npj | biofilms and microbiomes

Article

Published in partnership with Nanyang Technological University

https://doi.org/10.1038/s41522-025-00705-x

Extensively acquired antimicrobial-resistant bacteria restructure the individual microbial community in post-antibiotic conditions

Jae Woo Baek[1], Songwon Lim[1], Nayeon Park[1], Byeongsop Song[1], Nikhil Kirtipal[1], Jens Nielsen[2,3], Adil Mardinoglu[4,5], Saeed Shoaie[4], Jae-il Kim[1], Jang Won Son[6], Ara Koh[7] & Sunjae Lee[1] ✉

In recent years, the overuse of antibiotics has led to the emergence of antimicrobial-resistant (AMR) bacteria. To evaluate the spread of AMR bacteria, the reservoir of AMR genes (resistome) has been identified in environmental samples, hospital environments, and human populations, but the functional role of AMR bacteria and their persistence within individuals has not been fully investigated. Here, we performed a strain-resolved in-depth analysis of the resistome changes by reconstructing a large number of metagenome-assembled genomes from the gut microbiome of an antibiotic-treated individual. Interestingly, we identified two bacterial populations with different resistome profiles: extensively acquired antimicrobial-resistant bacteria (EARB) and sporadically acquired antimicrobial-resistant bacteria, and found that EARB showed broader drug resistance and a significant functional role in shaping individual microbiome composition after antibiotic treatment. Our findings of AMR bacteria would provide a new avenue for controlling the spread of AMR bacteria in the human community.

Recently, the misuse and overuse of antibiotics in medicine and food production have led to the emergence of antibiotic-resistant bacteria[1,2]. For example, the increased use of antibiotics during the COVID-19 pandemic has accelerated the development of multidrug-resistant bacteria[3], posing a new threat to modern society. Importantly, the human gut microbiome serves as a reservoir for antimicrobial resistance (AMR) genes[4–6], spreading these genes to the community. Initially, monitoring the spread of AMR genes was conducted using environmental samples, such as urban sewage[7] and samples from hospital environments[8], as well as by studying normal human populations in different countries[9,10] and vertical transmission from mothers to infants[11]. Such surveillance approaches[12] have been successful in the identification of AMR reservoirs, called resistomes, in the human community; however, a deep understanding of the taxonomical origins of AMR gene carriers and the impact of AMR bacteria on the gut commensal community is lacking. Some initial in-depth metagenomic studies have found that individual resistomes can persist for at least one year[9], but the dynamic changes in AMR bacteria and their impact on the community have been poorly studied at species or strain resolution. Therefore, functional analysis of AMR bacteria in the individual gut microbiome will be a key to understand their functional niche in the human community and managing their unintended consequences in the human gut microbiomes.

Here, we performed a strain-resolved, in-depth analysis of resistome changes for the identification of the driver microbial species of resistome changes and their functional niche in a given community. To this end, we analyzed publicly available shotgun metagenomic samples of healthy adults who underwent 4-day treatments with antibiotic cocktails[13], focusing on resistome dynamics at the species/strain level through de novo assembly of the metagenome-assembled genome (MAG). Interestingly, we identified bacteria that had acquired extreme resistance (EARB), those that had acquired sporadic resistance (SARB), and non-carriers, based on their AMR gene counts. Based on further functional analysis, we found that EARB displayed broad drug resistance and a significant functional role in shaping

[1]Department of Life Sciences, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea. [2]Department of Biology and Biological Engineering, Kemivägen 10, Chalmers University of Technology, Gothenburg, Sweden. [3]BioInnovation Institute, Copenhagen, Denmark. [4]Centre for Host-Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London, London, UK. [5]Science for Life Laboratory, KTH-Royal Institute of Technology, Stockholm, Sweden. [6]Division of Endocrinology and Metabolism, Department of Internal Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea. [7]Department of Life Sciences, Pohang University of Science and Technology, Pohang, South Korea. ✉e-mail: leesunjae@gist.ac.kr

individual microbiome compositions with the consequential strain switching, validated in multiple independent cohorts, including those with recurrent urinary tract infections (RUTI) and liver cirrhosis, and preterm infants. Therefore, our findings provide insights into the distribution and enrichment of AMR genes within bacterial strains and highlight the functional significance of multi-resistant bacteria in the antibiotic-treated gut community.

Methods

De novo metagenomic assembly

We obtained 57 metagenomic samples from the original paper[13] and generated high-quality (HQ) non-human reads. These data were collected before and after a 4-day treatment with antibiotic cocktails once a day (meropenem, gentamicin, and vancomycin). To observe longitudinal changes following the cessation of antibiotics, additional samples were obtained at day 8, day 42, and day 180. Note that among the possible 60 samples, three samples were excluded due to the failure in library construction from the original study. Human contaminations were removed from metagenomic samples again using GRC38 and bowtie2 (default option; version 2.3.5.1)[14]. Host removed reads were used for MAG construction using metaWRAP (default option; version 1.3.2) pipeline[15]. First, FASTQ files were assembled by MEGAHIT (v1.1.3)[16]. The assembled files were binned using three different representative binning tools (MetaBat2, MaxBin2, CONCOCT). The initial bins were refined to obtain final MAGs. CheckM (version 1.0.12)[17] was used for quality control of MAGs, and those with >70% completeness and <5% contamination were considered HQ.

Antimicrobial resistance gene analysis

AMR genes in the assembled genome were searched by Resistance Gene Identifier software (RGI; version 5.2.1) with Comprehensive Antibiotic Resistance Database (CARD) database (3.2.2)[18] as a reference (--input_type contig --alignment_tool DIAMOND --split_prodigal_jobs –clean). MAGs that contain more than 17 AMR genes were defined as Extremely Acquired Resistant Bacteria (EARB) in this study, and MAGs containing AMR genes <17 were considered as SARB; MAGs having no AMR genes were considered as non-carriers.

To analyze AMR dynamics, we counted the number of MAGs and AMR genes at different sampling points. In addition, we tracked the average number of AMR genes per MAG (by dividing the total number of AMR genes by the number of MAG) to identify individual variations in response to antibiotic treatment. We categorized individuals into 'Susceptible' and 'Tolerant' groups by considering the timing of the peak of AMR gene increases and the degree of the increase. To confirm microbial composition displacement among the tolerant group and the distances between different time points of given individual, we used metagenomic-based operational taxonomic unit (mOTU) relative abundance, from the original paper[13], and R package *vegan* (version 2.6-4) to calculate Bray-Curtis dissimilarity between day 0 (baseline) and day 180 of all subjects. We obtained multi-dimensional scaling coordinates and calculated the displacement.

Drug classes and mechanisms of AMR genes in each MAG were obtained from the RGI analysis results. To analyze the distribution of various drug classes, the number of occurrences of each drug class was counted in EARB and SARB. To examine the differences in drug classes between EARB and SARB, the frequency of each drug class was calculated, and the number of MAGs carrying each drug class was counted. To get a prevalence score, the proportion of MAGs carrying a given drug class (carrier ratio) was multiplied by the frequency of the drug class.

To analyze the sharing of drug classes and AMR genes among EARBs and SARBs, we counted the frequency of drug classes or AMR genes at a certain taxonomic level (genus level for EARB and phylum level for SARB). Specifically, for EARB, we considered a drug class or AMR gene to be "shared" if it was found in at least three out of the four total EARB genera. Similarly, for SARB, it was considered shared if it was detected in at least six out of the eight total SARB phyla.

To explore potential correlations between specific drug classes or AMR genes and taxonomies, we calculated the frequency of each drug class within every phylum or species. This was done by dividing the number of times an AMR gene associated with a drug class appeared within a given taxonomy by the total occurrences of that taxonomy level. We performed this calculation separately for EARB and SARB. Following the frequency determination, we added one to the mean frequency value to adjust for any instances of zero occurrence, then applied a logarithmic transformation to normalize the data distribution. We visualized the results using a pheatmap to provide an intuitive understanding of the frequency and distribution patterns. We also checked the AMR signature from MAGs based on non-negative matrix factorization (NMF) technique using R package *NMF*. By setting the factorization rank to two, we extracted basis and coefficient vectors from the given NMF factors.

Taxonomy annotation of MAGs

The taxonomy of MAGs, assembled in this study, was assigned using GTDB-TK classify_wf (version 2.1.0) with release 207_v2 database[19]. A phylogenetic tree for EARB was constructed from 12 healthy adults using GTDB-TK de_novo_wf. Since all EARBs were assigned within *Enterobacteriaceae* family, following parameter was used (--bacteria --taxa_filter o__Enterobacterales --outgroup_taxon f__*Enterobacteriaceae*). The result was used for plotting a phylogenetic tree in iTOL website (https://itol.embl.de/)[20].

Relative abundance profiling of MAGs

The abundance of MAGs constructed from 12 healthy adults was quantified by the *quant* function of *salmon* program (version 0.13.1) in metaWRAP pipeline[15] and merged with the taxonomy table. Phylum 'Firmicutes_C', 'Firmicutes_A' were converted to 'Firmicutes' to remove clade annotations for clarity.

To avoid bias caused by differences in absolute abundance between samples, the abundance of the same phylum in the same host was added, and its relative abundance was calculated and added based on the sampling point. Relative abundance was recalculated to obtain the relative abundances at the phylum level for each sampling point. To avoid bias caused by differences in absolute abundance between samples, we first summed the absolute abundance of each phylum within each sample and converted it into a sample-level relative abundance. Next, we grouped these sample-level relative abundances by sampling point, summed them again for each phylum, and recalculated the phylum-level relative abundances at the sampling-point level. To compare the composition of the top-5 abundant phyla between sampling points, minor phyla, whose abundance was less than the top-5, were converted to 'others' at each sampling point. The abundance table of the EARB-containing community was further investigated at the species level and categorized into 'EARB,' 'top-5' (the top-5 abundant species, excluding EARB), and 'others'. Lastly, to generate the relative abundance of *Enterobacteriaceae* and *Veillonellaceae*, family-level taxonomy information of MAGs was categorized into '*Enterobacteriaceae*', '*Veillonellaceae*', or 'others'.

