Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Gut resistome, mobile genetic elements and virulence-factor profiles across developmental stages of Erinaceus amurensis, a potential cultivated medicinal hedgehog species in Northeast China

Xiuqi Jia
Northeast Forestry University
Xiang Li
Northeast Forestry University
Xinru Lv
Northeast Forestry University
Yi Li
Northeast Forestry University
Zhimin Jin
Mudanjiang Normal University
Xinyu Yang
Mudanjiang Normal University
Dianwei Li
Mudanjiang Normal University
Zhensheng Liu
Northeast Forestry University
Hongliang Chai
Hongliang_chai@hotmail.com
Northeast Forestry University

Research Article

Keywords: Erinaceus amurensis, ARGs, MRGs, BRGs, VFs, metagenomics

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11628529/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Gut resistome, mobile genetic elements and virulence-factor profiles across developmental stages of *Erinaceus amurensis*, a potential cultivated medicinal hedgehog species in Northeast China

Xiuqi Jia[1,2†], Xiang Li[1†], Xinru Lv[1], Yi Li[1],Zhimin Jin[2],Xinyu Yang[2],Dianwei Li[2],Zhensheng Liu[1,3*], Hongliang Chai[1,3*]

1 College of Wildlife and Protected Area, Northeast Forestry University, Harbin 150040, China

2 College of Life Science and Technology, Mudanjiang Normal University, Mudanjiang 157000, China

3 Key Laboratory of Wildlife Conservation, National Forestry and Grassland Administration, Harbin 150040, China

*Corresponding authors

Hongliang Chai (Hongliang_Chai@hotmail.com),

Zhensheng Liu (zhenshengliu@163.com).

†These authors contributed equally to this work.

Abstract

Erinaceus amurensis is a valuable indigenous species with long-recognized medicinal value and emerging artificial cultivation prospects in Northeast China. Gut antimicrobial resistance genes, mobile genetic elements and virulence factors are closely associated with animal health and biosecurity in captive breeding, yet their baseline profiles across host developmental stages remain largely uncharacterized in this species. Here, we conducted an exploratory metagenomic baseline survey of the gut microbiota in 13 wild hedgehog (*Erinaceus amurensis*) from Northeast China, comprising five males, five females, and three foetuses from a single litter. Using deep short-read metagenomic sequencing and a de novo assembled nonredundant gene catalog, we annotated antibiotic resistance genes (ARGs), metal and biocide resistance genes, mobile genetic elements (MGEs), and virulence factors via the CARD, BacMet, custom mobile-genetic-element reference, and VFDB databases, respectively. Statistical comparisons employed Kruskal-Wallis tests with ε^2 (epsilon-squared effect size) effect sizes and bootstrap 95 % confidence intervals. Efflux-mediated resistance dominated the ARG pool across all groups (61.7-66.4%), and between-group differences reflected abundance redistribution rather than acquisition of novel gene families. MGE composition exhibited stronger group-level separation than ARGs, with foetuses showing transposase-dominated profiles and depletion of recombinases and integrases, consistent with a simplified subset of the adult MGE repertoire. Adult females displayed co-enrichment of type III/IV secretion system virulence factors and metal resistance genes (*glpF*, *frnE*), supported by sample-level Spearman co-occurrence correlations. Foetuses consistently showed reduced functional diversity across all four annotation modules. These preliminary metagenomic data indicate developmental-and sex-associated variation in gut resistome, mobilome, and virulence profiles. Detection of homologous DNA fragments does not establish gene expression, phenotypic resistance, or pathogenic risk, and sample-level cooccurrence cannot confirm physical linkage or horizontal transfer. Furthermore, our observational data cannot disentangle contributions from confounding variables including sex, maternal effects, and diet. Larger cohorts, long-read metagenome-assembled genomes, paired environmental controls, and functional validation are needed to confirm these correlative patterns. These

baseline data from wild populations provide an important reference for intestinal health management, rational antimicrobial use and resistance risk control in the upcoming standardized farming of medicinal hedgehogs.

Keywords: *Erinaceus amurensis*; ARGs; MRGs; BRGs; VFs; metagenomics

Introduction

Antimicrobial resistance (AMR) has become a major threat to global public health and ecosystem stability and is now recognized by the World Health Organization as a “slow pandemic[Error! Unknown switch argument. REF _Ref2935 \r \h * MERGEFORMAT Error! Unknown switch argument.].” Beyond clinical and agricultural settings, sustained human activities have released antibiotics, heavy metals, and biocides into natural environments over long time scales. These pressures have reshaped microbial communities and driven the persistence and spread of resistance genes (RGs), leading to the formation of a large and heterogeneous environmental resistome. Growing evidence shows that resistance genes originating from environmental reservoirs can enter clinical pathogens through multiple ecological routes, contributing to the emergence of hospital-acquired resistance[Error! Unknown switch argument. REF _Ref3108 \r \h * MERGEFORMAT REF _Ref3115 \r \h * MERGEFORMAT REF _Ref3118 \r \h * MERGEFORMAT Error! Unknown switch argument.]. Despite this risk, resistance dynamics in natural ecosystems, especially those involving wildlife, remain poorly characterized.

Within the One Health framework, wildlife is increasingly viewed as a biological indicator of environmental transmission of resistance determinants[Error! Unknown switch argument.]. A wide range of wild animals, including birds and small mammals, have been reported to carry diverse antibiotic resistance genes (ARGs) and mobile genetic elements (MGEs), suggesting their potential role in mediating gene exchange across environmental, animal, and human compartments[Error! Unknown switch argument.Error! Unknown switch argument.]. Notably, methicillin-resistant *Staphylococcus aureus* and other multidrug-resistant bacteria have been detected in European hedgehogs, with evolutionary origins that may predate the clinical use of antibiotics[Error! Unknown switch argument.]. These findings point to a long-standing association between wildlife hosts, microbial communities, and resistance traits.

The Amur hedgehog (*Erinaceus amurensis*) is widely distributed across Northeast Asia and occupies habitats that range from forests to agricultural and peri-urban landscapes. These habitats overlap extensively with areas of human activity. As an omnivorous and nocturnal species, *Erinaceus amurensis* frequently contacts soil, surface water, and diverse food sources, all of which are key reservoirs of environmental microbes. This ecological niche places the species at the interface between natural ecosystems and anthropogenic pressures, making its intestinal microbiome a suitable system for assessing environmental exposure to antibiotics, heavy metals, and associated resistance factors.

In recent years, *E. amurensis* has received growing attention as a special economic animal for medicinal utilization in China, and pilot artificial domestication and captive breeding practices have been gradually carried out. With the potential expansion of medicinal hedgehog farming, gut health, antimicrobial resistance dissemination and pathogen-related risks will become key constraints for the healthy development of this emerging industry. Baseline information on gut resistome, mobile genetic elements and virulence factors across different developmental stages is a fundamental prerequisite for scientific breeding and biosecurity management of farmed hedgehogs.

Horizontal gene transfer plays a central role in the dissemination of resistance genes, with mobile genetic elements acting as key mediators. The co-occurrence of plasmids, integrons, and transposons is often associated with increased transfer potential of ARGs and their movement from environmental reservoirs into pathogenic bacteria[Error! Unknown switch argument.]. In parallel, long-term contamination by heavy metals and biocides can impose co-selection pressures that favor the joint enrichment of antibiotic resistance genes and metal or biocide resistance genes. These processes may increase the stability and persistence of resistance networks in natural microbial communities[Error! Unknown switch argument.Error! Unknown switch argument.]. However, in wild mammals, the joint distribution of RGs, virulence factors(VFs), and MGEs has rarely been examined at the metagenomic level.

Host-related factors further complicate these dynamics. Habitat alteration, Spatial and temporal axes, and physiological condition are known to influence gut microbiome structure and function[Error! Unknown switch argument. REF _Ref19262 \r \h * MERGEFORMATError! Unknown switch argument.], yet their combined effects on resistance and virulence gene networks in wildlife remain unclear. In particular, foetal and juvenile stages may exhibit distinct patterns due to immature immune systems and limited environmental exposure, but systematic evidence from natural populations is lacking.

In this study, we analyzed wild Erinaceus amurensis individuals from Heilongjiang Province in Northeast China, including females, males, and foetal samples. Using high-depth metagenomic sequencing, we integrated annotations from CARD, BacMet, VFDB, and multiple MGE databases to characterize ARGs, metal and fungicide resistance genes, VFs, and MGEs in the *Erinaceus amurensis* gut microbiome. We compared their composition, abundance, and association patterns across sex and developmental stages. By doing so, we aimed to characterize variation in resistance and virulence networks associated with host developmental stages and sex in *E. amurensis*, and to provide fundamental baseline data for intestinal health management and resistance risk prevention in future medicinal hedgehog cultivation.

Results

3.1 Antibiotic resistance genes

Abundance redistribution rather than novel gene acquisition primarily accounts for gut resistome differences across developmental stages

To contextualize the functional potential of the gut microbiome in wild *Erinaceus amurensis*, we first performed functional annotation based on the KEGG database. Core metabolic pathways (including global and overview maps, carbohydrate metabolism, and amino acid metabolism) dominated the functional landscape across the whole cohort, while the “Drug resistance: antimicrobial” pathway also accounted for a substantial fraction of total reads (Figure 1a). The overall distribution pattern of KEGG pathways was largely congruent across adult male (M), adult female (FM), and fetal (FE) groups. However, read counts assigned to antimicrobial resistance pathways followed a clear abundance gradient: resistance-associated reads were most abundant in the M group, intermediate in the FM group, and markedly reduced in the FE group (Figure 1b–d).

