npj | biofilms and microbiomes

Article

Published in partnership with Nanyang Technological University

https://doi.org/10.1038/s41522-026-00986-w

Tick-vectored mobilization of antibiotic resistance genes: transboundary dissemination across wildlife-livestock-vector-environment interfaces

Zhenhua Lu[1,3], Ruishan Li[1,3], Kaichun Zhou[2], Shiyu Li[1], Shiwei Sun[1], Jiacheng Liu[1], Lele Zhao[1], Shijie Chen[1], Kun Liu[1] ✉, Xiang Yuan[2] ✉ & Zhongjun Shao[1] ✉

Antibiotic resistance genes (ARGs) are emerging as critical environmental contaminants across diverse ecological interfaces. To dissect evidence of microbiome and resistome in the different interconnected interfaces of ecotone, we conducted a field investigation of the microbiome and resistome of marmots, along with coexisting domestic sheep, ticks and their cave soils within the same ecological habitat. We used shotgun metagenomics with metagenome-assembled genomes (MAGs), species-resolved binning, ARG identification, source-tracker analyses, and horizontal gene transfer (HGT) network analysis to examine potential cross-interface dissemination. The composition of the mammalian gut microbiome was primarily comprised of *Firmicutes*, while ticks and soils exhibited distinct clusters that were predominantly dominated by *Proteobacteria*. The observed resistance mechanisms manifested niche-specific patterns, with target alteration predominating in mammals, whereas ticks exhibited elevated antibiotic inactivation/efflux strategies, and soils prioritized efflux mechanisms. Metagenomic assembly from these four groups yielded 5339 metagenome-assembled genomes (MAGs), of which 1481 met medium- or high-quality standards. Ticks exhibited 72% species similarity and 52% ARG concordance with marmots, while soils conserved 32% ARGs and >86% toxin genes with mammals. Our findings demonstrate that the transboundary dissemination of ARGs across different ecological interfaces, necessitates integrated surveillance of antimicrobial resistance at ecological boundaries to mitigate public health risks.

Antimicrobial resistance (AMR) has emerged as one of the most pressing global health challenges of the 21st century[1], with projections estimating up to 1.91 million annual deaths by 2050 if current trends continue[2]. This silent pandemic, driven by microbial evolution through genetic mutation, gene duplication, and horizontal gene transfer (HGT)[3–5], has made once-treatable infections increasingly lethal, threatening advancements in modern medicine, agriculture, and ecosystem stability[6–8]. While intrinsic bacterial resistance mechanisms are well documented[9], the ecological networks that facilitate the circulation of antibiotic resistance genes (ARGs) across interconnected biological systems remain poorly mapped—a critical oversight given the complex interactions among human, animal, and environmental health[10].

Transmission of AMR extends far beyond clinical settings and thrives at the intersection of wildlife, livestock, vectors, and environmental reservoirs[11]. Ticks, as obligate blood-feeding ectoparasites, epitomize this interconnectedness. Occupying diverse ecological niches, they parasitize wildlife, livestock, and humans, while their microbiota—a dynamic consortium of symbiotic and pathogenic microbes—serves as both a reservoir and a conduit for ARGs[12]. Recent metagenomic evidence reveals that ticks harbor multidrug-resistant bacteria[13], yet their role in spreading resistance

[1]Department of Epidemiology, The Ministry of Education Key Lab of Hazard Assessment and Control in Special Operational Environment, The Shaanxi Provincial Key Laboratory of Environmental Health Hazard Assessment and Protection, School of Public Health, The Fourth Military Medical University, Xi'an, China. [2]Wuwei City Center for Disease Control and Prevention, Wuwei, China. [3]These authors contributed equally: Zhenhua Lu, Ruishan Li. ✉e-mail: liukun5959@fmmu.edu.cn; 15168008@qq.com; shaozhj@fmmu.edu.cn

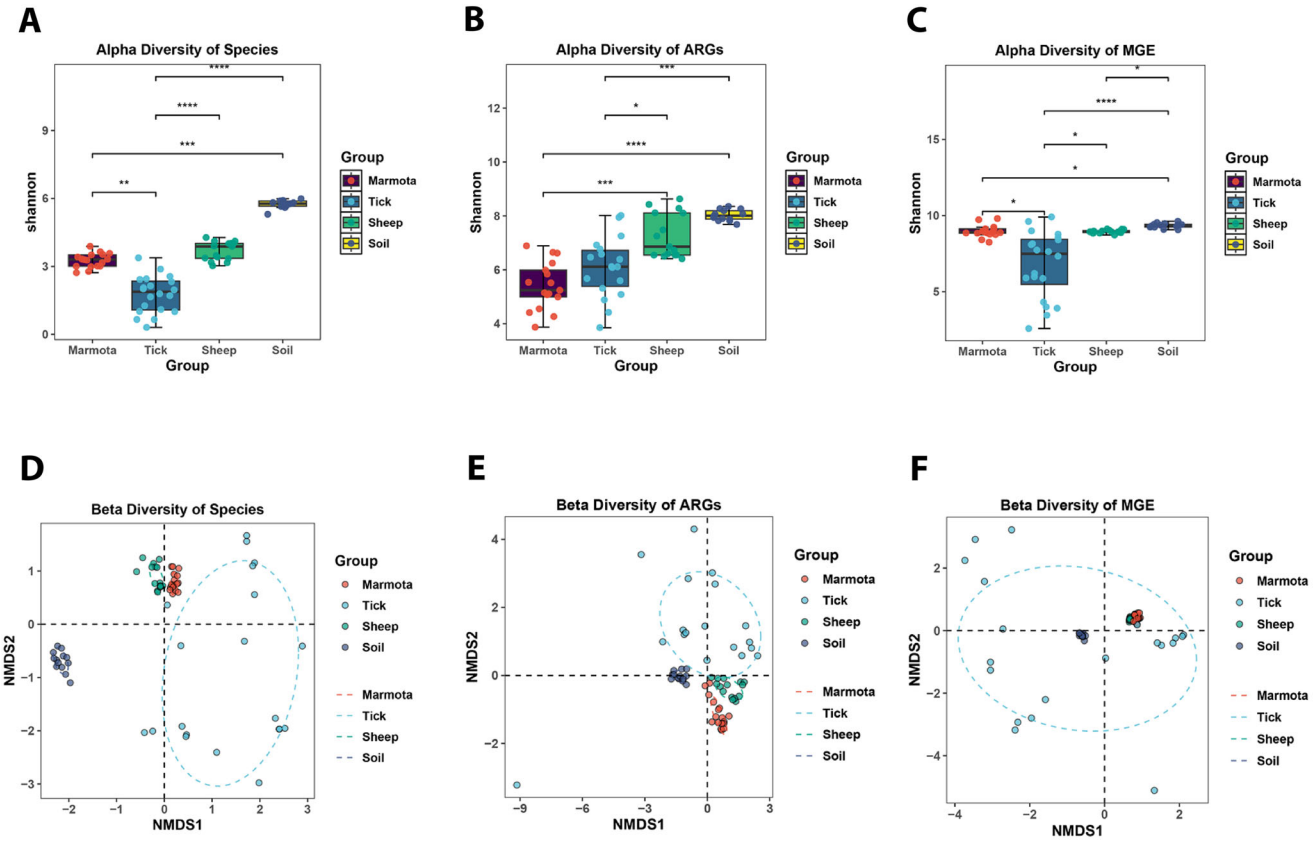

Fig. 1 | The diversity of species, ARGs, and MGE in different interfaces. **A**–**C** Alpha diversity of species, ARGs, and MGE in different interfaces (based on Shannon). **D**–**F** Beta diversity of species, ARGs, and MGE in different interfaces (based on Bray–Curtis distance); $^{**}p < 0.01$, $^{***}p < 0.001$. Significant difference of Beta diversity was calculated by NMDS using the Bray–Curtis Index of four groups.

across host species and ecosystems remains understudied. This gap reflects broader uncertainties in understanding how environmental reservoirs, such as antibiotic-contaminated soils and wildlife with minimal anthropogenic exposure, contribute to the global resistance—the collective ARG repertoire within microbial communities.

These complexities underscore the importance of the One Health framework, which acknowledges human health as inextricably linked to animal and environmental well-being. By considering the dynamic relationships between these domains, the framework provides a comprehensive approach to understanding the transmission of health risks, including those related to AMR, across different ecological interfaces[14]. Soils, for example, act as historical archives of resistance genes. Millennia of microbial warfare have forged ARGs, which are now being mobilized by anthropogenic pollution[15]. Despite lacking direct exposure to antibiotics, wildlife species harbor resistomes that are comparable to those of intensively farmed livestock, suggesting cryptic pathways for the acquisition of ARGs in the environment[16–18]. Even companion animals that share homes with humans participate in this network, blurring the boundaries between clinical and ecological transmission routes[19,20]. Within this intricate web, vectors like ticks may serve as unrecognized bridges, transferring ARGs between species through HGT during blood meals—a hypothesis supported by fragmented evidence but requiring systematic validation.

