. *J. Antimicrob. Chemother.* **76**, 101–109 (2021).
29. Guo, J. *et al.* VirSorter2: A multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* **9**, 37 (2021).
30. Feldgarden, M. *et al.* Amrfinderplus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* **11**, 1–9 (2021).
31. Clausen, P. T., Zankari, E., Aarestrup, F. M. & Lund, O. Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data. *J. Antimicrob. Chemother.* **71**, 2484–2488 (2016).
32. Clausen, P. T., Aarestrup, F. M. & Lund, O. Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinform.* **19**, 1–8 (2018).
33. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2021).
34. Antunes, N. T., Frase, H., Toth, M. & Vakulenko, S. B. The class A $\beta$-lactamase FTU-1 is native to *Francisella tularensis*. *Antimicrob. Agents Chemother.* **56**, 666–671 (2012).
35. Friendly, M. A fourfold display for 2 by 2 by k tables. Tech. Rep., Technical Report 217, Psychology Department, York University (1994).
36. Cycoń, M., Mrozik, A. & Piotrowska-Seget, Z. Antibiotics in the soil environment-degradation and their impact on microbial activity and diversity. *Front. Microbiol.* **10**, 338 (2019).
37. Zainab, S. M., Junaid, M., Xu, N. & Malik, R. N. Antibiotics and antibiotic resistant genes (args) in groundwater: A global review on dissemination, sources, interactions, environmental and human health risks. *Water Res.* **187**, 116455 (2020).
38. Kim, D.-W. & Cha, C.-J. Antibiotic resistome from the one-health perspective: Understanding and controlling antimicrobial resistance transmission. *Expe. Mol. Med.* **53**, 301–309 (2021).
39. Rogers, S. W., Shaffer, C. E., Langen, T. A., Jahne, M. & Welsh, R. Antibiotic-resistant genes and pathogens shed by wild deer correlate with land application of residuals. *EcoHealth* **15**, 409–425 (2018).
40. Dias, D., Fonseca, C., Mendo, S. & Caetano, T. A closer look on the variety and abundance of the faecal resistome of wild boar. *Environ. Pollut.* **292**, 118406 (2022).
41. Allen, H. K. *et al.* Call of the wild: Antibiotic resistance genes in natural environments. *Nat. Rev. Microbiol.* **8**, 251–259 (2010).
42. Sacristán, I. *et al.* Antibiotic resistance genes as landscape anthropization indicators: Using a wild felid as sentinel in Chile. *Sci. Total Environ.* **703**, 134900 (2020).