Profiling of antibiotic resistance genes (ARGs) using the CARD database demonstrated that the gut resistome of wild *Erinaceus amurensis* encompassed multiple classes of antibiotic resistance. Across all samples, genes conferring resistance to beta-lactams, aminoglycosides,

macrolide-lincosamide-streptogramins (MLS), and multidrug agents were the most dominant categories in terms of gene count (Figure 1e). This rank order was largely consistent across groups: the M group harboured the highest number of ARGs in nearly all classes, followed by the FM group; the FE group had the lowest total ARG count but still retained detectable resistance genes against classes including aminocoumarin, beta-lactam, and fluoroquinolone (Figure 1f–h).

When categorised by functional resistance mechanism (legend in Figure 1p), antibiotic efflux was the overwhelmingly dominant mechanism across all groups, accounting for 61.7 % to 66.4 % of the total resistome. This pattern is consistent with intrinsic chromosomally encoded microbial defences rather than clinically acquired resistance. Analysis of antibiotic resistance ontology (ARO) abundance profiles further showed that the M and FM groups had relatively diverse resistome composition, with multiple ARO families jointly contributing to the overall resistance profile (Figure 1m, n). In contrast, the resistome of the FE group was highly dominated by a single predominant ARO, which accounted for 100 % of annotated resistance gene abundance (Figure 1o).

Statistical note (applicable to all subsequent analyses): Differences in gene abundance among groups were tested using the non-parametric Kruskal–Wallis H test, with epsilon-squared (ε^2) reported as the effect size alongside bootstrap-derived 95 % confidence intervals. All reported *P*-values are nominal and uncorrected for multiple testing. Given that fetal samples were derived from a single litter and do not constitute fully independent biological replicates, all inter-group comparisons are interpreted as preliminary observational patterns rather than robust developmental effects.

Observed differences in two resistance mechanism categories are shown in Figure 1k: antibiotic target replacement and antibiotic inactivation showed higher relative abundance in the FE group ($\varepsilon^2 = 0.695$, nominal P= 0.016 and $\varepsilon^2 = 0.502$, nominal $P = 0.049$, respectively). These nominal differences should be interpreted as exploratory observational patterns rather than confirmatory developmental effects.No individual AMR gene family reached nominal statistical significance, although RND-type and ABC-type efflux pumps showed marginal trends with large effect sizes.

Alpha diversity analysis indicated highly conserved ARG richness across all samples, with a Chao1 index of 6 for every individual; no nominal group-level differences were observed for the Shannon diversity index or Pielou evenness index. Beta diversity analyses were performed to assess compositional differences in resistome profiles among groups. Principal coordinates analysis (PCoA) based on Bray–Curtis distances showed partial compositional separation among the three groups, with substantial inter-group overlap (Figure 1i). The first principal axis explained 56.03 % of total variance, and the second axis explained 15.71 %. ANOSIM showed an ordination pattern of overall compositional differentiation among groups (R = 0.608,nominal $P = 0.001$), consistent with PERMANOVA results (nominal P= 0.002).Rank-based distance analysis further supported this compositional pattern (Figure 1j).

Venn diagram analysis was conducted to examine the overlap of ARO gene families across the three groups (Figure 1l). A large proportion of ARO families was shared by all three groups, accounting for 50.96 % of the total identified ARO repertoire. By contrast, group-specific unique ARO families represented only a small fraction of the total (M: 30.93 %; FM: 13.62 %; FE: 0.16 %). Taken together, these results indicate that inter-group differences in

the gut resistome arise primarily from abundance redistribution within a shared core gene pool, rather than from the gain or loss of distinct resistance gene families. Notably, metagenomic read abundance of ARGs only reflects the relative coverage of corresponding DNA fragments in the dataset; higher abundance does not directly confirm in situ gene expression, active horizontal gene transfer events, or elevated pathogenic risk to the host.

3.2 Metal and biocide-resistance genes

Abundance redistribution rather than novel gene acquisition primarily accounts for gut metal–biocide resistome differences across developmental stages

Building on the antibiotic resistome profiling, we next employed the BacMet database to characterize biocide-and metal-resistance determinants (BMRGs) in the gut microbial communities of wild *Erinaceus amurensis*. Heatmap visualization at the BMRG level revealed a broad repertoire of biocide and metal resistance genes across all samples, with functional categories covering metal resistance, biocide resistance, and dual metal/biocide cross-resistance (Figure 2a). At the compound level, community abundance profiling showed heterogeneous distribution of putative resistance determinants targeting multiple heavy metal ions (including nickel, arsenic, copper, zinc, tungsten/molybdenum, and iron/gallium) and commonly used biocides such as triclosan (Figure 2b).

Venn diagram analysis was conducted to examine the overlap of BMRG repertoires among the three groups. The majority of BMRG families were shared across all three groups, accounting for 71.20 % of the total identified gene set, while group-specific unique genes represented only a small proportion (M: 21.30 %; FM: 5.88 %; FE: 0.00 %) (Figure 2c). This pattern mirrors the observation for antibiotic resistance genes, indicating that compositional differences in metal and biocide resistance profiles also arise primarily from abundance shifts within a shared core gene pool. Circos plots further illustrated the proportional composition of functional resistance categories within each group. Overall, the proportional distribution of resistance functional classes was broadly comparable across groups, yet fetal samples exhibited a relatively simplified compositional profile relative to adult groups (Figure 2d–f).

Beta diversity analysis based on BMRG abundance profiles was performed to assess compositional differences among groups. Principal coordinates analysis (PCoA) showed partial compositional separation among the M, FM, and FE groups, with PC1 explaining 62.12 % of total variance and PC2 explaining 4.74 % (Figure 2g). This pattern suggests that the composition of metal- and biocide-resistance gene pools varies in association with host physiological status. Nominal inter-group differences in multiple individual resistance genes, including genes involved in metal transport and biocide efflux, were observed via Kruskal–Wallis H tests (Figure 2h–i). These patterns are presented as exploratory observations given the non-independence of foetal samples.

Collectively, these findings provide baseline evidence that the gut microbiota of wild *Erinaceus amurensis* harbour functional genetic potential to cope with environmental metal and biocide stressors in their natural habitats, and that developmental variation in this resistome is driven by abundance redistribution rather than *de novo* gene acquisition.

3.3 Mobile genetic elements

Subset filtering and functional skewing rather than independent de novo assembly primarily accounts for gut mobilome differences across developmental stages

Mobile genetic elements (MGEs) represent the primary vectors of horizontal gene transfer

(HGT) and play a central role in the dissemination of resistance determinants within microbial communities. To explore the mechanistic basis underlying the observed resistome variation, we characterized MGE profiles in the gut microbiota of wild *Erinaceus amurensis* across the three groups.

As presented in Figure 3a, transposase constituted the overwhelmingly dominant MGE category across all groups, accounting for 63.0–89.3% of all annotated MGE terms, followed by recombinase (6.5–21.9%) and integrase (3.6–11.1%). This hierarchical pattern aligns with the consensus that DNA transposition, particularly insertion sequence (IS)-mediated cut-and-paste mobility, represents the most prevalent form of genetic mobility in mammalian gut microbiomes.

At the functional-category level, four out of seven MGE classes showed nominally significant inter-group differences (Kruskal–Wallis H test, all nominal $P<0.05$; Figure 3f)-a proportion notably higher than that observed for antibiotic resistance mechanisms (2 of 7) in the parallel CARD-based analysis. These nominal differences are interpreted as exploratory patterns.The most pronounced nominal difference was observed for transposase, which was markedly higher in the foetal group (FE, 89.28 ± 4.73%) relative to both adult groups (M, 63.05 ± 3.68%; FM, 63.68 ± 5.00%; $\varepsilon^2=0.541$, nominal $P=0.039$). Conversely, both recombinase (FE, 6.50 ± 2.79% vs M, 21.90 ± 3.85%; FM, 21.58 ± 1.93%; $\varepsilon^2 = 0.536$, nominal $P = 0.040$) and integrase (FE, 3.57 ± 1.61% vs M, 10.38 ± 1.33%; FM, 11.12 ± 1.18%; $\varepsilon^2=0.563$, nominal $P=0.034$) showed lower abundance in foetal samples. Transposon-related sequences were more abundant in adult males (M, 1.15 ± 1.01%) compared with females and foetuses (FM, 0.32 ± 0.18%; FE, 0.21 ± 0.14%; $\varepsilon^2 = 0.530$, nominal $P=0.041$). Conjugative transfer proteins showed a similar trend of reduction in foetuses but did not reach nominal statistical significance (nominal $P=0.089$). The class 1 integron marker *IntI1* was detected at very low abundance across all groups, with no nominal inter-group difference.

Subtype-level analysis, visualized in the abundance heatmap (Figure 3d), reinforced these overall patterns. Seven out of ten examined subtypes showed nominal differences with large effect sizes ($\varepsilon^2=0.54–0.69$). IS481 showed uniquely higher abundance in foetal samples (FE, 1.01 ± 0.40% vs M, 0.18 ± 0.10%; FM, 0.10 ± 0.10%; $\varepsilon^2 = 0.629$, nominal $P = 0.023$), whereas IS110, IS256, and site-specific integrase all showed lower abundance in FE relative to adult groups. The recombinase-associated subtypes RmuC ($\varepsilon^2=0.695$, nominal $P= 0.016$) and RecR ($\varepsilon^2=0.659$, nominal $P= 0.019$) exhibited the largest effect sizes, with a graded reduction from M through FM to FE, compatible with a developmental attenuation of homologous recombination machinery. ISL3 showed higher abundance specifically in adult males ($\varepsilon^2=0.579$, nominal $P=0.031$).