Advancements in next-generation sequencing (NGS) have revolutionized our capacity to dissect these interactions. Metagenomics, by circumventing cultivation biases, enables comprehensive profiling of unculturable microbial communities, revealing novel ARGs, mobile genetic elements (MGEs), and HGT hotspots across diverse matrices[21]. The Qilian Mountain National Nature Reserve represents a substantial marginal mountain range, situated in the northeast of the Qinghai-Tibet Plateau. As an important ecotone, it is a biodiversity hotspot and an emerging anthropic zone. This study investigates the ecological patterns of ARGs dissemination across the interfaces of wildlife-livestock-vector-environment in the Ecotone of Qinghai-Tibet Plateau margin. Using metagenomic approaches, we characterized resistomes from host-associated fecal samples, environmental soils, and ticks to evaluate cross-species transmission dynamics. By reconstructing HGT networks and resistome mechanisms, we identified how multi-host connectivity and habitat fragmentation jointly govern resistance propagation at ecological boundaries.

Results

Host-associated divergence in microbiome composition

The identity of *Marmota himalayana* was established via morphological and molecular phylogeny, alongside the concurrent taxonomic confirmation of *Ixodes canisuga* (Supplementary Fig. 4). Subsequent to rigorous host sequence removal and quality filtering, the de novo assembled contigs (>300 bp in length) were subjected to taxonomic annotation using Kraken2-derived taxonomic classifications with Bracken-based abundance calibration. Subsequent computational comparisons revealed statistically significant disparities ($p < 0.01$) in both Shannon and Chao1 richness indices across four groups (Figs. 1A and S1A). Notably, the Soil group exhibited the highest alpha diversity, while the Tick group showed the lowest microbial richness and evenness, suggesting a highly specialized microbial community within the vector. The final NMDS configuration ($k = 3$ dimensions) achieved a stress value of 0.08 (Fig. 1D), indicating adequate ordination fidelity according to Clarke's stress interpretation guidelines (stress < 0.2 means meaningful interpretation). The NMDS plot demonstrated clear sample-type-based clustering, wherein the proximity between the Marmota and Sheep groups indicated a shared microbial profile that contrasted with the more isolated Soil and Tick samples.

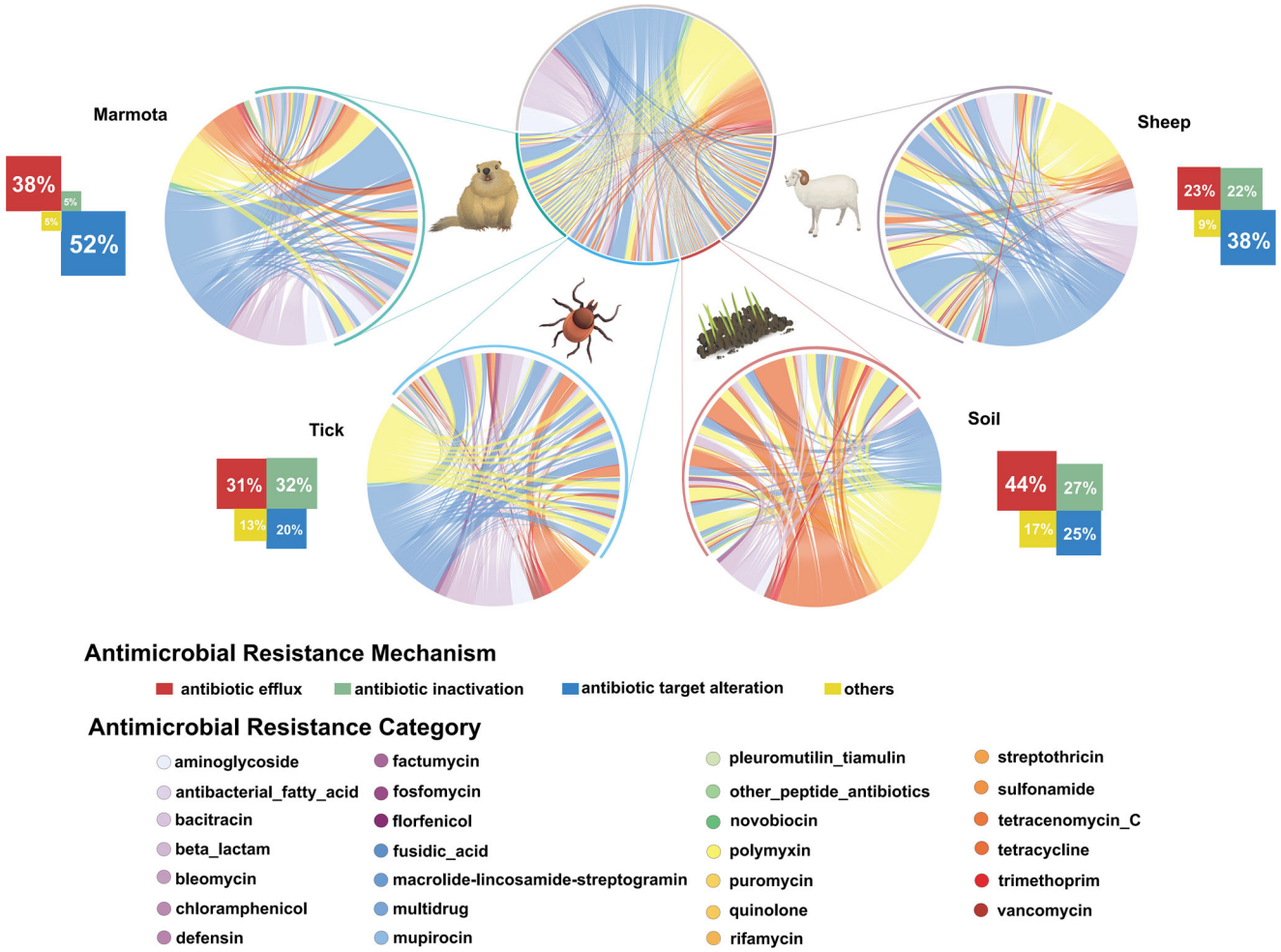

Fig. 2 | The patterns of resistome in different interfaces. The proportion area diagram represents the overall proportion of the AMR mechanism in each group. The chord diagram shows the overall and group correspondence of the AMR category.

Taxonomic analysis at the phylum level (Supplementary Fig. 1B) revealed conserved similar profiles between mammalian cohorts (Marmot and sheep groups), demonstrating a tripartite dominance hierarchy of *Firmicutes* (42.4% and 43.2%), *Proteobacteria* (28.3% and 27.2%), and *Bacteroidetes* (20.0% and 12.6%). The analysis revealed striking *Proteobacteria* dominance in both tick-associated microbiota (62.3%) and cave soil microbiomes (64.6%), while the composition of other taxonomic groups at the phylum level exhibited significant divergence. At the family taxonomic level (Supplementary Fig. 1C), *Enterobacteriaceae* maintained predominance in both mammalian digestive ecosystems (21.7% marmot vs 17.0% sheep). Tick-associated microbiota exhibited *Moraxellaceae* dominance (15.3%), while edaphic communities were characterized by *Sphingobacteriaceae* preeminence (10.2%). Furthermore, the *Enterobacteriaceae* family constitutes 11.7% of the tick-associated microbial community, thereby reflecting the similarity between hemophilic arthropods and their mammal hosts.

The patterns of resistome in different interfaces

Diversity analysis delineated specific patterns in both ARG and MGE diversity across different groups. Significant inter-group differentials ($p < 0.01$) emerged across all pairwise comparisons except Marmota-Tick and Sheep-Soil dyads for ARG profiles (Fig. 1B), contrasting with the Marmota-Sheep exception observed in MGE distribution (Fig. 1C). Both NMDS ordination models demonstrated robust structural validity through stress coefficients of 0.13 (ARG, Fig. 1E) and 0.10 (MGE, Fig. 1F).

Furthermore, the presence of ARGs associated with resistance to at least 25 antimicrobial classes was identified in fecal samples from wild *Marmota himalayana*, domesticated sheep, and parasitic ticks. It is noteworthy that the feces of the wild *Marmota himalayana* specimen exhibited a higher prevalence of specific ARGs, which are associated with resistance to fusidic acid. Conversely, adjacent burrow soil samples exhibited ARGs corresponding to a mere 24 antibiotic classes (Fig. 2). The absence of bleomycin- and defensin-related resistance genes in soil samples indicates discrepancies in the ARG profiles between biological specimens and their surrounding environment. A thorough analysis of the patterns of resistance mechanisms (Fig. 2) across diverse interfaces reveals that target alteration is predominant in marmot (51.9%) and sheep samples (37.7%), while ticks demonstrate higher proportions of antibiotic inactivation (32.0%) and antibiotic efflux mechanisms (31.4%). The soil composition is predominantly characterized by the presence of antibiotic efflux (44.2%), antibiotic inactivation (27.1%), and alterations to the antibiotic target (24.8%).