Functional analysis of MAGs

Gene prediction using *Prodigal* (default mode; version 2.6.3)[21] was conducted for further functional and community power analyses. Functions of predicted genes were annotated with KEGG using *hmmsearch* function of HMMER program (default option; version 3.3.1)[22] using pre-trained hidden Markov Models (prok90_kegg94) from Raven toolbox (https://github.com/SysBioChalmers/RAVEN/wiki/Use-Pre-Trained-HMMs). KEGG pathways/modules and the list of KEGG orthology (KO) terms were manually parsed from the KEGG database (release 103; 2022/07)[23] and used for functional analysis. For comparison between non-EARB and EARB, samples that contain EARB were selected, and the KO pathway and module table were generated by counting the number of KO genes for the pathway or module in each MAG. A significant difference pathway or module was identified using a two-sided *Wilcoxon* rank-sum test with a confidence level

of 0.95. Significant pathways were filtered based on the criteria: pathway gene coverage >30% and fold change >0.5 compared to the non-EARB. Significant modules were filtered based on the criteria: module gene coverage >0.8 and fold change >0.5 compared to the non-EARB.

Community power analysis

To investigate the functional importance of EARB in a community, a community power analysis was conducted. Genes of MAGs in the communities with EARB existing were predicted using prodigal (default mode; version 2.6.3) and KEGG annotated by *hmmsearch* function of HMMER (default option version 3.3.1). Therefore, KEGG enzyme count table was generated for each community. Functional uniqueness was measured by counting number of unique KEGG enzyme in a MAG when comparing others in the same community by leave-one-out approach. To calculate the functional importance of a microbe, the KEGG enzyme counts of a microbe were divided by the sum of the same enzyme counts of the entire community (proportion of KEGG enzymes for the microbe). The sum of the proportion of all KEGG enzymes of a microbe was considered as the community power score of the given microbe. Community power scores of three categorized microbes (EARB, SARB, non-carrier) were also calculated.

Co-enrichment analysis

Co-enrichment matrix was generated using mOTU relative abundance rarefied table downloaded from the original paper[13]. First, we choose three mOTU species corresponding to EARB (Escherichia coli, Klebsiella oxytoca, Klebsiella pneumoniae). Four non-EARB species (*Bacteroides thetaiotaomicron*, *Bacteroides ovatus*, *Bilophila wadsworthia*, *Parabacteroides distasonis*) were manually picked based on high prevalence and high community power score for comparison. Co-enrichment correlation between each selected microbes' abundance and other mOTUs' were calculated by Spearman test. mOTU species that were annotated ambiguously (annotated as 'motu_linkage_group') and showed correlation between $-0.3 <$ and < 0.3 with all selected species were excluded. The correlations that both microbes co-existed in at least 5 communities and showed significant p value (>0.05) were selected and plotted using the pheatmap R package. To count positively or negatively related mOTUs, filtering spearman correlation test by >0.3 or < -0.3 for each comparing bacteria. To verify the relationship between the relative abundance of EARB or non-EARB with mOTU richness, we sum up the relative abundance of all EARB mOTUs or non-EARB in each sample. Then we counted mOTUs, which are non-zero in relative abundance in each sample. We calculated the sum of relative abundance as log2 and then divided the samples into three categories (Rare: log2 relative abundance sum < -10, Normal: $-10 \leq$ log2 relative abundance sum < -5, Rich: $5 \leq$ log2 relative abundance sum). Correlation between relative abundance sum of EARB or non-EARB and number of non-zero mOTUs (mOTU richness) was measured by spearman's correlation coefficients.

Tracking endogenous strains among metagenomic samples

To check genomic similarity between basic EARBs, we created an ANI matrix using fastANI[24] with –matrix parameter with default options (version 1.33). For strain analysis, gene prediction was conducted for 12 EARB *E. coli* using *Prodigal*[21]. All predicted genes were concatenated and then clustered based on similarity using CD-HIT (version 4.8.1)[25] with -aS 0.9 -c 0.98 -n 10 -M 0 -d 0 -T 0 -G 0 parameters. From CD-HIT clustering information, extract the longest 20 homologous genes that are present in all 12 *E. coli* strains with the same length using cdbfasta (version 1.00). 20 homologous genes in most HQ *E. coli* genome (ERR1995253 bin.3) were extracted and built as a reference by bwa (version 0.7.17-r1188) for variant call and SNP search. Variant call was done by following steps. Paired FASTQ files from EARB *E. coli* containing samples were mapped against a previously built *E. coli* homologous gene reference. Generated bam file was sorted, indexed, and mpileup (mpileup -ugf) using samtools (version 1.9)[26]. Variants were called using bcftools (bcftools call -cv) (version 1.9)[27], and filtered to include only variants with a quality score of 20 or higher. Bam files were converted tabula form using sam2tsv version (d29b24f2b). The frequency of nucleotides in every variant point was counted using the converted tsv file. Nucleotide bases that existed in variant position <5% were discarded because of the possibility of sequencing error. SNP ratio of each variant position was counted using basic R functions. For SNP analysis for AMR gene, we did the same process except using AMR gene from ERR1995248 bin.3 (host 5 & day 8) for reference because it is the dominant *E. coli* of the host. Using the dominant strain AMR gene for variant calling, we were able to find a nucleotide ratio clustered into two in the day 4 sample. Comparing variant position and SNP information between longitudinal data, we could track the SNP ratio (SNP nucleotide frequency/total nucleotide frequency) of the variant position through sampling points. To confirm the existence of two *E. coli* strains in host 5 and the day 4 sample, all variant positions of the 54 AMR gene reference host 5 at day 4 were investigated for their sequence depth.

Validation using a different cohort dataset

To confirm the reproducibility of EARB in other antibiotic-exposed human metagenome data, we downloaded raw data (i.e., FASTQ files) of shotgun metagenomic samples, or isolate strain genome samples, from three different study dataset, such as RUTI-causing *E. coli* isolates ($n = 556$), liver cirrhosis patient ($n = 592$), and preterm infant ($n = 399$). We constructed MAG from the raw data from liver cirrhosis patient and preterm infant data and search AMR gene using same pipeline, such as MetaWRAP pipeline, as we described above. For liver cirrhosis patients data, *Megahit*[16] was used with -r parameter (for single-end input). For *E. coli* isolates data, we assembled contigs using *Megahit* then predicted genes using *prodigal* for AMR gene searching. We also investigated the *E. coli* isolates' SNP profile of variant positions, which were found in SNP analysis between the major strain and the minor strain of host 5. We calculated the AMR prevalence value and compared it with the healthy adults' EARB. To verify the community power score of EARB in live cirrhosis and preterm infant data, we selected communities that contain EARB and more than 5 other MAG (liver cirrhosis = 28, preterm infant = 24 communities). Communities for healthy adults, liver cirrhosis patients, and preterm infants were arranged by sampling point, MELD score (Model for End-Stage Liver Disease), and host day of life, respectively.

Results

De novo assembly of shotgun metagenomics revealed the enrichment of multi-AMR genes within specific bacterial strains

To investigate the longitudinal changes in the AMR gene repertoire of the gut microbiome after antibiotic treatment, known as the resistome, we obtained and processed publicly available fecal shotgun metagenomics data from 12 healthy adults with 5 time points (total 57 samples were used among possible 60 samples, with three samples excluded due to the library construction failure). These data were collected before and after a 4-day treatment with an antibiotic cocktail (meropenem, gentamicin, and vancomycin), as well as on days 8, 42, and 180 after the cessation of antibiotics (Fig. 1)[13]. First, we performed de novo genome reconstructions of 57 metagenomic samples using the metaWRAP pipeline[15], which integrates three binning tools (see Methods). A total of 7858 initial bins from the three binning tools were filtered into 2585 HQ MAGs, which passed the quality criteria of completeness (>70%) and contamination (<5%) based on CheckM estimations[17] (Fig. 1b and Supplementary Fig. 1a, b) and refined these into 1358 MAGs, which were merged from the consensus bins of the three binning tools (Supplementary Data 1–2).

We identified the AMR gene repertoire of the MAGs using CARD[28] as a reference (See Methods). We traced the changes in the number of MAGs and AMR genes identified in each sample and summarized them by time point. The number of MAGs was highest at baseline (day 0), significantly reduced from day 4 to day 42, and then recovered on day 180. We also found a significant increase in the number of AMR genes by day 8, suggesting that 4 days after the cessation of antibiotic treatment, resistant bacteria persisted and thrived (Fig. 1c). Interestingly, individual examinations of changes in