Beta-diversity analyses showed an ordination pattern consistent with group-level structuring of MGE communities. Analysis of similarities (ANOSIM) yielded an R statistic of 0.619 (nominal $P=0.001$; Figure 3e), suggesting that between-group dissimilarity tends to exceed within-group variation - a separation pattern more apparent than that observed for the ARG pool. Principal coordinates analysis (PCoA) ordination was consistent with this result, showing clear partitioning of foetal samples from adult groups along the primary axis (Figure 3c).

Venn analysis further illuminated the nature of these compositional differences (Figure 3b). The foetal group harboured the fewest total MGE terms (134) compared with adult males (279)

and females (244), and possessed only one unique term (0.34%), whereas adult males had 49 unique terms (16.44%). The core MGE set shared by all three groups comprised 127 terms (42.62%), and the M–FM pairwise overlap (99 terms,33.22%) far exceeded either adult–foetus overlap (M–FE: 4 terms; FM–FE: 2 terms).

In summary, the foetal MGE pool is not a distinct community assembled independently but rather a depauperate subset of the adult MGE repertoire with strong functional skewing toward IS-element-mediated transposition, This pattern is compatible with vertical or early-life acquisition filtering, and may also reflect developmental filtering of the gut microbial community rather than de novo environmental assembly.The coherent depletion of recombinase, integrase, and their subtype derivatives in the foetal gut further indicates comparatively limited capacity for homologous recombination and site-specific integration in early life.

The contrast between MGE and ARG results carries interpretive value. MGE composition exhibited more numerous and larger nominal group differences (4/7 categories, 7/10 subtypes) than ARG composition (2/7 mechanisms, 0/10 gene families), despite both datasets deriving from the same sample set. This dissociation suggests that host-associated physiological variation — particularly across developmental stages — correlates with stronger compositional shifts in the mobile genetic framework relative to the resistance-gene cargo itself. One plausible interpretation is that MGEs, being more directly tied to microbial lifestyle and genome plasticity, respond sensitively to the distinct ecological conditions of the foetal gut (e.g., reduced microbial diversity, altered nutrient availability, immature immune selection), whereas ARG abundances may be more constrained by the basal resistome of the shared founding microbial pool. The dominance of transposase in foetal samples further raises the possibility that IS-element-mediated transposition may serve as a potential route for resistance gene mobilization in early life, although short-read metagenomics cannot confirm physical linkage between specific MGEs and resistance genes.

3.4 Virulence-factor profiles

Virulome compositional shifts parallel resistome and mobilome remodeling across sex and developmental stages

To extend the functional profiling beyond resistance and mobility, we characterized the gut virulome profiles of wild *Erinaceus amurensis* across the three groups, to assess whether developmental and sex-associated variation observed in the resistome and mobilome extends to virulence-related functional repertoires.

Alpha diversity

Alpha diversity indices showed a marginal trend toward inter-group differences for both the Shannon index (H=4.993, nominal *P*= 0.082, ε^2=0.299) and Simpson index (H=5.705, nominal *P*= 0.058, ε^2= 0.371). Foetal samples exhibited lower evenness than adult groups, suggesting a simplified virulence repertoire in the foetal gut, consistent with the pattern observed for MGEs and resistance genes.

Beta diversity and community structural differentiation

Principal coordinates analysis (PCoA) based on Bray–Curtis distances revealed clear compositional structuring of the gut virulome across groups (PC1=62.31%,PC2=5.66%; Figure 4k). Female samples clustered separately from the other two groups, while male and foetal samples showed substantial overlap. Analysis of similarities (ANOSIM) showed a pattern of global divergence among groups (R=0.583, nominal P= 0.001, 999 permutations), as

visualized in the between-group rank distance boxplot (Figure 4l).

Functional classification of the virulome

At the Level 1 functional category (Figure 4a), offensive virulence factors dominated the gut virulome across all three groups. The relative abundance of offensive factors was higher in females (~61%) compared with males and foetuses (~51%), and this difference approached nominal statistical significance (H=5.626, nominal P=0.060, ε^2=0.363). This elevation in female samples was mainly driven by increased abundance of secretion system–related genes at Level 2. In contrast, defensive virulence factors associated with antiphagocytosis were depleted in females.

At the Level 2 functional subclass (Figure 4b), the overall functional landscape was further resolved, with abundance patterns across major categories (adherence, iron uptake, secretion system, toxin, invasion, etc.) visualized in the heatmap (Figure 4f). Subtype-level breakdowns showed distinct compositional shifts within key functional modules:

Adherence-related factors (Figure 4c): Polar flagella, type 1 fimbriae, type IV pili, and capsule components dominated the adherence profile, with group-specific enrichment patterns.

Secretion systems (Figure 4d): Diverse secretion apparatus (T2SS, T3SS, T4SS, T6SS, Dot/Icm) were detected across groups, with differential abundance driving the Level 1 offensive factor differences.

Total toxin abundance was comparable across groups, but the composition of toxin subclasses showed apparent restructuring among developmental stages, with a nominal *P* value of 0.043 (H= 6.286, ε^2=0.429).Beta-hemolysin/cytolysin, alpha-hemolysin, and RTX toxin represented the most abundant toxin classes.

Differentially abundant virulence factors

Ten virulence factors with nominal differential abundance among groups were observed via Kruskal–Wallis H tests (nominal $P<0.05$; Figure 4m). Specifically, type III secretion system, LOS (lipooligosaccharide), MOMP (major outer membrane protein), and T4SS effectors showed nominally higher abundance in females. T6SS, capsule, and lateral flagella showed nominally higher abundance in foetal samples. Polar flagella, capsule variants, and the *PhoP* regulatory system were enriched in adult males.

Core and unique virulome components

Venn analysis revealed that 611 virulence factors (55.19% of the total) were shared across all three groups, constituting the core gut virulome of *Erinaceus amurensis*.(Figure 4g). Males harboured the highest number of group-unique virulence factors (124,11.20%), whereas foetuses and females had far fewer unique entries (14 and 3, respectively).

Group-specific virulome composition was further illustrated by pie charts of the top abundant factors within each group:

Adult males(Figure 4h): The male virulome was highly diverse, with dominant components including *Pht*, *EVP*, *Dot/Icm* type IVB secretion system, *SodC*, and *Lvh* type IVA secretion system, alongside multiple toxin and adherence factors.

Adult females(Figure 4i): The female virulome was overwhelmingly dominated by BFT (Bacteroides fragilis toxin) subtype 1, which accounted for 91.68% of the total, with only minor contributions from other factors such as *EspP* and *YapE*.

Foetuses(Figure 4j):The foetal virulome was dominated by secretion system–related

components, with *Trw* type IV secretion system (41.65%) and *Rvh* T4SS effectors (41.10%) together representing the majority, followed by type IV pili (17.25%).

The Circos plot (Figure 4n) further visualized the distribution and relative abundance of key virulence determinants across groups, highlighting the distinct enrichment patterns of secretion systems, adherence factors, and toxins in each developmental stage.

Overall, the gut virulome exhibits clear sex-and developmental-stage-associated compositional shifts that parallel the remodeling patterns observed in the resistome and mobilome, supporting a coordinated response of microbial functional repertoires to host physiological status. Notably, metagenomic read-based abundance reflects only homologous DNA coverage and cannot represent *in situ* gene expression or actual pathogenic risk for the host.

Discussion

4.1 Independent interpretation of each functional repertoire

It should be noted that observed differences in resistance-, mobility-, and virulence-related gene profiles may be partially confounded with underlying taxonomic turnover of the gut microbial community across groups. A broad phylum-level compositional overview of the gut bacterial community is provided in Supplementary Figure S1; a more systematic taxonomic analysis of gut microbiota in this host species has been reported in our previous work[Error! Unknown switch argument.].

4.1.1 Antibiotic-resistance genes

Consistent with other wild-mammal gut resistome studies[Error! Unknown switch argument.Error! Unknown switch argument.], intrinsic efflux-dominated ARG profiles prevailed in wild *Erinaceus amurensis*. Foetuses showed higher relative abundance of target-replacement and antibiotic-inactivation mechanisms without acquiring unique resistance genes, demonstrating that intergroup variation arose via abundance reshuffling within shared gene repertoires. Partial compositional separation indicated host-related physiological status exerts partial but not deterministic effects on the resistome. Owing to the absence of matched mother-foetus pairs, definitive conclusions concerning vertical transmission cannot be drawn. Homology-based resistome annotation predicts genetic potential and does not guarantee phenotypic antibiotic resistance, which requires culture-based validation.

4.1.2 Metal and biocide-esistance genes

Wild *Erinaceus amurensis* gut microbiota carry genetic repertoires for coping with metal and biocide stressors. Several metal-resistance genes were enriched in female adults. Without paired environmental reference samples, we cannot disentangle host-intrinsic filtering from external environmental acquisition; Based on current research findings environment-driven selection remains a plausible but unproven hypothesis[Error! Unknown switch argument.Error! Unknown switch argument.]. Sequence-based functional predictions require phenotypic assays for confirmation.