The reconstruction of metagenomically assembled genomes

A total of 5339 metagenome-assembled genomes (MAGs) were reconstructed, 1481 of which met or exceeded the medium-quality level of the minimum information about a metagenome-assembled genome standard (completeness ≥ 50% and contamination < 10%). A total of 339 (9.1%) of these MAGs were identified as high-quality genomes, meeting criteria for completeness > 90% and contamination < 5%.

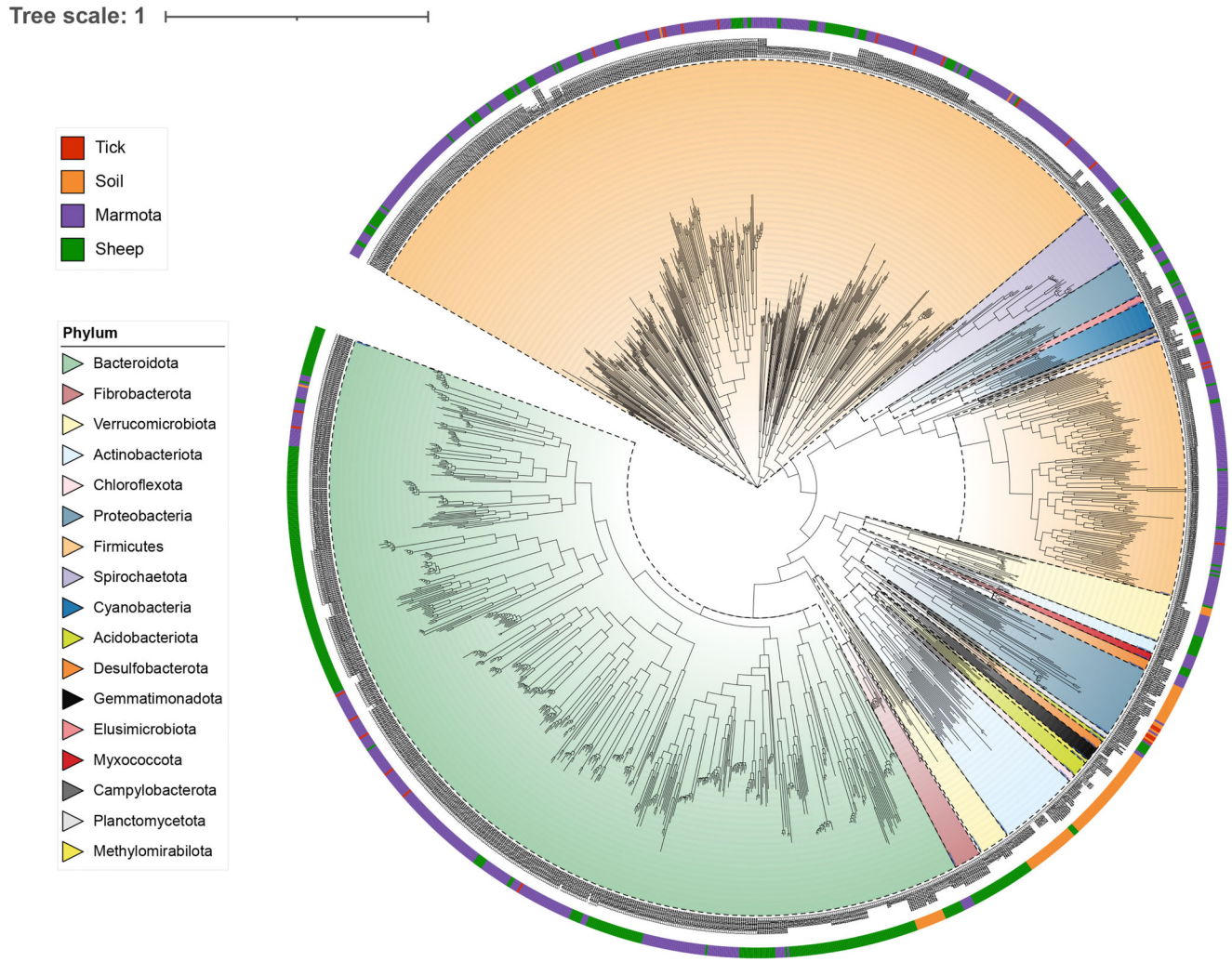

Fig. 3 | Phylogeny and taxonomy of reconstructed MAGs. A maximum-likelihood phylogenomic tree was constructed using 1481 MAGs and 255 GTDB reference genomes. The color-coded bars represent the four sample types (red: tick; orange: soil; purple: marmota; green: sheep), while colored boxes are used to highlight various clades representing distinct bacterial phyla.

Taxonomic classification analysis across host and environmental niches revealed distinct microbial phylogenetic architectures at the MAG level (Fig. 3). We further combined all high- and medium-quality level MAGs and dereplicated at 99.9 and 95% average nucleotide identities (ANIs) to produce 834 SGBs (Supplementary Fig. 2). We identified 263 uSGBs (31.5% of SGBs) and 571 known SGBs (68.4% of SGBs) based on the threshold of 95% ANI and 30% alignment fraction (AF).

Cross-host shared resistome and functional gene pools

A comparative analysis of shared microbial species, ARGs, and toxin genes across four sample groups was conducted, resulting in the identification of 18 ARGs and 78 toxin genes common to all groups (Fig. 4A–C). Additionally, three groups shared 4 microbial species, 27 ARGs, and 41 toxin genes. Interspecific analysis demonstrated distinct sharing patterns: Ticks exhibited the highest overlap with marmots (71.9% species similarity, 51.7% ARG concordance, and 90.1% toxin gene similarity), followed by more modest sharing with sheep (12.5% species, 47.1% ARGs, and 87.6% toxin genes). It is noteworthy that while ticks did not share any species with soil samples, they exhibited 31.0% ARG and 71.9% toxin gene commonality. Soil samples exhibited complete species divergence from both marmots and sheep, yet retained 32.1% ARGs with each mammalian host. This was coupled with high toxin gene preservation, with 86.1% in marmots and 87.0% in sheep. A comparative analysis between marmots and sheep revealed limited species overlap (10.4%), contrasting with substantial genetic exchange (44.8% ARG sharing and 86.3% toxin gene conservation).

A SourceTracker-style mixing analysis based on ARG feature profiles estimated the relative contributions of tick- and soil-associated sources to the resistomes of marmota and sheep. Across marmota samples (Fig. 4D), the inferred tick contribution was consistently high (0.46–0.63 across individuals), whereas the soil contribution was generally smaller and more variable (0.01–0.37); the remaining fraction was assigned to the "unknown" category (0.18–0.45). Across sheep samples (Fig. 4E), the inferred tick contribution was also substantial (0.29–0.56), with a minor contribution from soil (0.01–0.11) and a comparatively larger proportion assigned to "unknown" (0.43–0.65).

Host composition characteristics and horizontal transfer risk of ARG

Following an evaluation of the risk of ARGs in MAGs, it was determined that sheep exhibited the highest risk of ARGs, harboring both Rank II and Rank I ARGs in the sheep group (Fig. 4F). This elevated risk profile was found to be associated with the presence of certain bacterial lineages, namely those belonging to the *Bacteroidaceae* and *Paludibacteraceae* families, which are specific to sheep hosts. Marmota and tick populations exhibited moderate risks, predominantly characterized by Rank III/IV classifications. In contrast, soil samples predominantly manifested low-risk Rank IV profiles.