43. Finley, R. L. *et al.* The scourge of antibiotic resistance: The important role of the environment. *Clin. Infect. Dis.* **57**, 704–710 (2013).
44. D'Costa, V. M., Griffiths, E. & Wright, G. D. Expanding the soil antibiotic resistome: Exploring environmental diversity. *Curr. Opin. Microbiol.* **10**, 481–489 (2007).
45. D'Costa, V. M. *et al.* Antibiotic resistance is ancient. *Nature* **477**, 457–461 (2011).
46. Sims, D., Sudbery, I., Ilott, N. E., Heger, A. & Ponting, C. P. Sequencing depth and coverage: Key considerations in genomic analyses. *Nat. Rev. Genet.* **15**, 121–132 (2014).
47. Papp, M. & Solymosi, N. Review and comparison of antimicrobial resistance gene databases. *Antibiotics* **11**, 339 (2022).
48. Matei, I. A. *et al.* A review on the eco-epidemiology and clinical management of human granulocytic anaplasmosis and its agent in Europe. *Parasites & Vectors* **12**, 1–19 (2019).
49. Atif, F. A. *Anaplasma* marginale and *anaplasma* phagocytophilum: Rickettsiales pathogens of veterinary and public health significance. *Parasitol. Res.* **114**, 3941–3957 (2015).
50. Ismail, N., Bloch, K. C. & McBride, J. W. Human ehrlichiosis and anaplasmosis. *Clin. Lab. Med.* **30**, 261–292 (2010).
51. Botelho-Nevers, E., Socolovschi, C., Raoult, D. & Parola, P. Treatment of *Rickettsia* spp. infections: A review. *Expert Rev. Anti-infective Ther.* **10**, 1425–1437 (2012).
52. Brennan, R. E. & Samuel, J. E. Evaluation of *Coxiella burnetii* antibiotic susceptibilities by real-time PCR assay. *J. Clin. Microbiol.* **41**, 1869–1874 (2003).
53. Kersh, G. J. Antimicrobial therapies for Q fever. *Expert Rev. Anti-infective Ther.* **11**, 1207–1214 (2013).
54. Koetsveld, J. *et al.* In vitro antimicrobial susceptibility of clinical isolates of *Borrelia miyamotoi. Antimicrob. Agents Chemother* **62**, e00419-18 (2018).
55. Stanek, G., Wormser, G. P., Gray, J. & Strle, F. Lyme borreliosis. *The Lancet* **379**, 461–473 (2012).
56. Biswas, S. & Rolain, J.-M. *Bartonella* infection: Treatment and drug resistance. *Futur. Microbiol.* **5**, 1719–1731 (2010).
57. Caspar, Y. & Maurin, M. *Francisella tularensis* susceptibility to antibiotics: A comprehensive review of the data obtained in vitro and in animal models. *Front. Cell. Infect. Microbiol.* **7**, 122 (2017).
58. Maurin, M., Abergel, C. & Raoult, D. DNA gyrase-mediated natural resistance to fluoroquinolones in Ehrlichia spp. *Antimicrob. Agents Chemother.* **45**, 2098–2105 (2001).
59. Branger, S., Rolain, J. & Raoult, D. Evaluation of antibiotic susceptibilities of Ehrlichia canis, *Ehrlichia chaffeensis*, and *Anaplasma* phagocytophilum by real-time PCR. *Antimicrob. Agents Chemother.* **48**, 4822–4828 (2004).
60. Raoult, D., Torres, H. & Drancourt, M. Shell-vial assay: Evaluation of a new technique for determining antibiotic susceptibility, tested in 13 isolates of *Coxiella burnetii. Antimicrob. Agents Chemother.* **35**, 2070–2077 (1991).
61. Caspar, Y., Hennebique, A. & Maurin, M. Antibiotic susceptibility of *Francisella tularensis* subsp. holarctica strains isolated from tularaemia patients in France between 2006 and 2016. *J. Antimicrob. Chemother.* **73**, 687–691 (2018).
62. Heine, H. S., Miller, L., Halasohoris, S. & Purcell, B. K. In vitro antibiotic susceptibilities of *Francisella tularensis* determined by broth microdilution following CLSI methods. *Antimicrob. Agents Chemother.* **61**, e00612-17 (2017).
63. Kreizinger, Z. *et al.* Antimicrobial susceptibility of *Francisella tularensis* subsp. holarctica strains from Hungary, Central Europe. *J. Antimicrob. Chemother.* **68**, 370–373 (2013).
64. Dantas, G. & Sommer, M. O. Context matters-the complex interplay between resistome genotypes and resistance phenotypes. *Curr. Opin. Microbiol.* **15**, 577–582 (2012).
65. McArthur, A. G. & Tsang, K. K. Antimicrobial resistance surveillance in the genomic age. *Ann. N. Y. Acad. Sci* **1388**, 78–91 (2017).
66. Boolchandani, M., D'Souza, A. W. & Dantas, G. Sequencing-based methods and resources to study antimicrobial resistance. *Nat. Rev. Genet.* **20**, 356–370 (2019).