4.1.3 Mobile genetic elements

Insertion-sequence-based transposition constituted the primary mode of genetic mobility. Group-level differentiation was more pronounced for MGEs than for ARGs, implying stronger compositional shifts in mobility-related machinery correlated with developmental-stage-associated variation relative to resistance-gene cargo.The foetal mobilome was characterised by transposase predominance alongside depletion of recombinase and integrase-encoding components, pointing toward a reliance on short-range

IS-mediated transposition with limited capacity for large-scale homologous recombination or conjugative exchange.Similar coupling between resistance and mobility has been reported in human wastewater, livestock systems, and polluted aquatic environments[Error! Unknown switch argument. REF _Ref21166 \r \h * MERGEFORMAT Error! Unknown switch argument.].Short-read metagenomics cannot resolve chromosomal versus plasmid locations or physical linkage between MGEs and functional genes, nor can homology annotation confirm active transposition.Notably, all results regarding mobility-associated genes reflect differences in the relative abundance of annotated gene categories, and do not represent direct measurements of active transposition or horizontal gene transfer.

4.1.4 Virulence-factor profiles

Offensive virulence determinants dominated the gut virulome. Female adults exhibited an offensive-biased virulence signature marked by elevated secretion-system components, while foetuses were enriched for factors mediating interbacterial competition. All female-unique virulence factors were T4SS/T4P-related, molecular machineries with dual roles in virulence effector delivery and conjugative horizontal gene transfer. All ordination results were interpreted descriptively; detection of virulence-gene homologues does not correspond to active pathogenic phenotypes.

4.2 Integrated cross-module interpretation

4.2.1 Cross-module signatures of simplified functional gene repertoires in foetuses

Integrating the ARG, BacMet, MGEs and VFDB datasets, consistent cross-module patterns were identified in foetal samples. Foetuses exhibited reduced alpha-diversity and fewer unique genes across all four functional datasets, and their overall functional gene pools could be regarded as a depauperate subset of adult gene repertoires. At the MGE level, genes associated with recombination and integration were depleted in foetal samples. The foetal virulome was dominated by genes mediating interbacterial bacterial competition, whereas no foetus-specific resistance genes were detected in the resistome, with only abundance shifts of partial functional pathways. Such cross-module consistencies indicate an overall simplified status of the functional gene repertoire in foetal gut microbiota.

4.2.2 Potential co-enrichment of T4SS-associated virulence factors and metal-resistance genes in adult females

Concurrent enrichment of T3SS/T4SS virulence factors and elevated abundance of metal-resistance genes including *glpF* and *frnE* were observed in adult female samples. All unique virulence factors detected in females belonged to the T4SS/T4P mobile-element family. Spearman's rank-order correlation based on sample-wise abundance matrices revealed nominal positive correlations between T4SS components and *glpF/frnE*. RTX toxin genes were strongly positively correlated with metal-resistance genes, whereas capsule-related defensive virulence factors showed negative correlations with metal-resistance genes, indicating an offensiv-defensive functional trade-off at the community level. Three non-mutually-exclusive mechanistic hypotheses were proposed to interpret these statistical associations: (1) co-resistance, where T4SS resides on conjugative plasmids or *ICEs* and is physically linked to metal-resistance genes[Error! Unknown switch argument.]; (2) cross-resistance via shared multi-stress efflux machineries[Error! Unknown switch argument.]; (3) co-regulation, in which metal-sensing two-component systems modulate both metal tolerance and virulence gene expression[Error! Unknown switch argument.].

4.2.3 Mobile genetic elements are more responsive to host physiological shifts than resistance and virulence repertoires

Among the four functional datasets, MGEs exhibited the highest magnitude of compositional differentiation across groups. By contrast, almost no intergroup differences were detected at the ARG gene-family level, and only partial group-specific variations were observed for virulence factors, with most core virulence genes shared by all three groups. These observations suggest that mobile genetic elements, as key machinery for microbial genome plasticity, may be more responsive to host physiological and developmental state, while core resistance and virulence genes remain relatively conserved. Foetal samples showed predominance of IS-driven short-range transposition-associated genes and reduced abundance of genes related to large-scale homologous recombination and conjugation. This pattern is consistent with reduced horizontal transfer potential of resistance and virulence genes during early-life stages, but does not represent direct functional evidence. The clinical class-1 integron *IntI1* was present at extremely low abundance across all samples, while insertion sequence-related genes accounted for a larger proportion of annotated MGEs. The relative contribution of different MGE types to horizontal gene transfer in this population requires further investigation.

4.2.4 Greater numbers of unique functional elements within mobilome and virulome in adult males

Venn-diagram analyses demonstrated that male adults harboured the largest number of unique functional elements in both the mobilome and virulome datasets, which was consistent with the trend of higher ARG diversity in males. We hypothesise that this pattern may be associated with broader home-ranging behaviour and greater exposure to environmental microorganisms in male *Erinaceus amurensis*.

4.2.5 Implications for biosecurity and health management in medicinal hedgehog cultivation

The gut resistome, mobilome and virulome profiles of wild *E. amurensis* obtained in this study represent the natural microbial background without artificial feeding and antibiotic intervention, which can serve as a critical baseline reference for subsequent artificial domestication and large-scale breeding. Once hedgehogs are transferred to captive farming systems, these baseline data can be used to identify abnormal enrichment of resistance genes and virulence factors caused by breeding conditions, feed and antimicrobial use. The co-occurrence patterns between metal resistance genes, MGEs and virulence factors also remind us that continuous surveillance of gut resistance and virulence elements is necessary during medicinal hedgehog farming, to mitigate potential biosecurity risks and support sustainable healthy development of the special medicinal animal breeding industry. Notably, all findings are based on wild populations, and further validation in captive-bred individuals is required.

Limitations

This study represents an exploratory metagenomic survey of gut functional gene repertoires in wild Erinaceus amurensis. Several constraints arising from field-sampling conditions and analytical pipelines should be acknowledged:

1.Sample size and biological independence: A total of 13 samples were included, with only $n = 3$ foetal specimens. As all foetuses were derived from a single litter, they cannot be

treated as fully independent biological replicates. All inter-group comparisons therefore yield observational patterns. To mitigate statistical bias from small sample size, effect sizes and bootstrap confidence intervals were reported for cautious interpretation, and strong causal inferences were avoided throughout this work.

2.Constraints in sampling design: Due to practical challenges in wildlife field sampling, matched mother-foetus pairs, post-natal longitudinal specimens and environmental reference samples (soil, food and habitat-associated microbiota) were not obtained. Using the present dataset, the relative contributions of vertical transmission, host-intrinsic filtering and environmental input cannot be quantitatively disentangled.

3.Technical constraints of metagenomics: Short-read sequencing and homology-based annotation only provide abundance information of DNA fragments, and cannot directly reflect *in-situ* gene transcription, protein activity or phenotypic traits such as antimicrobial resistance, pathogenicity and transposition activity. Furthermore, sample-level Spearman correlations only reflect community-wide co-occurrence patterns and do not prove physical linkage at the contig or MAG level; co-localisation of MGE-resistance-virulence cassettes within individual microbial genomes cannot be directly confirmed.

4.Statistical constraints: Kruskal-Wallis tests were applied for exploratory candidate-feature screening in this study. Nominally significant P-values from multiple tests were not adjusted by FDR, which introduces potential false-positive risk. To avoid over-interpretation, statistical significance alone was not used as the exclusive filtering criterion; we prioritised interpretation of features with plausible biological context. Ordination outputs including PCoA and ANOSIM are presented solely as descriptive compositional patterns rather than confirmatory evidence.

5.Inseparable confounding variables: Host developmental stage, sex, maternal effects, diet and baseline microbial community composition are mutually confounded. The current study design cannot fully partition their individual effects; hence, developmental stage is not claimed as the sole predominant driver.

Future perspectives

1. Sample-cohort expansion: Enlarge sampling cohorts, collect foetuses from multiple litters and matched mother-offspring pairs, together with corresponding environmental controls including habitat soil and food-associated microbiota, to disentangle environmental-input effects from host-intrinsic filtering.

2. Technological improvement: Apply long-read metagenomics to generate high-quality MAGs, resolve physical linkage among ARG-MGE-virulence cassettes, identify their microbial hosts, and test co-selection-related hypotheses.

3. Multi-omics integration and functional validation: Integrate metatranscriptomics to evaluate in-situ expression of resistance-, MGE-and virulence-related genes. Perform bacterial isolation and culture-based assays including antimicrobial susceptibility testing, virulence phenotyping and transposition-activity measurement, to advance DNA-based gene-inventory evidence towards phenotypic biological evidence.

4. Ecological surveillance extension: Conduct multi-site wildlife sampling across heavy-metal-pollution gradients to characterise responses of gut resistome-mobilome-virulome in wild mammals, and generate baseline datasets for wildlife microbial-risk surveillance under the One-Health framework.

5. Farming practice-oriented research: Carry out comparative studies between wild and captive-bred hedgehog populations, explore the effects of artificial diet, breeding environment and antimicrobial exposure on gut resistome and virulence profiles, and formulate targeted intestinal health management schemes for medicinal hedgehog cultivation.