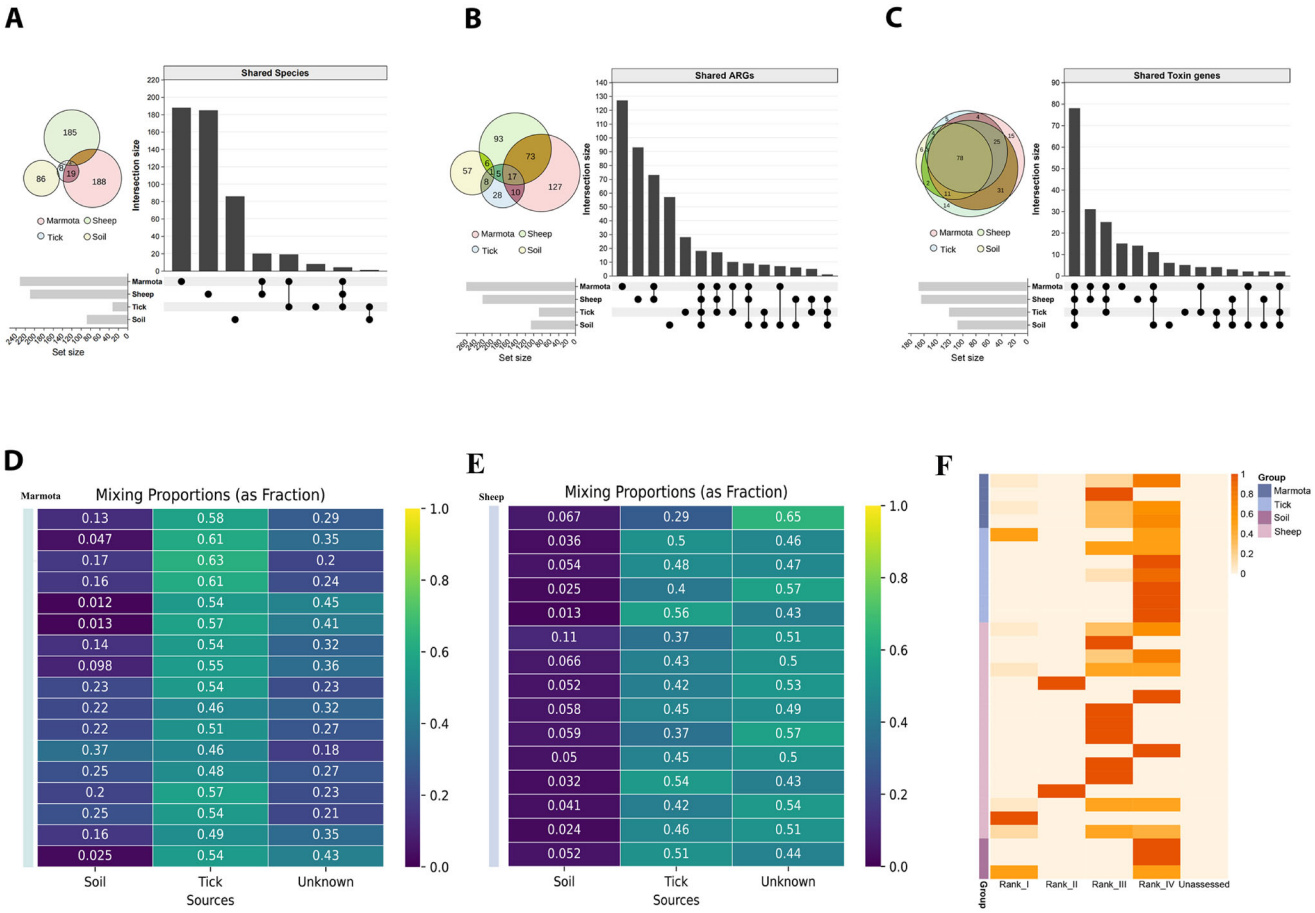

Fig. 4 | Cross-host shared function genes are pooled in the same ecosystem. A–C Number of species, args, and toxin genes of MAGs shared in different interfaces. **D-E** Source Tracker analysis of resistome in (**D**) marmot and (**E**) sheep samples. **F** Rank of risk ARGs in different interfaces.

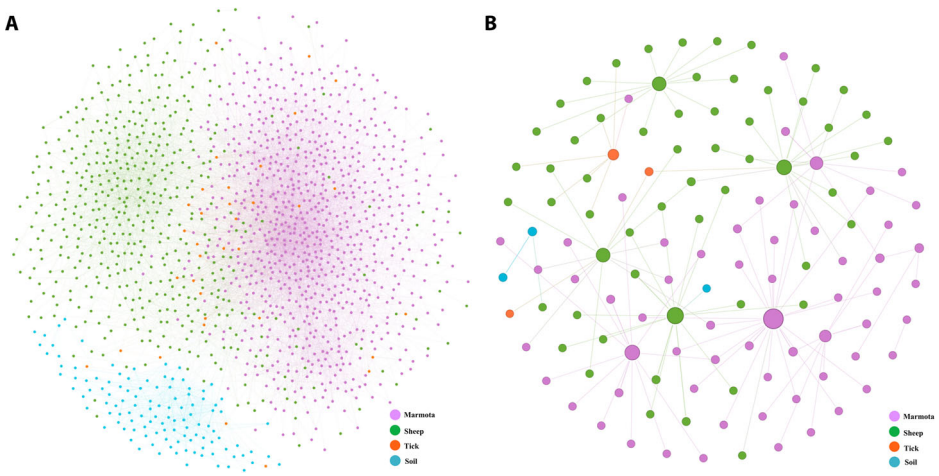

Fig. 5 | HGT Interaction networks at microbial genomic resolution. A The co-occurrence network of HGT among MAGs. **B** The co-occurrence network of HGT among MAGs with high-risk ARGs.

Enterobacteriaceae emerged as a shared ARG reservoir across sheep, marmota, and tick, whereas *Pseudomonadaceae* and *Moraxellaceae* bridged tick-soil transmission. Our comprehensive analysis of MAGs and high-risk MAGs revealed an active ARG shared network connecting multiple species (Fig. 5A and B). The network analysis showed that Marmota and Sheep-derived MAGs (represented by purple and green clusters) functioned as central hubs, exhibiting significantly higher degree centrality compared to other groups. This structural layout suggests that these hosts serve as critical reservoirs for ARG dissemination. The presence of tight clustering among divergent species indicates that HGT events are not taxonomically restricted but occur frequently across broad phylogenetic boundaries. Furthermore, the peripheral but connected position of Tick-derived MAGs (orange) implies a potential route for resistance gene flow from environmental sources to animal hosts. The data suggest the presence of a potential exchange of genes between Marmota and sheep-associated MGEs and tick populations (Fig. 6). The notable transmission link has been identified specifically between different interfaces and phyla (Supplementary Fig. 3).

Discussion

AMR represents a growing threat to global public health, as there is an absence of effective and safe alternative treatments[22]. The propagation of the pathogen occurs not only in healthcare facilities, but more persistently at

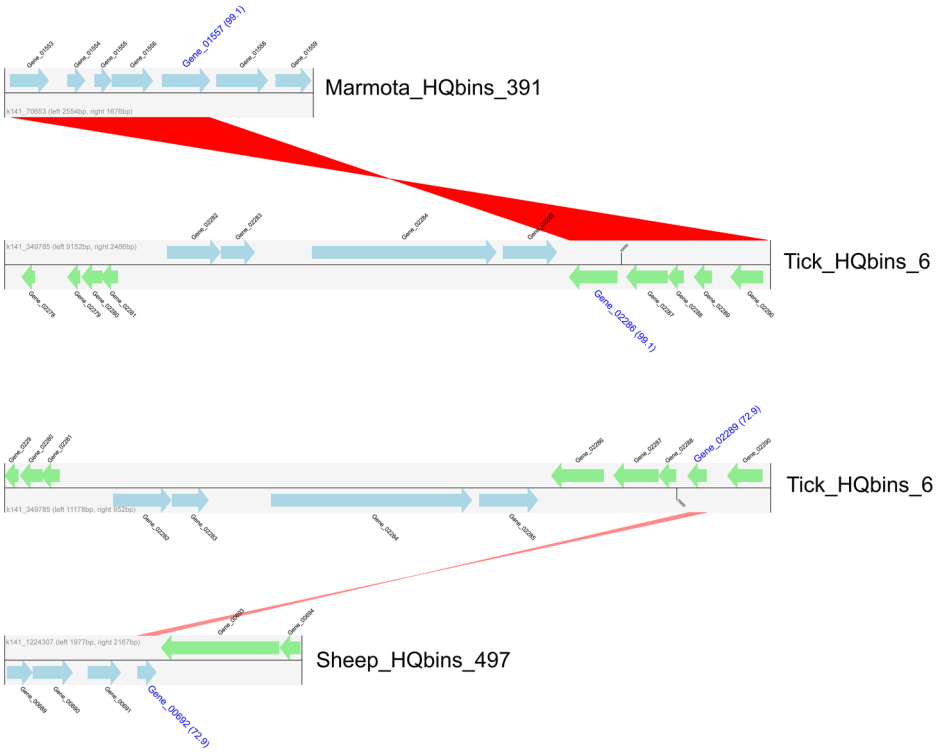

Fig. 6 | Typical HGT occurring in MAGs between different interfaces.

zoonotic transmission nodes where human, animal, and environmental reservoirs interact synergistically[10]. This study employs metagenomic approaches to delineate the microbiome and functional gene interactions—particularly ARGs—across wildlife (*Marmota himalayana*), vector species (*Ixodes canisuga*), livestock (*Ovis aries*), and their shared environment within a regional ecosystem. By reconstructing HGT networks at microbial genomic resolution, we reveal fine-scale genetic flux dynamics among these ecological compartments. This, in turn, advances a One Health-aligned framework to address AMR dissemination at the wildlife–livestock-vector-environment interfaces.