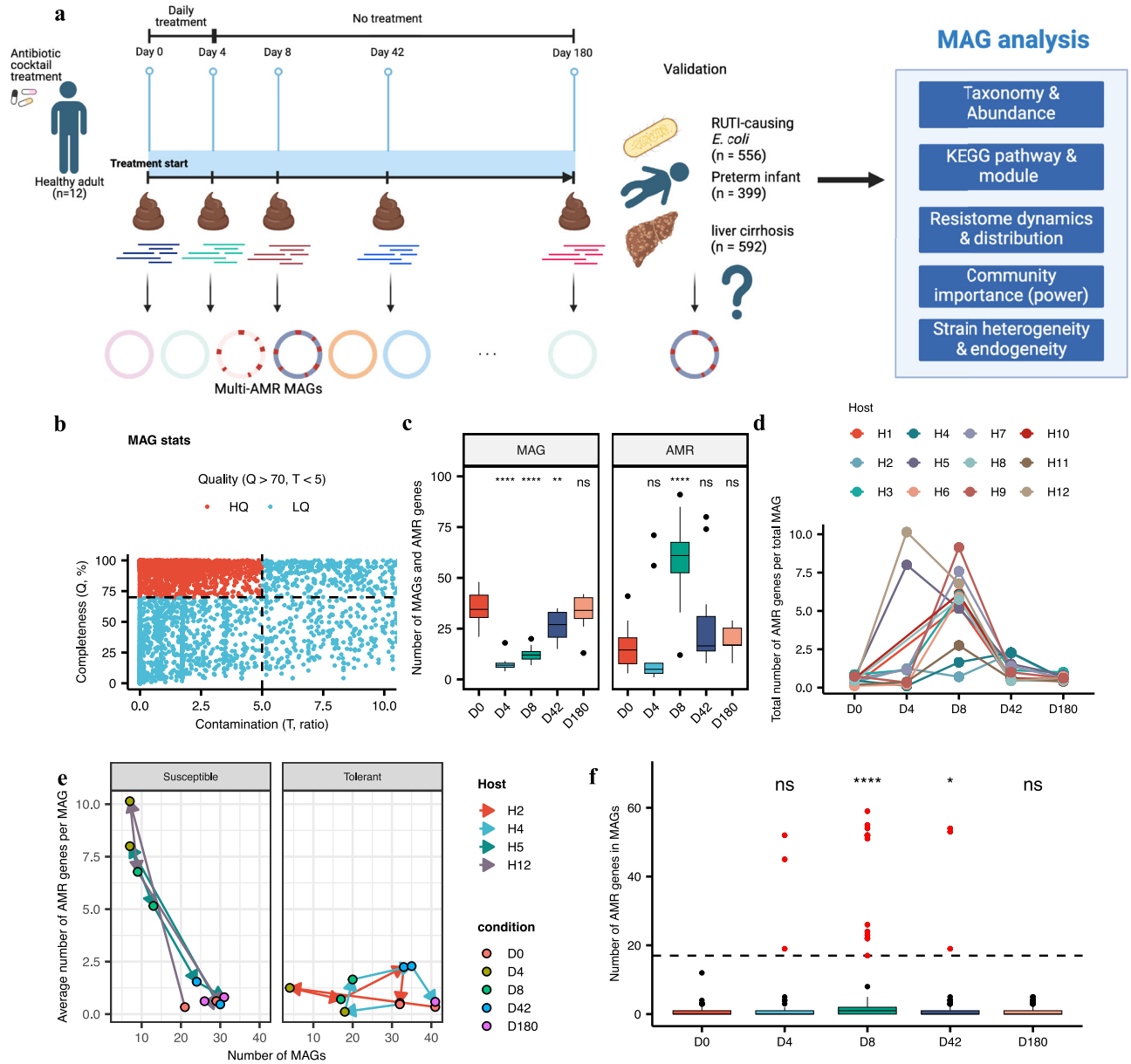

Fig. 1 | Metagenome-assembled genomes revealed individual resistome changes and discovered bacteria that had acquired extensive antimicrobial resistance. **a** Overview of the shotgun metagenome analysis conducted in this study. From 12 healthy adults, 56 metagenomics datasets were sampled at five different time points, and three additional datasets (recurrent urinary tract infection [RUTI]-causing *Escherichia coli* genomes, preterm infant metagenomes, and liver cirrhosis metagenomes) were used for further validation. **b** Scatter plot of metagenome-assembled genome (MAG) quality score of three different binning tools. Using CheckM algorithm, we selected high-quality MAGs (total number of high-quality (HQ) MAGs = 2585 [blue], total number of low-quality (LQ) MAGs = 5268 [red]) based on a completeness (Q) >70% and contamination (T) less than 5% (block dotted lines). **c** Boxplot of the number of *refined* MAGs based on the consensus of three binning tools (completeness > 70%, contamination < 5%; $n = 1358$) and the number of antimicrobial resistance (AMR) genes found in subjects at each sampling point. AMR genes were discovered by Resistance Gene Identifier (RGI) tool using the Comprehensive Antibiotic Resistance Database (statistical significance measured by Student's *t* test comparing each time point group with the day (D)0 group as reference group. $^{****}p \le 0.0001$, $^{**}p \le 0.01$). **d** Individual dynamics of the number of AMRs per MAG. Each line indicates AMR changes in a given host. **e** Line plots tracking the changes in the number of MAGs and average AMR gene burden per MAG of susceptible (Host 5 and 12) and tolerant subjects (Host 2 and 4). **f** AMR gene distribution within each MAG ($n = 1358$). MAGs containing more than 17 AMR genes (black dotted line) were regarded as bacteria that had acquired extreme resistance (EARB) (red dots, $n = 20$) (statistical significance measured by Student's *t* test comparing each time point group with the D0 group, $^{****}p \le 0.0001$, $^{*}p \le 0.05$).

the AMR gene ratio revealed diverse resistance responses of the gut microbiome to antibiotic treatment. For example, some individuals (host 5 and 12; susceptible group) showed drastic changes in their resistome from baseline, while other individuals (host 2 and 4; tolerant group) showed a similar level of AMR gene burden. This suggests different susceptibilities of individuals to antibiotic-induced dysbiosis (Fig. 1d, e and Supplementary Fig. 1c). In a previous study, administering antibiotics to healthy subjects not only altered their gut microbiomes to resemble those of intensive care unit patients, but also resulted in a protracted recovery in some subjects, as confirmed by the maximum displacement between baseline and final data points in principal component analysis[29]. In line with these findings, our study revealed that individuals in the susceptible group, who experienced a rapid enrichment of AMR genes following antibiotic treatment, demonstrated the longest distance from baseline (day 0) to the final sampling point (day 180), indicating a delayed recovery of the microbiome (Supplementary Fig. 1d).

Next, we examined the correlation between MAGs and the increased resistome by quantifying the prevalence of AMR genes within each MAG at different time points (Fig. 1f). In our temporal analysis, we found that a specific set of 20 MAGs contained a notably high number of AMR gene, reaching a saturation point in the cumulative distribution of AMR counts, with each MAG containing at least 17 AMR genes (Supplementary Fig. 2a). Therefore, we defined bacteria with more than 17 AMR genes as extensively acquired antimicrobial resistance bacteria (EARB). The EARB showed comparable assembly quality compared to other MAGs (Supplementary Fig. 2b, c), indicating that the higher number of AMR genes in EARB was not an artifact of poor metagenome assembly. Notably, we found that individuals with different resistome changes, i.e., susceptible and resistant groups, showed distinct emergence timings of EARB species following antibiotic treatment (Supplementary Fig. 1c). Therefore, to better understand the changes in individual microbiomes after antibiotics treatment, the dynamics of the abundance and functions of EARB species will be the key to understanding the effects of antibiotics on gut microbiome community.

Characteristics of EARB strains at the taxonomical, functional, and community level

We further explored the characteristics of EARB by comparing their taxonomy and abundance with those of other microbes within the same microbial community. We first annotated the taxonomy of all MAGs identified using Genome Database Taxonomy Toolkit (GTDB-TK)[19] and constructed a phylogenetic tree for the EARB species (Supplementary Fig. 3a, see Methods). Of note, all EARB belonged to the Proteobacteria phylum (also known as Pseudomonadota), predominantly within the genera *Escherichia*, *Klebsiella*, *Enterobacter*, and *Cronobacter*, which are widely recognized as pathobionts associated with various chronic diseases and bloodstream infections[30]. We also examined the taxonomy of MAGs containing fewer than 17 AMR genes, referred to as SARB. The majority of SARB belonged to the phyla Bacteroidetes (also known as Bacteroidota) (62.8%) and Firmicutes (also known as Bacillota) (27.5%), while a small proportion belonged to Proteobacteria (3.8%). AMR non-carriers mostly belonged to Firmicutes (82.6%), implying different taxonomic preferences for the acquisition of antimicrobial resistance genes (Supplementary Fig. 4a, b).

Next, we estimated the relative abundance profiles of all MAGs at the phylum and family levels (Supplementary Fig. 3b, c). Interestingly, the abundance of *Enterobacteriaceae*, to which all EARB were assigned, and *Veillonellaceae*, known pathobionts associated with intestinal inflammation[31,32], significantly increased when antibiotic-induced dysbiosis peaked (i.e., day 4 and 8). However, the abundance of the known commensal phylum, Bacteroidetes, significantly decreased on day 4 and day 8 (Student's *t* test p value = 0.0017 on day 8, compared with baseline). The pattern of changes in microbial composition was also confirmed using kraken2[33], which showed resembled result (Supplementary Fig. 4c). We observed a shift in the microbial community composition over the short-term (day 0 to 42) following antibiotic treatment. This shift was characterized by a decreased relative abundance of EARB and an increased relative abundance of various minority species, suggesting a pattern of recovery from dysbiosis in the hosts (Supplementary Fig. 3d and Supplementary Fig. 4d).

Differential resistomes between EARB and SARB strains

Next, we identified differences in the resistomes of EARB and SARB based on their annotated AMR gene information using CARD[28]. Based on the resistome prevalence values (that is, the observed frequency of a given AMR drug class normalized to the proportion of AMR gene-carrying bacteria; see Methods), we compared the differences in resistome composition at the drug class level between EARB and SARB strains (Fig. 2a). Of note, the resistome carried by EARB strains was more evenly distributed for many drug classes, whereas the resistomes carried by SARB were highly enriched for fluoroquinolone and tetracycline classes, which were not relevant to the antibiotics used in this study. Interestingly, a previous report showed that tetracycline-resistant genes are spread over many different geographical regions, reflecting the same tendency as observed in the SARB resistome[34]. These distinct resistomes between EARB and SARB were also confirmed in the pheatmap using the relative frequency of drug class matrix (Supplementary Fig. 5a, b) and they were clustered again into EARB and SARB groups by unsupervised clustering (k-means) (Supplementary Fig. 6a). Among the three drug classes used in this study (carbapenem, aminoglycoside, and glycopeptide), carbapenem resistance was predominantly found in EARB (20 out of 23 MAGs of the given resistome were EARB) (Supplementary Data 3). Similarly, aminoglycoside resistomes were found mostly in EARB (20 of 40 MAGs of a given resistome were EARB). However, the glycopeptide resistome was found in SARB (44 MAGs), but was absent in EARB. The original study pointed out that gram-negative bacteria are naturally resistant to glycopeptide antibiotics, such as vancomycin, and EARB (all gram-negative) might have natural resistance to such glycopeptide antibiotics, thereby not necessitating a glycopeptide resistome[35].