Methods

Study area and sample collection

Wild *Erinaceus amurensis* individuals were captured using live-cage trapping from April to July 2023 in Sishan Forest Farm, Jidong County, Heilongjiang Province, China (131°08′18″E–131°54′47″E, 44°51′12″N–44°58′53″N), a typical forest-farmland mosaic region. All healthy adult individuals were live-trapped for non-invasive fecal sample collection. Notably, one pregnant female was found in a moribund state due to spontaneous wild traumatic injury at capture. This individual was transported to the laboratory for emergency assessment and died shortly after arrival. To avoid wasting valuable biological material from this rare wild pregnant specimen, fecal samples of this maternal female were first collected upon arrival, and gastrointestinal tract contents from three late-gestation intrauterine foetuses were subsequently collected via anatomical dissection under ethical approval. Importantly, all foetuses were retrieved directly from the uterus with no exposure to the external environment or maternal milk, eliminating confounding from postnatal microbial colonization and lactation for early-developmental baseline profiling. After all sample collection was completed, all healthy captured adult individuals (excluding the moribund pregnant female) were released back to their original field habitat.

Therefore, a total of 13 individuals were included in the present study, including five adult males, five adult females, and three foetal individuals derived from this single pregnant female. Samples were grouped by sex and developmental stage: adult male group (M), adult female group (FM), and foetal group (FE). It should be clarified that the three foetal samples originated from a single litter and thus were not fully independent biological replicates; this limitation was taken into consideration in all subsequent statistical analyses and result interpretation.

Fresh fecal material and foetal gastrointestinal contents were collected following standardized protocols; non-invasive fecal sampling for adult individuals was performed as previously described[Error! Unknown switch argument.]. All samples were immediately placed on dry ice, transported to the laboratory, and stored at −80 °C until subsequent experimental processing. All sampling procedures were approved by the Heilongjiang Provincial Forestry Authority and complied with relevant wildlife protection regulations and ethical guidelines. No wild animals were actively hunted, sacrificed, or euthanized for the purpose of this study, and opportunistic sampling of the moribund individual was conducted under the approved ethical protocol.

DNA extraction and library construction

Total community genomic DNA was extracted using the Mag-Bind® Soil DNA Kit (Omega Bio-tek, USA) following the manufacturer’s instructions. DNA concentration and purity were assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA), and DNA integrity was verified by 1% agarose gel electrophoresis. Qualified genomic DNA was fragmented to an average insert size of approximately 350 bp using a Covaris M220 ultrasonicator. Paired-end metagenomic sequencing libraries were constructed using the NEXTFLEX Rapid DNA-Seq Kit (Bioo Scientific, USA), including adapter ligation, PCR

amplification, and magnetic bead purification steps.

Metagenomic Sequencing and Bioinformatic Analysis

Metagenomic sequencing was performed using the Illumina NovaSeq 6000 platform (Illumina, USA) with a PE150 paired-end strategy. A total of 87.22 Gb of raw metagenomic data was generated from the 13 samples. All raw sequencing data were processed through a standardized three-step pipeline including quality control, host sequence removal, *de novo* assembly, gene prediction, non-redundant gene catalog construction, and abundance quantification for subsequent functional and taxonomic analyses.

Step 1: Quality control and host sequence removal. Raw reads were quality-filtered using fastp (v0.20.0) to remove low-quality reads with Phred quality values < 20, ambiguous bases, and adapter-contaminated sequences. Trimmed reads shorter than 50 bp were discarded. High-quality clean reads were further mapped to the reference genome of *Erinaceus amurensis* using BWA (v0.7.17) to eliminate host-derived sequences. After filtering, 86.17 Gb of clean data was retained, with individual sample data sizes ranging from 6.10 Gb to 7.76 Gb and an average of 6.63 Gb.

Step 2: De novo assembly and open reading frame prediction. Clean reads from each sample were independently assembled *de novo* using MEGAHIT (v1.2.9), and only contigs with a minimum length of 300 bp were retained. Open reading frames (ORFs) were predicted from the assembled contigs using Prodigal (v2.6.3), and ORFs shorter than 100 bp were removed. Importantly, the assembled contigs in the present study served only as intermediate raw sequences for gene prediction and were not used for gene abundance quantification or community statistical analyses, exclusively supporting the construction of the subsequent non-redundant gene catalog.

Step 3: Non-redundant gene catalog construction and gene abundance quantification. All qualified ORFs predicted from the 13 samples were integrated and clustered using CD-HIT (Cluster Database at High Identity with Tolerance) (v4.6.1) with 90% sequence identity and 90% alignment coverage thresholds to construct a project-specific non-redundant gene catalog. Clean reads of each sample were realigned to the non-redundant gene catalog using SOAPaligner (v2.21) with a minimum sequence identity of 95% to calculate and normalize gene abundance matrices. All community-level analyses in this study, including α / β diversity analysis, PCoA, ANOSIM, Venn analysis, Kruskal–Wallis H test, and correlation analysis, were entirely performed based on the gene abundance matrix of the non-redundant gene catalog. No direct alignment of clean reads against external databases or contig-based statistical analysis was implemented.

Nucleotide sequences of the non-redundant gene catalog were translated into protein sequences andfunctionally annotated against multiple public databases using DIAMOND (v2.0.13) with an E-value cutoff of $\leqslant 1 \times 10^{-5}$ **Error! Unknown switch argument.** REF _Ref23325 \r \h * MERGEFORMAT REF _Ref23328 \r \h * MERGEFORMAT REF _Ref23335 \r \h * MERGEFORMAT REF _Ref23338 \r \h * MERGEFORMAT REF _Ref23341 \r \h * MERGEFORMAT**Error! Unknown switch argument.**. Assembly statistics for each sample, including total contig number, total assembled length, N50, N90, and sequence length distribution, are summarized in Table S1 to fully evaluate the assembly quality.

Functional and resistance gene annotation

Functional annotation of non-redundant protein sequences was performed using DIAMOND (v2.0.13; E-value $\leqslant 1 \times 10^{-5}$) **Error! Unknown switch argument.**.General biological

functions were annotated against the KEGG database. Antibiotic resistance genes (ARGs) were identified and classified based on the Comprehensive Antibiotic Resistance Database (CARD)[Error! Unknown switch argument.]. Metal and biocide resistance genes were annotated using the Antibacterial Biocide and Metal Resistance Genes Database (BacMet), with annotations classified at the gene, compound, and functional class (metal resistance, biocide resistance, dual resistance) levels[Error! Unknown switch argument.]. Mobile genetic elements (MGEs), including insertion sequences (IS), transposons, integrons, conjugative plasmids, genomic islands and phage-associated elements, were annotated using an in-house curated MGE reference database (MGEs90). This database integrates MGE-related protein sequences from the NCBI-NR database and class 1 integron integrase IntI1 sequences from the I-VIP database. Detailed construction and annotation procedures are provided in Supplementary Text S1. (see Supplementary Text S1 for construction details). Virulence factor annotation was strictly performed using only the Virulence Factor Database (VFDB) core dataset, and the predicted dataset was excluded to ensure annotation accuracy and reliability[Error! Unknown switch argument.]. For each predicted gene, only the best alignment hit was retained as the final annotation result. The relative abundance of each functional category was calculated by summing the normalized abundance of all corresponding annotated genes.

Statistical analysis

All statistical analyses were performed in R (v4.2.2). The normality of all datasets was firstly tested to select appropriate statistical methods. Given the limited sample size and the commonly non-normal distribution of metagenomic gene abundance data, we uniformly adopted the non-parametric Kruskal-Wallis H test for all inter-group differential analyses of ARGs, metal/biocide resistance genes, MGEs, and virulence factor profiles. Considering the non-independent attribute of foetal samples from a single litter, we additionally report epsilon-squared (ε^2) effect sizes with 10,000 bootstrap-derived 95% confidence intervals (95% CIs) to quantify the magnitude of inter-group differences. All reported P values are nominal and uncorrected for multiple testing.

Non-metric multidimensional scaling (NMDS) and analysis of similarities (ANOSIM) were applied to evaluate inter-group dissimilarity of microbial functional profiles. NMDS ordination with stress value < 0.2 was considered statistically reliable, and ANOSIM $R > 0$ with nominal $P < 0.05$ indicated a pattern of group separation. Alpha diversity indices (Shannon, Simpson) of resistome and virulome were calculated at the functional feature level. A nominal P-value < 0.05 was defined as the threshold for exploratory screening of differential features in all analyses[Error! Unknown switch argument.Error! Unknown switch argument.].

Data availability

All raw metagenomic sequencing data generated in this study have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database under BioProject accession number PRJNA1146738, and are publicly accessible as of the date of submission. The one-to-one correspondence between individual samples, experimental groups, and SRA run accessions is provided in Table 1. Custom bioinformatic scripts and intermediate analytical datasets supporting the findings of this study are available from the corresponding author upon reasonable request.

Table 1 should be placed here.

Ethics and biosafety statement

All sampling procedures of wild Erinaceus amurensis were reviewed and approved by the Experimental Animal Welfare Ethics Committee of Mudanjiang Normal University prior to field work (approval No. IACUC-MNU-2023-15, date: March 15, 2023). An official supplementary confirmation of ethical approval for academic publication was issued on January 26, 2026 (approval No.: 2026-011).

Biosafety Statement: All experimental procedures were conducted in an Animal Biosafety Level 2 (ABSL2) laboratory in compliance with the Regulations on Biosafety Management of Pathogenic Microorganism Laboratories in China.