The substantial compositional divergence observed across mammalian hosts (wild marmots and sheep), ticks, and soil microbiomes underscores the profound influence of host physiology and ecological niches in shaping microbial communities. Our analyses revealed distinct alpha-diversity patterns. Specifically, soil exhibited the highest diversity in different interfaces, while mammalian gut microbiomes exhibited higher Shannon entropy and Chao1 richness compared to the tick group ($p < 0.01$). The observed disparity in microbial diversity across interfaces reflects distinct ecological constraints and niche specialization. Soil's superior alpha diversity likely stems from its heterogeneous physicochemical gradients and resource availability, fostering niche partitioning and coexistence of phylogenetically divergent taxa[23]. The diversity of mammalian gut microbiomes reflects the nutrient-rich, stable intestinal environments that support diverse microbiota[24,25]. In contrast, ticks—constrained by their blood-feeding ecology—harbored reduced diversity, dominated by *Proteobacteria* (62.3%), likely reflecting selective pressures from host immune factors and limited nutritional inputs[26,27]. The NMDS ordination (stress = 0.08) demonstrated clear spatial separation of the four groups, with marmot and sheep microbiomes clustering distantly from ticks and soils. It is noteworthy that tick-associated communities exhibit high dispersion across other groups, suggesting a potential role as ecological bridges in both host-parasite and environmental interface. This finding is consistent with their ecological function as ectoparasites.

At the phylum level, the conserved multi-phylum structure in mammals contrasts sharply with the *Proteobacteria* supremacy in ticks. *Firmicutes* specialize in complex carbohydrate metabolism critical for herbivorous and hibernation-driven energy demands[28,29], while *Proteobacteria* in ticks may thrive on host-derived lipids and heme[30–33]. Soil communities exhibited distinct niche partitioning, the dominance of *Proteobacteria* aligns with their well-documented roles in organic matter decomposition and carbon cycling[34], while *Actinobacteria* were known for secondary metabolite production and sporulation[35,36]. The predominance of the *Sphingobacteriaceae* family in the soil (10.2%) is indicative of oligotrophic adaptation, a characteristic that is well-documented within the scientific literature. This family is renowned for its ability to degrade recalcitrant organic matter in nutrient-poor environments[37].

The distribution of ARGs across host-associated and environmental compartments is characterized by a stratification that reveals distinct ecological drivers shaping resistome architecture. Our analysis demonstrates that biological hosts harbor significantly higher ARG diversity compared to soils ($p < 0.01$, Kruskal–Wallis). Specifically, marmot feces are uniquely enriched in fusidic acid resistance genes, a pattern potentially linked to their hibernation physiology, which may select for Staphylococcus-derived resistance traits in the gut microbiome[38,39]. The absence of bleomycin- and defensin-related ARGs in soils underscores the role of environmental filtering mechanisms, which may be attributable to a diminished selective pressure from these antimicrobials within non-host ecosystems. A diversity analysis was conducted, which confirmed that the resistome is strongly partitioned by ecological niche. Although marmot-tick ($p = 0.067$) and sheep-soil ($p = 0.112$) pairs demonstrated non-significant dissimilarity, these trends imply nuanced transmission dynamics. Ticks' close physical association with mammalian hosts may facilitate the exchange of ARGs, whereas sheep-soil convergence could reflect pasture grazing-mediated gene flow[40–42]. The prevalence of target alteration mechanisms in mammals (51.9% in marmots; 37.7% in sheep) is consistent with host-driven selection for chromosomal mutations in core gut symbionts (*Bacteroidetes*), which frequently evolve target-modifying resistance over prolonged periods of host-microbe coadaptation[43]. For marmot feces, the prominence of target alteration/target protection-type mechanisms may be consistent with a resistome shaped more by host ecological filtering and intrinsic

community structure under relatively limited direct anthropogenic antibiotic selection, compared with managed livestock settings. Finally, for sheep, stronger anthropogenic selection pressures and higher host-density/contact rates in managed contexts could favor mechanisms commonly carried on MGEs (such as enzymatic inactivation and associated dissemination potential).

In contrast, ticks and soils exhibited a preference for energy-dependent resistance strategies. Ticks demonstrated a clear predilection for antibiotic inactivation (32.0%) and efflux pumps (31.4%), likely adaptations to transient exposure to host-derived antimicrobial peptides during blood feeding[44]. Based on these, we hypothesize that tick interface represents a highly connected ecological "junction" linking wildlife, livestock, and the environment; such source mixing may increase the probability of acquiring mobile-element–associated resistance functions, leading to a more balanced mechanism profile rather than a single dominant category. Soils' efflux-dominated profile (44.2%) mirrors microbial survival strategies in chemically heterogeneous environments, where broad-spectrum efflux systems provide fitness advantages against natural antimicrobials and pollutants[37]. Genomic analysis of reconstructed MAGs reveals that microbial persistence in the tick is supported by specialized physiological adaptations. We identified sophisticated heme homeostasis machinery—including Hemerythrin (Hr) sensors and the HmuVUT ABC transport system—which likely mitigates host-derived oxidative stress and prevents the toxic Fenton reaction[45]. Furthermore, the presence of the Mla pathway[46] and cell envelope reinforcement genes (MraY and GT4) suggests a robust capacity to maintain membrane integrity against host surfactants[47,48]. These adaptive traits facilitate the stable colonization of these taxa, potentially allowing them to serve as persistent reservoirs for the diverse resistome identified in this study.

While genomic databases have expanded rapidly, they remain heavily biased toward human-related or clinical isolates, leaving the microbial "dark matter" in ecological niches like soil and tick-host systems largely unexplored. The dereplication of MAGs at 99.9% and 95% average nucleotide identity (ANI) thresholds produced 834 SGBs, including 263 uSGBs (31.5%), underscoring the vast unexplored microbial diversity in wildlife-livestock interfaces. The 263 uSGBs, predominantly derived from soil and tick samples, represent novel microbial lineages that may be adapted to extreme or understudied niches. Their presence in environmental samples (78% of uSGBs) underscores the inadequacy of existing reference databases for soil and arthropod-associated taxa. Conversely, the high proportion of known SGBs in mammals (89% of sheep MAGs) reflects better characterization of host-associated microbiomes. Our identification of these USGBs serves as a critical reminder of the limitations of existing repositories and demonstrates the indispensable value of systematic field-based metagenomic surveys in uncovering hidden biodiversity.

The asymmetric sharing patterns of microbial species, ARGs, and toxin genes across marmots, sheep, ticks, and soil underscore distinct mechanisms of genetic exchange at ecological interfaces. The striking discordance between taxonomic overlap and functional gene conservation—exemplified by the fact that ticks share 71.9% species and 51.7% ARGs with marmots, yet no species and 31.0% ARGs with soils—identifies two key transmission pathways: The first topic is host-parasite cohabitation, which facilitates the transfer of species and genes. The second topic is environment-mediated HGT, which enables the persistence of ARGs across taxonomically disconnected niches. Ticks' dual role as biological vectors and environmental interfaces is evident in their hybrid gene pool. The high species similarity with marmots reflects bloodmeal-driven microbial colonization, whereas the overlap of ARG/toxin genes with soils suggests the acquisition of extracellular DNA from burrow substrates. The near-complete toxin gene conservation between ticks and mammals (87–90%) may indicate strong selective pressure to maintain virulence factors for host immune evasion, contrasting with the niche-specific enrichment of ARGs. Soil exhibits a paradoxical profile, characterized by complete species divergence from hosts yet a high prevalence of ARGs and toxin gene sharing, reaching 32% and greater than 86%, respectively. This phenomenon underscores its function as a non-living reservoir for MGEs. The preservation of host-associated ARGs (β-lactamases) in soils is likely facilitated by the stabilization of plasmids via biofilm communities or adsorption to mineral particles[49,50], thereby enabling prolonged environmental persistence. The high degree of toxin genes' resilience, as indicated by their 86–87% retention, may be indicative of their utility in microbial competition, independent of host interactions[51].