We further investigated the resistome profiles of EARB and SARB based on the shared AMR drug classes at the genus level, and also by the broader taxonomic rank (e.g., the phylum level for the SARB strains due to their phylogenetic diversity) (Fig. 2b, d and Supplementary Fig. 7a). Of note, all EARB shared resistance to similar antimicrobial drug classes, whereas SARB did not, except for fluoroquinolone and tetracycline. Therefore, the extensive antibiotic resistance capacity of EARB provides a fitness advantage when treated with antibiotics, whereas SARB shows different survival outcomes according to the type of antibiotic used. However, at the gene level, neither EARB nor SARB strains shared AMR genes with other phylogenetic groups (genera or phyla), implying that similar drug class resistomes were conferred by unique or narrowly distributed AMR genes (Fig. 2c, e and Supplementary Fig. 7b). This result also suggests that horizontal gene transfer (HGT) of AMR genes between bacteria of different taxa is uncommon. Rather, a specific taxonomic group harboring an extensive resistome and AMR gene spread through HGT might only occur in specific genes.

We also traced the changes in resistome profiles at different time points after extracting resistome signatures using NMF (Supplementary Fig. 8a, see Methods). By applying an unsupervised NMF decomposition to the resistome profiles, we identified two distinct signatures in the resistome, *sig1* and *sig2*. As expected, these two signatures resembled the EARB and SARB resistome prevalence scores, further confirming our findings through independent methods. Of note, EARB-like resistome signature (*sig1*) was significantly increased on days 4 and 8, whereas SARB-like resistome signature (*sig2*) was decreased on days 4 and 8 and then recovered at days 42 and 180. Such resistome signatures imply that EARB drives overall resistome changes after the antibiotic treatment, as observed in the data presented in Fig. 1.

EARB strains showed functional dominance in the post-antibiotic gut microbial community

In our previous findings (Supplementary Figs. 3 and 4), we identified EARB strains co-enriched with specific bacterial groups, suggesting that EARB play a significant role in shaping microbial communities under antibiotic treatment. Therefore, to further investigate the impact of EARB within the community, we devised a new scoring method that estimates functional dominance, termed as the *community power score* (Fig. 3a).

First, we examined the annotated functions (i.e., Kyoto Encyclopedia of Genes and Genomes [KEGG] orthologs) of all the species detected in a given metagenomic sample (i.e. functional profile of sample 1 in Fig. 3a). Next, we gave different weights to the functions of target species in the profile matrix by normalizing their functions by the total number of species having the given functions. If a given function is present only in a target species, that species would have a higher weight for that function. On the contrary, if a given function is present in all species detected in the given samples, the target species would have a lower weight for that function. Lastly, by summarizing all the functional weights of the target species, a community power score of the given species was calculated. After this calculation, the community power score of given species can be compared with other species

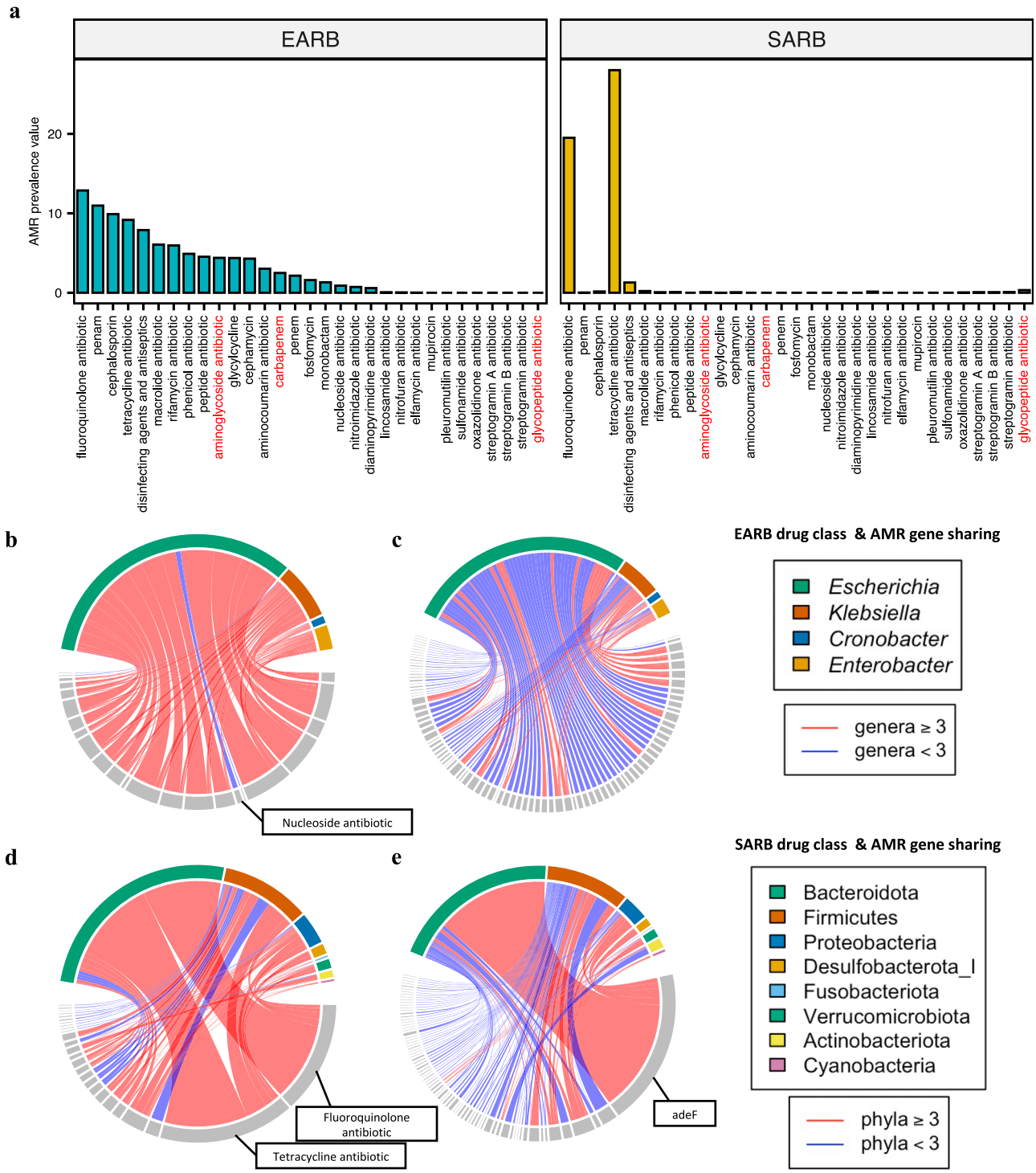

Fig. 2 | Differential resistomes between EARB and SARB strains. a Drug classes resistant to EARB and SARB strains identified by AMR prevalence scores. We calculated AMR prevalence scores based on the observed frequency of given drug classes in EARB or SARB strains, while normalizing them according to their carrier ratios. We show the AMR prevalence scores from EARB (left) and SARB (right) strains based on the descending orders of EARB prevalence scores (the class of antibiotics used for treatment in this study is colored red). **b**, **c** Circos plots of (**b**) shared drug classes ($n = 23$) and (**c**) shared AMR genes ($n = 803$) (upper bounds) of given genera of EARB ($n = 4$) (gray, lower bounds). Red links indicate drug classes or AMR genes shared by more than three genera, whereas blue links indicate those shared by less than three genera. **d**, **e** Circos plots of (**d**) shared drug classes ($n = 26$) and (**e**) shared AMR genes ($n = 832$) of given phyla of SARB ($n = 8$). Red links indicate drug classes or AMR genes shared by more than three phyla, whereas blue links indicate those shared by fewer than three phyla.

detected in a given metagenomic sample (see boxplot for community power scores for sample 1 in Fig. 3a). Therefore, community power score can assess the microbe's functional uniqueness and importance within the community (see Methods).

Remarkably, we found that, in all metagenomic samples, EARB showed the highest community power compared to the other microbes, indicating its significant functional importance in the community. However, SARB and AMR non-carriers showed significantly lower community