67. Pawlowski, A. C. *et al.* A diverse intrinsic antibiotic resistome from a cave bacterium. *Nat. Commun.* **7**, 1–10 (2016).
68. Stratton, C. W. In vitro susceptibility testing versus in vivo effectiveness. *Med. Clin.* **90**, 1077–1088 (2006).
69. Mehta, H. H., Ibarra, D., Marx, C. J., Miller, C. R. & Shamoo, Y. Mutational switch-backs can accelerate evolution of *Francisella* to a combination of ciprofloxacin and doxycycline. Front. Microbiol. **13** (2022).
70. Aminov, R. I. & Mackie, R. I. Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol. Lett.* **271**, 147–161 (2007).
71. Oz, T. *et al.* Strength of selection pressure is an important parameter contributing to the complexity of antibiotic resistance evolution. *Mol. Biol. Evol.* **31**, 2387–2401 (2014).
72. Martínez, J. L., Coque, T. M. & Baquero, F. What is a resistance gene? Ranking risk in resistomes. *Nat. Rev. Microbiol.* **13**, 116–123 (2015).
73. Battilani, M., De Arcangeli, S., Balboni, A. & Dondi, F. Genetic diversity and molecular epidemiology of *Anaplasma. Infect. Genet. Evol.* **49**, 195–211 (2017).
74. Rikihisa, Y. *Anaplasma* phagocytophilum and *Ehrlichia chaffeensis*: Subversive manipulators of host cells. *Nat. Rev. Microbiol.* **8**, 328–339 (2010).
75. Paddock, C. D. & Childs, J. E. *Ehrlichia chaffeensis*: A prototypical emerging pathogen. *Clin. Microbiol. Rev.* **16**, 37–64 (2003).
76. Cutler, S. *et al.* A new *Borrelia* on the block: *Borrelia miyamotoi*-a human health risk?. *Eurosurveillance* **24**, 1800170 (2019).
77. Wagemakers, A., Staarink, P. J., Sprong, H. & Hovius, J. W. *Borrelia miyamotoi*: A widespread tick-borne relapsing fever spirochete. *Trends Parasitol.* **31**, 260–269 (2015).
78. Satjanadumrong, J., Robinson, M. T., Hughes, T. & Blacksell, S. D. Distribution and ecological drivers of spotted fever group *Rickettsia* in asia. *Ecohealth* **16**, 611–626 (2019).
79. Tomassone, L., Portillo, A., Nováková, M., De Sousa, R. & Oteo, J. A. Neglected aspects of tick-borne rickettsioses. *Parasites & Vectors* **11**, 1–11 (2018).
80. Schotthoefer, A. M. & Frost, H. M. Ecology and epidemiology of Lyme borreliosis. *Clin. Lab. Med.* **35**, 723–743 (2015).
81. González-Barrio, D. & Ruiz-Fons, F. *Coxiella burnetii* in wild mammals: A systematic review. *Transbound. Emerg. Dis.* **66**, 662–671 (2019).
82. Telford, S. R. III. & Goethert, H. K. Ecology of *Francisella tularensis. Annu. Rev. Entomol.* **65**, 351 (2020).
83. Hayes, B. M. *et al.* Ticks resist skin commensals with immune factor of bacterial origin. *Cell* **183**, 1562–1571 (2020).
84. Forestal, C. A. *et al. Francisella tularensis* has a significant extracellular phase in infected mice. *J. Infect. Dis.* **196**, 134–137 (2007).
85. Yu, J.-J. *et al.* The presence of infectious extracellular *Francisella tularensis* subsp. novicida in murine plasma after pulmonary challenge. *Eur. J. Clin. Microbiol. Infect. Dis.* **27**, 323–325 (2008).
86. Durão, P., Balbontín, R. & Gordo, I. Evolutionary mechanisms shaping the maintenance of antibiotic resistance. *Trends Microbiol.* **26**, 677–691 (2018).
87. Martinez, J. L. General principles of antibiotic resistance in bacteria. *Drug Discov. Today Technol.* **11**, 33–39 (2014).

## Acknowledgements

## Author contributions

N.S. takes responsibility for the data's integrity and the data analysis's accuracy. M.P. and N.S. conceived the concept of the study. M.P., N.S. and S.Á.N. participated in the bioinformatic analysis. M.P. and N.S. participated

in the drafting of the manuscript. A.G.T., GV, L.M., MP, N.S., R.F. and S.Á.N. carried out the critical revision of the manuscript for important intellectual content. All authors read and approved the final manuscript.

### Competing interests
The authors declare no competing interests.

### Additional information
**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-35356-5.

**Correspondence** and requests for materials should be addressed to N.S.

**Reprints and permissions information** is available at www.nature.com/reprints.

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