Declarations

Funding

This study was supported by the Young Academic Backbone Project of Mudanjiang Normal University (MNUQN202305); the Prevention and Control of Emerging and Major Infectious Diseases-National Science and Technology Major Project (2025ZD01900100); the National Natural Science Foundation of China (32500447); the Major Project of Guangzhou National Laboratory (GZNL2023A01001); the Heilongjiang Provincial Natural Science Foundation Joint Fund Key Project: Research into AI and eDNA-based High-Precision Monitoring and Intelligent Prediction Technologies for the Aquatic Ecosystem of Mirror Lake (ZL2025C004); the sub-project under the National Key Research and Development Programme: Establishment of a Common Black-striped Field Mouse Population (2023YFF0724603-4). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] Makumbi J P, Leareng S K, Bezuidt O K, et al. Persistence of high-risk antimicrobial resistance genes in extracellular DNA along an urban wastewater-river continuum[J]. Cell Reports, 2026, 45(4): 117128. DOI:10.1016/j.celrep.2026.117128.

[2] Zhu L, Gao M, Li J, et al. Why Antimicrobial Resistance Does Not Perpetually Expand in an Antibiotics-Free Environment: Insight from Quorum Sensing[J]. Environmental Science & Technology, 2026, 60(15): 11653-11664. DOI:10.1021/acs.est.5c16848.

[3] Singh N, Golicha R O, Thakur C S, et al. Aeromonas in South Asia: genomic insights into an environmental pathogen and reservoir of antimicrobial resistance[J]. Nature Communications, 2026, 17(1). DOI:10.1038/s41467-026-68712-w.

[4] Despotović M, de Nies L, Busi S B, et al. Reservoirs of antimicrobial resistance in the context of One Health[J]. Current Opinion in Microbiology, 2023, 73: 102291. DOI:10.1016/j.mib.2023.102291.

[5] Torres R T, Cunha M V, Araujo D, et al. A walk on the wild side: Wild ungulates as potential reservoirs of multi-drug resistant bacteria and genes, including Escherichia coli harbouring CTX-M beta-lactamases[J]. Environmental Pollution, 2022, 306: 119367. DOI:10.1016/j.envpol.2022.119367.

[6] Dominguez J E, Lovera R, Juliana S, et al. Rats as sources of multidrug-resistant Enterobacteriaceae in animal production environments[J]. Zoonoses and Public Health, 2023, 70(7): 627-635. DOI:10.1111/zph.13071.

[7] Vikesland P J, Garner E, Gupta S, et al. Differential Drivers of Antimicrobial Resistance across the World[J]. Accounts of Chemical Research, 2019, 52(4): 916-924. DOI:10.1021/acs.accounts.8b00643.
[8] Blanco-Peña K, Quesada-Alvarado F, Salas-González D, et al. A multidisciplinary approach to analyze the antimicrobial resistance in natural ecosystems[J]. Environmental Research, 2024, 251(Pt 1): 118549. DOI:10.1016/j.envres.2024.118549.
[9] Essack S. Antibiotic resistance and One Health: a mapping project[J]. The Lancet Global Health, 2018, 6: S27.
[10] Luo Y, Tan L, Zhang H, et al. Characteristics of wild bird resistomes and dissemination of antibiotic resistance genes in interconnected bird-habitat systems revealed by similarity of bla TEM polymorphic sequences[J]. Environmental Science & Technology, 2022, 56(21): 15084-15095.
[11] Wang Y, Qian Y, Shi C, et al. Seasonal Dynamics and Stability of Gut Antibiotic Resistance Genes in Plateau Pika (Ochotona curzoniae) and Plateau Zokor (Eospalax baileyi)[J]. Journal of Hazardous Materials, 2026: 141112.
[12] Larsen J, Raisen C L, Ba X, et al. Emergence of methicillin resistance predates the clinical use of antibiotics[J]. Nature, 2022, 602(7895): 135-141.
[13] Cross B J, Partridge S R, Sheppard A E. Impacts of mobile genetic elements on antimicrobial resistance genes in gram-negative pathogens: Current insights and genomic approaches[J]. Microbiological Research, 2025: 128340.
[14] Li L G, Xia Y, Zhang T. Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection[J]. The ISME journal, 2017, 11(3): 651-662.
[15] Gillieatt B F, Coleman N V. Unravelling the mechanisms of antibiotic and heavy metal resistance co-selection in environmental bacteria[J]. FEMS Microbiology Reviews, 2024, 48(4): fuae017.
[16] Amato K R, Yeoman C J, Kent A, et al. Habitat degradation impacts black howler monkey (Alouatta pigra) gastrointestinal microbiomes[J]. The ISME journal, 2013, 7(7): 1344-1353.
[17] Holt C C, van der Giezen M, Daniels C L, et al. Spatial and temporal axes impact ecology of the gut microbiome in juvenile European lobster (Homarus gammarus)[J]. The ISME Journal, 2020, 14(2): 531-543.
[18] Greene L K, Williams C V, Junge R E, et al. A role for gut microbiota in host niche differentiation[J]. The ISME Journal, 2020, 14(7): 1675-1687.
[19] Jia X, Li Q, Yang X, et al. Comparative analysis of the gut microbiota composition and diversity in Erinaceus amurensis from the Wandashan Mountain range area based on metagenomics[J]. Frontiers in Microbiology, 2024, 15: 1479352.
[20] Alcock B P, Huynh W, Chalil R, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database[J]. Nucleic acids research, 2023, 51(D1): D690-D699.
[21] Muloi D M, Wee B A, McClean D M H, et al. Population genomics of Escherichia coli in Livestock-keeping households across a rapidly developing urban Landscape [J]. Nature Microbiology, 2022, 7(4): 581–589.
[22] Pal C, Bengtsson-Palme J, Rensing C, et al. Antibacterial biocide and metal resistance genes database., 2014, 42, pp[J]. D737-D743. DOI: https://doi. org/10.1093/nar/gkt1252.
[23] Demian W L, Cormier O, Mossman K. Immunological features of bats: Resistance and

tolerance to emerging Viruses [J]. Trends in Immunology, 2024, 45(3): 198–210.
[24] Benavides J A, Salgado-Caxito M, Opazo-Capurro A, et al. ESBL-Producing Escherichia coli Carrying CTX-M Genes Circulating among Livestock, Dogs, and Wild Mammals in Small-Scale Farms of Central Chile [J]. Antibiotics, 2021, 10(5): 510.
[25] Treskova M, Kuhlmann A, Freise F, et al. Occurrence of Antimicrobial Resistance in the Environment in Germany, Austria, and Switzerland: A Narrative Review of Existing Evidence [J]. Microorganisms, 2022, 10(4): 728.
[26] Karami Y, López-Castilla A, Ori A, et al. Computational and biochemical analysis of type IV pilus dynamics and stability[J]. Structure, 2021, 29(12): 1397-1409.e6. DOI:10.1016/j.str.2021.07.008.
[27] Lill Z, Thongchol J, Solis D, et al. The cryo-EM structure of bacteriophage PRR1 and its role in conjugation inhibition[J]. Journal of Virology, 2026: e00515-26.
[28] Qin Z, Wang Q, Yang Q, et al. Agricultural fungicides shape soil and sediment reservoirs of multidrug-resistant fungi[J]. Nature Communications, 2026, 17: 7179.
[29] Shen S, Zhao S, He Z, et al. The Efflux–Two-Component System (TCS)–Virulence Axis Drives Resistance–Virulence Convergence in Aquatic “Superhost Precursors” under Pollution Stress[J]. Environmental Science & Technology, 2026, 60(15): 11253-11266.
[30] FasoliS, FerlizzaE, And reariG, SandriC, Dond iF, Is aniG. Noninvasives amplingmethod forur inalysisandurineprote inprofileincaptive giraffes. JournalofVeterinaryDiagnosticIn vestigation, 2021, 33 (1): 25-34.
[31] Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i890. doi:10.1093/bioinformatics/bty560
[32] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-1760. doi:10.1093/bioinformatics/btp324
[33] Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31(10):1674-1676. doi:10.1093/bioinformatics/btv033
[34] Hyatt D , Chen G L , Locascio P F , et al. Prodigal: prokaryotic gene recognition and translation initiation site identification[J]. BMC Bioinformatics, 2010, 11(1):119-119. doi:10.1186/1471-2105-11-119
[35] Noguchi H, Park J, Takagi T. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. Nucleic Acids Res. 2006;34(19):5623-5630. doi:10.1093/nar/gkl723
[36] Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28(23):3150-3152. doi:10.1093/bioinformatics/bts565
[37] Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. Bioinformatics. 2008;24(5):713-714. doi:10.1093/bioinformatics/btn025
[38] Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12(1):59-60. doi:10.1038/nmeth.3176
[39] Edalatmand A, Ta T E, Zhao C, et al. CARD: Epi-Contextualizing Antimicrobial Resistance Determinants Using Deep Learning Language Models[J]. bioRxiv, 2026: 2026.08. 14.744850.
[40] Pal,C.;Bengtsson-Palme,J.;Rensing,C.; Kristiansson, E.; Larsson, D. G. J. (2014). BacMet: antibacterial biocide and metal resistance genes database. Nucleic Acids Research, 42(D1), D737–D743. doi:10.1093/nar/gkt1252

[41] Chen L, Zheng D, Liu B, et al. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on[J]. Nucleic acids research, 2016, 44(D1): D694-D697.

[42] Kilroy C, Stephens T, Greenwood M, Wech J, Brown L, Matthews A, Patterson M, Patterson M. Improved predictability of peak periphyton in rivers using site-specific accrual periods and long-term water quality datasets. The Science of the Total Environment, 2020, 736: 139362.

[43] Clarke K R, Somerfield P J, Gorley R N. Clustering in non-parametric multivariate analyses. Journal of Experimental Marine Biology and Ecology,2016, 483: 147-155.