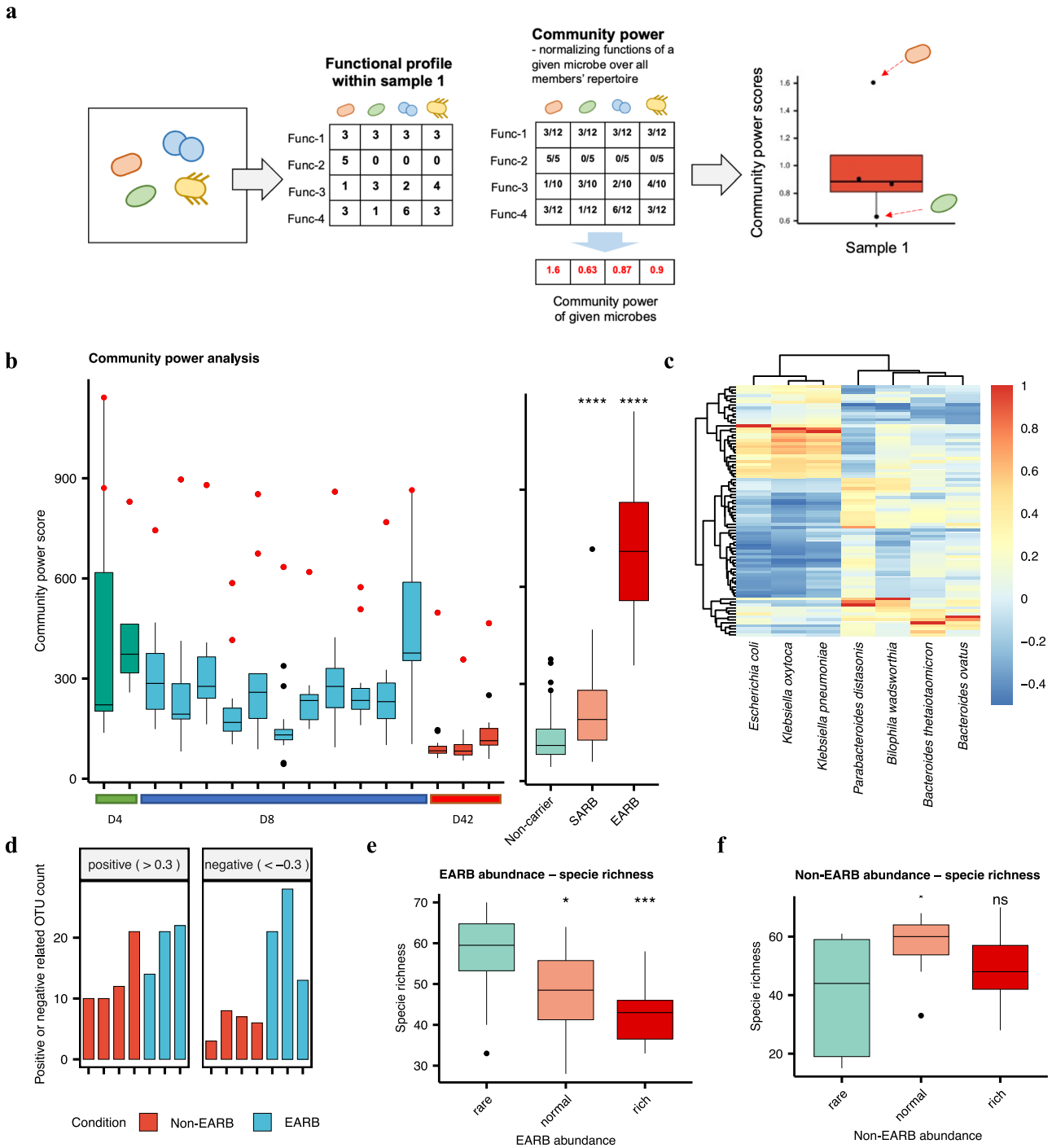

Fig. 3 | EARB strains harbored a greater functional repertoire capable of shaping the community. a The outline of community power analysis. We calculated the functional repertoire of given microbial strains, herein termed community power scores, based on the number of genes harboring specific Kyoto Encyclopedia of Genes and Genomes orthology (KO) terms. In short, we counted the total functional repertoire of given microbial strains based on their genes harboring specific KO terms, after normalizing them according to the total number of KO terms in all the microbial strains in the given metagenomic samples. Therefore, all community power scores for microbes in a given metagenomic sample could be calculated and compared with each other. The microbes of the highest community power score had the highest functional repertoire in a given metagenomic sample. **b** Community power scores of metagenomic samples that carried EARB strains ($n = 16$). We compared the community power scores among EARB, SARB, and non-carriers (right panel) and found that EARB outscored the others (Student's t test p values < 0.0001 [****]). **c** Co-abundance (i.e., Spearman's correlation coefficients) of three EARB strains (*Escherichia coli*, *Klebsiella oxytoca*, *Klebsiella pneumoniae*) and four non-EARB strains with the highest community power score in one of the metagenomic samples (*Bacteroides thetaiotaomicron*, *Bacteroides ovatus*, *Bilophila wadsworthia*, *Parabacteroides distasonis*). Microbial strains significantly correlated with at least one of the seven selected strains (absolute correlation coefficients >0.3) are shown, and unknown metagenomic-based operational taxonomic unit (mOTU) species were excluded. Additionally, mOTUs that were not present in more than five metagenomic samples were also excluded (rows in the heatmap, $n = 84$). **d** The number of significantly correlated strains (absolute Spearman's correlation coefficients >0.3) with EARB and non-EARB strains. **e** Negatively correlated between EARB abundance and species richness, based on the sum of log2-transformed relative abundances (rare < −10, −10 ≤ normal < −5, rich ≥ −5) and species richness. **f** Less-correlated patterns between sample groups of different non-EARB abundances, based on the sum of log2-transformed relative abundances (rare < −10, −10 ≤ normal < −5, rich ≥ −5) and species richness (relative abundance > 0). Differences between sample groups were tested using Wilcoxon rank-sum tests.

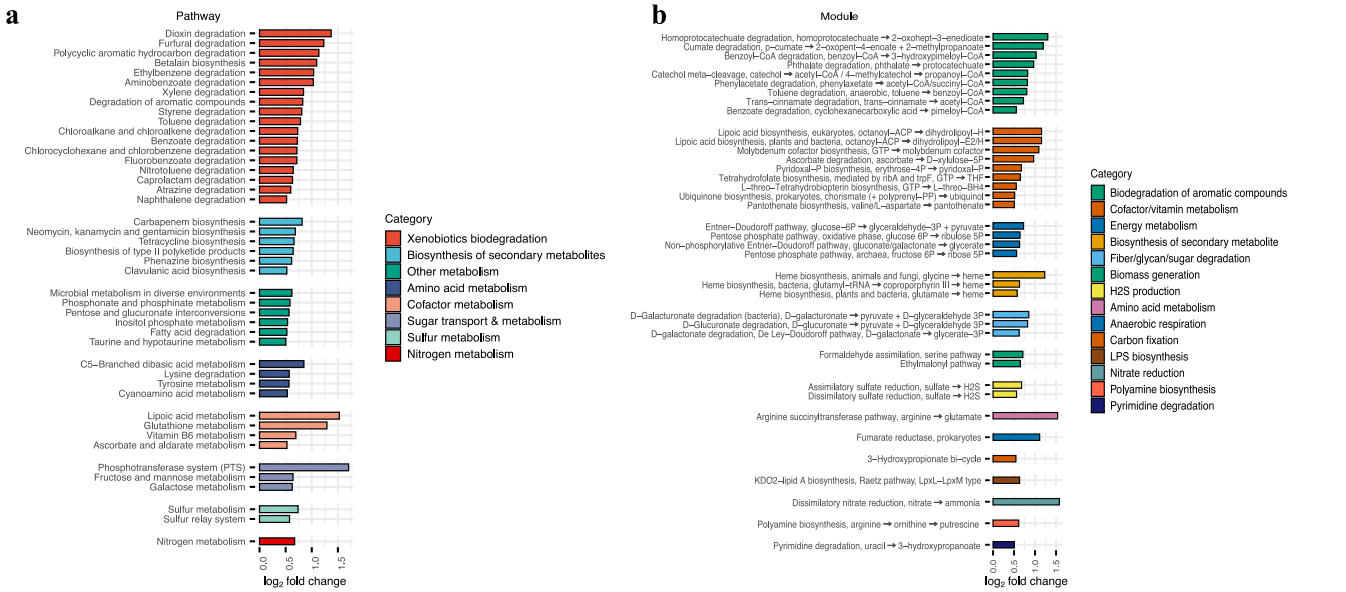

Fig. 4 | The enriched metabolic pathways of EARB strains in a given community. **a** KEGG metabolic pathways or **b** KEGG modules significantly enriched among EARB strains compared to other microbes in given metagenomic samples (total number of EARB = 20 and total number of non-EARB compared = 217, Wilcoxon rank-sum test p values < 0.05, log2 fold change >0.5, and pathway coverage >0.3 or module coverage >0.8).

power scores compared to EARB strains (Student's t test p values < 1×10^{-4}) (Fig. 3b, right).

To further identify which community members were influenced by EARB strains, we examined microbes that were co-abundant with EARB (*Escherichia coli*, *Klebsiella oxytoca*, and *Klebsiella pneumoniae*) and other non-EARB (i.e., SARB and non-carriers) with the highest community power scores (*Bacteroides thetaiotaomicron*, *Bacteroides ovatus*, *Bilophila wadsworthia*, and *Parabacteroides distasonis*) (|Spearman's correlation coefficients |>0.3) based on mOTU-based species abundance profiles (see Methods). Interestingly, co-abundant microbes with EARB strains were mostly pathobionts (Supplementary Data 4), including *Clostridium* spp., *Fusobacterium* spp., and *Veillonella* spp., which promote intestinal inflammation[31]. In contrast, members that were co-abundant with non-EARB were mostly commensal bacteria, including *Bacteroides* spp., *Roseburia* spp., and *Alistipes* spp. When we analyzed the negative bacterial interactions between EARB and non-EARB, we found that EARB showed a much higher number of negatively correlated microbes than non-EARB, implicating EARB outcompeting in given microbial community (Fig. 3d). We substantiated our findings using correlations between EARB abundance and species richness and identified negative trends between these two parameters (Fig. 3e). However, the abundance of non-EARB with high community power correlated with increased species richness in the middle-range abundance level (Fig. 3f). Therefore, we identified two distinct patterns of functionally impactful bacteria: 1) microbes with the high community power that might influence a small set of bacteria and 2) microbes with high community power that are co-abundant with a large number of bacteria, behaving like commensal species.

Next, to better characterize the functional impact of EARB on other interacting microbes, we performed functional enrichment analysis of EARB compared to non-EARB based on the KEGG pathways and modules[36] (Wilcoxon rank-sum test p values < 0.05, log2 fold change >0.5, pathway coverage >0.3 or module coverage >0.8; Fig. 4a, b). Interestingly, we found that EARB were significantly enriched in amino acid metabolism, sugar metabolism, and reductive transformations of xenobiotics, including reductions in aromatic compounds, such as processes supplemented with enriched cofactor biosynthesis (e.g., glutathione, vitamin B6, and ascorbate). Notably, xenobiotic degradation is known to support anaerobic respiration through nitrogen and sulfur reduction, promote biomass generation of the given microbiota, and provide a fitness advantage[37], and is also likely to provide the transformation capacity for drug compounds, including antibiotics. In addition, we found multiple biosynthesis pathways for antibiotic-like compounds in EARB, which may include bacteriocin, a narrow-spectrum antimicrobial substance[38–40]. In summary, we found that the EARB strains were equipped with biosynthetic pathways, xenobiotic-reducing metabolism, and anaerobic respiration capacity.