Marmot-sheep comparisons further illuminate HGT dynamics: limited species overlap (10.4%), but substantial ARG (44.8%) and toxin (86.3%) sharing suggest grazing-mediated gene flow via shared pastures[40–42]. The hierarchical distribution of antimicrobial resistance gene (ARG) risks across hosts and environments, coupled with active HGT networks, reveals critical linkages between microbial community structure and resistance dissemination. Sheep exhibited the highest ARG risk burden, uniquely harboring high-risk Rank I/II genes (*Bacteroidaceae*-associated MAGs), likely due to their exposure to veterinary antibiotics[52] and nutrient-rich gut environments that favor ARG enrichment in commensal taxa[53,54]. By bridging human activity and the environment, livestock in semi-natural systems could act as stable reservoirs that sustain and spread the resistome dissemination. In contrast, marmots and ticks showed moderate risks (Rank III/IV), reflecting limited anthropogenic antibiotic pressure in wild ecosystems, while soils retained low-risk profiles dominated by efflux-mediated ARGs—a signature of environmental persistence rather than recent selection. We emphasize that the ranking reflects a prioritization of potential public health relevance rather than a direct measure of clinical outcomes; epidemiological linkage would require additional surveillance and clinical data. Co-occurrence network analysis identified *Enterobacteriaceae* as a cross-host ARG reservoir, underscoring its genomic plasticity and capacity to harbor transmissible plasmids across phylogenetically distant hosts (sheep, marmots, ticks). This aligns with prior studies implicating Enterobacteriaceae in global multidrug resistance dissemination[55]. Notably, *Pseudomonadaceae* and *Moraxellaceae* bridged tick-soil ARG exchange, suggesting their dual adaptability to host-associated and free-living niches. These taxa, known for biofilm formation and broad-spectrum efflux systems, likely stabilize extracellular DNA in soil, facilitating HGT through transformation or phage-mediated transduction[56,57]. The marmot-soil transmission link, visualized in network topology, highlights soil as a transient reservoir for wildlife-shed ARGs. This pathway may involve Marmota burrow deposition of fecal microbes, followed by soil-mediated plasmid stabilization and subsequent uptake by grazing sheep.

The study has two primary limitations: Firstly, it is imperative to acknowledge the limitations imposed by spatial-temporal and sampling constraints. These constraints include single-region sampling, restricted sample sizes, and an absence of longitudinal data. Thus, our data demonstrate cross-interface similarity and potential dissemination under the long-term ecological accumulation, but do not resolve directionality or causality as snapshot transfer events, rapid fluctuations in ARG abundance, or time-resolved gene mobilization processes. Future longitudinal designs with higher temporal resolution are needed to directly resolve these momentary processes and mechanistically validate ARG activity and transfer pathways. Secondly, as co-occurrence networks and metagenome-based risk stratification cannot confirm active HGT or validate ARGs functionality without experimental evidence. A multi-layer validation framework using advanced genomic linkage methods (long-read and Hi–C), expression/activity profiling (meta-omics), and targeted experimental verification (cultivation or microcosm conjugation assays) is required to establish active HGT and functional ARGs.

Methods

Experimental design and sample collection

A total of 54 Fecal samples from wild *Marmota himalayana* were sampled from May 2023 to August 2024 as part of the animal health surveillance program conducted by the Centre for Disease Control and Prevention of Wuwei, Gansu Province, China. To investigate microbiome co-occurrence patterns and potential transmission routes among wild *M.himalayana*, vectors, livestock, and environmental reservoirs, we additionally collected

fecal samples of sheep, parasitic ticks of *M.himalayana* and soil samples from associated habitats. Marmot fecal samples were collected simultaneously with their parasitic ticks and burrow soil. Soil was excavated from the burrow walls at a depth of 20–30 cm from the entrance, where three subsamples of equal mass were collected per burrow and homogenized into a single composite sample. Sheep fecal samples were collected from individuals located in close proximity to the marmot sampling sites. The identification of species was conducted by experienced professionals through morphological observation and molecular biological methods. Specifically, Marmota species were identified using the mitochondrial cytochrome b (cytB) gene[58], while ticks were identified through the mitochondrial DNA-encoded cytochrome coxidase subunit I (COX I) and the nucleus-encoded 16S ribosomal DNA (16S rDNA), employing PCR (Supplementary Table 1) and Sanger sequencing techniques[59]. Tick specimens were surface-sterilized through sequential immersion in 5% sodium hypochlorite and 70% ethanol, followed by thorough rinsing with sterile Milli-Q water. Subsequently, the salivary glands, Malpighian tubules, and ovaries were isolated via fine microdissection in ice-cold 1x PBS. To effectively eliminate residual host blood components, the excised tissues underwent at least five consecutive washes in fresh PBS prior to being processed for nucleic acid extraction. Upon arrival at the laboratory, the composite soil samples were processed under aseptic conditions to ensure representativeness and remove non-microbial components. The soil was then thoroughly homogenized and passed through a sterile 2 mm mesh sieve to achieve a uniform particle size and eliminate coarse organic matter. For each sample, approximately 0.25–0.5 g of the sieved soil was aliquoted into sterile 2 mL screw-cap bead-beating tubes specifically designed for soil DNA extraction. To prevent cross-contamination, all processing tools, including sieves and spatulas, were surface-sterilized with 70% ethanol and rinsed with sterile Milli-Q water between individual samples. Triplicate tick specimens derived from identical sites were then pooled and subjected to high-throughput nucleic acid isolation protocols. All specimens were transported on dry ice and subsequently stored at -80°C until processing. Ultimately, we selected 17 fecal samples from marmots parasitized by ticks in non-repeating locations, corresponding to 60 female parasitic ticks (Triplicate samples from the same host source were pooled for library construction), 13 soil samples from marmot burrows (only one sample was collected from the duplicate cave), and 15 fecal samples from nearby semi-grazing management sheep for subsequent metagenomic sequencing (Details are provided in Supplementary Data 1).

DNA extraction and metagenomic sequencing

DNA was extracted using sample-specific commercial kits: The QIAamp DNA Mini Kit for tissue samples, Fast DNA Stool Mini Kit for fecal material, and PowerSoil DNA Isolation Kit for soil samples (Qiagen, Germany), following manufacturers' protocols. DNA concentration was quantified using a Qubit 3.0 fluorometer (Thermo Fisher Scientific, USA). The integrity of library fragments was assessed using an Agilent 2100 Bioanalyzer, confirming an optimal OD260/280 ratio of 1.8–2.0 and an OD260/230 ratio within the range of 2.0–2.2. Sequencing libraries were prepared using the NEBNext Ultra DNA Library Prep Kit (New England Biolabs, USA) and sequenced on an Illumina HiSeq X-ten platform (150-bp paired-end reads) through a commercial service provider (Guangdong Magigene Biotechnology Ltd., China). Library sequencing generated a minimum output of 15 Gb per sample to ensure sufficient depth and accuracy for downstream metagenomic analysis.

Microbial community analysis and function gene annotation

Raw reads were processed through adapter trimming, quality filtering, and host DNA removal using KneadData v0.7.6 (https://huttenhower.sph.harvard.edu/kneaddata). Host genome removal was performed via Bowtie2 alignment against reference genomes of *Marmota himalayana* and tick species. Quality-controlled reads were assembled using MEGAHIT v1.0.3 (https://github.com/voutcn/megahit) with default parameters. Then the assembled contigs were taxonomically classified using Kraken2[60] with abundance estimation via Bracken[61]. The PathoFact pipeline[62] was employed for concurrent prediction of ARGs (performed by DeepARG[63] and RGI[64]), MGEs, and virulence factors. Open reading frames were predicted using Prodigal v2.6.3 (https://github.com/hyattpd/Prodigal), with ARG/MGE abundances normalized to estimated prokaryotic cell counts (ARGs-OAP v3.2).

SourceTracker analysis and DNA-level validation of shared ARGs

We implemented a SourceTracker analysis using ARG feature profiles derived from our cross-group ARG dataset[65,66]. Briefly, tick and soil were treated as source environments, and each marmota and sheep sample was treated as a sink. We then estimated the mixture proportions for each sink sample, reporting contributions from (i) tick, (ii) soil, and (iii) an "unknown" component capturing sources not represented in the training set. To address the evidence of potential transboundary dissemination across different interfaces at the DNA level, we conducted a nucleotide-level analysis to determine whether shared ARG signals correspond to identical or highly similar sequences. We constructed a tick-derived AMR gene nucleotide reference catalog from tick metagenome contigs. Briefly, assembled contigs were searched against the CARD protein database[64] (v3.0) using DIAMOND blastx[67] (v0.9.34) and filtered using stringent criteria (identity ≥ 90%, query coverage ≥ 90%, *e*-value ≤ 1e − 10, sensitive mode). For high-confidence matches, the corresponding coding sequences (CDSs) were precisely extracted from the contigs based on the alignment coordinates (qstart and qend) reported by DIAMOND. Extracted regions were deduplicated by retaining the best match (highest identity) for each unique contig region. The resulting nucleotide sequences were compiled to build a non-redundant, position-specific AMR gene database. Then this tick AMR nucleotide catalog was employed as the reference for BLASTn comparisons against three other metagenomic contig datasets (nucleotide identity ≥ 99%, *e*-value ≤ 1e − 5, alignments ≥ 300 bp).