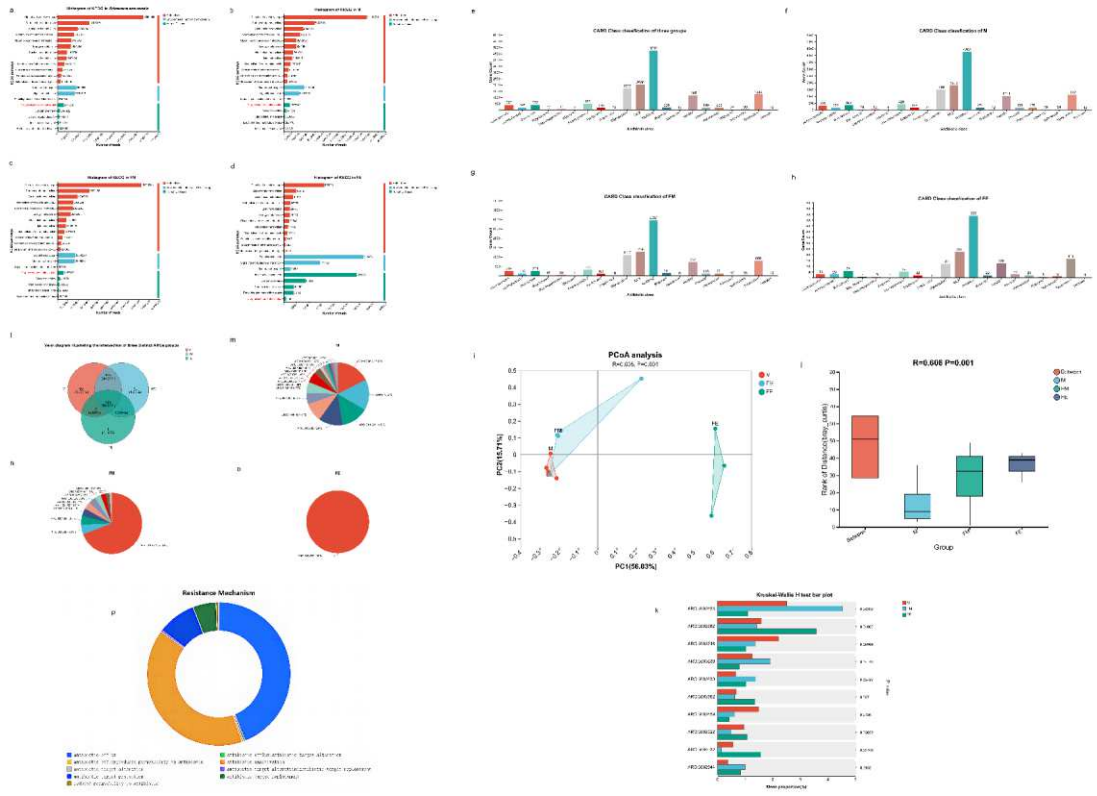


Figure 1. Profiles of antibiotic-resistance genes (ARGs) across M, FM and FE groups. **(a)** Horizontal bar chart showing the read counts of KEGG pathways across all samples. Pathways are grouped into three main categories: Metabolism (red), Environmental Information Processing (blue), and Human Diseases (green). The “Drug resistance:antimicrobial” pathway is highlighted in red text.**(b)**Horizontal bar chart illustrating KEGG pathway read counts for the M group. Colours correspond to the three major KEGG functional categories described in panel (a).**(c)**Horizontal bar chart showing KEGG pathway read counts for the FM group. Colour coding follows the same functional grouping as panel (a).**(d)**Horizontal bar chart presenting KEGG pathway read counts for the FE group. Functional-group colour scheme is consistent with panel (a).**(e)**Bar plot of CARD-classified antibiotic-resistance gene counts across all samples. X-axis represents antibiotic classes, y-axis indicates the total gene count of ARGs for each antibiotic class.**(f)**Bar plot displaying CARD-classified ARG gene counts in the M group. X-axis denotes antibiotic classes, y-axis shows the gene count of ARGs.**(g)**Bar plot for CARD-classified ARG gene counts in the FM group. X-axis indicates antibiotic classes and y-axis represents ARG gene count.**(h)**Bar plot of CARD-classified ARG gene counts in the FE group. X-axis shows antibiotic classes, y-axis

represents ARG gene count.**(i)** Principal-coordinate analysis (PCoA) based on Bray-Curtis dissimilarity of ARG profiles. PC1 and PC2 explained 56.03 % and 15.71 % of total variance, respectively. Analysis of similarities (ANOSIM) result is shown in the top-left corner (R=0.608, nominal P=0.001).Polygons enclose samples from each experimental group.**(j)**Boxplot derived from ANOSIM, showing the rank of Bray-Curtis distance for between-group and within-group (M, FM, FE) comparisons.**(k)** Kruskal-Wallis H test bar plot for ARGs with nominal inter-group differences. Bars indicate the mean relative proportion (%) of each ARO entry in M, FM and FE groups. Asterisks denote nominal $P < 0.05$, ** nominal $P < 0.01$.**(l)**Venn diagram depicting shared and unique ARGs among M, FM and FE groups. Numbers and percentages represent the count and proportion of ARGs in each partition.**(m)**Pie chart illustrating the relative proportion of dominant ARO entries in the M group. Each sector corresponds to one ARO accession with its percentage annotated.**(n)**Pie chart showing the relative proportion of dominant ARO entries in the FM group. Sectors represent individual ARO accessions with percentage values labelled.**(o)**Pie chart presenting the relative proportion of ARO entry in the FE group.**(p)**Donut chart illustrating the classification of ARG resistance mechanisms. Different colours represent distinct antibiotic resistance mechanisms annotated by the CARD database.

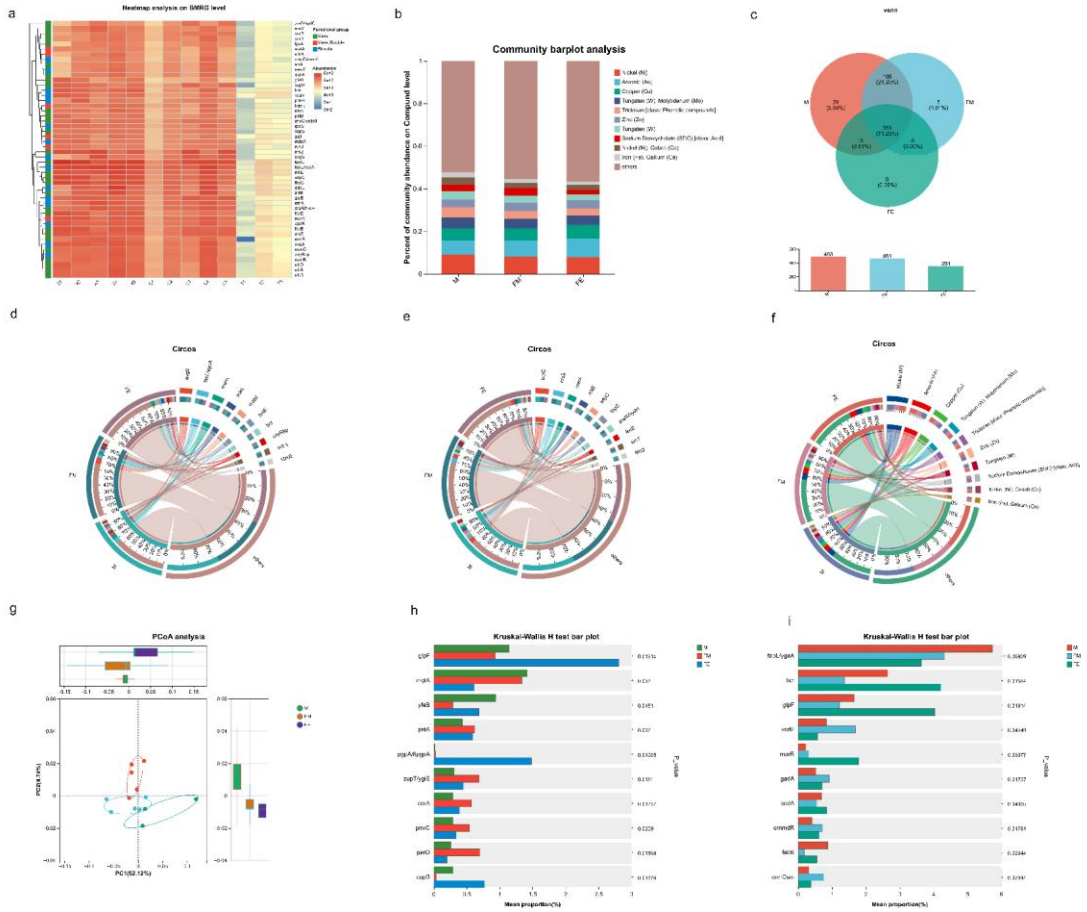


Fig2. Overview of metal-and biocide-resistance genes (BMRGs) in the M, FM and FE groups. **(a)**Heatmap illustrating the abundance profiles of BMRGs across individual samples. The colour gradient corresponds to gene abundance. The sidebar on the right classifies genes into three functional groups: Metal, Biocide, and Metal/Biocide. Hierarchical clustering is performed on the row side. **(b)**Stacked bar plot showing the community-level relative abundance of

metal- and biocide-related compounds. The y-axis represents the percentage of community abundance at the compound level; each colour denotes one compound category. **(c)**Venn diagram displaying shared and unique BMRGs among the three groups. Numbers and percentages indicate the count and proportion of genes in each section. The bar chart below summarises the total detected BMRG numbers for each group. **(d)**Circos plot showing the distribution of biocide-resistance genes across groups. Outer arcs represent sample groups, right-side labels are biocide-resistance gene names, and inner ribbons reflect the proportional contribution of each gene to different groups. **(e)**Circos plot visualising the distribution of metal-resistance genes in M, FM and FE groups. Outer arcs stand for experimental groups, right labels correspond to metal-resistance genes, and inner ribbons indicate gene-group proportional relationships. **(f)**Circos plot depicting the distribution profiles of resistance-related compounds across three groups. Outer arcs represent sample groups, right-side labels represent different metal- and biocide-associated compounds, and inner ribbons show compound contribution proportions among groups. **(g)**Principal-coordinate analysis (PCoA) based on BMRG composition. PC1 and PC2 account for 62.12 % and 4.74 % of total variation, respectively. Ellipses represent 95 % confidence intervals for each group. Marginal boxplots show the distribution of sample scores along PC1 and PC2 axes. **(h)**Kruskal-Wallis H-test bar chart for metal-resistance genes with nominal inter-group differences. Bars represent the mean relative proportion (%) of each gene in M, FM and FE groups. Asterisks indicate nominal $P < 0.05$. **(i)**Kruskal-Wallis H-test bar chart for biocide-resistance genes showing nominal group-wise variations. Asterisks denote nominal P levels (* nominal $P < 0.05$; ** nominal $P < 0.01$).