Moreover, based on CAZyme and KEGG profiles, principal component analysis revealed distinct clustering of EARB, SARB, and non-carrier, highlighting their functional differences (PERMANOVA p values = 0.001; Supplementary Fig. 9a–d and Supplementary Data 5, 6). Interestingly, unlike EARB species, SARB were enriched among polysaccharide pathways, including N-glycan biosynthesis (M00055: N-glycan precursor biosynthesis; hypergeometric test p value = 5.4×10^{-4}) and carbohydrate degradation enzymes (148 GH and 10 CE families for SARB species vs 7 GH and none CE families for EARB species)

Multiple EARB strains were carried endogenously and alternated in individual hosts

Next, we checked EARB strain changes after antibiotic treatment to determine whether they were carried endogenously from the same hosts or re-colonized spontaneously from outside the host. To this end, we first selected EARB strains (n = 20) from different hosts and checked their genomic similarities by the average nucleotide identity (ANI) scores[41] (see Methods). The ANI scores showed distinct patterns between *E. coli* and other EARB of different genera (Supplementary Fig. 10a). When we focused on the ANI scores between the EARB *E. coli* strains (n = 12), we observed considerable strain diversity, ranging from 96.56 to 99.99% (Fig. 5a). Of note, EARB *E. coli* strains from the same host (i.e., host 12) on day 4 and 8 showed a substantially high ANI score of 99.99% (Fig. 5b). Given that identical strains were defined based on an ANI score of 99.8%[42], the two EARB *E. coli* strains from the host 12 could be regarded as identical, implying same strains endogenously carried between day 4 and day 8. In contrast, EARB *E. coli* strains from host 5 on day 4 and 8 showed distinct ANI scores (99.40% and 98.78%) (Fig. 5b and Supplementary Fig. 10b). When compared with other EARB *E. coli* strains, host 12 *E. coli* strains showed identical ANI scores whereas host 5 *E. coli* strains showed different ANI score patterns. Therefore, we found substantial strain heterogeneity of EARB strains of *E. coli*, even between those originating from the same hosts.

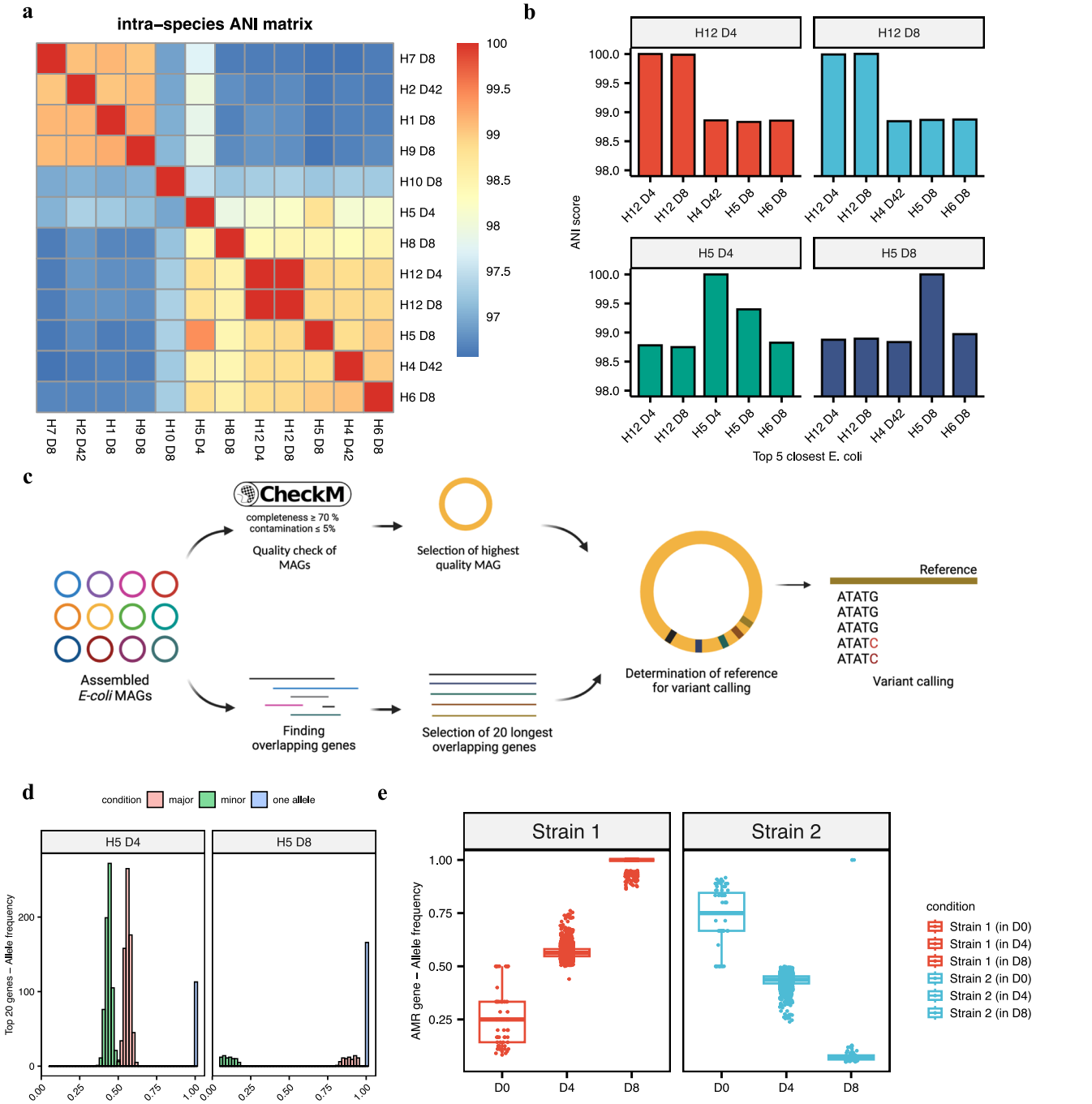

Fig. 5 | EARB strains were carried endogenously by some hosts and major strains alternated after antibiotic treatment. a Intra-species average nucleotide identity (ANI) scores of *E. coli* strains belonging to EARB ($n = 12$). **b** Top 5 closest EARB *E. coli* strain ANI scores with two specific strains identified from the same host (H5, H12). **c** The workflow for the identification of single-nucleotide polymorphisms (SNPs) from a given metagenome sample. First, we selected 20 homologous genes with the longest lengths from all EARB *E. coli* MAGs as the on-target genes for SNP analysis. Next, we used the 20 gene set from the best-quality MAG for building reference and performed variant calling for metagenomics samples which contain EARB *E. coli* strains. **d** Allele frequency distributions of SNPs identified in metagenomic samples of host 5 at day 4 (left, $n = 805$) and day 8 (right, $n = 240$). We identified two different patterns of major and minor alleles, and also one allele existing in both strains. **e** SNP analysis was conducted using *E. coli* AMR genes of host 5 at day 8 as a reference set. We found that two *E. coli* strains shared same variant positions and different SNP profiles (nucleotide) under different conditions (median of allele frequencies for strain 1 at day 0, 4, and 8 = 0.250, 0.563, 1.000, respectively, and those for strain 2 at day 0, 4, and 8 = 0.750, 0.437, 0.0694, respectively).

To further identify strain changes in the longitudinal samples, we performed SNP calling of homologous genes present in host 5 EARB *E. coli* strains (Fig. 5c and Supplementary Data 7, see Methods). Briefly, we identified 20 homologous genes, possessed by all *E. coli* strains, then built a reference using the homologous gene set from the best-quality *E. coli* strains (H6 D8). Next, we called the variant positions (SNPs) by mapping the host 5 metagenome reads, and checked the shared variant positions and calculated allele frequencies of all variant positions. As a result, we found 804 and 240 SNPs of homologous gene reference from day 4 and 8, respectively (Fig. 5d). Analysis of the allele frequency of the identified variants revealed two separate clusters on day 4, with the major cluster being close to 55% and the minor cluster being close to 45%. However, on day 8, one allele converged to

almost 100% (median: 98%). This suggests that one of the two strains present on day 4 became dominant on day 8.

Since, we identified the existence of two distinct strains in host 5, we tried to track down the changes of strains using EARB *E. coli* AMR gene set of host 5 on day 8 (number of AMR genes = 54) as a reference due to the high dominance of a single strain. Our analysis identified 1841 variants on day 4 across 49 AMR genes out of 54 AMR reference gene sets. Considering that a single strain was dominant on day 8, while two strains were present on day 4, we hypothesized that the variants identified on day 4 were due to the presence of a minor strain. Consequently, the five genes without variants may either be due to perfect matching in either strain or may only be found in the major strain. Examination of the read depth of the AMR genes revealed that one of the five genes presented a depth similar to that of the other 49 genes carrying variants, suggesting that it was shared between the two strains and perfectly matched. The remaining four genes displayed significantly lower depths, indicating that they were exclusive to the major strain (Supplementary Fig. 10c). Allele frequency analysis of the 1,841 variant positions on day 4 revealed a distribution similar to that of the homologous genes, with approximately 55% major strains and 45% minor strains (Fig. 5e and Supplementary Data 8). Comparing the major and minor nucleotides of these variant positions longitudinally by tracking the same variant position with the same SNP nucleotide between day 4 and day 8, we found that the major strain on day 4 became dominant on day 8. To determine whether these AMR genes were also present at baseline, we also tracked SNP profile on day 0 (Fig. 5e). Although only a single allele was confirmed at many variant positions owing to the shallow sequencing depth (single allele position = 333), the major and minor strains were still identifiable at certain positions (two-allele position = 76). Remarkably, the strain that was predominant on day 8 was observed as a minor strain on day 0, suggesting that the emergence of this EARB may be attributed to strain alterations. As the endogenous nature of these *E. coli* strains has been established in a previous study[43], our results confirmed that the occurrence of EARB in an antibiotic-exposed environment is caused by opportunistic pathogenic bacteria that already inhabit the host.