MAGs reconstruction

The MAGs were reconstructed using SemiBin[68] under a dual-phase assembly strategy involving initial single-sample assembly complemented by co-assembly-based phylogenetic reconstruction. CheckM v1.1.2 (https://github.com/Ecogenomics/CheckM) with the options 'lineage_wf -c 50 -x 10' was used to assess the completeness and contamination of all MAGs with stringent quality thresholds (>50% completeness, <10% contamination). Taxonomic classification utilized GTDB-Tk v2.1.1[69] against the GTDB R207_v2 database. In short, the GTDB-Tk classifies each genome based on ANI to a curated collection of reference genomes, placement in the bacterial reference genome tree, and relative evolutionary distance (RED). Species-level genome bins (SGBs) were formed from MAGs at a 95% ANI[70] threshold using the 'cluster' function in dRep v2.2.3[71]. The SGBs that include at least one reference genome or metagenome-assembled genome in the Genome Taxonomy Database were classified as known SGBs. Those lacking reference genomes were labeled as unknown SGBs (uSGBs). Phylogenomic analysis was conducted using the PhyloPhlAn 3.0 tool[72] based on a set of 400 conserved bacterial protein sequences. ARG/MGE host linkages were established by detecting these elements within MAGs through the PathoFact pipeline. MetaCHIP pipeline[73] was used to assess the horizontal transfer risk of these ARGs among different groups. ARG risk ranking was performed using ARG-RANKER, which classifies ARGs into four ranks (I–IV) based on anthropogenic enrichment, mobility, and presence in ESKAPE pathogens (decision-tree framework). High-risk ARGs were defined as Rank I–II[74]. Function annotated using eggNOG-mapper[75] v5 against the eggNOG database[76] with default parameters.

Statistical analysis and data visualization

Statistical comparisons were conducted using nonparametric Kruskal–Wallis tests among disparate groups. In the event that the overall test was found to be significant ($p < 0.05$), pairwise comparisons between groups were subsequently performed with the Wilcoxon rank sum test to ascertain the underlying cause of the observed difference. The multiple test

correction was executed through the implementation of the Bonferroni correction method. The statistical significance of the findings was established at $p < 0.05$, employing conventional inferential thresholds. Alpha and beta diversity of taxonomic profiles, ARGs, and MGEs were calculated using the vegan package (https://github.com/vegandevs/vegan) in the R (v4.0.2) software. Non-metric Multidimensional Scaling (NMDS) was implemented based on Bray–Curtis dissimilarity matrices to ordinate microbial community structures. The boxplot and bubble plot were executed using the ggplot2 package in R v4.0.2. The phylogenetic tree was plotted using Interactive Tree of Life (iTOL) v5 online[77]. The co-occurrence networks of args were visualized using Gephi software 0.9.2 (https://github.com/gephi/gephi). The others were all performed in the R language.

Data availability

The sequence data in this study have been deposited into the CNGB Sequence Archive of China National GeneBank Database (CNGBdb) with accession number CNP0009025.

Received: 10 September 2025; Accepted: 30 March 2026;
Published online: 14 April 2026

References

1. World Health Organization. *Global Antibiotic Resistance Surveillance Report 2025: Summary* (WHO, 2025).
2. GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *Lancet* **404**, 1199–1226 (2024).
3. Lin, H. et al. Epigenetic modifications and metabolic gene mutations drive resistance evolution in response to stimulatory antibiotics. *Mol. Syst. Biol.* **21**, 294–314 (2025).
4. Durão, P., Balbontín, R. & Gordo, I. Evolutionary mechanisms shaping the maintenance of antibiotic resistance. *Trends Microbiol.* **26**, 677–691 (2018).
5. Arnold, B. J., Huang, I. T. & Hanage, W. P. Horizontal gene transfer and adaptive evolution in bacteria. *Nat. Rev. Microbiol.* **20**, 206–218 (2022).
6. Larsson, D. G. J. & Flach, C.-F. Antibiotic resistance in the environment. *Nat. Rev. Microbiol.* **20**, 257–269 (2022).
7. Lugagne, J. B. & Dunlop, M. J. Anticipating antibiotic resistance. *Science* **375**, 818–819 (2022).
8. Groussin, M. et al. Elevated rates of horizontal gene transfer in the industrialized human microbiome. *Cell* **184**, 2053–2067.e2018 (2021).
9. Shi, X., Xia, Y., Wei, W. & Ni, B. J. Accelerated spread of antibiotic resistance genes (ARGs) induced by non-antibiotic conditions: roles and mechanisms. *Water Res.* **224**, 119060 (2022).
10. Hassell, J. M. et al. Clinically relevant antimicrobial resistance at the wildlife-livestock-human interface in Nairobi: an epidemiological study. *Lancet Planet. Health* **3**, e259–e269 (2019).
11. Hassell, J. M. et al. Socio-ecological drivers of vertebrate biodiversity and human-animal interfaces across an urban landscape. *Glob. Change Biol.* **27**, 781–792 (2021).
12. Chavarría-Bencomo, I. V., Nevárez-Moorillón, G. V., Espino-Solís, G. P. & Adame-Gallegos, J. R. Antibiotic resistance in tick-borne bacteria: a one health approach perspective. *J. Infect. Public Health* **16**, 153–162 (2023).
13. Wei, N., Lu, J., Dong, Y. & Li, S. Profiles of microbial community and antibiotic resistome in wild tick species. *mSystems* **7**, e0003722 (2022).
14. Winkler, A. S. et al. The Lancet One Health Commission: harnessing our interconnectedness for equitable, sustainable, and healthy socioecological systems. *Lancet* **406**, 501–570 (2025).
15. D'Costa, V. M. et al. Antibiotic resistance is ancient. *Nature* **477**, 457–461 (2011).
16. Arnold, K. E., Williams, N. J. & Bennett, M. 'Disperse abroad in the land': the role of wildlife in the dissemination of antimicrobial resistance. *Biology Lett*. https://doi.org/10.1098/rsbl.2016.0137 (2016).
17. Torres, R. T. et al. Wild boar as a reservoir of antimicrobial resistance. *Sci. Total Environ.* **717**, 135001 (2020).
18. Brealey, J. C., Leitão, H. G., Hofstede, T., Kalthoff, D. C. & Guschanski, K. The oral microbiota of wild bears in Sweden reflects the history of antibiotic use by humans. *Curr. Biol.* **31**, 4650–4658.e4656 (2021).
19. Branck, T. et al. Comprehensive profile of the companion animal gut microbiome integrating reference-based and reference-free methods. *ISME J*. https://doi.org/10.1093/ismejo/wrae201 (2024).
20. Yang, Y. et al. Pet cats may shape the antibiotic resistome of their owner's gut and living environment. *Microbiome* **11**, 235 (2023).
21. Taş, N. et al. Metagenomic tools in microbial ecology research. *Curr. Opin. Biotechnol.* **67**, 184–191 (2021).
22. Upadhayay, A. et al. Resistance-proof antimicrobial drug discovery to combat global antimicrobial resistance threat. *Drug Resist. Updat.* **66**, 100890 (2023).
23. Tecon, R. & Or, D. Biophysical processes supporting the diversity of microbial life in soil. *FEMS Microbiol. Rev.* **41**, 599–623 (2017).
24. Ma, J. et al. The interaction among gut microbes, the intestinal barrier and short chain fatty acids. *Anim. Nutr.* **9**, 159–174 (2022).
25. Gregor, R. et al. Mammalian gut metabolomes mirror microbiome composition and host phylogeny. *ISME J.* **16**, 1262–1274 (2022).
26. Buysse, M., Binetruy, F., Leibson, R., Gottlieb, Y. & Duron, O. Ecological contacts and host specificity promote replacement of nutritional endosymbionts in ticks. *Microb. Ecol.* **83**, 776–788 (2022).
27. Douglas, S. S. Jr, Hunter, W. B., Qureshi, J. A. & Cano, L. M. Transcriptomic characterization of Wolbachia endosymbiont from *Leuronota fagarae* (Hemiptera: Psylloidae). *Microbiome Res. Rep.* **4**, 19 (2025).
28. da Silva Pereira, M. et al. Microbial Rumen proteome analysis suggests Firmicutes and Bacteroidetes as key producers of lignocellulolytic enzymes and carbohydrate-binding modules. *Braz. J. Microbiol.* **30**, 241–253 (2025).
29. Mathur, H., Beresford, T. P. & Cotter, P. D. Health benefits of lactic acid bacteria (LAB) fermentates. *Nutrients*. https://doi.org/10.1016/j.crmeth.2025.101005 (2020).
30. Kengmo Tchoupa, A., Eijkelkamp, B. A. & Peschel, A. Bacterial adaptation strategies to host-derived fatty acids. *Trends Microbiol.* **30**, 241–253 (2022).
31. Allen, P. E. & Martinez, J. J. Modulation of host lipid pathways by pathogenic intracellular bacteria. *Pathogens.* https://doi.org/10.1016/j.tim.2021.06.002 (2020).
32. Nolan, S. J., Romano, J. D. & Coppens, I. Host lipid droplets: an important source of lipids salvaged by the intracellular parasite *Toxoplasma gondii*. *PLoS Pathogens* **13**, e1006362 (2017).
33. Dos Santos, P. T., Larsen, P. T., Menendez-Gil, P., Lillebæk, E. M. S. & Kallipolitis, B. H. *Listeria monocytogenes* relies on the heme-regulated transporter hrtAB to resist heme toxicity and uses heme as a signal to induce transcription of lmo1634, encoding listeria adhesion protein. *Front. Microbiol.* **9**, 3090 (2018).
34. Song, J. et al. Bacterial necromass as the main source of organic matter in saline soils. *J. Environ. Manag.* **371**, 123130 (2024).
35. Katz, L. & Baltz, R. H. Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* **43**, 155–176 (2016).
36. Núñez-Montero, K. et al. Advances in Antarctic Research for Antimicrobial Discovery: a comprehensive narrative review of bacteria from Antarctic environments as potential sources of novel antibiotic compounds against human pathogens and microorganisms of industrial importance. *Antibiotics* **7**, 123130 (2018).
37. Zhang, L. et al. Nitrifiers drive successions of particulate organic matter and microbial community composition in a starved macrocosm. *Environ. Int.* **157**, 106776 (2021).