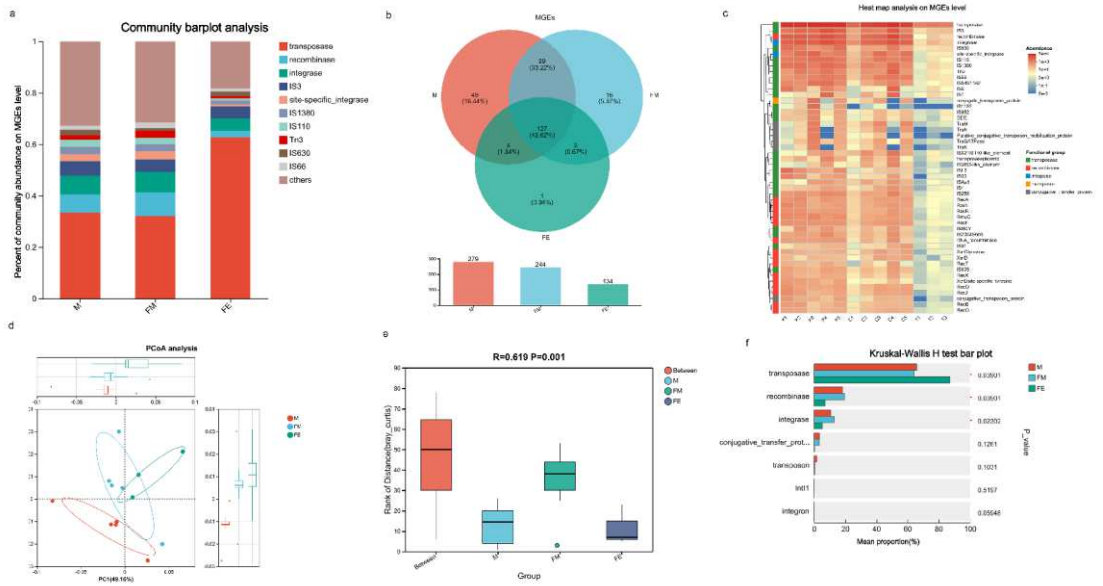


Fig 3. Figure 3 Mobile-genetic-element (MGE) profiles in gut microbiota of wild Erinaceus amurensis.**(a)**Stacked bar plot showing relative community abundance of major MGE functional categories across three groups. Groups: M, adult male; FM, adult female; FE, foetus.**(b)**Heatmap displaying normalized abundance patterns of MGE subtypes across samples. Colour gradient represents gene abundance.**(c)**Venn diagram showing shared and group-unique MGE entries among M, FM and FE groups. Bar plot below the Venn diagram shows the total number of detected MGE features within each group. Values represent counts

and percentage proportions of total detected MGE homologues.**(d)**Principal-coordinate analysis (PCoA) based on Bray-Curtis dissimilarity of MGE profiles. Ellipses denote 95 % confidence intervals for each group. Marginal boxplots show sample-score distributions along PC1 and PC2 axes.**(e)**ANOSIM rank-distance box plot for MGE community dissimilarity. "Between" indicates pairwise dissimilarity across different groups; M, FM, FE denote within-group pairwise distances for each cohort. ANOSIM statistic = 0.619, nominal P= 0.001(999 permutations).**(f)**Bar plot of nominally differential MGE categories derived from Kruskal-Wallis H-tests. Bars represent mean relative proportion (%), nominal P-values are labelled on the right side.

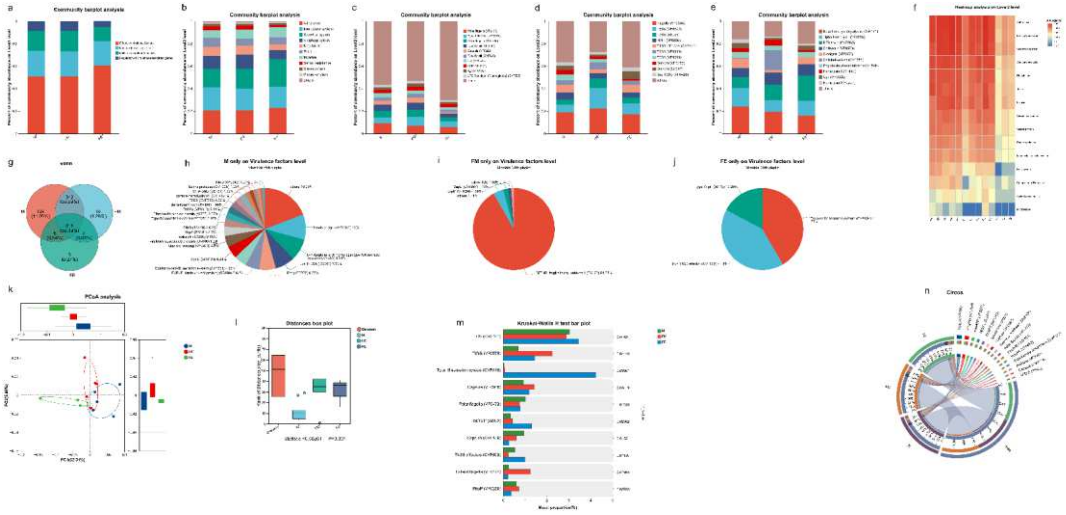


Fig 4. Sex-and developmental-associated remodelling of gut bacterial virulome profiles in wild *Erinaceus amurensis* .**(a)** Stacked bar plot showing relative community abundance of Level-1 virulence-factor categories, including offensive virulence factors, nonspecific virulence factor, defensive virulence factors and regulation of virulence-associated genes. Groups: M, adult male; FM, adult female; FE, foetus. **(b)** Stacked bar plot illustrating relative abundance of all Level-2 virulence-factor functional categories across three groups. **(c)** Stacked bar plot depicting the relative abundance of adherence-related Level-2 virulence-factor sub-categories. **(d)**Stacked bar plot for secretion-system-associated Level-2 virulence-factor sub-categories. **(e)** Stacked bar plot showing compositional profiles of toxin-related Level-2 virulence factors. **(f)**Heatmap displaying normalized abundance patterns of Level-2 virulence-factor functional categories across all individual samples. Colour gradient represents gene abundance. **(g)**Venn diagram showing shared and group-unique virulence factor entries among adult male (M), adult female (FM), and foetal (FE) groups. Values represent counts and percentage proportions of total detected VFDB homologues. **(h)** Principal-coordinate analysis (PCoA) based on Bray-Curtis dissimilarity of virulence-factor profiles. Ellipses denote 95 % confidence intervals for each group. Marginal boxplots show distribution of sample scores along PC1 and PC2 axes. **(i)** ANOSIM rank-distance box plot for virulence-factor community dissimilarity. "Between" indicates pairwise dissimilarity across different groups; M, FM, FE denote within-group pairwise distances for each cohort. ANOSIM statistic = 0.583, nominal P = 0.001 (999 permutations). **(j)** Pie chart showing composition of virulence factors uniquely detected in the adult-male group. **(k)**Pie chart showing composition of virulence factors uniquely detected

in the adult-female group. **(l)**Pie chart showing composition of virulence factors uniquely detected in the foetal group.**(m)** Bar plot of nominally differential virulence-factor features derived from Kruskal-Wallis H-tests. Bars represent mean relative proportion (%), nominal P-values are labelled on the right side. **(n)** Circos plot visualizing the relative abundance distribution of major virulence-factor homologues across M, FM and FE groups. Outer circular tracks represent sample groups, inner ribbons link virulence-factor entries to corresponding host groups according to their relative abundance.

Supplementary Files

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

- Table1SRAaccessioninformationforallmetagenomicsamples.doc
- TableS1Metagenomicdenovoassemblystatisticsacrossthe13studysamples.csv
- graphicalabstract.tif
- SupplementaryTextS1.ConstructionoftheinhousecuratedMGEs90referencedatabaseformobilegeneticelementannotation.docx
- References.docx
- SupplementaryFigureS1.Phylumlevelcompositionofgutbacterialcommunitiesacrossthefoetaladultfemaleandadultmalegroups..pdf
- Highlights.doc