EARB strains identified in diverse antibiotic-induced dysbiotic conditions

We further investigated other populations with frequent antibiotic exposure to determine if their microbiota contained EARB (Fig. 6). For example, we selected shotgun metagenomic cohorts and the genomes of bacterial isolates showing antibiotic resistance, such as those causing RUTI[43], and performed de novo assembly to identify EARB strains harboring more than 17 AMR genes (see Methods). First, we found that all *E. coli* isolates from RUTI were EARB strains (mean number of AMR genes they carry = 61; Supplementary Data 9). Interestingly, based on SNPs that defined major and minor strains in a previous analysis (Fig. 5), we found that most *E. coli* isolates (65%) resembled the major strain's SNP profile (i.e. having more than half of same nucleotides of 1,841 SNPs that define major strains; Supplementary Fig. 11), whose abundance was increased by antibiotics. This implies that, even within RUTI-EARB *E. coli*, the SNP patterns of the major strain were more prevalent in antibiotic-exposed environments. Next, we checked other MAGs assembled from antibiotic-exposed cohorts (i.e., preterm infant[44] [$n = 399$] and liver cirrhosis cohorts[45] [$n = 592$]) and found EARB strains among the selected cohorts (303 of 399 infant samples and 60 of 592 liver cirrhosis samples) (Fig. 6). More interestingly, when we checked their resistome profiles (Fig. 6a), we found that they showed highly similar resistome profile to the EARB strains we found in a previous analysis (Fig. 2a). Based on further functional investigations, such as community power scores (Fig. 6b), we found that EARB had the greatest functional impact on a given community likewise. Of note, preterm infants showed a higher prevalence of EARB strains, whereas patients with liver cirrhosis had a higher average number of assembled MAGs than preterm infants (Fig. 6c, Supplementary Data 10, 11), implying that preterm infants are more susceptible to EARB colonization, as they have a simple microbial community with less colonization resistance[46]. Additionally, metagenomic analysis of samples from patients with liver cirrhosis from two different samples, stool ($n = 264$) and saliva ($n = 328$), showed that the prevalence of EARB strains was higher in stool samples ($n^{stool} = 57$ and $n^{saliva} = 7$). These results showed that EARBs prevalently exist within the population who have been treated with antibiotics multiple times and show the same characteristics as our findings. Lastly, to investigate EARB strain sharing across different cohorts, we performed ANI analysis on all EARB *E. coli* strains generated in this study. The results revealed a wide range of ANI values, from 93.78% to 99.99%, indicating substantial strain diversity among EARB *E. coli* strains (Supplementary Fig. 12a). Nevertheless, we identified identical strains (ANI similarity ≥99.5%) shared across different cohorts, suggesting that some strains have the capacity for inter-cohort transmission (Supplementary Fig. 12b). Interestingly, not all *E. coli* strains were shared between cohorts. This observation implies the existence of both globally shared *E. coli* strains and cohort (or person)-specific *E. coli* strains (Supplementary Fig. 12c, d). Therefore, we confirmed the existence of EARB in diverse populations based on the newly discovered MAGs from different data sources (i.e., isolate genomes, single-end and paired-end sequencing data).

Discussion

The need for research on antibiotic-resistant bacteria at both the species and subspecies levels has increased. In this study, we performed de novo assembly of MAGs to investigate the dynamics of AMR genes at the species and strain levels within microbiomes altered by antibiotics. First, using the MAG-based approach, we observed different responses of individual microbiomes to antibiotics, which were classified into susceptible and tolerant groups. Second, based on AMR gene repertoire analysis, we identified two distinct groups of antibiotic-resistant bacterial strains in individual microbiomes: EARB and SARB. Notably, pathway enrichment analysis showed that EARB harbored specific functions that enhanced their survival and competitive advantage, such as the degradation of xenobiotic substances, including aromatic antibiotics, and the promotion of biosynthetic pathways. Third, we found that EARB harbored a wide range of AMR genes for different drug classes compared to SARB, which were mostly enriched for only tetracycline and fluoroquinolone antibiotic resistance. Fourth, we found that EARB played a significant role in altering post-antibiotic microbial community dynamics based on the community power score we devised. Finally, we identified multiple resistant lineages of EARB strains and their intra-species competition within the same host during antibiotic treatment, highlighting the importance of species-specific strategies to manage antibiotic resistance.

Notably, functional profiles of MAGs indicated enhanced xenobiotic degradation pathways within the EARB, suggesting their potential role in the persistence of antibiotic-susceptible bacteria via degradation enzymes, such as beta-lactamases. In addition, we identified an enrichment of secondary metabolite and antibiotic-like compound biosynthesis pathways among the EARB strains. This suggests that EARB strains possess a competitive advantage, facilitating their survival in the intestines of healthy individuals[47]. These findings elucidate the complex interplay between EARB and the gut microbiota, potentially driving the persistence of antibiotic resistance. Therefore, further functional analysis is essential to fully understand how EARB promotes the co-enrichment of antibiotic-susceptible bacteria and supports their persistence in the intestines of healthy individuals.

Importantly, we validated the existence of EARB using data from different antibiotic-exposed microbiomes. We consistently identified EARB in RUTI-causing *E. coli* isolates and in metagenomic data from patients with liver cirrhosis and preterm infants. Moreover, these EARB exhibited characteristics consistent with our previous findings. Furthermore, we identified high ANI scores (≥99.5%) for EARB *E. coli* strains from independent cohorts, implying that the same EARB strains spread in different geographical regions like Denmark, the US, and the UK (Supplementary Fig. 12). This necessitates further in-depth analysis of emerging EARB strains and their functional roles in dysbiotic conditions of many other chronic diseases. In addition, we observed a higher

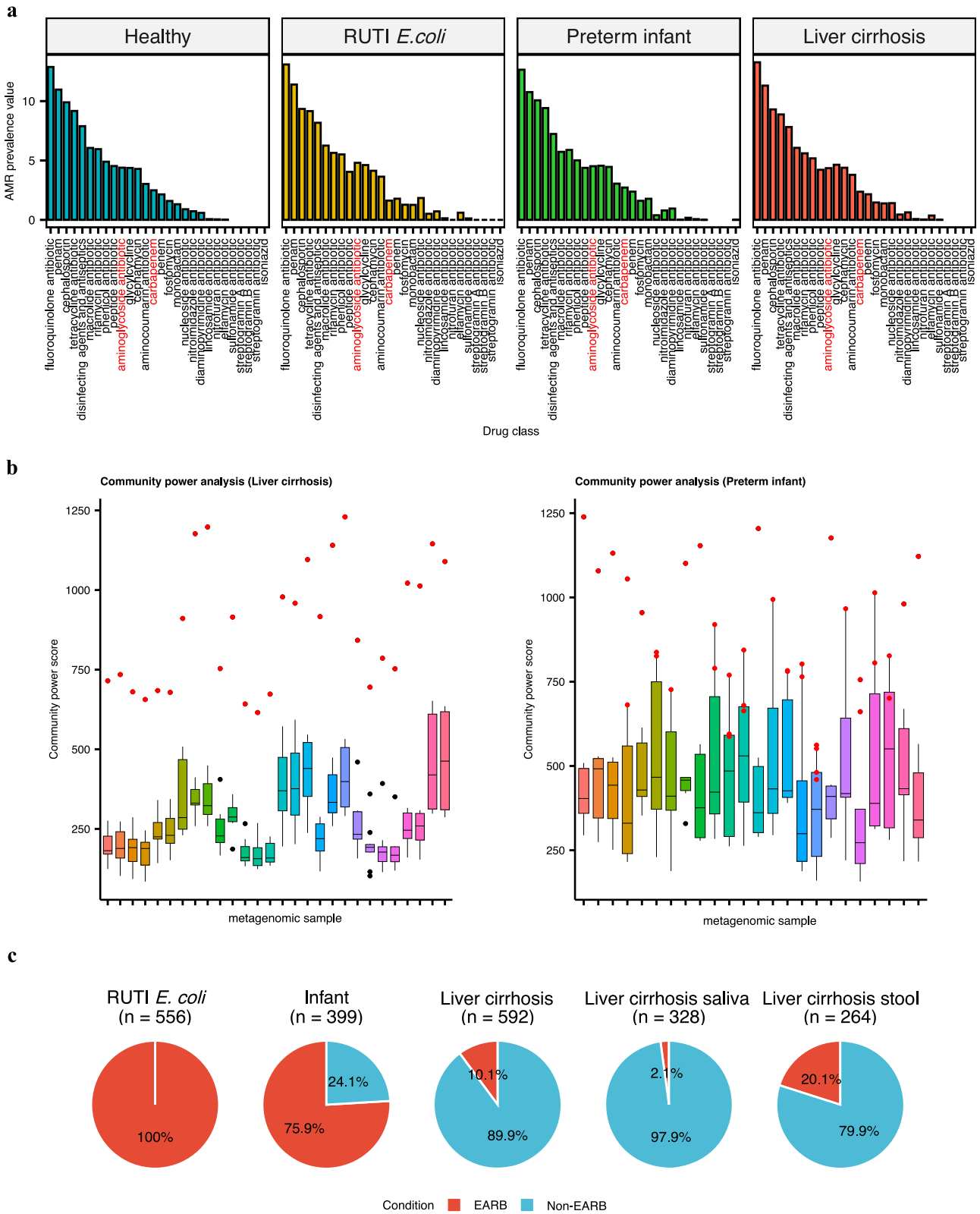

Fig. 6 | EARB strains found in different antibiotic-exposed cohorts with the greatest functional impact on the community. a AMR prevalence comparison between data from healthy adults and data from three different cohorts (recurrent urinary tract infection [RUTI]-causing *E. coli*, preterm infants, and patients with liver cirrhosis). **b** Community power analysis applied to the metagenomics samples of liver cirrhosis (left, $n = 592$) and preterm infant (right, $n = 399$) study data. Community power analysis was conducted on samples harboring five MAGs and at least one EARB strain (red dot indicates the EARB strain). **c** Proportion of samples containing EARB strains in three different datasets (*E. coli* isolate, preterm infant, and liver cirrhosis [saliva and stool]).

prevalence of EARB in the gut microbiota than in the oral microbiota. This may be due to the differences in the impact and duration of antibiotics in the gut and the oral cavity. This suggests that the emerging patterns of EARBs are affected by specific environments or host-related characteristics.