38. Elazhari, M. et al. Characterization of fusidic acid-resistant Staphylococcus aureus isolates in the community of Casablanca (Morocco). *Int. J. Med. Microbiol. IJMM* **302**, 96–100 (2012).
39. Yu, F. et al. Dissemination of fusidic acid resistance among Staphylococcus aureus clinical isolates. *BMC Microbiol.* **15**, 210 (2015).
40. Tang, T., Chen, Y., Du, Y., Yao, B. & Liu, M. Effects of functional modules and bacterial clusters response on transmission performance of antibiotic resistance genes under antibiotic stress during anaerobic digestion of livestock wastewater. *J. Hazard. Mater.* **441**, 129870 (2023).
41. Wee, B. A., Muloi, D. M. & van Bunnik, B. A. D. Quantifying the transmission of antimicrobial resistance at the human and livestock interface with genomics. *Clin. Microbiol. Infect.* **26**, 1612–1616 (2020).
42. Wang, F. et al. Neglected drivers of antibiotic resistance: survival of extended-spectrum β-lactamase-producing pathogenic *Escherichia coli* from livestock waste through dormancy and release of transformable extracellular antibiotic resistance genes under heat treatment. *Environ. Sci. Technol.* **57**, 9955–9964 (2023).
43. Hulse, S. V., Antonovics, J., Hood, M. E. & Bruns, E. L. Host-pathogen coevolution promotes the evolution of general, broad-spectrum resistance and reduces foreign pathogen spillover risk. *Evol. Lett.* **7**, 467–477 (2023).
44. Chatterjee, D., Sivashanmugam, K., Delgado, L. F. & Andersson, A. F. Immunomodulatory peptides: new therapeutic horizons for emerging and re-emerging infectious diseases. *Front. Microbiol.* **15**, 1505571 (2024).
45. Graça-Souza, A. V. et al. Adaptations against heme toxicity in blood-feeding arthropods. *Insect Biochem. Mol. Biol.* **36**, 322–335 (2006).
46. Ekiert, D. C. et al. Architectures of lipid transport systems for the bacterial outer membrane. *Cell* **169**, 273–285.e217 (2017).
47. Chung, B. C. et al. Crystal structure of MraY, an essential membrane enzyme for bacterial cell wall synthesis. *Science* **341**, 1012–1016 (2013).
48. Whitfield, C., Wear, S. S. & Sande, C. Assembly of bacterial capsular polysaccharides and exopolysaccharides. *Annu. Rev. Microbiol.* **74**, 521–543 (2020).
49. Molin, S. & Tolker-Nielsen, T. Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Curr. Opin. Biotechnol.* **14**, 255–261 (2003).
50. Di Iorio, E. et al. Environmental implications of interaction between humic substances and iron oxide nanoparticles: a review. *Chemosphere* **303**, 135172 (2022).
51. Nachmias, N. et al. Systematic discovery of antibacterial and antifungal bacterial toxins. *Nat. Microbiol.* **9**, 3041–3058 (2024).
52. Hu, J. et al. Animal production predominantly contributes to antibiotic profiles in the Yangtze River. *Water Res.* **242**, 120214 (2023).
53. Yan, K. et al. Diffusion and enrichment of high-risk antibiotic resistance genes (ARGs) via the transmission chain (mulberry leave, guts and feces of silkworm, and soil) in an ecological restoration area of manganese mining, China: Role of heavy metals. *Environ. Res.* **225**, 115616 (2023).
54. Chai, J., Zhuang, Y., Cui, K., Bi, Y. & Zhang, N. Metagenomics reveals the temporal dynamics of the rumen resistome and microbiome in goat kids. *Microbiome* **12**, 14 (2024).
55. Feng, Y. et al. Regional antimicrobial resistance gene flow among the One Health sectors in China. *Microbiome* **13**, 3 (2025).
56. Wise, M. G. et al. Global trends in carbapenem- and difficult-to-treat-resistance among World Health Organization priority bacterial pathogens: ATLAS surveillance program 2018-2022. *J. Glob. Antimicrob. Resist.* **37**, 168–175 (2024).
57. Rubiola, S. et al. Shotgun metagenomic sequencing of bulk tank milk filters reveals the role of Moraxellaceae and Enterobacteriaceae as carriers of antimicrobial resistance genes. *Food Res. Int.* **158**, 111579 (2022).
58. Kocher, T. D. et al. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* **86**, 6196–6200 (1989).
59. Yu, D. W. et al. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods Ecol. Evol.* **3**, 613–623 (2012).
60. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
61. Lu, J., Breitwieser, F. P., Thielen, P. & Salzberg, S. L. Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **3**, e104 (2017).
62. de Nies, L. et al. PathoFact: a pipeline for the prediction of virulence factors and antimicrobial resistance genes in metagenomic data. *Microbiome* **9**, 49 (2021).
63. Arango-Argoty, G. et al. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome* **6**, 23 (2018).
64. Alcock, B. P. et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **51**, D690–d699 (2023).
65. Knights, D. et al. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **8**, 761–763 (2011).
66. McGhee, J. J. et al. Meta-SourceTracker: application of Bayesian source tracking to shotgun metagenomics. *PeerJ* **8**, e8783 (2020).
67. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
68. Pan, S., Zhu, C., Zhao, X.-M. & Coelho, L. P. SemiBin: incorporating information from reference genomes with semi-supervised deep learning leads to better metagenomic assembled genomes (MAGs). *bioRxiv* https://doi.org/10.1101/2021.08.16.456517 (2021).
69. Chaumeil, P. A. et al. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* **36**, 1925–1927 (2019).
70. Goris, J. et al. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evolut. Microbiol.* **57**, 81–91 (2007).
71. Butler, G. et al. Evolution of pathogenicity and sexual reproduction in eight Candida genomes. *Nature* **459**, 657–662 (2009).
72. Segata, N., Börnigen, D., Morgan, X. C. & Huttenhower, C. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat. Commun.* **4**, 2304 (2013).
73. Song, W., Wemheuer, B., Zhang, S., Steensen, K. & Thomas, T. MetaCHIP: community-level horizontal gene transfer identification through the combination of best-match and phylogenetic approaches. *Microbiome* **7**, 36 (2019).
74. Zhang, A. N. et al. An omics-based framework for assessing the health risk of antimicrobial resistance genes. *Nat. Commun.* **12**, 4765 (2021).
75. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
76. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–d314 (2019).
77. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256–w259 (2019).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant nos. 82273689 and 82473689), Key Research and Development Program of Shaanxi (grant no. 2024SF-YBXM-289), and Wuwei City Science and Technology Plan Project (No. WW25Z01SF027).

Author contributions

Zhenhua Lu: writing—original draft, software, methodology, investigation, formal analysis, and visualization. Ruishan Li: investigation, data curation, and writing—original draft. Kaichun Zhou, Shiyu Li, Shijie Chen, and Shiwei Sun: methodology and investigation. Jiacheng Liu and Lele Zhao: methodology. Xiang Yuan: resources, investigation, and supervision. Kun Liu and Zhongjun Shao: writing—review and editing, project administration, and funding acquisition.

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/s41522-026-00986-w.

Correspondence and requests for materials should be addressed to Kun Liu, Xiang Yuan or Zhongjun Shao.