In addition, we performed strain-level analysis of the EARB in individual microbiomes, particularly *E. coli*, and identified remarkable genomic diversity within strains isolated from the same host over time, indicating the existence of multiple strains or lineages. Longitudinal data analyses allowed us to track the development of these EARB *E. coli* strains, uncovering two key patterns: the presence of native strains within the host gut and the occurrences of a "strain sweep", where certain strains outcompete others, potentially due to superior resistance capabilities or other fitness advantages[48]. This advantage can also be explained by the observed weighted SNP patterns in RUTI-causing *E. coli*, suggesting that these differences may influence the adaptability of bacteria to antibiotic environments. These findings indicate that the gut serves as a reservoir for diverse antibiotic-resistant strains and demonstrate the dynamic nature of microbial evolution influenced by selective pressures. Such observations have significant implications for therapeutic strategies, emphasizing the need for in-depth research to elucidate the determinants of strain variability and improve approaches to effectively manage antibiotic resistance. However, the current findings of strain changes after antibiotic treatment were based on a relatively small number of observations (two individuals), and further validations in a large-scale intervention cohort may be needed. In addition, emerging patterns of EARB and SARB strains can be dependent on the types of antibiotics. Therefore, it is also tempting to speculate that another type of EARB or SARB, having different resistome profiles from the ones we found, can be identified by cohorts of different antibiotics treated.

In summary, our study provides a comprehensive investigation into the dynamics of antibiotic-resistant bacteria through MAG analysis. Our MAG-based approach provides a detailed understanding of the entire resistome at the individual species level. Here, we found that EARB became a dominant force in microbial communities following antibiotic treatment, and that their genetic diversity and evolutionary pathways present notable challenges in managing antibiotic resistance. These findings not only deepen our understanding of antibiotic resistance dynamics but also pave the way for more targeted strategies to combat the spread of antibiotic-resistant bacteria.

Data availability

All metagenome and *Escherichia coli* isolated genome data were obtained from the European Nucleotide Archive (ENA) and the Sequence Read Archive (SRA). The metagenome data of antibiotic-treated individuals in this study were obtained from ENA under accession number ERP022986 (only paired samples were used). Preterm infant metagenome data were obtained from ENA under BioProject ID PRJNA301903 (both paired and genome data were filtered for library layout and strategy, respectively). Liver cirrhosis patient metagenome data were obtained from ENA under the identifier number: PRJEB38481. *Escherichia coli* isolate genomic data were obtained from SRA under accession number PRJNA786867, with a release date of 2022-01-20. All MAGs constructed in this study were stored and are publicly available on figshare: https://figshare.com/projects/EARB_Project/193070 (Supplementary Data 1).

Code availability

R script for Community Power Analysis is available in our github repository link: https://github.com/LifeMiningLab/CommunityPowerAnalysis. R script for figure generation is also provided in the github repository link : https://github.com/LifeMiningLab/EARB_project/tree/main.

Received: 6 September 2024; Accepted: 15 April 2025;
Published online: 14 May 2025

References

1. Browne, A. J. et al. Global antibiotic consumption and usage in humans, 2000–18: a spatial modelling study. *Lancet Planet Health* **5**, e893–e904 (2021).
2. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P T* **40**, 277–283 (2015).
3. Abubakar, U., Al-Anazi, M., Alanazi, Z. & Rodriguez-Bano, J. Impact of COVID-19 pandemic on multidrug resistant gram positive and gram negative pathogens: a systematic review. *J. Infect. Public Health* **16**, 320–331 (2023).
4. Sommer, M. O. A., Dantas, G. & Church, G. M. Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **325**, 1128–1131 (2009).
5. Gasparrini, A. J. et al. Persistent metagenomic signatures of early-life hospitalization and antibiotic treatment in the infant gut microbiota and resistome. *Nat. Microbiol.* **4**, 2285–2297 (2019).
6. Campbell, T. P. et al. The microbiome and resistome of chimpanzees, gorillas, and humans across host lifestyle and geography. *ISME J.* **14**, 1584–1599 (2020).
7. Hendriksen, R. S. et al. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* **10**, 1124 (2019).
8. Chng, K. R. et al. Cartography of opportunistic pathogens and antibiotic resistance genes in a tertiary hospital environment. *Nat. Med.* **26**, 941–951 (2020).
9. Forslund, K. et al. Country-specific antibiotic use practices impact the human gut resistome. *Genome Res.* **23**, 1163–1169 (2013).
10. Hu, Y. et al. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat. Commun.* **4**, 2151 (2013).
11. Parnanen, K. et al. Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements. *Nat. Commun.* **9**, 3891 (2018).
12. Smith, S. D. et al. Diversity of antibiotic resistance genes and transfer elements-quantitative monitoring (DARTE-QM): a method for detection of antimicrobial resistance in environmental samples. *Commun. Biol.* **5**, 216 (2022).
13. Palleja, A. et al. Recovery of gut microbiota of healthy adults following antibiotic exposure. *Nat. Microbiol.* **3**, 1255–1265 (2018).
14. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
15. Uritskiy, G. V., DiRuggiero, J. & Taylor, J. MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**, 158 (2018).
16. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
17. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
18. 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).
19. Chaumeil, P. A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* **36**, 1925–1927 (2019).
20. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
21. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* **11**, 119 (2010).
22. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–W37 (2011).

23. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).
24. Jain, C., Rodriguez, R. L., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
25. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
26. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
27. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
28. Alcock, B. P. et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48**, D517–D525 (2020).
29. Anthony, W. E. et al. Acute and persistent effects of commonly used antibiotics on the gut microbiome and resistome in healthy adults. *Cell Rep.* **39**, 110649 (2022).
30. Ecker, D. J. et al. New technology for rapid molecular diagnosis of bloodstream infections. *Expert Rev. Mol. Diagn.* **10**, 399–415 (2010).
31. Zhan, Z. et al. Overabundance of Veillonella parvula promotes intestinal inflammation by activating macrophages via LPS-TLR4 pathway. *Cell Death Discov.* **8**, 251 (2022).
32. Al-Otaibi, F. E. & Al-Mohizea, M. M. Non-vertebral Veillonella species septicemia and osteomyelitis in a patient with diabetes: a case report and review of the literature. *J. Med. Case Rep.* **8**, 365 (2014).
33. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
34. Carr, V. R. et al. Abundance and diversity of resistomes differ between healthy human oral cavities and gut. *Nat. Commun.* **11**, 693 (2020).
35. Sim, C. K. et al. A mouse model of occult intestinal colonization demonstrating antibiotic-induced outgrowth of carbapenem-resistant enterobacteriaceae. *Microbiome* **10**, 43 (2022).
36. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
37. Koppel, N., Maini Rekdal, V. & Balskus, E. P. Chemical transformation of xenobiotics by the human gut microbiota. *Science* **356**, eaag2770 (2017).
38. Sassone-Corsi, M. et al. Microcins mediate competition among Enterobacteriaceae in the inflamed gut. *Nature* **540**, 280–283 (2016).
39. Baquero, F., Lanza, V. F., Baquero, M. R., Del Campo, R. & Bravo-Vazquez, D. A. Microcins in enterobacteriaceae: peptide antimicrobials in the eco-active intestinal chemosphere. *Front. Microbiol.* **10**, 2261 (2019).
40. Markovic, K. G. et al. Colicins and microcins produced by enterobacteriaceae: characterization, mode of action, and putative applications. *Int. J. Environ. Res. Public Health* **19**, 11825 (2022).
41. Yoon, S. H., Ha, S. M., Lim, J., Kwon, S. & Chun, J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* **110**, 1281–1286 (2017).
42. van Dijk, L. R. et al. StrainGE: a toolkit to track and characterize low-abundance strains in complex microbial communities. *Genome Biol.* **23**, 74 (2022).
43. Stracy, M. et al. Minimizing treatment-induced emergence of antibiotic resistance in bacterial infections. *Science* **375**, 889–894 (2022).
44. Gibson, M. K. et al. Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat. Microbiol.* **1**, 16024 (2016).
45. Patel, V. C. et al. Rifaximin-alpha reduces gut-derived inflammation and mucin degradation in cirrhosis and encephalopathy: RIFSYS randomised controlled trial. *J. Hepatol.* **76**, 332–342 (2022).
46. Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. Development of the human infant intestinal microbiota. *PLoS Biol.* **5**, e177 (2007).
47. Wen, X. et al. Metabonomics reveals an alleviation of fitness cost in resistant E. coli competing against susceptible *E. coli* at sub-MIC doxycycline. *J. Hazard Mater.* **405**, 124215 (2021).
48. Li, J. et al. Antibiotic treatment drives the diversification of the human gut resistome. *Genom. Proteom. Bioinformatics* **17**, 39–51 (2019).

Acknowledgements

This work was supported by grants of the Basic Science Research Program (2021R1C1C1006336) and the Bio & Medical Technology Development Program (2021M3A9G8022959, RS-2024-00419699) of the Ministry of Science, ICT through the National Research Foundation; by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare (HR22C141105), South Korea; by the " National Institute of Health (NIH)" research project (2024-ER2108-00; 2024-ER0608-00), Korea; and also by a GIST Research Institute (GRI) GIST-MIT research Collaboration grant by the GIST in 2024. This work was also supported by the Korea Bio Data Station (K-BDS) with computing resources including technical support.

Author contributions

J.W.B., N.P., B.S., N.K., J.N., A.M., S.S., J.-I.K., J.W.S., A.K., and S.L. reviewed the manuscript and provided the critical comments. J.W.B. and S.L. wrote the manuscript. J.W.B. and S.L. performed the bioinformatics analysis. J.W.B. and S.L. conceived the project.

Competing interests

J.W.B., N.P., B.S., N.K., J.N., A.M., S.S., J.-I.K., J.W.S., A.K., and S.L. declared no financial competing interests. A.K. serves as an associate editor of this journal but had no involvement in the peer review or decision-making process for this manuscript. J.N. is a member of the advisory board of the journal, but he had no influence on the peer review process or decision to publish this paper.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41522-025-00705-x.

Correspondence and requests for materials should be addressed to Sunjae Lee.